Proceedings of the IRCS Workshop on Prosody in Natural Speech, August 5-12 1992

PROSODIC PHRASE AS A PROTOTYPE

Stephan Schuetze-Coburn

Department of Linguistics University of California, Los Angeles Los Angeles, CA 90024-1543

ABSTRACT

The linguistic unit 'prosodic phrase' has an underlying if not overt syntactic basis in many phonological and descriptive accounts of prosodic structure. On the other hand, phonetically oriented definitions are usually too limited or vague, so that they fail in the analysis of natural, connected speech. The basis for avoiding phonetic substance or for not providing adequate phonetic detail is the apparent lack of a clear set of invariant phonetic cues with which the category 'prosodic phrase' may be defined. It is suggested that while this may indeed be the case, there are alternatives to searching for criterial attributes. Viewing the category 'prosodic phrase' as a prototype is one way of shifting the perspective away from the expectation of necessary and sufficient conditions and towards a characterization of 'prosodic phrase' which more accurately reflects even the variation found in spontaneous speech. Properties of prototypes in linguistic theory are examined, and the implications of considering a prosodic phrase category as a prototype are explored in the context of a German conversational narrative which has been analyzed auditorily into 'intonation units'.

1. INTRODUCTION*

In this paper, I will be continuing the discussion of the prosody of 'intonation units' that opened the workshop with the presentations of Du Bois and Chafe. Initially, though, I will be framing the points that I have to make in general terms, rather than in terms which are relevant only to intonation units, so I prefer to call the units that I will be talking about in the first two sections of my paper 'prosodic phrases'. I'd like to think of 'prosodic phrase' as a cover term for the various phrase-length prosodic units that are commonly found both in discourse and phonologically oriented studies. Examples include the 'tone group' of Halliday (1963, 1967a, 1985), the 'tone unit' of Crystal &

Quirk (1964) and Crystal (1969, 1975), and Brazil (1975, 1978, 1985), the 'intonation group' of Cruttenden (1970, 1986) and Fox (1984), the 'intonation phrase' of Pierrehumbert (1980), the 'intonational phrase' of Selkirk (1981) and Nespor & Vogel (1983), the ' γ -frame' of Gibbon (1984), the 'major phrase' of Ladd (1986), and many other loosely related expressions (for example, the units of Pike 1945, Trager & Smith 1951, and of other earlier works such as Palmer 1924). By grouping these units together, I don't want to imply that any pair of units is actually equivalent.¹ But there are obvious similarities, and I do think that they all are trying to get at a certain kind of prosodic organization—what can be called the basic phrasing of utterances.

My focusing exclusively on this structural level shouldn't be taken to mean that I believe other possible levels of prosodic organization aren't worth investigating-either in a phonological hierarchy, where especially smaller units are of interest (e.g. word, clitic group), or in a prosodic account of discourse structure (where the larger units, e.g. major paratone, pitch sequence, and so forth, find their place). But I do believe that prosodic phrases form a particularly interesting level of organization for a variety of approaches to the structure of speech, especially that which is natural, connected, and spontaneous. Units at this level have been claimed to function as a domain, for instance, for the information structure of discourse (e.g. Halliday 1967b; Kreckel 1981; Chafe 1987), and for speech production (e.g. Laver 1970; Svartvik 1982), as well as for various intonational features (e.g. declination, Pierrehumbert 1980) and phonological rules (e.g /t/-flapping in English, Nespor & Vogel 1982). They have also been shown to have interactional significance by contributing to turn organization (Oreström 1983; Ford & Thompson 1992).

1.1. Prosodic phrase and data type

My point of departure for examining the nature of the category 'prosodic phrase' is a methodological one. How do linguists study intonation, and how does this bear upon the definition of prosodic units? While it is seldom acknowledged, when we construct phonological models or perform

^{*} I wish to thank Jack Du Bois for helpful suggestions and the members of the workshop for their comments. Special thanks to Mark Liberman and Cindie McLemore for the use of the Linguistics Dept. Phonetics Lab at the University of Pennsylvania. I am especially grateful to Felicia Hurewitz for digitizing and pitch tracking the conversational excerpt discussed in this paper. The conversation was recorded in 1988 in Bielefeld, Germany, during a stay funded by the *Deutscher Akademischer Austauschdienst.* I also thank Marian Flaherty, Hartmut Kreft, Marlene Marlow, and Silvia Rode for their advice and help with the transcription; all responsibility for the 'final' stage is of course mine.

¹ Many researchers have assumed (or asserted) equivalence. As each of the above-mentioned units is defined differently (and serves various purposes), I prefer to be more cautious on this point (cf. Couper-Kuhlen 1986: 76).

perceptual experiments, the extent of a domain such as 'prosodic phrase' can be easily manipulated ('extent' as measured in terms of some linguistic construct). Example data that are to be accounted for in a model and stimulus materials in an experiment are not independently provided, but are of course selected specifically for their suitability in realizing the stated goals. But because the linguistic structures that these kinds of analyses are based on are fundamentally syntactic in origin, not prosodic, the syntactic structure ends up influencing-or even determining-the prosodic domain as well.² Typically, what one finds is a single grammatical phrase, simple sentence, or sentence pair with a 'normal' (or special) prosody projected onto it, rather than a natural prosodic structure, with the syntax only secondary or incidental. Another way to think about this state of affairs is to consider the number of intonational accounts available in which the syntactic structure is assumed and a prosody is subsequently assigned, versus those where the prosodic structure is given and a syntactic form is then selected (or alternately, accompanies the prosody). There is good reason to view such data with mistrust.³

Momentary reflection on data type is important because in working with extensive amounts of continuous, spontaneous speech, it becomes apparent that the extent of the domain 'prosodic phrase' does not correlate very well with how it is usually illustrated, but rather it exhibits much greater (unpredictable?) variation. Broadly speaking, each instantiation of a prosodic phrase is the product of a set of interlocutors—a speaker and the hearer(s)—in a specific context, so its scope emerges from the interplay of an array of factors. In other words, the extent of a prosodic domain is not given in advance in terms of only syntactic (or other such) constraints imposed by the researcher's uniform idea of what the domain should be; instead, it varies as required for the relevant communicative purposes.⁴ The point here is that the effect of nonlinguistic parameters on the shape of prosodic phrases has been-it can be said-subverted entirely, and the influence of nonsyntactic ones has been largely ignored in linguistic description.

Crucially, what is evident in the first place is that the challenge of segmenting speech into prosodic phrases is frequently not met, but is neatly finessed by placing certain nonprosodic restrictions on the data. The homogeneous, sentence-oriented language that is often encountered in theoretical and descriptive accounts of intonation does not reflect the variable conditions and pressures that exist in natural language use: speakers and hearers regularly create phrasing that transcends the usual paradigms. Yet, as many of us will agree, if any concept of prosodic phrase is to be truly viable as a linguistic category, it must be possible to specify how all phrases are to be reliably identified in a prosodic analysis of natural, in addition to idealized, speech.

1.2. Defining prosodic phrase

A second orientation point is to consider previous attempts at defining the extent of prosodic phrases. The myriad units encountered in the literature provide one indication that there is little agreement on how prosodic phrasing is to be achieved, either theoretically or in the context of the analysis of particular language data. Many characterizations of prosodic phrase—whether phonologically or descriptively oriented-involve the identification of some unifying pitch sequence, at least indirectly. Two examples: Pierrehumbert (1980) considers the primary prosodic configuration of her phrasal unit to be the 'tune', i.e. a series of pitch accents followed by a phrase accent and a boundary tone. For Altenberg (1987), the defining configuration is 'a coherent intonation contour optionally bounded by a pause and containing (among other things) a salient pitch movement (the nucleus), normally at the end of the unit' (p. 47). While the conceptual notions contained in these characterizations might be clear enough, delimiting all actual instances of intonational coherence in an extended stretch of connected speech—and not just picking out the clear cases, or constructing their equivalent-proves to be a somewhat elusive task. It would seem that by specifying prosodic phrases largely in phrase-internal terms, one is able to cover only a subset of a given text. Yet I take it that a goal of an adequate prosodic description is to account for all utterances. Without a doubt, many-for some texts perhaps even a majority-of contours can be identified relatively unproblematically, either on a perceptual basis or with the help of instrumentation. But some phrases invariably defy identification using the definitive criteria that a focus on the notion 'intonation contour' demands.⁵ Unfortunately,

 $^{^2}$ In some cases, the assumption is that syntactic structure is prior, since it serves as the input to a phonological component.

³ Cf. Gibbon's (1988) discussion of data types and his skepticism of accounts based on isolated, invented items for ad hoc illustration (which he classifies as 'anecdotal').

⁴ To be fair, some phonologists have recognized the inherent variability in basic phrasing (in the sense that they have consider it 'free'), e.g. Selkirk (1981: 130). Still, prosodic boundaries are held to align with syntactic boundaries.

⁵ Most phrase-sized units do seem to have an identifiable coherent contour, but not all do. (Of course, whether the presence of a 'coherent contour' alone, i.e. without any concomitant features of the type mentioned below, is enough for a prosodic phrase to be perceived is an open question.) Three main classes of phrases regularly lack a contour in the usual sense:

⁽a) Uncompleted phrases. These are considered separate units whether a contour is present or not; i.e., they are not treated as 'residue' which can be incorporated into some other phrase, or ignored. They may exhibit some relative structure despite their lack of a contour. This is most clearly illustrated in sequences of uncompleted phrases.

attending only to highly prominent aspects of the prosody—i.e. the pitch accents—does not give adequate indication of where the boundary between phrases is to be drawn.

Faced with the difficulty of unambiguously distinguishing the component parts of 'coherent intonation contours' (e.g. boundary tones or nuclear accents), it is tempting to abandon the standard line of defining prosodic phrase altogether and to adopt instead a more reliable and 'objective' measure. Brown, Currie & Kenworthy (1980) took such an approach when they rejected a well-known phrasal unit based on nuclear tone in favor of a pause-based unit. But the simplicity of this kind of 'definition' of prosodic phrase is plainly spurious. Although pauses in both spontaneous and read speech are readily measurable (in relative terms), they do not all reflect a common origin (cf. Chafe 1980; Deese 1980; Goodwin 1981), so that strictly pause-based units are not necessarily meaningful when it comes to characterizing prosodic structure as a whole. More importantly, we should stop to consider whether there is any a priori reason to believe whether a single parameter, be it pause or pitch accent or whatever, might unambiguously identify coherent phrasing, either in principle or in practice. Given the complexity of the phenomenon, surely the answer must be no. While getting a firm grip on prosodic phenomena is notoriously difficult, concerns about how to make a category operational should not force us to abandon theoretical (and empirical) substance.

We might ask then: Why not make use of a wide range of prosodic features in defining prosodic phrases? Now, researchers have long noted that factors in addition to specific pitch patterns correlate with phrase boundaries to varying degrees. Prosodic features like silent and 'filled' pauses and other such vocalizations, 'anacrusis' and 'final' lengthening (and other features tied to the 'timing' of the speech), and certain voice quality features are often used to facilitate phrase boundary identification, even if their exact status vis-à-vis these boundaries remains unclear. (Other features such as local variations in pitch width and intensity, as well as largely segmental features like aspiration, also suffer from similar limitations.) Crystal (1969: 205), for instance, states: 'In fact, any process of intonation analysis will take simultaneous account of both boundary cues and

(b) Nonlinguistic vocalizations. A variety of audible vocal sounds are treated as separate phrases when they are not perceived as part of a larger phrase gestalt, including laughter, inhalation, and coughing.

(c) Short responses or backchannel utterances. Especially when low pitched, these frequently have no clearly identifiable primary prominence or contour, but are nevertheless perceived as phrases. These prosodic phrases need not contain any prominent ('accented') syllables at all.

The extent to which these 'deviant' phrases can be ignored depends on the assumptions and goals of one's approach. internal structures ... and any comprehensive definition of the tone-unit must also have recourse to a complementarity of cues'. He then gives pitch reset and pause as the two primary criteria he uses. Yet, while these and other 'boundary cues' undoubtedly exist to some degree, they have been viewed by many as relatively marginal.

There are several reasons for this lack of interest. One difficulty in making serious use of such correlates is the 'optionality' of all of the above-mentioned phonetic features. In classical definitions, optional features cannot be defining. Thus, it would seem that the main problem to overcome in determining the phrasing of connected speech-which is the focal point here and one of the topics of the workshop-is the lack of a single, ever-present identifiable cue (or invariant set of cues) in the acoustic signal (or the perception thereof), either at phrasal boundaries or in conjunction with intonation contours. Given that received approaches have not been genuinely successful, it may prove fruitful to look at the problem of delimiting prosodic phrases from a fresh vantage point. Instead of searching for the 'correct' invariant components which could be forged into an viable definition, it is perhaps worth considering whether the customary way of defining linguistic units is suitable in this instance. Specifically, while necessary and sufficient criteria may seem adequate or appropriate for defining linguistic categories when some forms of language are examined (e.g. decontextualized language or spoken language which originates from written form, as in reading aloud or in text-to-speech systems), it may, however, also be the case that adherence to a rigid definition of prosodic phrase will never capture the variability which is an integral part of spontaneous discourse.

2. PROTOTYPES AND PROSODIC PHRASE

In recent years, the notions of 'fuzzy' categories and 'prototypes' have been exploited to account for an impressive range of linguistic data (Rosch 1978; Lakoff 1987; Rudzka-Ostyn 1988; Tsohatzidis 1990), including phonological categories (Jaeger 1986; Nathan 1986; Taylor 1989). I would like to suggest that the best way to treat the category 'prosodic phrase' is as a prototype along the lines of these previous studies. In doing so, I believe we come closer to balancing our desire to formulate explicit models of prosody with the practical concerns that arise in dealing with natural speech.⁶ The general thrust of this proposal is, of course, not entirely new. Precursors to this idea include Chafe (1987), who describes intonation units using a schematic 'general format'. In subsequent work, Chafe

⁶ While it is also desirable that the psychological validity of this type of model be demonstrated for this category, such a claim must await later experimental confirmation; here, the precise cognitive representation is not an issue. I wish only to examine the plausibility of this type of model with regard to the characteristics of prosodic phrase.

(1992) calls this format 'the structure of a prototypical intonation unit'. Schuetze-Coburn, Shapley & Weber (1991) also discuss intonation units in terms of deviation from an (abstract) prototype, but there has been to date no detailed treatment of prosodic phrase in terms of a full range of characteristics attributed to prototype models.

2.1. Characteristics of prototype models

In order to evaluate how well a prototype model might define 'prosodic phrase', it is instructive to look at some primary characteristics of prototype categorization and consider their general relevance to the category. Geeraerts (1989) compiles a set of four properties which he says are typical of prototypicality. These include the notions of 'criterial attributes', 'family resemblance', 'centrality', and 'gradience'. Lakoff (1987) covers additional properties, including 'embodiment', and 'basic-level categorization', which I examine below as well.

The notion 'criterial attributes' has to do with the requirement of a 'checklist' of features, each of which must obtain for a definition to apply (cf. Fillmore 1975). The lack of such attributes defining a category is a salient feature of prototypicality. This notion is, of course, the catalyst to the present discussion, and so plays an obvious role here. The optional status of EVERY phonetic feature in a characterization of prosodic phrase undermines a normal definition in terms of necessary and sufficient conditions. If no feature is criterial, then no obvious distinction between essential and incidental features can be established. But in a definition constructed around a prototype, this property would pose no problem; each feature may serve as an attribute in phrase boundary production and perception. However, this is not to say that all features carry equal weight; some may seem more important than others, which the model captures through the property 'centrality' (see below).

'Family resemblance' refers to the idea that exemplars of a category may share individual category features with just a subset of members, yet the subsets of a category overlap so that there is enough similarity among members for each to be included in the category (cf. Wittgenstein 1953). Given the lack of criterial attributes mentioned above, this property is directly relevant to 'prosodic phrase' in that there is no feature which all instances of the category share. When a range of prosodic (and not just strictly intonational) features is taken into consideration, it is apparent that individual prosodic phrases cannot be structurally identical, but instead must resemble one another to greater or lesser degrees, depending on the features realized in any given phrase.

'Centrality' can be summarized in a maxim 'All phrases are not created equal'. That is, there are core and peripheral members of a category, with core members being more 'salient'. Alternately, centrality can be measured in terms of the frequency which member characteristics occur, with predominating features being salient. Centrality is thus a gradient notion concerning the relative degree of membership in a category or the importance of individual features in characterizing membership.⁷ As applied to prosodic phrase, if it can be shown that phrases which are clearly instances of the category make better exemplars than others, or if the balance of features differs from phrase to phrase, then the prosodic phrase category would exhibit this prototype property.

'Gradience' refers to the idea that category boundaries are indeterminate, or 'blurred at the edges' (Geeraerts 1989: 593).⁸ If we could point to peripheral cases where it could not be decided whether a given token was an instance of a (specified) category or not, then we could say that the category has no inherent clear-cut boundaries. This is one property which seems to be inapplicable to a prosodic phrase category. In order to speak of a category which contains actual prosodic phrase tokens as members, it will be necessary to draw boundaries to delineate one phrase from another. The physical (and psychological) integrity of prosodic phrases must be established in this way, unlike that of concrete objects, like 'cup', which I presume are more obviously independent entities. Note that the lack of (membership) gradience does not preclude gradience in terms of centrality, but it must be possible to decide (in principle at least) whether or not a phrase has been produced.9

Additional characteristics of prototype models have been summarized by Lakoff (1987). Two apparently relevant properties which provide necessary epistemological links in a cognitive model are the notions 'embodiment' and 'basiclevel categorization. As these notions have less to do with the structure of a category, remarks here are only meant to be suggestive. Lakoff distinguishes two types of embodiment. The first, 'category (or conceptual) embodiment', refers to the biological and experiential grounding of categories, i.e. to their fundamentally nonautonomous nature. Regarding prosodic phrase, this grounding would mean that such a category could not be an independent linguistic (phonological) construct, but would have clear cognitive and social bases. I view the establishment of a prosodic phrase category founded in phonetic substance as a first step in the confirmation of this kind of grounding. In this regard it is interesting to note that some phonetic features used to identify prosodic phrases have been claimed to be 'language independent': they have been found to occur in a range of languages, so that there is reason to suspect that

⁷ Lakoff (1987) differentiates two aspects of centrality: 'centrality' proper, and 'centrality gradience'.

⁸ For Lakoff (1987), this property is 'membership gradience', to be distinguished from 'centrality gradience'.

⁹ The question of 'intonational sandhi', in which two (abstract) prosodic phrases are intonationally merged, is an interesting one. On the surface, though, there is but one phrase.

more than structural factors are involved. Pause, F_0 declination, F_0 reset, diminution of F_0 range, prosodic lengthening, and intensity decrease are all such features (Vaissière 1983). Yet, while the inventory of phonetic features may be similar, their relative importance probably differs from language to language, so that a certain arbitrary (i.e. linguistic) component remains.

The second type of embodiment, 'functional embodiment', refers to the psychological status of the category. If a category can be shown to be employed by language users automatically, without conscious effort, then the category exhibits this prototype property. The observation that prosodic phrases are highly relevant to interactional behavior, without speakers and hearers making direct reference to them, points in this direction.

'Basic-level categorization' has to do with the hierarchical organization of categories; as the name implies, basic-level categories are claimed to be cognitively basic. That phraselevel units do not constitute simply one level in a prosodic hierarchy is evidenced indirectly by their falling near the middle of the hierarchy, rather than being the top or bottom level. On the other hand, very suggestive evidence can be found in the area of language acquisition. From my own preliminary observations, it is apparent that in the course of development from a one to two-word stage, prosodic properties of individual words-which constitute complete phrases (cf. Menn 1983)—are concatenated along with the lexical material. That is, the integrity of the prosodic phrase is at first respected, reflecting its fundamental nature. Only later, as language use becomes more sophisticated, are the prosodies melded into a single 'coherent contour', with standard accompanying phonetic cues.

C: ...(1.0) </dads/> ((looking))(1)1 2 ... </d3u/d3u/> tòp 3 ... òn 4 ...(.7) ((taking top off)) tòp </aba/> 5 .. òff 6 ... top .. òff 7 P: ... top's òff 8 .. mhm

An instance of this process is given in example (1).¹⁰ Each line of (1) constitutes a prosodic phrase, as indicated by the intonation contour gestalt, the distribution of pitch prominences, and pausing. The phrases of interest here are 4, 5, and 6, which contain two instances of the same 'topic + predication' structure *top off* by speaker C (age 25 months). Phrases 4 and 5 contrast with phrase 6. In the

first instance, the topic + predication is distributed over two prosodic phrases, i.e. each part has its own intonation contour. This reflects the earlier one-word stage of speech production (cf. the similar pattern in phrases 2-3), even though the syntactic and semantic structure is now more complex. In the second instance, the topic + predication is contained in one prosodic phrase. Here, the prosodic structures of the two words have been integrated, mirroring the syntax and semantics of the construction.

In sum, on a preliminary assessment, the category 'prosodic phrase' appears compatible with a prototype model. Three of the four properties of prototypes discussed by Geeraerts (1989) seem applicable; two other nonstructural characteristics outlined by Lakoff (1987) seem to apply as well. These considerations, I believe, are promising in the development of a model of prosodic phrase that is anchored in phonetic substance.

2.2. Notions of prototype

At this point it will help clarify matters to mention that there are two general orientations when describing a prototype, as Cruse (1990) points out. Under one perspective, the relations between the members of a given category are in view: the focus is on the 'prototypical exemplar' (either as an idealization or as the 'most representative' member), against which the other (actual and potential) members of the category may be evaluated. In other words, the prototype serves as a cognitive reference point for the categorization of nonprototypical tokens. This is the notion of prototype advanced in Schuetze-Coburn et al. (1991) and Chafe (1992). Under the other perspective, the category as a whole is in view: the focus is instead on the prototypical characteristics of the category, and the properties which serve to define it in an intensional sense.¹¹ It is this latter perspective that is of primary interest here, as the ultimate goal is to advance a phonetically based characterization of the category prosodic phrase in terms of prototypical features. Regarding the applicable prototype properties discussed above, the phonetic features that will satisfy this goal will (a) be realized unpredictably (lack of criterial attributes), (b) cluster into reoccurring subsets (family resemblance), and (c) be associated with varying degrees of importance (centrality).

Numerous phonetic attributes occur with sufficient regularity, especially at the edges of phrases, so that they can be considered prototypical features of the prosodic phrase. These features include silent pauses (the absence of vocalization or 'offtime'), which occur BETWEEN phrases; pitch reset and accelerated speech at the BEGINNING of a phrase; lengthened speech and laryngealization or other low pitch phenomena at the END of a phrase; and various vocaliza-

¹⁰ Transcription conventions: three dots indicate a silent pause (.3-.6 second); two dots, a shorter pause; durations for longer pauses are given in parentheses (). Utterances with no obvious lexical correspondence are enclosed in </ />>. Prominence is indicated with a grave accent. Transcriber comments are given in double parentheses (()).

¹¹ Cf. Geeraerts' (1989) feature analysis of prototype characteristics.

tions which generally indicate some sort of hesitation ('filled pauses'), such as *uhm*, usually occurring at the beginning of a phrase, but occasionally constituting a separate, delimiting phrase. Other features, such as overall intensity peak timing, intensity diminution, and voice quality modulation may have a scope which extends over the course of a phrase.¹² These phonetic features all play a role in addition to the phrase-internal feature 'coherent intonation contour', as manifested by a particular configuration of accents. All these cues have been discussed individually in some detail at one time or another by various authors. How they each make a contribution in the production and perception of prosodic phrases in connected speech is the idea worth exploring further.

3. INTONATION UNITS AND THE PROTOTYPE MODEL

I wish now to give a brief indication of the descriptive potential of a prototype-based model of prosodic phrase by examining its flexibility in a concrete application. A short (4' 20") exchange by two speakers of colloquial Standard German was selected from a longer recording of a spontaneous conversation. The text was transcribed and the utterances were divided into speaker turns. Subsequently, turn units were segmented into the prosodic phrases called 'intonation units' using the system of auditory analysis presented and outlined in Du Bois, Schuetze-Coburn, Cumming, and Paolino (1991; 1992); cf. also Chafe (this volume). The phonetic basis of intonation units makes this unit an appropriate selection as a representative of prosodic phrases as a whole.

3.1. Prosodic cues to intonation units

During the segmentation process, various phonetic cues are taken into consideration, primarily those mentioned above. In the transcription, systematic attention is devoted to four prosodic cues—silent pause, accelerated speech, lengthened speech, and laryngealization—which constitute the most important phonetic features for phrasing that we feel can be auditorily noted with adequate reliability.¹³ In addition, some use is made of pitch reset (here, a marked shift in pitch, generally on a nonprominent syllable, at the beginning of an intonation unit). Where marked, it is derived from an inspection of the pitch tracks of the excerpt, together with a comparison of the pitch periods before and after the unit boundary by measuring the frequency of the glottal pulse from the digitized waveform. The gestalt perception of a 'coherent intonation contour' itself is difficult to quantify directly, but empirical evidence suggests that this is not necessary. In other words, while not all of the prosodic features that may be present in the signal are individually attended to in the transcription system, the segmentation process IS sensitive to factors which contribute to the perception of a phrasal-level contour gestalt, and cues that are important for the segmentation of each phrase are noted.¹⁴

3.2. Prosodic cue patterns

The transcribed excerpt contains 269 intonation units. Of these, 113 were not included in the tabulations presented below: 30 constitute solely 'nonlinguistic' vocalizations (primarily laughter); 83 include backchannel utterances, overlapping turns, and turn-initial intonation units—in short, all units that were already delimited by turn boundaries. The remaining 156 turn-internal intonation units were examined for the presence of the four main prosodic cues listed above; the results are given in Table 1. From the table, it it clear that the frequency of occurrence for the tabulated cues varies greatly in the selected excerpt. While about two-thirds of the intonation units are preceded by silent pauses, only one in ten is bounded by laryngealized speech.

| Silent | Acceler'd | Length'd | Laryng'd |
|--------|-----------|----------|----------|
| Pause | speech | speech | speech |
| 105 | 74 | 34 | 15 |
| (68%) | (47%) | (22%) | (10%) |

Table 1. Occurrence of four prosodic cues in a conversational excerpt.

With regard to their place in a prosodic phrase model, it is noted that each feature exhibits the expected prototypical properties. While silent pauses of various lengths occur with some frequency, not every pair of intonation units is separated by a pause. Furthermore, phrase-internal pauses, though less common, also occur. Concerning accelerated and lengthened speech, while there is a tendency for the rate of speech to decelerate through the course of an intonation unit, tempo does not always vary in this way, and patterns are complex. That is, although stretches of accelerated speech often occur at the beginning of an intonation unit, rather than elsewhere, acceleration does not only correspond to unit beginnings: some units are perceived as consisting entirely of accelerated speech. Similarly, while segment and syllable lengthening does occur at the end of phrases, this cue is not limited to this position, and marked lengthening

¹² This list is not exhaustive, of course; one could point to additional regularly occurring phonetic cues as well.

¹³ Pause durations, however, were checked instrumentally in conjunction with another study (Schuetze-Coburn, in progress); Estimated pauses proved to be valid; 96% of the transcribed pauses varied no more than \pm .1 second from their acoustically measured counterparts (the claimed accuracy of the auditory judgements).

¹⁴ Of course, other aspects of prosodic notation are not included here either; detailed representation of the pitch accents, for example, is left for other systems.

is much less common overall. Finally, while vocalization sometimes becomes laryngealized at the end of a phrase, or glottal constrictions occur phrase initially (or both the end of one phrase and the beginning of another are so marked), this feature also occurs within phrases, and its occurrence is fairly limited.

For the purposes of evaluating a prototype model, an aspect of feature occurrence more interesting than the frequency of individual features is the way which features pattern with each other. Combinations of prosodic cues for the tabulated intonation units are given in Table 2. In the table, feature combinations are read horizontally, with a minus sign indicating the absence of a feature, and a plus sign, its presence. Thus, the top row gives the number of cases where none of the four prosodic cues is present, i.e. [-PAUSE, -ACCELERATION, -LENGTHENING, -LARYNGEALIZA-TION], of which there are 12. That is, (reading across the table) out of the 51 intonation units which lack a preceding silent pause, 22 lack in addition initial accelerated speech; out of these 22 cases, 13 lack final lengthening; and out of these 13 cases, 12 have no laryngealization. In the second row, the number of cases where laryngealization is the only cue are given, i.e. [-PAUSE, -ACCELERATION, -LENGTHENING, +LARYNGEALIZATION], of which there is 1. And so on until the last row, which gives the number of cases where all four cues are present (i.e. 1 case).

| Sile Pau | ent ise | Accel spec | ler'd xh | Leng spee | th'd ch | Laryi spee | ng'd xch | |
|-------------|------------|---------------|-------------|--------------|------------|---------------|-------------|-----------------------|
| 51 | - | 22 | - | 13 | - | 12 | - | ⇐ No cues |
| | | | | | | 1 | + | |
| | | | | 9 | + | 8 | - | • |
| | | | | | | 1 | + | |
| | | 29 | + | 21 | - | 20 | - | (13%) |
| | | | | | | 1 | + | . , |
| | | | | 8 | + | 8 | - | |
| | | | | | | 0 | + | |
| 105 | + | 60 | - | 49 | - | 43 | _ | (28%) |
| | | | | | | 6 | + | |
| | | | | 11 | + | 10 | - | |
| | | | | | | 1 | + | |
| | | 45 | + | 39 | - | 36 | - | (23%) |
| | | | | | | 4 | + | |
| | | | | 6 | + | 5 | - | |
| | | | | | | 1 | + | \Leftarrow All cues |

Table 2. Cooccurrence patterns for four prosodic cues in a conversational excerpt.

Feature combinations are read off horizontally; the top row is all cues absent; the bottom, all present. Percentages for combinations occurring in over 10% of tabulated intonation units are given to the right. Certain configurations of cues are clearly much more common than others. Pause alone turns out to be the most frequent pattern, found in 28% of the tabulated intonation units. Pause plus accelerated speech is found in an additional 23%, and accelerated speech alone marks 13%. Together, these three feature configurations cover almost twothirds of the tabulated units. Perhaps somewhat surprisingly, given the past emphasis on THE prototype for all prosodic phrases, it is also evident that the presence of all four features in a unit is a rare event (1%), just as it is relatively uncommon for none of these features to be present (8%). Instead, fully 85% of the tabulated units exhibit 1 or 2 prosodic cues. Thus, a TYPICAL intonation unit has at least one prosodic feature, but its boundaries are not marked to an extreme degree.

The best way to illustrate the variation in feature combinations which cue intonation units is to present an excerpt from the transcription and discuss the combination present in each instance. Example (2) is a stretch of nine intonation units by speaker A, who is talking about a friend planning to write funeral marches for rich Americans. (As before, each line in the transcription represents a separate intonation unit.)¹⁵

- (2) 151 A: ~%also er würde er so~/+ so he would he like
 - 152 (H) «.3» ^ vòrher so 'n hàlbes Jàh=r /+ ◊ before like a half year
 - 153 auf so (%) <% uh %> /-on like uh
 - 154 ^ bei dem lèb'm müss'n /+ by him live must
 - 155 .. (H) «.4» ^ ~d'mit er sie richtig~ kènn'nlernt /+ so.that he them correctly gets.to.know
 - 156 ^ bevor die Leute sterb'n /+ before the people die
 - 157 ^ das wéiß [ja man] ja mèi=st'ns /+ that knows yes one yes mostly
 - 158 B: [(CLEARS THROAT)]

¹⁵ Additional transcription conventions: accelerated speech is bracketed by tildes, lengthened segments are followed by an equal sign; laryngealization is indicated with percent signs. Pitch reset is indicated by a raised caret; 'focus' accent, by an acute accent mark. Inhalation is represented by (H), with its duration following in double angled brackets. Intonation units are marked for transitional continuity class as '+' (continuing) or '•' (final). Perceived terminal pitch direction is indicated with '\' (fall) or '/' (nonfall). Uncompleted units end in double hyphen '--'. Questionable unit boundaries are indicated with '\$\Overlapping speech is enclosed in square brackets. Elision of segments is indicated by a "''. See Du Bois et al. (1991; 1992) for more details.

| 159 | A: | bevor man so stirbt /+ before one like dies |
|-----|----|--|
| | | |

- 160 .. ~jedenfalls~ bèi% ◊ .. Lèuten die so länger krànk [sind \ nè] /+ in.any.case by people that like longer ill are ok
- 161 B: [jà] \+ yes

(A: 'So he would like / half a year before / on like uh- / have to live at his [house] / so that he really gets to know them / before people die / you usually know'/ B: (CLEARS THROAT)/ A: 'before one dies / at least with people who have been sick a long time' / B: 'yeah' /)

In this excerpt, there are eight intonation unit transition points that need to be discussed with respect to the prosodic cues signaling the phrase boundaries. Point 1 is the boundary between units 151 and 152. Here there is a short (.3 second) break in vocalization in conjunction with the inbreath between the units. Such a break is comparable in its timing to a silent pause; however, after an inhalation, there is a very strong tendency to reset one's pitch, and pitch reset is an important feature for cuing a new intonation unit boundary.¹⁶ In this case, there is an obvious shift in pitch on *vorher* (of about 60 Hz, see Figure 1 @ time 149.25).¹⁷ Point 2 is the boundary between units 152 and 153. One feature occurs here, lengthening on the final syllable of unit 152. While usually a fairly robust cue, the boundary in this instance is perceived to be rather weak, which is indicated by the diamond at the end of the line.

Point 3 is the boundary between units 153 and 154. Glottal constriction and creaky voice through uh is found at the end of unit 153 (which is perceived as uncompleted), followed by pitch reset on *bei* (not visible in Figure 1 @ time 150.75, as no pitch values registered for the preceding low pitched segments). Point 4 is the boundary between the next pair of units, 154 and 155. Here there is slight pause and then again an inbreath with a following pitch reset (of about 150 Hz, in this case a shift down, see Figure 1 @ time 152). In addition, there is an initial stretch of accelerated speech in unit 155.

Point 5 is the boundary between units 155 and 156. Of the prosodic features considered here, only pitch reset (of about 60 Hz, see Figure 1 @ time 153.5) is evident. Point 6 lies between units 156 and 157. Once again, the only cue

present is pitch reset (of about 80 Hz, see Figure 1 @ time 154.5).

Point 7 is the boundary between units 157 and 159. Here there is both lengthening at the end of unit 157 and a slight pause between units. Point 8 is the boundary between units 159 and 160. There is a slight pause between units and accelerated speech at the beginning of unit 160 (one of the very common feature combinations). Note that in unit 160, glottal constriction occurs after *bei*, and a slight pause is perceived. An intonation unit boundary was considered at this point (indicated by the diamond in the line), but was ultimately rejected. (This situation contrasts with boundary point 2 above, where a boundary was indeed marked.)

4. SUMMARY

In the first section of my paper, I discussed some of the shortcomings of the usual approaches to a (general) linguistic unit 'prosodic phrase'. In the second section, I suggested that this category be treated as a prototype category, and I outlined typical characteristics that such a model might exhibit. In the third section, I examined the distribution of four prosodic features in an excerpt from a spontaneous German conversation which had been segmented into 'intonation units' in order to illustrate that these features meet expectations placed on prototypical characterizations of the category. It was found that an intonation unit will exhibit a variety of cues, yet rarely, if ever, do all phonetic features actually occur in any given instance. Nevertheless, it is the case that a conjuncture of cues is usually identifiable before a prosodic phrase is perceived. While a set of features may cooccur at particular points in a utteranceclearly identifying a prosodic boundary-it is also the case that individual features are found elsewhere, as illustrated above. That is, prosodic boundaries are manifested more or less strongly, depending in part on how many features are present, but the presence of a feature apparently does not guarantee a boundary. One problem, then, is that the feature threshold used to determine whether a boundary is identified-or is not selected-could (with our current understanding) be arbitrarily set by the researcher. The larger open question with phonetic cues is thus not their inventory, but their relative weight (e.g. pitch reset is arguably more central than laryngealization), and the interpretation of their interaction in influencing the perception of prosodic phrase boundaries. These are empirical questions which can be answered, but it will require a close analysis of a large amount of connected, preferably spontaneous, natural speech.

¹⁶ Schuetze-Coburn, Shapley & Weber (1991) found in a corpus of American English conversation that a new intonation unit was perceived every time the speaker's pitch was fully reset (but that this reset occurred on average only once every other intonation unit).

¹⁷ Octave errors in the pitch tracking have been adjusted manually.



Figure 1. F_0 tracings corresponding to example (2). Scale is Hz as a function of time in seconds.

REFERENCES

- Altenberg, B. 1987. Prosodic Patterns in Spoken English. Lund: Lund University Press.
- Brazil, D. 1975. Discourse Intonation. Birmingham Univ.

— 1978. Discourse Intonation II. Birmingham Univ.
— 1985. The Communicative Value of Intonation in

- English. Birmingham Univ. Brown, G.; K. L. Currie & J. Kenworthy. 1980. Questions of Intonation. London: Croom Helm.
- Butterworth, B. 1983 (ed.) Language Production II. London: Academic Press.
- Chafe, W. L. 1980. Some reasons for hesitating. In Dechert & Raupach (eds.), 169-80.
- 1987. Cognitive constraints on information flow. In Tomlin (ed.), 21-51.
- ----- 1992. Discourse, Consciousness, and Time. Ms., UCSB.
- This volume. Prosody in English and Seneca natural discourse.
- Couper-Kuhlen, E. 1986. An Introduction to English Prosody. Tübingen: Max Niemeyer.
- Cruse, D. A. 1990. Prototype theory and lexical semantics. In Tsohatzidis (ed.), 382-402.
- Cruttenden, A. 1970. On the so-called grammatical function of intonation. *Phonetica* 21.182-92.
- ----- 1986. Intonation. Cambridge: Cambridge Univ. Press.

Crystal, D. 1969. Prosodic Systems and Intonation in English. Cambridge: Cambridge University Press.

Crystal, D. & R. Quirk. 1964. Systems of Prosodic and Paralinguistic Features in English. The Hague: Mouton.

- Cutler, A. & D. R. Ladd. 1983 (eds.) Prosody: Models and Measurements. Berlin: Springer-Verlag.
- Dechert, H. W. & M. Raupach. 1980 (eds.) Temporal variables in speech. The Hague: Mouton.
- Deese, J. 1980. Pauses, prosody, and the demands of production in language. In Dechert & Raupach (eds.), 69-84.
- Du Bois, J. W.; S. Schuetze-Coburn; D. Paolino & S. Cumming. 1991. Discourse Transcription. Ms., UCSB.
- Edwards, J. A. & M. D. Lampert. 1992 (eds.) Talking Data. Hillsdale, NJ: Lawrence Erlbaum.
- Enkvist, N. E. 1982 (ed.) Impromptu Speech: A symposium. Abo: Abo Akademi.
- Fillmore, C. J. 1975. An alternative to checklist theories of meaning. BLS 1, 123-31.
- Ford, C. E. & S. A. Thompson. 1992. Projectability in conversation. Ms., UCSB.
- Fox, A. 1984. German Intonation. Oxford: Clarendon Press.

Fretheim, T. 1981 (ed.) Nordic Prosody II. Trondheim: Tapir.

Geeraerts, D. 1989. Introduction: Prospects and problems of prototype theory. *Linguistics* 27.587-612.

Gibbon, D. 1984. Intonation as an adaptive process. In Gibbon & Richter (eds.), 165-192.

- 1988. Intonation and discourse. In Petöfi (ed.), 3-25.

- Gibbon, D. & H. Richter. 1984 (eds.) Intonation, Accent and Rhythm. Berlin: de Gruyter.
- Goodwin, C. 1981. Conversational Organization. New York: Academic Press.
- Halliday, M. A. K. 1963. The tones of English. Archivum Linguisticum 15.1-28.
- ----- 1967a. Intonation and Grammar in British English. The Hague: Mouton.

- ----- 1967b. Notes on transitivity and theme in English, Part 2. Journal of Linguistics 3.199-244.
- 1985. An Introduction to Functional Grammar. London: Edward Arnold.
- van der Hulst, H. & N. Smith. 1982 (eds.) The Structure of Phonological Representations I. Dordrecht: Foris.
- Jaeger, J. J. 1986. Concept formation as a tool for linguistic research. In Ohala & Jaeger (eds.), 211-37.
- Kreckel, M. 1981. Tone units as message blocks in natural discourse. Journal of Pragmatics 5.459-76.
- Ladd, D. R. 1986. Intonational phrasing. Phonology Yearbook 3.311-40.
- Lakoff, G. 1987. Women, Fire, and Dangerous Things. Chicago: Chicago University Press.
- Laver, J. 1970. The production of speech. In Lyons (ed.), 53-75.
- Lyons, J. 1970 (ed.) New Horizons in Linguistics. Harmondsworth: Penguin.
- Menn, L. 1983. Development of articulatory, phonetic and phonological capabilities. In Butterworth (ed.), 3-50.
- Nathan, G. S. 1986. Phonemes as mental categories. BLS 12, 212-23.
- Nespor, M. & I. Vogel. 1982. Prosodic domains and external sandhi rules. In van der Hulst & Smith (eds.), 225-55.
- 1983. Prosodic structures above the word. In Cutler & Ladd (eds.), 123-40.
- Ohala, J. J. & J. J. Jaeger. 1986 (eds.) Experimental Phonology. Orlando: Academic Press.
- Oreström, B. 1983. Turn-taking in English Conversation. Lund: CWK Gleerup.
- Palmer, H. E. 1924. English Intonation with Systematic Exercises (2nd edition). Cambridge: W. Heffer.
- Petöfi, J. 1988 (ed.) Text and Discourse Constitution. Berlin: de Gruyter.
- Pierrehumbert, J. B. 1980. The Phonology and Phonetics of English Intonation. Ph.D. diss., MIT.
- Pike, K. L. 1945. Intonation of American English. Ann Arbor: University of Michigan Press.
- Rosch, E. 1978. Principles of Categorization. In Rosch & Lloyd (eds.), 27-48.
- Rosch, E. & B. Lloyd. 1978 (eds.) Cognition and Categorization. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rudzka-Ostyn, B. 1988 (ed.) Topics in Cognitive Linguistics. Amsterdam: John Benjamins.
- Schuetze-Coburn, S. In progress. The measurement and representation of pauses in natural discourse.
- Schuetze-Coburn, S.; M. Shapley & E. G. Weber. 1991. Units of intonation in discourse. Language and Speech 34.207-234.
- Selkirk, E. O. 1981. On prosodic structure and its relation to syntactic structure. In Fretheim (ed.), 111-40.
- Svartvik, J. 1982. The segmentation of impromptu speech. In Enkvist (ed.), 131-145.
- Taylor, J. R. 1989. Linguistic Categorization. Oxford: Clarendon Press.
- Tomlin, R. S. 1987 (ed.) Coherence and Grounding in Discourse. Amsterdam/Philadelphia: Benjamins.
- Trager, G. L. & H. L. Smith. 1951. An Outline of English Structure. Norman, OK: Battenburg Press.
- Tsohatzidis, S. L. 1990 (ed.) Meanings and Prototypes. London: Routledge.
- Vaissière, J. 1983. Language-independent prosodic features. In Cutler & Ladd (eds.), 53-66.
- Wittgenstein, L. 1953. Philosophical Investigations. Oxford: Blackwell.