

The Interplay of Syntax and Prosody in the Expression of Thoughts

Author(s): Wallace Chafe

Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Pragmatics and Grammatical Structure (1997), pp. 389-401

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

The Interplay of Syntax and Prosody in the Expression of Thoughts

Wallace Chafe
Santa Barbara

There has always been a tendency in linguistics to pay more attention to aspects of language that are present for everyone to observe and less attention to its more hidden aspects: more attention to form and less to meaning. But behind the form lie thoughts, and thoughts are as much a part of language as what can be heard or (when it has been transcribed) seen. Language is fundamentally a way of organizing both thoughts and sounds, more or less independently, and associating the thoughts and sounds with each other, as suggested in figure 1.

Thoughts \longleftrightarrow Sounds

Figure 1. Basic Components of Language

There is an important sense in which the thoughts have priority over the sounds. It seems obvious, once one stops to think about it, that it is the flow of thoughts that keeps language moving, not the flow of sounds, whose function is only to express the thoughts. Linguistic form exists in the service of the thoughts, not vice versa.

The situation just described creates a problem for linguistic research. Because sounds are publicly observable and thoughts are not, it is the overt form of language that has greater research tractability, that makes researchers feel they are doing real science. But if the thoughts have priority as the driving force of language, linguistics faces a situation in which the greater functional importance belongs to the part of language that is less easy to investigate. It is a problem that has always kept us from understanding language as well as we might if things were reversed.

The sound side of language has itself two aspects. One dominates the kinds of things that are usually written down, the "segments" of language that are represented in writing systems: vowels, consonants, and syllables. The other aspect is slighted in writing systems. It falls under the heading *prosody*: the variations in pitch, loudness, duration, and voice quality that provide the support on which the vowels and consonants and syllables ride along. Curiously, we don't have a fully satisfactory term for sound minus prosody. Perhaps we are so accustomed to thinking of syntax and grammar independently of prosodic

considerations that it has not seemed necessary to have a term for non-prosodic form. Like others I will speak of *segmental* sounds, but a better term is needed. It is interesting to reflect on the fact that prosodic sounds are produced largely by the lungs and larynx, while segmental sounds are produced mainly in the mouth. Prosodic sounds have a longer evolutionary history, associated especially with the basic organization of thoughts and with emotions, and having more in common with animal calls and music, for example. Segmental sounds are evidently associated more closely with the unique evolution of the human cortex.

Again there is a difference in research tractability. It has always seemed easier to deal with the segmental aspects of sound than the prosodic. Just as writing systems have done a better job of representing the former than the latter, more linguistic research has gone into aspects of language that are represented segmentally. One reason may be that segmental phenomena are mostly discrete, whereas prosodic phenomena are more likely to be continuous. But contributing to this picture has also been the fact that the segmental and prosodic aspects of sound express different aspects of thought. There seems to be a natural affinity between segmental sounds and certain components of thought, between prosodic sounds and certain other components. This paper is an attempt at exploring these different contributions of segmental and prosodic sounds to the expression of different aspects of thoughts.

If sounds can be divided into their segmental and prosodic aspects, is there an analogous dichotomy in the realm of thoughts? The answer appears to lie in distinguishing the *content* of thoughts on the one hand from what I will call their *infrastructure* on the other. Content involves above all the kinds of things we experience through our senses in our interactions with the world around us: perceptions of events (things that happen) and of states (the way things are), as well as of the people, objects, and abstractions that participate in events and states. I have found it useful to use the word *idea* as a technical term to refer to our experiencing of events, states, and their participants (Chafe 1994: 80). So far as *content* is concerned, then, thoughts involve the manipulation of ideas in that sense.

The infrastructure of thoughts includes their organization and evaluation. It includes, for example, an organization of thoughts into content *units* of various sizes and properties, along with what I will call *signposts* that guide the flow of thoughts from one unit to another. It includes as well the evaluation of thoughts: their assignment to

different *weights* and different manifestations of *affect*.

Just as prosodic sounds have been less tractable to research than segmental sounds, the infrastructure of thoughts has been less easy to deal with than their content. We have not succeeded as well in understanding how thoughts are organized and evaluated as we have in understanding the semantics of ideas, in the sense described above. It is then interesting to note that the content aspect of thoughts is associated largely with the segmental aspect of sounds, whereas their infrastructure is expressed to a large extent by prosody. Language uses pitch, loudness, duration, and voice quality to convey the organization and evaluation of ideas, though often in concert with segmental phenomena. Figure 2 is an attempt to sort out these components of language. Those with greater research tractability have been placed in boldface on the right of each branch.

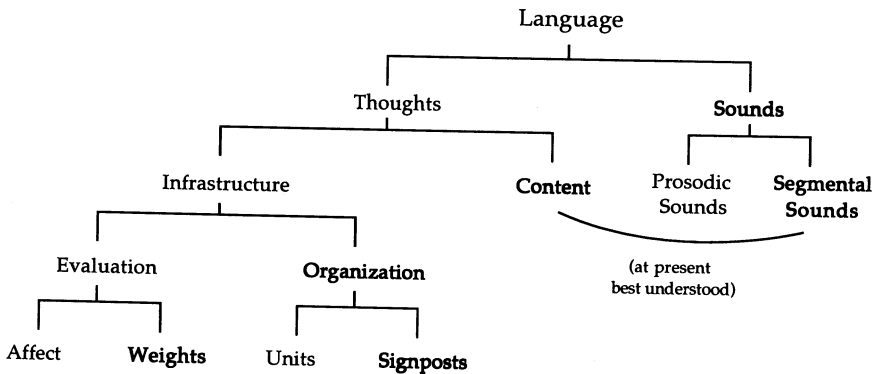


Figure 2. More Detailed Components of Language

I will limit the discussion here to the nature of the infrastructure, and particularly to the question of how its ultimate components (signposts, units, weights, and affect) are expressed by prosody, while at the same time considering some of the contributions made by segmental sounds. In order to illustrate some of the ways in which these two aspects of sound are distributed, I will make use of the following excerpt from a conversation that I (speaker W) recently had with a woman (speaker M) who was talking about her childhood on the Allegany Reservation in New York State. Divisions into topics and subtopics (discussed below) are indicated with heavy and light horizon-

tal lines respectively. M had just been talking about how her father managed a Studebaker dealership.

1	M	Yeah,	
2	M	[he did] thát,	
3	W	[/??/?/]	
4	M	.. and he .. rùn the garáge,	
5	W	.. yeah,	
6	M	and= Róscœ was the mechánic,	
7	M	.. cause he jùst frèsh out of cóllege,	
8	W	wéll.	
9	M	... and he .. he tòok càre of àll the bók work.	
10	W	... Yéah.	
11	M	.. Uh= Róscœ did.	
12	W	.. Well góod.	
13	M	My óldest bróther.	End Subtopic INTRODUCE ROSCOE
14	M	... He was réal .. réally níce,	Begin Subtopic ROSCOE'S NICENESS
15	M	and wéll réad and,	
16	W	Mm,	
17	M	... and èverybody líked hím and,	
18	M	... é èverything went over góod.	
19	M	... And hè was the búsiness mánager,	
20	M	... and so=,	
21	M	.. hím and Dád got along réal fine~	
22	M	réal góod~	End Subtopic ROSCOE'S NICENESS
23	W	Hmh.	and End Topic STUDEBAKER DEALERSHIP
24	M	... So hé was a Sny'der,	Begin Topic MYRTLE'S SIBLINGS
25	M	shé=	and Begin Subtopic FIRST HUSBAND
26	M	.. My móther was màrried twíce.	
27	W	... Yéah.	
28	M	.. Her fírst hùsband was-	
29	M	.. Jèwett Sny'der from-	
30	M	... Cattaráugus Reservàtion.	
31	W	.. Mhm,	End Subtopic FIRST HUSBAND
32	M	.. So she had uh-	Begin Subtopic FIRST SIBLING SET
33	M	... she had uh-	
34	M	Róscœ,	
35	M	... and= Jím,	
36	M	... and my síster,	
37	M	... and .. Mílford.	
38	M	She had four chídren.	
39	M	... From her fírst hùsband [.. see?]	
40	W	[Mhmh.]	
41	W	Mhmh,	End Subtopic FIRST SIBLING SET
42	M	... So thén uh,	Begin Subtopic SECOND HUSBAND
43	M	... after hé-	AND SECOND SIBLING SET
44	M	... after hé passed away,	
45	M	why then,	
46	M	when shé got acquainted with Dád,	
47	M	why then,	

- 48 M ... she had just-
- 49 M ... mé,
- 50 M and Bábe,
- 51 M and Vín;
- 52 M .. there's thrée of us.
- 53 M .. [So],
- 54 W [mhmh,]
- 55 M .. the Símcx. End Subtopic SECOND HUSBAND
- 56 W ... Mhmh, AND SECOND SIBLING SET
-
- 57 M So uh- Begin Subtopic HARMONY
- 58 M ... but wé never thought we were any différent.
- 59 M ... We álwáys thought we were a=ll-
- 60 M ... [all the] sáme,
- 61 W [No différent.]
- 62 M you know,
- 63 W Mhmh,
- 64 M .. wé never thought we had différent fáther or différent-
- 65 M .. anything élse,
- 66 W mhmh,
- 67 M .. Néver thóught of that.
- 68 W .. mhmh,
- 69 M .. Néver néver thought of it.
- 70 W .. mhmh,
- 71 M Cause we áll= ... were háppy and,
- 72 W .. mhmh,
- 73 M and .. we were ráised togéther,
- 74 M and [éverything],
- 75 W [mhmh],
- 76 W well thát was góod.
- 77 M .. It was áll ríght;
- 78 M we néver thought ánything dífferent of;
- 79 M ány of the chýldren;
- 80 M you know.
- 81 W Yéah;
- 82 W that was góod.
- 83 M .. But it was níce,
- 84 M .. we all got along good together,
- 85 W ... Hmh, End Subtopic HARMONY
-
- 86 M .. but thát was a long long ... tíme agó. Begin Coda LONG TIME AGO
- 87 W ... hmh.
- 88 M .. I thínk of it nów, End Coda LONG TIME AGO
- 89 M it's been a ló=ng tí=me. End Topic MYRTLE'S SIBLINGS
-
- 90 W Whát was that about a- Begin Topic AB ROTH PLACE
- 91 W ... remember there was a-
- 92 W ... some mán that,
- 93 W ... showed úp at the dóor thére,
- 94 W ... at the Ab Róth place.
- 95 M Oh thát was up thére.
- 96 M At that Ab Róth place.

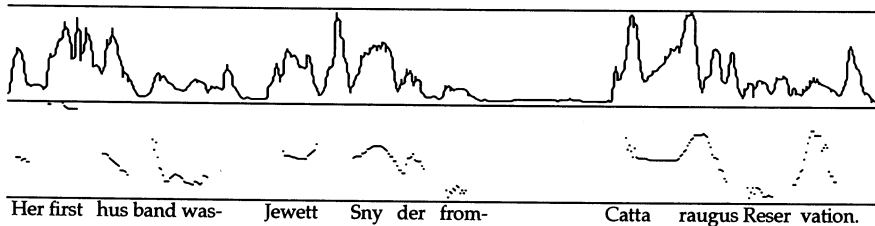
Organization: Units. Thoughts are organized into units at several levels. Space limitations will force me to discuss and illustrate other aspects of the infrastructure in a sketchier fashion, but this aspect demands a more extended treatment that will provide a background against which the others can be more briefly placed.

In Chafe (1994) and elsewhere I have emphasized the way the stream of speech naturally segments itself on prosodic grounds into short phrases of a type I have been calling *intonation units*, the boundaries of which are identifiable on the basis of various prosodic criteria: pausing, distinctive final intonation contours, baseline pitch changes, initial acceleration and final deceleration, boundary changes in voice quality, etc. These intonation units are usefully hypothesized to be expressions of individual foci of consciousness, minimal units of thought organization. Each line in the excerpt above represents a separate intonation unit.

If each intonation unit expresses a momentary focus of consciousness, the content of such a focus is variable, ranging from a brief regulation of the interaction as in 1 (*yeah*), to an event as in 4 (*he run the garage*), or a state as in 6 (*Roscoe was the mechanic*), or a participant in either an event or state as in 13 (*my oldest brother*). Every language has its own way of assembling the content elements within an intonation unit, and the result provides the basis for what is traditionally regarded as *syntax*. There is no single syntactic structure that is associated with a focus of consciousness, as can be seen by comparing 1 with 4 or 13. But when the content elements organize an event or state, as often happens, they are expressed in a syntactic *clause*, as in 4 or 6. More than half the intonation units in a typical sample of English conversation form clauses, but the proportion varies with speakers and genre, as well as across languages. Chinese, for example, appears to devote fewer intonation units to clauses (Tao 1996).

Instead of being constrained by a single syntactic structure, the foci of consciousness that are expressed in intonation units are constrained by an inability of speakers to verbalize more than one new idea at a time, more than one idea that has been activated from a previously inactive state (Chafe 1994, chapter 9). This constraint is well illustrated by the sequence 28-30, represented here with amplitude above and pitch below (displayed with the help of Eric Keller's program "Signalize"):

28-30



In the entire sequence there were three new ideas: that of the first husband, of his name, and of his place of origin. Because of the one-new-idea constraint they were necessarily expressed in three intonation units, none of which was a clause. A syntactic analyst would not segment the sequence in this way. 28-30 provide an excellent example of the difference between the grammar of ordinary spoken language and the kind of grammar we have learned to expect from several millennia of studies based on written language.

As people talk they constantly face the problem that individual foci of consciousness have a capacity too small to accommodate the amount of information they might wish to focus on, an amount comprising what I have been calling a *center of interest*: a kind of superfocus of consciousness that attempts to go beyond the constraints imposed by the evolution of human information-processing abilities (Chafe 1994: 139-44). When speakers decide, as they proceed from one focus of consciousness to the next, that a larger center of interest has been adequately verbalized, they signal that fact with a falling pitch perceived as sentence-final prosody. At the same time, ideally, they close off the syntax too, ending a complete sentence. This closure is often briefly acknowledged by an interlocutor, as it was, for example, with the *yeah* in 10.

It happens frequently, however, that these two devices, the prosodic closure and the syntactic closure, fail to coincide. Evidence from repeated tellings of the same experience suggests that sentences do not reflect a consistent unit of cognitive processing, but rather an on-line judgment that something has been completed (Chafe 1994: 143-44). It is a judgment that has both prosodic and syntactic consequences, but there may be difficulty in pulling the two together. This speaker finished a sentence prosodically and syntactically at the end of 9 (*he took care of all the book work*), but then tacked on two clarifications in 11 (*Roscoe did*) and 13 (*my oldest brother*), having decided she had not quite verbalized everything necessary. Functionally, the entire se-

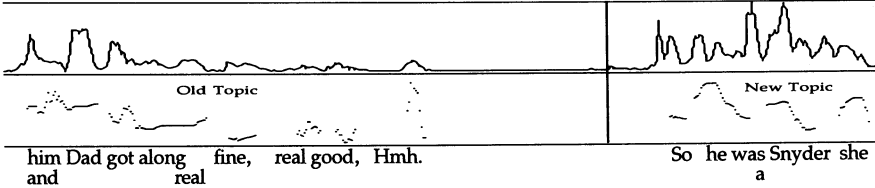
quence 1-13 expressed a single center of interest, but the syntactic result was less coherent. This kind of mismatch between prosody and syntax is typical of ordinary speech.

Aside from syntactic closure, there can be other manifestations of the closure of centers of interest. For example, listeners often acknowledge their awareness of such closures by inserting backchannel expressions. I mentioned above the *yeah* in 10, but this listener was particularly fond of saying *mhm* or *mhmh*, as in 31, 40-41, 54, 56, and so on. Other segmental sounds may be used to initiate new centers of interest, for example the word *so* that occurred at the beginning of 24, 32, 42, and 57.

Although a speaker decides on the run, as it were, to close a center of interest, larger discourse *topics* constitute more stable units of mental representation. They qualify as another kind of *idea*, in the sense that they too are a way of organizing content that remains stable throughout successive activations and verbalizations of the same material. They differ from the ideas expressed in the smaller intonation units in that they cannot be in fully active consciousness all at once. Most of spoken language consists of speakers navigating through these semiactive topics, illuminating first one area and then another with a series of fully active foci expressed in intonation units.

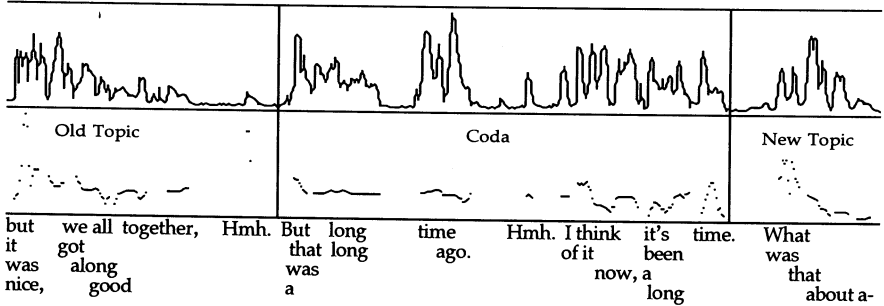
Like foci of consciousness and centers of interest, topics also have prosodic manifestations. They typically begin with heightened volume, pitch, and tempo, and peter out when speakers decide they have been adequately verbalized. Often they end with creaky voice. Often they are followed by significant pausing before a new topic is introduced. There are two topic boundaries in this excerpt. The first insertion of a new topic occurred at 23. The following display suggests the petering out that preceded this boundary, and shows the significant pause that occurred before the next topic was taken up with renewed vigor in 24:

21-25



The other introduction of a topic in this excerpt occurred at 90. It can be noted that the prosody of the portion labeled Old Topic here (83-85) matched closely the portion with the same label in 21-23, differing only in the more compressed scale of the display. This time the petering out of the previous topic was followed by wistful coda, where there was an iconic lengthening of the words *long time* in 89. At this point the listener recognized an opportunity to jump in with a topic-changing question:

83-90



The entire excerpt contained within it one complete topic, extending from 24 through 89. As speakers develop a topic each focus of consciousness does not jump around randomly, but follows a trajectory based on some already familiar schema. In this case the highest level divisions of the schema can be seen as a sequence of thesis, antithesis, and synthesis. First came a subtopic involving the mother's first husband and the children of that marriage, then a second subtopic involving her second husband and the children of that marriage, and finally a subtopic dealing with the harmony that existed among the step-siblings.

The first-husband subtopic ended with 39, followed by the listener's segmental recognition of the boundary with two *mhmh*'s in 40 and 41. The second-husband subtopic ended with the afterthought in 55, preceded and followed by two *mhmh*'s in 54 and 55. The harmony subtopic had a long drawn-out ending, with much repetition and elaboration of a single idea that was first verbalized in 58 at the beginning of the topic: *we never thought we were any different*. Everything that followed was a repetition of that idea, viewed from different angles. It is interesting to see how many *mhmh*'s were produced by the interlocutor during this sequence, with six repetitions that began in 63

and were supplemented by two monosyllabic *hmh*'s in 85 and 87. The listener had already caught the gist of the subtopic in 63, and continued to acknowledge his understand of it throughout the elaborations that followed.

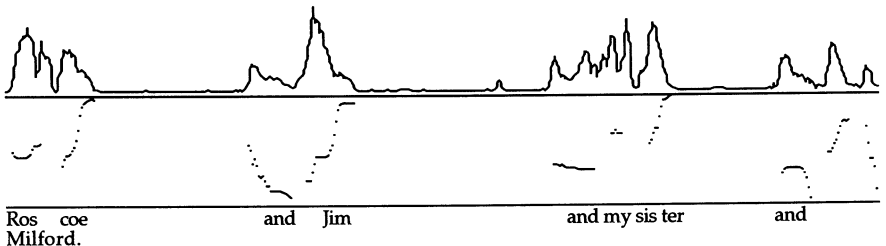
Topics often begin with a temporal setting, orienting what is to follow. This one *ended* with a temporal orientation in 86-89 (*but that was a long long time ago. I think of it now, it's been a long time*). The speaker had already established a frame involving the family of her childhood, and here she provided a metacomment on the experience of reactivating a remote past.

In summary, the organization of thoughts into foci of consciousness, centers of interest, subtopics, and topics is to a large extent expressed prosodically, but is supported by segmental phenomena such as the use of *discourse markers* (Schiffrin 1987) like *so*, of backchannel responses, and of syntactic closure.

Organization: Signposts. The organization of thoughts involves not only the presence of thought units at various levels of organization, but also connections between those units. The most obvious way way in which a connection is marked by prosody appears at the ends of intonation units, where amplitude, duration, and voice quality all play a role, but where the terminal pitch contour is especially salient. This contour may be either forward-looking or backward-looking; that is, it may anticipate something to come, or may show that something has arrived at closure.

The sequence 34-37 provides a simple example of how prosodic and segmental sound can support each other in showing connections between foci of consciousness:

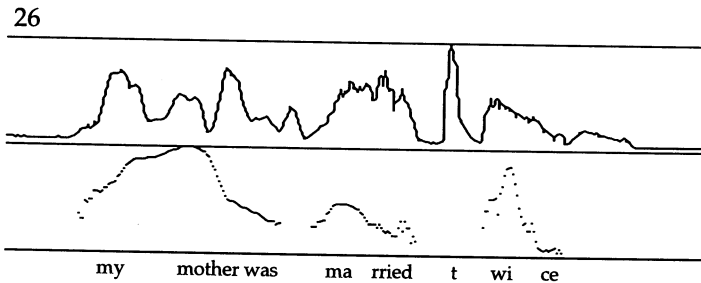
34-37



The rising pitch on all but the last of these four intonation units indicated that "this is one in a sequence of parallel items" (a list intonation), followed by closure with a falling pitch on *Milford* when the list

was complete. There was a symbiosis between the prosody (anticipating more to come) and the conjunction *and*, which, looking back rather than forward, connected its own focus to the one that immediately preceded.

Evaluation: Weights. Turning now from organization to evaluation as another component of the infrastructure of thoughts, we can note briefly how both prosodic and segmental sounds support each other in the attachment of differing degrees of prominence to different thought elements. In 26 a new referent, *my mother*, was given higher pitch as shown in the following display, but was also expressed segmentally with a full noun phrase. Subsequent mentions of the same referent were pronounced at a lower pitch, and were expressed segmentally with the pronoun *she* as in 6 (*she had four children*).

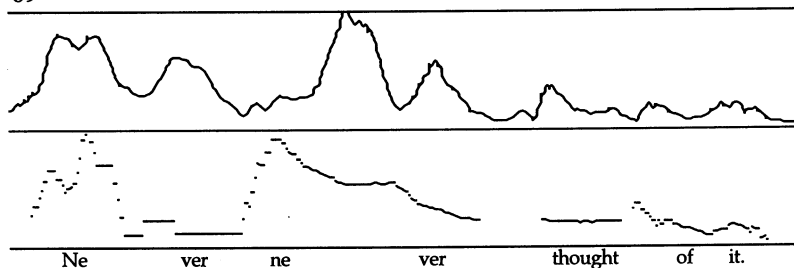


Evaluation: Affect. The other aspect of evaluation included in figure 1 was affect. It is obvious that one of the major functions of prosody is to express emotions and attitudes, but those are properties of thought that remain the least tractable of all as suggested by their placement on the far left in figure 1. We still don't know very well how to describe emotions, much less how to relate specific prosodic patterns to them. One reason may be that emotions are gradient and not discrete. Another may be that they are contextually determined, so that, for example, the same physiological state might be interpreted as excitement in one context, fear in another. Although all speech may have some affective content, one wonders if there is for each speaker a baseline emotional level from which he or she may sometimes depart. I will focus here on what seem to be two cases of such a departure in the excerpt above.

In 69 there was an expression of strong commitment that was expressed by an increased pitch range extending up to 263 Hz (this speaker's peak was usually around 215 Hz), along with a falsetto voice

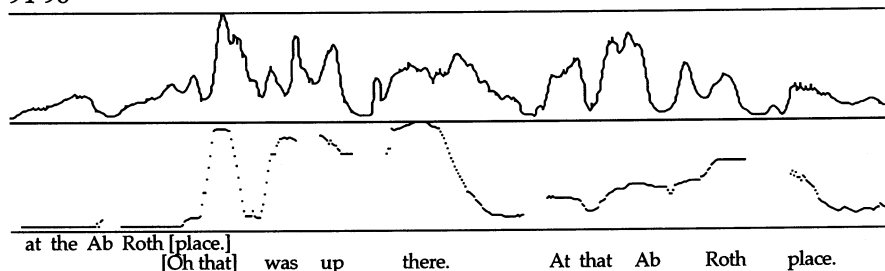
quality. Segmental phenomena also contributed to the effect of heightened emotion, especially in the repetition of the word *never*. (In fact, *never* occurred a total of six times during this subtopic.)

69



A different emotion was expressed in 95 at the introduction of a new topic, where the speaker voiced her recognition of that topic along with surprise at its introduction. Again there was an increased pitch range, this time extending up to 268 Hz. At the same time the emotion was manifested segmentally in her use of the word *oh*:

94-96



This has been a quick look at a few examples of ways in which prosodic sound expresses the infrastructure of thought, while interacting in that function with segmental sound. The total contour of pitch, amplitude, duration, and voice quality that emerges is determined in multiple ways, all of which operate in concert with segmentally expressed aspects of thought to produce the total effect that language achieves. Understanding better how these phenomena are distributed will require careful attention to details of real language, as well as an open mind and a willingness to explore multiple pathways of investigation.

References

- Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: The University of Chicago Press.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Tao, Hongyin. 1996. *Units in Mandarin Conversation: Prosody, Discourse, and Grammar*. Amsterdam/Philadelphia: John Benjamins.