PROSODIC ELEMENTS AND PROSODIC STRUCTURES IN NATURAL DISCOURSE

Anthony C. Woodbury

Dept. of Linguistics
University of Texas at Austin
Austin, TX 78712

ABSTRACT

Although usually taken for granted, it is anything but clear that prosodic elements are organized into autonomous prosodic structures such as intonational phrases. A framework is outlined within which the structural and communicative organization of prosodic elements in samples of natural discourse might be discovered inductively. The framework assumes that the structural organization of a stretch of speech consists of the set of recurrent patterns it contains (including prosodic patterns), and that such patterns are recognizable to speakers. It is further hypothesized that in the normal or usual case, logically independent patterns (e.g., the placement of pauses vs. the placement of intonational cadences) will converge or unify; and that if they do not unify, speakers may draw special pragmatic inferences from this fact. Three samples of natural speech are analyzed in order to present the approach and demonstrate three key properties of the prosodic structure that it uncovers: (a) the potential independence of prosodic patterns and thematic structure; (b) the potential for bundles of prosodic elements to recur as prosodic 'macrostructures,' often associated by speakers with particular styles, contexts, and social personas; (c) the potential for prosodic patterns (and elements) to carry meaning that is iconic in character, but regulated by culturally specific conventions and practices.

0. PROSODIC ELEMENTS

When we speak of prosody, we are concerned with phonological and phonetic elements such as the following. Let us call them PROSODIC ELEMENTS:

- Pausing. (Including structural pauses after whole utterances, rhetorical pauses, micropauses, apparent hesitations and disfluencies.)

- Other durational modulations. (Including final lengthening, anacrusis, 'rhetorical lengthening,' local rhythmic and arrhythmic patterning.)

- Stress and related features.

- Pitch targets. (H,L, possibly rises and falls.)

- Pitch alignments (to prominent syllables, word edges, and other sites, cf. Pierrehumbert and Beckman 1988).

- Pitch scaling. (Including initialization, initial raising, final lowering, downstep, and catathesis (Hirschberg & Pierrehumbert 1986).)

- Segmental sandhi processes. (E.g., American English Flapping, French Liaison (Nespor & Vogel 1986).)

- Prosodic reshaping. (Including 'Rhythm rules' (Liberman and Prince 1977), phrasal truncation (Sapir 1949), refooting rules (Woodbury 1987b, 1992).)

- Voice quality modifications. (Including falsetto, breathy and creaky voice, pharyngealization, vibrato (Miller 1992), phrase-final devoicing (Michelson 1991).)

- Others?

Perhaps the three leading questions about prosodic elements are:

- THE IDENTITY QUESTION. How are they to be identified and described, and how are they perceived by humans?

- THE DISTRIBUTION QUESTION. How are they distributed in discourse, and what cognitive faculties govern those distributions?

- THE MEANING QUESTION. What do they mean, or do, in discourse?

1. THE PROSODIC HIERARCHY

An answer to the Distribution Question is offered by PROSODIC HIERARCHY THEORY (Selkirk 1980, Nespor & Vogel 1986, Hayes 1989, Inkelas and Zec 1990), and to some extent by any approach assigning autonomy to

prosodic phrases, intonational phrases, breath groups, and the like (Halliday 1967, Beckman and Pierrehumbert 1988, Chafe 1980). It also gives, implicitly, a partial answer to the Meaning Question.

Prosodic Hierarchy Theory claims that prosodic elements refer for their distribution to a hierarchy of abstract, autonomous, discrete units like that in Figure 1. Furthermore, it claims that these abstract units—rather than individual prosodic elements—map in certain ways to syntax, pragmatics, and thematic discourse structure. Thus it answers the Distribution Question by asserting massive coordination among a variety of aspects of grammar and speech. And it answers the Meaning Question with the implicit assertion that abstract phrase breaks of the hierarchy, rather than individual prosodic elements such as pauses, intonational cadences, and the like, will bearers of (pragmatic) meaning.
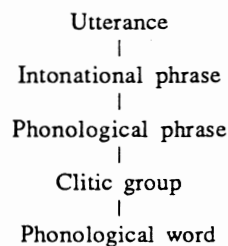
Utterance
|
Intonational phrase
|
Phonological phrase
|
Clitic group
|
Phonological word

Figure 1: The Prosodic Hierarchy according to Hayes 1989.

2. CRITIQUE

Two colleagues and I have criticized this approach on empirical grounds (Liberman, McLemore, & Woodbury 1991). Our critique has several main points.

First, (most) effects of prosodic constituency are gradient with respect to junctural strength, and thus do not provide any clear evidence for qualitative constituent types.

Second, the hypothesized prosodic structure is extremely ambiguous in practice, so that determinate, intersubjectively valid descriptions are not generally possible if the hypothesis of qualitatively distinct prosodic levels is maintained.

Third, even when a given prosodic effect is nongradient with respect to junctural strength (i.e., either present, or absent), it is still generally difficult to assign it successfully to just one constituent type (Woodbury 1992).

Fourth, (most) clear "prosodic" constituents seem to correspond to independently-needed units, generally from the domain of information structure.

And fifth, (some of the) key phonetic correlates of prosodic structure can be seen as natural solutions to the problems of presenting the (non-prosodic) structure of messages and of managing communicative interaction (McLemore 1991).

In the face of these doubts, we considered it reasonable to approach prosody anew with a NULL HYPOTHESIS, according to which the distributions of prosodic element refer directly to functions and structures that are outside of prosody and that are independently known to be part of discourse, including syntax, pragmatics, and thematic structure.

In principle, this null hypothesis implies no coordination at all among prosodic elements—let alone the massive coordination implied by the Prosodic Hierarchy—since it expects only that each prosodic element will bear some relationship to something elsewhere in discourse. Nevertheless, it is crucial to determine whether any such coordination obtains, even if less rigid or necessary than that projected by the Prosodic Hierarchy; that is, whether prosodic elements can assemble themselves into PROSODIC STRUCTURES of any kind. The question is no less one of meaning (or communicative function) than it is of distribution, since if such structures do exist, then on any theory of prosody at all they too should be expected to relate to syntax, pragmatics, or some other aspect of discourse.

The problem of prosodic structure is my focus in this paper. Section 3 argues that this structure should be sought inductively through observation of natural discourse. Section 4 outlines a framework for finding prosodic (and other discourse) structure which builds on work in poetics. The last three sections raise further issues while applying the framework to three discourse samples.

3. OBSERVING NATURAL DISCOURSE PROSODY

We know so little about the distribution and meaning of prosodic elements that we must first observe and describe them in natural discourse. This is not to say that experimentation, modeling, and even introspective study do not have their place. But they work best when they rest on an idea of the variety and diversity of speech prosody.

The term NATURAL DISCOURSE needs elaboration, for investigators seem to use it in at least two quite distinct senses. On the one hand, it is used to refer to any extemporaneous—not scripted—speech. That is perhaps the grammarian's sense, since the speaker is generating the forms using his/her own grammar. On the other hand, the term is also used to designate speech that is real—not simulated—social action. And perhaps that is

the anthropologist's sense of it, since anthropologists have often observed that when people simulate behavior, they do so with reference to stereotypes or IDEOLOGIES of social action, rather than the tacit models they rely on when performing or responding to social action in real life (see Silverstein 1979 for a careful review). As a result, simulated behavior is often recognizably different. To take just one example, consider the skill an actor must have in order to perform scripted dialog believably and effectively: such special skill would be unnecessary if there were no inherent gulf between real and simulated social action.

I would suggest that by following the anthropologists' lead in connecting natural discourse to social action, we can best appreciate its diversity. From that perspective, it is not enough to sample natural discourse simply by turning on the radio, for it spans every possible facet of social life. As workers in the ethnography of speaking and sociolinguistic pragmatics have emphasized (see Hymes 1974, Bauman & Sherzer 1974, Gumperz 1982, Levinson 1983, and the journals *Language in Society* and the *Journal of Linguistic Anthropology*), it encompasses narrative, conversation, and oratory. It includes ceremonial, ritual, formal, and institutional speech, in societies both with, and without, highly diversified institutional structures. It includes magical and religious speech. It includes prose, poetry, chanting, and singing. It includes speech with different purposes, from exhortation, to instruction, to description, to elicitation. And it includes scripted speech, whether read or recited from memory, along with extemporaneous speech.

But is it really necessary to sample so many kinds of speech? Is the prosody of a language not more or less uniform, regardless of the use to which it is put? The inductive perspective advocated here lets us see for ourselves. Even the small amount of description I have done convinces me that unlike syntax, morphology, and lexical phonology, prosody seems to vary not a little, but fundamentally, across genres, varieties, uses, and the like, even within a single language. If this is so, then it certainly is worth it to pursue diversity and to generalize only cautiously about the intonational systems of entire languages.

4. FRAMEWORK

To make consistent, useful observations, it is necessary to do so within an explicit framework making at least some basic theoretical assumptions. Such a framework should allow:

- 'Thick' description of the form, distribution, and meaning of prosody in individual texts

- Comparability across descriptions of different kinds of natural discourse, to allow for appropriate inductive generalization.

If our critique of it is justified, Prosodic Hierarchy theory is not such a framework: it is less than ideal for thick description since it may deflect attention from those aspects of prosody not crucially relevant to the hierarchy; and by focusing on the abstract units rather than individual prosodic elements, it may at times overstate some similarities across descriptions while missing others. Its problem in short is that it checks for a certain kind of all-encompassing order and coordination among prosodic elements, rather than gauging structure in whatever shape or form it may take.

4.1. Jakobson's Poetics

How then is structure to be gauged? Let us begin with a particularly useful notion of discourse structure from poetics. Despite its poetic origins, it can be extended beyond what we may wish to designate as 'poetry,' or value as verbal art. The basic idea, due to Roman Jakobson, is that there is a poetics to ALL discourse; and it is a fundamental key to discourse understanding. He gives his idea the following quite pungent formulation (Jakobson 1960:358):

> The poetic function projects the principle of equivalence from the axis of selection into the axis of combination. Equivalence is promoted to the constitutive device of the sequence. In poetry one syllable is equalized with any other syllable of the same sequence; word stress is assumed to equal word stress, as unstress equals unstress; prosodic long is matched with long, and short with short; word boundary equals word boundary, no boundary equals no boundary; syntactic pause equals syntactic pause, no pause equals no pause. Syllables are converted into units of measure and so are morae or stresses.

Essentially Jakobson is proposing a principle of recurrence, and claiming that it creates or reinforces structural equations (which then invite inference about content, given speakers' expectation that form can diagram content). The strongest instance of it is simple repetition. Illustrating from written poetry:

and miles to go before I sleep;
and miles to go before I sleep.

A weaker instance is in parallelism, a form of partial recurrence:

243

He called for his pipe
and he called for his bowl
and he called for his fiddlers three.

The recurrent elements or units may be of any kind: phonological, syntactic, lexical, morphological, thematic, and so on. All the examples above involved syntactic units. But phonological units recur in rime, alliteration, and meter. And morphological units recur in grammatical parallelism.

Recurrence establishes PATTERNS of various kinds. The most elementary is simple alternation (as, e.g., in trochaic meter); a very elaborate pattern is that of the Shakespearean sonnet, shown in Figure 2, where the pattern is global (it pervades the whole poem), it involves a hierarchy of fixed depth (i.e., a fixed number of qualitatively distinct levels), and it has counted parts (e.g., five feet to a line, two lines to a couplet). Obviously, different kinds of discourse are likely to show different degrees of pattern elaboration.
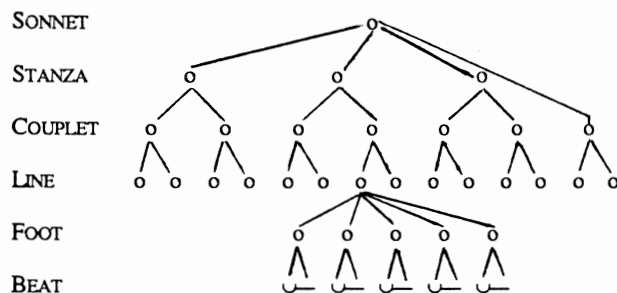

Figure 2: Shakespearean sonnet form.

4.2. The framework

Jakobson's principle can form the basis of a heuristic framework or method for determining how, and how much, a given instance of natural discourse is structured, and as part of that, the extent to which the prosodic elements within it are structured.[1] At minimum, any discourse instance should have some kind of thematic recurrence; a sequence of sentences (even if not parallelistic in any internal respects); regularly recurrent pitch cadences; and an alternation between pausing and silence. Since these different patterns need not all be congruent, our framework is most useful if it assumes at the outset that they form distinct parts the of overall discourse organization:

Assumption 1. Natural discourse is organized as a set of components, where a component is any well-defined patterning of recurrent elements that is present

---

[1] Hymes 1981 and Silverstein 1984 are some other explicit efforts to extend Jakobson's poetics to the analysis of natural discourse of various kinds.

in a stretch of discourse and distinct from other such patternings.

If this is to be of any interest, we must also assume:

Assumption 2. Speakers can recognize well-defined patterning of recurrent elements.

If true, this may hold by virtue of a general human capacity for pattern recognition, rather than any specifically linguistic faculty.

The following then is a likely minimal set of components:

- Thematic patterning

- Syntactic patterning

- Pitch patterning (pitch accents, cadences, scaling)

- Pause patterning

Of course, a sample of discourse would require more if other elements were also at work creating recurrent patterns. For example, if distinct prosodic elements establish distinct patterns, then these patterns must count as separate components. The notion of component, like the notion of recurrence, is entirely empirical and heuristic.

The above assumptions may approach the minimum needed for gauging structure in discourse. But the following theoretical claims might be added as assumptions for heuristic purposes:

Assumption 3. In the usual or default case, elements and structures of different components should converge or unify, standing in one to one correspondence, in the simplest cases (perhaps on general iconic grounds, rather than by virtue of any specifically linguistic faculty).

Assumption 4. Deviations from such convergence may be salient to speakers and may lead them to draw special pragmatic inferences.

These assumptions do not exclude any distributional configurations: rather they predict (rightly or wrongly) how certain distributional configurations will be processed and interpreted. For example, in many instances of discourse the general pattern is for sentence breaks, intonational cadences, and pauses to coincide regularly. This would constitute a set of defaults among three logically separate components, *i.e.,* syntactic, pitch, and pause patterning. Departures from this convergence—enjambment of sentences by pause suppression, rhetorical pauses and pitch falls within sentences, and the like—would then stand out as special

and invite certain special interpretations. On these assumptions, structuredness is a matter of degree that must be gauged, rather than only a quality that must be specified. Furthermore—at least by hypothesis—it serves as a norm in terms of which potentially significant departures are measured, rather than simply a defining characteristic of well-formed speech (as in grammar).

In the following sections, all four assumptions are used to track prosodic elements, gauge prosodic structure, and evaluate their communicative roles, in three samples of natural speech. Each analysis is focused on a key property or characteristic of the structure being sought. The first and third samples come from myth performances by Central Alaskan Yupik Eskimo elders; the second is a pair of routines from an American television comedian.

5. DOES PROSODY REFLECT THEMATIC PATTERNING?

According to a widely held view, prosody is there to reinforce the preexisting thematic or content patterning of speech. In terms of the framework discussed above, prosodic elements and structures would then cue the boundaries or the internal 'high points' of such units. But workers in sociolinguistic pragmatics (e.g., Silverstein 1976:33-35, Gumperz 1982:100ff) have argued that pragmatic markers, including prosodic cues, need not only reflect or reinforce preexisting elements of context—they can affect, shape, and create participants' constructions of context. Accordingly, a subtly different view of prosody and thematic patterning might hold that although prosodic elements may reinforce independently recognizable thematic units, they may also at times be used to propose or create novel constructions of thematic patterning that are not independently inferable, or that alloy or conflate thematic patterning with other considerations, or that are present only as one of many possible 'takes' on thematic patterning.

In connection with this last point, it is important to make clear just how varied thematic patterning can be. In narrative, it can involve patterns of character foregrounding, or of tense/aspect shifts, or of scene changes, or of parallelistic, recurrent episodes. In conversation (Levinson 1983), it can be based on adjacency pairs (like question and answer, offer and acceptance, request and denial), or conversational activity types (greetings, leave-takings, 'pre-sequences,' etc.), and can be highly ritualized, as in verbal dueling (Labov 1972). In oratory, it can involve parallelistic figures that frame a rhetorical progression or transformation; or by the speaker taking the parts of alternating participants in a simulated argument or conversation. In ceremonial speech, it can involve distinct sections corresponding to different stages of a ceremony, or the progression of special speech act types, or the alternation of fixed texts

and impromptu speech. At very least, this range and variety should caution that empirical result gotten for one kind of thematic patterning may well fail to predict results for other kinds of patterning; and it should indicate that for any one given text, there are many ways to conceive of thematic patterning. (In the terms of Assumption 1 above, a given discourse could have several orthogonal thematic (sub)components).

Figure 3 presents the opening of a myth performance in Central Alaskan Yupik Eskimo (CAY). Each line of the Figure shows a single word (or, when apparently sharing a single intonational contour, two words); its intonational profile, consisting of the initial pitch (if preceded by a pause), which is generally a low point; the pitch peak, which generally occurs on the first or second stressed syllable (boldface); and the pitch at the end, which generally is another low point. Following that is an indication of following pause length (0.0 if there is no pause), and then an English translation.

The myth is particularly interesting because it has a highly elaborate pattern of counted-out thematic parallelism among episodes. In the story, an orphan grandson paddles upriver in his kayak, meeting five successively more fearsome creatures and then taking them into his magical power. He returns home to his grandmother, denying having done anything special when she asks him. Then he goes into the communal men's house where he is harassed by bored, cruel shamans who want him to try his hand at conjuring. Under duress, he gives in and conjures, in turn, each of the five animals he had charmed, eliciting a successively greater reaction among his tormentors. Specifically, the whole myth divides into five Parts (Figure 4a). Of these, Parts II and IV then divide further into five Episodes, corresponding to the animals the Grandson charms (Figure 4b). Finally, each Episode divides into three Divisions according to the logic of the action (Figure 4c).[2]

Obviously, not all myths have such robust and formally elaborate patterns of thematic recurrence. Therefore, this myth presents a special opportunity to ascertain thematic patterning independently and gauge its relationship to prosody.

As it turns out, two prosodic elements, initial and final low pitch, mark off major units (e.g., Episodes in II, Divisions in IV). These units either begin with lower pitch, or end with lower pitch, or both. Thus in Figure 3, Parts I and II both begin below 80 Hz, while Part II ends somewhere below 90. (Sporadically, both of these elements also occur elsewhere.) They clearly reinforce independently recognizable thematic patterning. Even so, they still do not do this in the strictest possible way: for

---

[2]The terms Part, Episode, and Division are descriptively useful but theoretically arbitrary.

in Part II they mark Episodes, while in Part IV they mark Divisions. The explanation for this is that Episodes in II, and Divisions in IV, are all units of roughly the same length. Thus the distribution of low initial and final pitch is also partly rhythmic, or time dependent. By reflecting a particular mix of thematic and rhythmic factors, initial and final low pitch show a modicum of autonomy. Moreover, because they do so together, they operate not only as individual prosodic elements, but as prosodic STRUCTURES, and constitute, in the terms of Assumption 1 above, a COMPONENT in this stretch.

More sharply autonomous is a pattern involving the pitch peaks. Observe that within each sentence, the heights of successive pitch peaks generally increase, giving a culminative effect within the sentence. Moreover, peak heights also build from the beginning of the story to the end of Division A in Episode i, where 392 Hz is reached with the help of an astonishing narrative falsetto Mezak used in this and in many other of his performances. Thematically, the building pitch seems to mark the development of the action up to the point where the boy must perform masterfully. After

CAY	Init L	Peak	Final L	Pause	English
[I. INTRODUCTION]					
Nunat ^ukut	77	116	86	<0.0>	There was a village [of people]
uitaura'rqelriit		140	87	<0.0>	who lived
kuigem ^ceniini.		150	84	<2.8>	on a riverbank.
Tutgara'urluqelriigneg	104	140	97	<0.0>	A grandchild and grandparent
ilaluteng.		152	95	<0.0>	were among them.
Tutgara'urlurlua ^im',		146	113	<1.5>	And that grandchild,
tan'ga'urlull'rauluni;	82	162	112	<0.0>	was a boy;
Angutnguluni=w'.		150	<116	<2.5>	It was a male.
[II. GRANDSON GOES UPRIVER]					
Tua=i=ll'=am	78	156	128	<0.0>	Well
caqerluni,		151	120	<2.3>	one time
kuigkun ^e:--	142	227	186	<0.0>	*on the river--*
kuimegteggun=am		294	222	<0.0>	*on their river*
asgurtuq..		333	256	<1.5>	*he [paddled] upstream.*
[i. ENCOUNTERS PTARMIGANS]					
A Tua::=ih,	136	136	93	<1.4>	Well
asgurturaqerluni,	109	184	136	<3.1>	as he went upstream
qang:qiiregnek,	142	285	136	<1.7>	a pair of ptarmigan
callul:riignek	357	392	368	<0.0>	*who were fighting*
tekituq.		363	—	<1.8>	*he encountered.*
B Tua=i=ll'	285	285	123	<0.0>	Well,
ketairamikek		270	125	<0.0>	when he passed alongside them
cenami	[83]	226	117	<0.0>	on shore
piagnek:		133	80	<1.8>	he said to them:
"Aah!	140	140	113	<1.7>	"Hey!
Tua:=i!	113	156	123	<2.4>	Enough!
Pisqekumtek ^taugaam	128	217	120	<0.0>	But when I tell you
piniartutek!		262	127?	<1.8>	you shall [carry on some more]!"
Aa tua=i	151	204	180	<0.0>	Hey enough now
uter— utertek!"		214	192?	<1.6>	go— go home!"
C Aren im*kug=am*	123	263	222	<0.0>	My and sure enough the two
*qang*qiirek		294	144	<0.0>	the ptarmigans
niilluteg		184	125	<0.0>	they obeyed
ayagtuk.		126	<99	<1.7>	and went away.

Figure 3: Opening of a Central Alaskan Yupik Eskimo myth performance by Evon Mezak of Nunapitchuk, Alaska, recorded in about 1972. The text, and a detailed analysis of it, appear in Woodbury 1987a.

Myth
{
 I. Introduction
 II. Grandson goes upriver
 III. Grandson goes home
 IV. Grandson goes to men's house
 V. Closing
}

Figure 4a: Expansion of the Myth into five Parts.

Parts II and IV
{
 i. Grandson meets/calls ptarmigans
 ii. Grandson meets/calls dunlins
 iii. Grandson meets/calls cranes
 iv. Grandson meets/calls caribou
 v. Grandson meets/calls wolves
}

Figure 4b: Expansion of Parts II and IV into five Episodes.

Episode
{
 A. Grandson comes, and is challenged
 B. Grandson reacts masterfully (in quoted speech or song)
 C. Response to Grandson's reaction
}

Figure 4c: Expansion of Episodes into three Divisions.

this point there is more or less a denouement. The same building pattern then occurs again in following episodes.

On Assumption 1, these pitch fluctuations represent a unique pattern and hence call for a separate component. To be sure, the component correlates with an aspect of thematic patterning, but by picking out a climax in narrative development, it does so in a wholly different way than the low initial and final pitches. Indeed, it can even be seen as a device by which the narrator *proposes* to his hearers an interpretation of narrative development in this section.

A final interesting pattern is presented by pausing. As in much Native American discourse, the pauses are long and come at regular intervals, adding salience to pausing as a poetic feature (on Assumption 2, but also on the grounds of experience: see Tedlock 1983). This is best seen if we consider a reformatted version of the translation, Figure 5, in which line-breaks correspond to pauses. The default pattern for the whole text (Assumption 3) is for pauses to occur at sentence breaks, and, occasionally, for somewhat shorter pauses to occur at points in between. (Scrupulous observance of this default is one way that boring-sounding prose can be achieved in CAY!) Here however, the default is upset twice in the second line, where 'And the grandchild' is enjambed with the preceding sentence with no intervening pause, while simultaneously set off by a pause from it's sequel, 'was a boy'. Thematically

English	Pause
[I. INTRODUCTION]	
There was a village [of people] who lived on a riverbank.	<2.8>
A grandchild and grandparent were among them. And the grandchild	<1.5>
was a boy; He was a male.	<2.5>
[II. GRANDSON GOES UPRIVER]	
Well one time	<2.3>
on the river--on their river he [paddled] upstream.	<1.5>
[**i. ENCOUNTERS PTARMIGANS]**	
A Well	<1.4>
as he went upstream	<3.1>
a pair of ptarmigan	<1.7>
who were fighting he encountered.	<1.8>
B Well, when he passed alongside them on shore he said to them:	<1.8>
"Hey!	<1.7>
Enough!	<2.4>
But when I tell you you shall [carry on some more]!"	<1.8>
Hey enough now go— go home!	<1.6>
C My and sure enough the two ptarmigans they obeyed and went away.	<1.7>

Figure 5: English translation from Figure 3, reformatted so that line-breaks correspond to pauses.

speaking, the pause here marks the division between old and new information, adjoining 'And the grandchild' to the sentence that first introduced the grandchild.

Likewise, the default is upset when very long pauses occur in mid sentence, as in the first and fourth lines of Part II in Figure 5. Very long pauses are unexpected, and hence (on Assumption 4) invite special interpretation. As in all languages/speech communities with which I am familiar the unexpected delay at this point heightens narrative suspense. This effect presumably follows from basic, essentially non-linguistic strategies that all people have for dealing with expectations that fail (momentarily) to materialize.

In this section we have seen where prosodic elements reflect thematic patterning but still alter it slightly on rhythmic grounds; where prosodic elements not only reflect thematic patterning, but propose an interpretation of it; and where distinct prosodic elements are coordinated loosely, in terms of a default. Several conclusions may be drawn. First, even when thematic patterning is independently recognizable, prosodic elements need not simply reflect it. Therefore, it is not safe to assume—as many investigators seem to do—that prosody will provide a perfect diagram of some fixed (abstract) thematic structure of speech in cases where thematic patterns are not patent or overt in any other way. Second, there are interesting distributional and functional relationships not only between individual prosodic elements and thematic structure, but among prosodic elements. That is, prosody can still be said to have structure, albeit of a far more diffuse and complex type than has generally been assumed.

6. PROSODIC MACROSTRUCTURES

Although individual languages and individual speakers have broad prosodic resources, it is striking how few are actually used in many natural instances of speaking. This is illustrated by two short routines by Jay Leno, a television comic (Figure 6).

Thematically, each routine has two major parts: one where the comedian recites something heard or seen elsewhere, taking on the voice of the source; and then one where he parries in his own voice with the punch line (and then feigns nonchalance when the applause comes). The transcript is broken into lines representing putative intonational phrases, i.e., domains implicated by prominent final pitch cadences and (usually) final rallentando or lengthening. Shown in boldface are those syllables having salient pitch accents (transcribed below in Pierrehumbert's 1980 notation), and in small caps, the one among them with the highest pitch. At right is an indication of the highest pitch peak; the pitch at the final boundary (two pitches for 'continuation rises,' corresponding to the trough and the boundary); and an indication of pause time, if any.

Of all the prosodic resources or options that Pierrehumbert describes for English, only a relative few appear here: nearly all pitch accents are LH* (rises to high); they occur densely, i.e., several to an intonational phrase; the heights of pitch peaks fluctuate considerably, showing great range (with the highest marking both new topics, and very salient foci); and nearly all the cadences are falls to low. Further, pausing is highly facultative (Figure 7): the first pause phrase of the first routine is enormous (four sentences, five intonational phrases), while later on, and in the second routine, they are extremely short.

Why just these resources? Leno's choices may at first seem determined solely by content and communicative purpose. LH* intonation is often associated with new information (Hirschberg and Pierrehumbert 1991), and that is appropriate since he casts himself here as a bringer of news. His wide pitch range and high pitch accent density add punch and vividness. And his syncopated use of pause seems the essence of comedic timing. Yet this is not the whole story, for many of these effects can be approximated with a different set of resources: Why should he not draw on them too from time to time?

I would suggest that Leno has constructed what might be called a PROSODIC MACROSTRUCTURE—a small set of resources that then become the material out of which the patterns of Assumption 1 largely are crafted. It certainly is true that Leno's macrostructure is well-suited to his task. But beyond that, it becomes associated through use partly with him and his comedic style in the minds of his audience.

Macrostructures can range greatly in their conventionality, from those that are traditionally tied to particular genres, to those which have become habitual for individual in particular settings, to those composed quite on the spur of the moment. Queen (1992) raises this issue in her discussion of the oratory of Martin Luther King. She shows that King made a highly distinctive set of prosodic choices in his oratory, and demonstrates that these choices partly continue traditions of African-American preaching, and partly constitute a unique personal style. In the case of Jay Leno, his macrostructure may to some extent follow a tradition in American stand-up comedy and in part be his own construct. In any case it is interesting that it is not identical to the prosodic choices of other stand-up comedians, nor to the choices he himself makes in other speech settings (e.g., interviewing guests).

To the extent that prosodic macrostructuring turns out to be a significant fact of natural speech, several points can be made about it. First, it represents another way in which prosodic elements coordinate with each other to form prosodic structures, albeit a kind of structuring quite different even from that suggested by the Prosodic Hierarchy. Second, it means that natural discourse may be more orderly, and therefore more amenable to systematic

[ROUTINE I]

Text			
An' HERE'S somp'm I got ou' th' paper today 　LH*　　　　　　　　LH*　　H*	285	131	<0.0>
a MAjor New York-- newspaper 　LH*　　H*　　　　LH*	158	109	<0.0>
I THINK i wz the New York Po:st. 　LH*　　　　　LH*　　　LH*	178	114	<0.0>
Now THESE a' their statisti:cs. 　LH*　　　LH*	243	102	<0.0>
NOT mine:. LH*	238	89-147	<0.3>
They said TWEn'y five percent: 　　　LH*　　HL*	322	123	<0.0>
of the homeless are alcoHOLi:cs, 　　　　　　　　LH*	128	79	<0.7>
TWEn'y five percen' are drug a:ddicts, (L)H*　　HL*　　　　LH*	228	106-217	<0.4>
and THIRty percen' 　　LH*	316	116	<0.0>
'ave been instiTUtionalized 　　　　　LH*	232	97	<0.0>
at ONE time 'r another 　LH*	149	94-119	<0.0>
for mental disaBIlities. 　LH* (L-)　　LH*	232	100	<1.0>
Now I know that seems like a pretty high perCEN'age. 　　　LH*　　　LH*　　　LH*　　　　LH*	200	109	<0.0>
But ya KNOW, LH*　　H*	166	166	<0.1>
when ya comPARE it to co:ngress 　LH*　　　LH*　　　LH*	312	111	<0.0>
gee it's NO:T tha:t high: really. LH*　　LH* (L)H* LH* L-　L*	370	105	<1.0>
Yknow.		--	
Applause, etc.			<...>

[ROUTINE II]

Text			
You KNOW what's grea:t:? 　LH*　　　HL*	400	--	<0.0>
SEE, 　H*	250	250	<0.5>
SEE, 　H*	243	198	<0.2>
I ALways like to watch politicians:: LH*　　LH*　　　　LH* LH*	400	106	<0.0>
try to JUStify:: H*　　LH*	185	102	<0.0>
their JO:BS. 　LH*	133	94	<0.0>
I SAW a senator LH*　LH*	200	96	<0.0>
on ONE a' those: 　LH*	116	(101)	<0.6>
sunday morning TALK shows the other day. H*　　　　LH*	163	86	<0.0>
An' 'E said LH* H*	149	89	<0.0>
th't the ACtions of the se:nate 　　LH*　　　　LH*	270	87	<0.5>
have creAted alot a jo:bs: 　LH*　　LH*　LH*	164	104-107	<0.3>
for alot a CItizens:. L*　LH*	208	86	<0.0>
YEAH but: H* L-　L*	133	90	<1.0>
LETS fa:ce it. H*　LH*	208	95	<0.0>
YOU can't make a career out o'jury duty:. LH*　　　　LH*　　　LH*	185	83	<0.2>
Ya KNOW? 　LH*	153	153	<1.0>
You know what I mean?	--	--	<0.0>
That--<0.5--<obs> THAT 'S five bucks a day: 　　　　　LH*　　LH* LH*　　LH*	169	97	<...>

Figure 6: Two American English standup comedy routines performed Jay Leno on NBC television, March 1990. The two routines occurred in immediate succession.

249

An' HERE'S somp'm I got ou' th' paper today % a MAjor New York-- newspaper % I THINK i wz the New York
    Po:st. % Now THESE a' their statisti:cs. % NOT mine:.%   <0.3>
They said TWEn'y five percent: % of the homeless are alcoHOLi:cs,%   <0.7>
TWEn'y five percen' are drug a:ddicts,%   <0.4>
and THIRty percen' % 'ave been instiTUtionalized % at ONE time 'r another % for mental disaBIlities.%   <1.0>
Now I know that seems like a pretty high perCEN'age % But ya KNOW,%   <0.1>
when ya comPARE it to co:ngress % gee it's NO:T tha:t high: really. %   <1.0>
Yknow. % *Applause, etc.*   <...>
You KNOW what's grea:t:? % SEE, %   <0.5>
SEE, %   <0.2>
I ALways like to watch politicians:: % try to JUstify:: % their JO:BS. % I SAW a senator % on ONE a' those: % <0.6>
sunday morning TALK shows the other day. % An' 'E said % th't the ACtions of the se:nate %   <0.5>
have creAted alot a jo:bs: %   <0.3>
for alot a CItizens:. % YEAH but: %   <1.0>
LETS fa:ce it. % YOU can't make a career out o'jury duty:. %   <0.2>
Ya KNOW? %   <1.0>
You know what I mean? %   <0.0>
That--   <0.5>
<obs> THAT 's five bucks a day: %   <...>

Figure 7: Transcript of Figure 6, reformated so that line-breaks correspond to pauses. '%' marks cadence locations.

study, than often thought. Third, it points up the need for 'thick' description in natural discourse study, since it is the particular mix of elements that gives a macrostructure its cultural and stylistic associations, rather than the abstract patterns they instantiate. At the same time, of course, such variation makes it clear that an appropriate notion of pattern must be formulated abstractly enough so as not to be locked to particular phonetic forms or choices. However practical they may be in some respects, this in fact is a weakness of many transcription oriented approaches to prosody (e.g., Hirschberg and Beckman 1992) since they tend to emphasize just certain elements, without regard to their centrality in forming salient distributional patterns or in conveying discourse meaning in the speech sample in question.

7. NATURAL AND CONVENTIONAL ASPECTS OF PROSODY

We last consider natural and conventional aspects of the distribution and interpretation of prosodic elements. The case in point is a stylized downtrending or deaccenting phenomenon that is salient in the narrative prose of many, but not all, speakers of the CAY dialect of Chevak and Hooper Bay, Alaska. This dialect is moderately different form that discussed in Sec. 5.

Shown in Figure 8 is a section from a myth told by Thomas Moses of Chevak. On thematic grounds, it constitutes a single episodic unit. Each line is a word (or two where intonation is continuous). On a pattern that is seldom violated by Chevakers, the unit on each line shows a clear pitch trough initially (on the first stress of the word), followed by a peak. The peak is at the end of the word, unless a low tone occurs there (moving the peak back to the last stress). Pragmatically, the final low

marks disjunction—the lower the tone, the greater the implied break. Generally these breaks correspond well, in both placement and degree, to syntactic constituency breaks (Woodbury 1989). Lower case 'w' indicates 'whisper'; these whispers count as very low tones and arise when the tone heads below a certain threshold. Timing of pauses, if any, follows next; and last is the English translation.

The phenomenon of interest, which I call ATTENUATION, is characterized by significantly lowered H pitch peaks and reduced amplitude. Attenuated sequences are marked in Figure 8 by plain type, while nonattenuated sequences are boldface. The high pitch values of the attenuated items are generally noticeably lower. I am not in a position to assert that the attenuated/nonattenuated distinction is categorical (rather than gradient), even though it usually sounds and looks quite distinctive. Categorical or not, it counts as a prosodic structure, rather than a prosodic element, since it involves a cluster of prosodic elements (pitch scaling and amplitude) which pattern together.

A first observation about attenuation in Figure 8 is that it accompanies all postposed constituents (which are underlined). Because CAY has very rich inflectional morphology, these postposed constituents can usually count as supplements to already-fully-formed sentences. In terms of Assumption 3, there is an apparently exceptionless default pattern holding between two components (syntactic patterning, and the patterning of prosodic attenuation):

Postposed constituents are prosodically attenuated.

This still leaves instances where nonpostposed material is attenuated. When the distribution of a prosodic structure

CAY	Init L	Peak	Final L	Pause	English
Piuraqerluni^taw'	84	135	96	<3.5>	Then once
caller'e'mini,	82	108		<1.3>	when she was doing things,
taun'^ar'e'naq,	81	92	76	<1.5>	that woman,
angutmeng^uumeng	103	149	120	<0.0>	this man
tang'elliuq:	89	105		<2.0>	appeared to her:
Kanaqliit=gguq=gg'^atk'ekui!	76	105	w	<6.0>	He had a parka all of muskrat!
Piluku^taw'	81	104		<3.5>	He said to her
amatngurrvakaami^taw'	100	140	126	<0.0>	that because he was so grateful
nuliq—nuliq—nuliqnaluk'	93	113		<0.0>	he had come to ask her
ullagyaaqniluku.	84	104	w	<5.5>	to be his wife.
Taw—Taw—Tawaten,	78	88		<1.4>	S—S—So [he spoke],
anautellermineng	81	105		<0.0>	because he was so grateful
amatngurpakaami.	84	92	w	<4.0>	that she had rescued him.
Tawa=gggur^taum	70	102		<0.0>	And so that one
civunran^taw',	84	102		<0.5>	who stood before him,
tupekluku,	74	84		<0.5>	accepted him,
uing—uing—uingyunrilami.	84/68	84	w	<5.5>	be—be—because she had no husband yet.
Tang—Tangnerrayauluni	76	138		<0.0>	He—He seemed a stranger
Tangnerrauluni=gguq^taw'	81	128		<0.0>	He looked a bit strange
taun'^angun.	78	85	w	<5.5>	that man.
Cuna=gggur^un'^taw'	86	94		<0.0>	So it was for this [woman]
nuliqluku^taum'^taw'	86	94	81	<0.0>	he married her this [man]
nuliqsagulluku.	79	93	w	<1.7>	He had her as his wife.
Piculliniluni=gggur^taun'	95	101	93	<0.0>	He was good at getting things this [man],
pissuraqami^tawaam.	82	99	w	<2.5>	when he hunted.
Maklagculuni=llu.	80	91	w	<3.5>	And good at getting bearded seal.

Figure 8: Opening of a Central Alaskan Yupik Eskimo myth performance by Thomas Moses of Chevak, Alaska, recorded in 1978 (text in Woodbury 1984).

(or element) appears not to depend on the distribution of something else, one must suspect (on the null hypothesis) that the fact of its placement alone may contribute new information. In light of the postposed attenuation cases, and certain features of the nonpostposed cases, I suspect that attenuation contributes the following pragmatic information:

> Attenuation defocuses syntactic/prosodic constituents and labels them as clarifications or SUPPLEMENTS to the interpretation that the speaker expects the audience to have constructed or deduced from the talk so far.

It is fairly clear how this applies to postposed constituents. However, it may still be reasonable to maintain the exceptionless default posited above in order to enforce the link between prosody and syntax explicitly (rather than suppose that all instances of postposing will function in context as supplements).

The pragmatic account also applies in the nonpostposed cases. An interesting one is the attenuation of nonpostposed *tang'elliuq* 'he appeared (to her)' in the fifth line of Figure 8. While 'appeared to her' cannot felicitously be cut from the English translation, it happens that CAY speakers routinely not only attenuate, but sometimes completely delete, verbs of seeing, saying, apparition, and the like. It is therefore plausible to treat *tang'elliuq* as defocused. Indeed, even for a related class of English verbs, there is a tendency place nuclear stress on the subject in preference to the verb (e.g., *THE COPS came, CLINTON spoke, TRUMAN died*, etc.) The parallel is quite striking.

Another interesting set of cases are the attenuated sentences near the bottom. Each of them contributes information which Native hearers could plausibly infer. For example, the last two sentences ascribe to the man certain abilities. Yet these are already inferable by the following logic. In CAY myth, animal transformers wear their fur or feathers as parkas; since the man's parka is all of muskrat, he is a transformed muskrat (and we learn that for sure later in the story); therefore, as a muskrat, he should be a good aquatic hunter. The

importance of these examples is that they show that the notion of supplementation described above holds not only within, but between, sentences. Because of that, it is most plausible to view supplementation as a discourse category that happens to have a conventional relationship to syntax (via the exceptionless default).

The description so far simply assumes that the pragmatic account of attenuation is a matter of convention in the relevant speech community. But this would hardly explain its similarities in both form and function to such English phenomena as postnuclear 'deaccenting' and tag intonation. In work on the form and meaning of cadences and related pitch figures in the speech of some University of Texas sorority members, McLemore (1991) has argued that natural, iconic, principles constrain the interpretation of prosodic elements. Accordingly, an iconic account for attenuation might then run as follows:

> Attenuated sequences are less prominent than nonattenuated sequences. When they follow nonattenuated sequences with otherwise similar pitch patterns, they become less prominent replicas. These properties make available the defocusing and supplementing functions, which in turn 'invite' certain postposed constituents.

At the same time, McLemore emphasizes that while iconic principles may direct interpretation, they cannot strictly determine it: a role is needed for cultural conventions. Applying this in the present case, it happens that attenuation (and the pragmatic category described as supplementation) are important tropes in much Chevak narrative. In other localities they are not. For example, the narrative in Sec. 5 contains just a few clear case of attenuation and supplementation. In yet other CAY communities, supplementation is frequent but it is not marked by attenuation (Woodbury 1992). Details like these are too particular ever to follow solely from iconic principles: they must continue historical patterns of actual use and interpretation.

8. CONCLUSION

We have made some basic assumptions about discourse structure and used them to gauge the extent to which prosodic elements in individual samples of natural discourse show structure and convey meaning. The advantage at very least is their generality and broad applicability to natural discourse prosody. Prosodic Hierarchy Theory presupposes too much structural convergence, while descriptively-oriented transcription systems presuppose the importance of particular prosodic elements regardless of their importance to the distributional or pragmatic structure of the particular speech in question.

Furthermore, I hope to have shown that investigation in these terms quickly points up interesting phenomena. Most importantly, we have been able to find prosodic structuring in natural discourse, albeit of a more diffuse and ramified type than that posited by Prosodic Hierarchy Theory. Also revealed were a complex relationship of prosodic structures and elements to thematic structure; the phenomenon of prosodic macrostructuring; and the influences of iconic principles and cultural conventions on the use of prosodic structures. Once noticed and formulated in general terms (however tentatively), such findings serve as guides in further investigation.

While I have claimed some progress on the Distribution and Meaning questions noted at the beginning of this paper, the Identity question has been left nearly untouched. One part of what is needed is simply to continue to improve our understanding of the phonetics and phonology of all potential prosodic elements. Another aspect—consonant with the natural discourse oriented approach described here—is to characterize as precisely as possible the phonetic (or phonological) correlates of the significant elements of distribution and meaning in samples of natural discourse prosody, regardless of their simplicity or complexity. That is, it is necessary to know in what ways logically separate phonetic elements of prosody might be bundled together and treated as single elements of prosodic distribution or prosodic meaning; which such bundles are common in particular cultures, languages, speech event types, or idiolects; and whether their meanings show family resemblances regardless of where they occur.

REFERENCES

1. Bauman, Richard, and Joel Sherzer. 1974. Explorations in the ethnography of speaking. Cambridge University Press.

2. Chafe, Wallace L. 1980. The deployment of consciousness in the production of a narrative. In Chafe, W.L. (ed.), The pear stories: cognitive, cultural and linguistic aspects of narrative production. Norwood, NJ: Ablex.

3. Gumperz, John J. 1982. Discourse strategies. Cambridge, UK: Cambridge University Press.

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

5. Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Kiparsky, P. & G. Youmans eds., Phonetics and phonology, Vol I: Rhythm and meter. New York: Academic Press. 201-260.

6. Hirschberg, Julia and Mary Beckman. 1992. Draft report on the TOBI conventions for prosodic labeling. Bell Labs MS.

7. Hirschberg, Julia and Janet Pierrehumbert. 1986. The intonational structuring of discourse.

Proceedings of the 24th Annual Meeting of the ACL. Morristown, NJ: ACL. 136-144.

8. Hymes, Dell. 1974. Foundations in sociolinguistics. U Pennsylvania Press.

9. Hymes, Dell. 1981. In vain I tried to tell you: Essays in Native American ethnopoetics. Philadelphia: U of Pennsylvania Press.

10. Inkelas, Sharon and Draga Zec. 1990. The phonology-syntax connection. Chicago: U of Chicago Press.

11. Jakobson, Roman. 1960. 'Concluding statement: linguistics and poetics. In T. A. Sebeok ed., Style in language. Cambridge: MIT P.

12. Labov, William. 1972. Rules for ritual insults. In W. Labov, Language in the inner city. Philadelphia: U Pennsylvania Press. 297-353.

13. Levinson, Stephen C. 1983. Pragmatics. Cambridge, UK: Cambridge University Press.

14. Liberman, Mark Y., Cynthia A. McLemore, & Anthony C. Woodbury. 1991. On the nature of prosodic phrasing. Grammatical foundations of prosody and discourse: the conference. LSA Linguistic Institute, Santa Cruz.

15. Liberman, Mark, and Alan Prince. 1977. On stress and linguistic rhythm. Linguistic Inquiry 89:249-336.

16. McLemore, Cynthia. 1991. The pragmatic interpretation of English intonation: sorority speech. Doctoral Dissertation. University of Texas, Austin.

17. Michelson, Karin. 1991. Semantic and discourse factors in Oneida utterance-final phonology. Ms.

18. Miller, Corey. 1992. Prosodic aspects of M. L. King's "I have a dream today" speech. This volume.

19. Nespor, Marina, & Irene Vogel. 1986. Prosodic phonology. Dordrecht: Foris.

20. Pierrehumbert, Janet B. 1980. The phonology and phonetics of English intonation. M.I.T. doctoral dissertation.

21. Pierrehumbert, Janet B., and Mary Beckman. 1988. Japanese tone structure. Cambridge: MIT Press.

22. Pierrehumbert, Janet and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, and M. Pollack eds., Intentions in communication. Cambridge: MIT Press. 271-311.

23. Queen, Robin. 1992. Prosodic organization in the speeches of Martin Luther King. This volume.

24. Sapir, Edward. 1949. Male and female forms of speech in Yana. In Mandelbaum, David, ed., Selected writings of Edward Sapir. Berkeley: U of California Press. 206-212.

25. Selkirk, Elisabeth O. 1980. The role of prosodic categories in English word stress. Linguistic Inquiry 11:563-605.

26. Silverstein, Michael. 1976. Shifters, linguistic categories, and cultural description. In Basso, Keith and Henry Selby, Meaning in anthropology. Albuquerque: U New Mexico Press. 11-56.

27. Silverstein, Michael. 1979. Language structure and linguistic ideology. In Paul R. Clyne et. al., eds., The elements: A parasession on linguistic units and levels. Chicago: Chicago Linguistic Society. 193-247.

28. Silverstein, Michael. 1984. On the pragmatic 'poetry' of prose: parallelism, repetition and cohesive structure in the time course of dyadic conversation. In Schiffrin, Deborah, ed., Meaning, form, and use in context: linguistic applications. Washington: Georgetown U Press. 181-199.

29. Tedlock, Dennis. 1983. The spoken word and the work of interpretation. Philadelphia: University of Pennsylvania Press.

30. Woodbury, Anthony C. 1984. Cev'armiut qanemciit qulirait=llu: Narratives and tales from Chevak, Alaska. Fairbanks: University of Alaska Press.

31. Woodbury, Anthony C. 1987a. Rhetorical structure in a Central Alaskan Yupik Eskimo traditional tale. In Sherzer and Woodbury eds., Native American discourse. Cambridge, UK: Cambridge U Press.

32. Woodbury, Anthony C. 1987b. Meaningful phonological processes. Language 63:685-740.

33. Woodbury, Anthony C. 1989. Phrasing and intonational tonology in Central Alaskan Yupik Eskimo: some implications for linguistics in the field. In John Dunn, ed. 1988 Mid-America Linguistics Conference papers. Norman: U. of Oklahoma. 3-40.

34. Woodbury, Anthony C. 1992. Utterance-final phonology and the prosodic hierarchy: a case from the Nunivak dialect of Central Alaskan Yupik Eskimo. Paper given at Linguistic Society of America Annual Meeting, Philadelphia.