

Subordinate Tones of Voice

Author(s): Carol Woodall

Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society (1984), pp. 414-424

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.

Subordinate Tones of Voice
Carol Woodall
University of New Mexico

1. Issues. For some years now the notion that grammar and intonation constitute separate sub-systems of English has been championed by Dwight Bolinger. In 1972 he wrote "Whether one tries to set up prosodic rules for syntax or syntactic rules for prosody, the result is the same: two domains are confused which should be kept apart" (p. 644). Elsewhere he proposed using the term "cooperation" in preference to "interdependence" as an appropriate descriptor for the interrelationship between intonation and grammar, adding that we could probably assume that "if a given intonation is possible at all, it will be possible with any syntactic category" (1982, p. 4). The number of supporters for this view has grown in recent years. D. R. Ladd and Mark Liberman, among others, have adopted modified versions of the position. Ladd has in addition recently stressed the critical importance of separating structural descriptions of intonational phenomena from the functional and semantic goals they effect (1983). He has built on the work of researchers such as Pierrehumbert, 't Hart (with colleagues) and others, particularly in adopting their standards of phonetic explicitness, while attempting to remedy the inability of these approaches to express phonological and functional generalizations. As I am sympathetic to both of these goals as well as to the maintenance of a hearty if not complete separation between intonation and grammar, these are the broad positions supported by this paper.

Numerous non-grammatical functions have been proposed for intonation, among them emotional expression, semantic contrast, definition of social relations, and discourse cohesion. It is the last of these we will explore here. We will in fact consider evidence that the production of discourse cohesion is the most pervasive use of pitch -- that discourse concerns are continuously and unavoidably expressed in speakers' intonational decisions. These concerns are not in competition with other pitch functions but define the ways those other functions relate to one another in the production of coherent messages. States of emotional excitation and highly informative speech are no more exempt from requirements of coherence than are other, less marked utterance types, and thus no conflict is involved in accepting these roles simultaneously. What we will attempt to do here is uncover a few of the structural devices called on to relate individual pitch gestures in larger segments of meaning. Many of our findings are consistent with those reported by Chafe and summarized in his hierarchical and flow models of narrative construction.

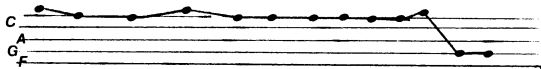
2. Method. Given the uncertainty in the field as to which entities are significant and thus which notational devices should be adopted, I have chosen a notational system which I believe begs as few theoretical questions as possible. The system that seemed capable of recording pitch levels as reliably and consistently as possible was a kind of musical transcription; the system also provides an intersubjectively comparable base. A group of four of us (two musicians and two

linguists) listened to recorded speech samples and then (humming along if necessary) submitted the pitch record to a musical staff, comparing perceptions with notes on a keyboard. We hoped in this way to retain the subjective essentials of pitch interpretation.

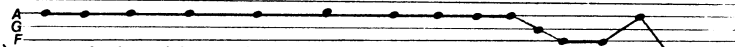
During the initial stages of data analysis I was surprised to find speakers or listeners (or both) regularizing pitch behaviors I had expected to be more random. Specifically, speakers were heard to rise and fall methodically to "notes" they had previously used for similar purposes (i.e., as rise or fall targets). These speakers were not returning to variable embodiments of a concept "high" or "low"; they were returning, for example, to F, F#, and G, and their chosen targets made semantic sense in relation to one another. An onset of G might introduce a new discourse direction, whereas an onset one step lower, at F, would announce follow-up material. Initially it was suspected that only very gifted speaker-listener pairs would be able to track tones in this orderly manner. That suspicion turned out to be unfounded, as all pairs produced similarly patterned regularities. (This does not mean that listeners agreed on every point of the transcriptions; we did not. Listener agreement/disagreement is currently under study.)

3. Upper bounds. Many of the regularities in pitch patterning seem best explained by hierarchical unit-building motivations. The highest, presumably most informative pitches produced by each speaker were located at sentence onsets — if not on the first syllable, then within the first contour. Even where speakers made emphatic points later in the sentence, pitch levels accompanying these points rarely surpassed the standard set at the onset of that sentence. Moreover, some speakers returned to the same note or one of a small set of notes repeatedly for initializing sentence-type utterances. Speaker 7, a female psychology lecturer who had received a student award for her lecturing, was extremely methodical in this respect. Over half of her sentence onsets were recorded precisely at D#, and the others fell within one half-step to either side of this level. Presumably the listeners appreciated systematicity in this form. Although not as consistently as speaker 7, all speakers exhibited preferences for particular notes in onset position. These were few in number (3-5) and were spread over no more than about two steps on the scale.

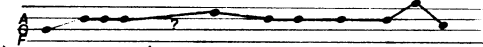
We would like to claim that these regularities stem from the speakers' use of pitch to denote major discourse sectors. We must acknowledge, however, an alternate explanation for high pitch at sentence onsets. It has been proposed that initial high pitch on sentences derives from increased subglottal pressure following inhalation. This makes high onsets an involuntary feature of the respiratory system and possibly not a semantic gesture. In fact, this explanation for high onsets no doubt has merit. On the other hand, it fails to acknowledge speakers' apparent abilities to discriminate and recall pitches with greater delicacy than simply recognizing them as "high". The following sequence of sentence onsets from an anthropology lecture illustrates the discrimination pattern found commonly in our data:



1) I'd say talk first about Eskimo economy....



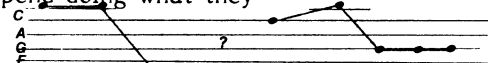
2) one of the things that stands out about aboriginal eskimos....



3) the areas_{a/} in rich areas like Alaska.... b/

Note that peaks rise to D in sentence 1 but only to A in sentences 2 and 3. This represents a 2.5-step lowering between 1 and 2 but no equivalent lowering between 2 and 3, ruling out the interpretation of drift. If the sample could have been longer, it would be obvious that onsets continue at and slightly to either side of A, exhibiting no further falls. Nothing as high as the initial D occurs in the transcript until the speaker begins summarizing his major point, and there we find:

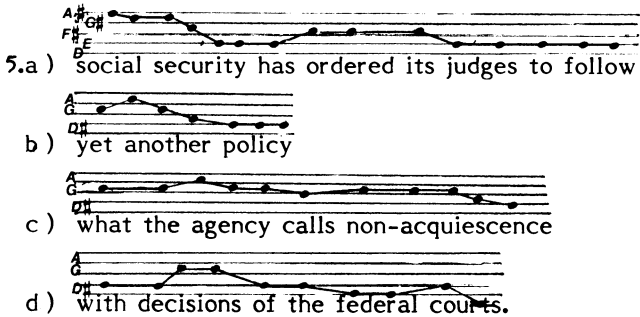
'Pygmies do not spend their lives working; they work no more than a three-and-one-half to four-day week, and the rest of the time they spend doing what they'



4) want to do_{a/} and and sleeping a lot_{b/}

These observations summarize a contrast developed throughout the lecture between our conception of the good life and the conceptions of certain "undeveloped" cultures elsewhere on the globe. As there were no instances of pitches higher than D in the transcript, we see clearly that this speaker chose to dispense his highest notes on initializing and summary accents. The choice implies a significance for major-sector onset marking which is almost equal (given correction for S-internal downdrift) to that of the lecture focus. The two demarcators together form upper bounds over this coherent sector of discourse. The specific functions of these markers would be undiscernable within systems that recognize all high pitches as equal -- and the respiratory explanation is among these. We can accept it as a partial explanation for onset heights while still recognizing semantic functions for these pitches at a more precise level of analysis.

4. The single drop. We noted that in the above example the extra-high paragraph marker was adjusted downward in a single step -- at the onset of the second sentence. Subsequent sentence onsets were approximately equal, indicating that the speaker had purposely deviated from a pattern in order to produce the one-time marker. A similar configuration was often found sentence-internally. We observe it in the following sentence produced by Jessica Savitch during a televised news report:



The initial pitch at A# is not repeated, although pitches on "another" and "agency" are only one-half step lower. The second contour peak is adjusted downward to a G, after which all remaining peaks are uttered at G or A. An abstract of peak heights illustrates clearly the adjustment pattern:

Peak abstract
sentence 5



The shape of this adjustment has in common at least two features with that among the onsets of a cluster — the single high pitch initially with subsequent highs that never quite reach the onset level again, and smaller variations between subsequent peaks. The presence of these variations suggest that the hierarchizing process might be repeated at lower levels. There is no downward drift of the type believed by many researchers to characterize English statements. A few examples of downdrifting sentences were recorded but they were much less common than those with a single downward adjustment. Our sentences did of course exhibit terminal falls, but these usually were realized on a single word or pair of syllables in the last contour. This finding was a surprise. I believe the phenomenon is related to the relatively long sentences produced by most of our speakers and the need to keep pitches in a viable production range. Downward drift was controlled largely by returns to normative levels at the ends of contour phrases.

5. Contours and phrases. "Contour" and "phrase" have evolved into technical terms during this analysis. Contour is the name given to any pair of directional pivots, such as a "rise-fall" or a "fall-rise". The rise-fall contour shape was used far more frequently in our data, and thus most of my norms and interpretations are derived from it. More data will have to be collected (probably from different discourse types) before claims can be made about the similarities between these two contour shapes. Here we will refer largely to rise-fall events when mentioning contours. Rises and falls are viewed as accomplishing different functions here as in many descriptions of intonation, although the functions are demarcative and relational rather than semantic in some other sense. Generally speaking, rises are associated with significance and falls with boundaries.

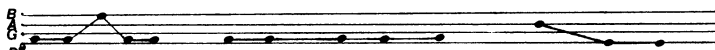
We have seen onset "highs" used to signal new discourse sectors and new sentence-type units. These highs may be viewed as rises from a norm

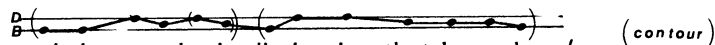
level which is calculated quickly during an utterance production. Classifying them as rises will allow a consistent treatment of these highs with other rises that occur sentence-internally. Rises within a sentence are correlated with informationality. As Bolinger (1983) says, we go up on what interests us. Or alternately, as suggested by Brazil, Coulthard, and Johns (1980) in their more socially oriented analysis, we use high key for information we want to be taken as contrastive (and thus informational). Falls are viewed as demarcating sectors of speech. These sectors come in many sizes; the basic ones can be as short as syllables (rare) or as long as several words. All are defined by the falls that terminate them. Basic sectors are called contours in this study. A contour occasionally rises and does not fall but levels off (as in Bolinger's B accent). These gestures are also counted as complete contours, and such flattening serves the same demarcative function as a fall. Flats or "planes" as I refer to them can be formed by any two or more syllables uttered on a common pitch.

Earlier we interpreted onset rises as demarcating sectors of speech such as sentences and sentence clusters. Now we must modify this interpretation slightly, since it is falls, not rises, that are principally demarcators. Onset rises, like others, should imply significance. They do, but they point to significance in environments where speakers' points almost never occur. The rise message is thus interpreted as applying to a string and not to the lexical item on which it occurs. The principle is similar to that of using capital letters both to signify personal names and sentence onsets.

The critical relational features of any contour are its onset level, its peak level, and the level to which it falls. We will consider below examples of how these pivot elements serve to define a subordinate relational network within a discourse.

6. Open and closed contours. First, however, we need to consider one functional property of the contour itself, separate from its individual rise and fall components — its open or closed status. A contour that begins on C, rises to E, and then falls again to C can be considered closed. The pitch has returned to the level from which the deviation occurred, signaling an intention to continue as before. Examples of this contour type are extremely common in our data. A few are included here:

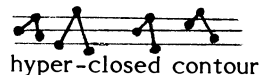
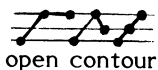
6)  in particular *a/* if we're dealing with *b/* tree structures *c/*

7)  and the very basic distinction that he makes *a/* (*contour*) is between the explicit and the implicit *b/* (*ex. 7: whole step scale*)

The contours in both 6a) and 6b) allow for local accents which appear as isolated events. There is a consistent starting level from which these deviations depart and return in serial fashion. Their status as separate gestures is made clear intonationally. In 7a), the planes and contours

are somewhat less transparent as separate gestures, but we can separate them from one another by noting the sequences of B's in initial position and then 'noting' where the note B occurs again in the sentence. It is at the onset of the next important contour. We further note that the NP completion on "makes" is marked intonationally by a return to the B on which it began. We can say that the NP has been marked as a closed discourse constituent by the application of a broad pitch contour over the entire string. The contour is counted closed on the basis of its return to the note on which it began. Note that the grammatical phrase is actually divided into two parts by its pitch contours but that the first part, which ends on the word "basic", does not return to B. This contour is considered "open" — pending additional material for its completion. Applying a closed contour to a string of material gives in a kind of autonomous status. It announces that the speaker wishes the former constituent to be considered finished, and that a new element will follow. I think it effects on a smaller scale the same type of closure that characterizes sentence completion.

The phrase is a string of contours bounded in much the same way as individual closed contours except that it is hyper-closed. The fall that forms a phrase boundary must be to below the onset level of the unit it is closing. Phrases formed by these hyper-closures allow still another hierarchical level of organization to discourse components. Phrases seem to me to be the entities referred to as "contours" by many linguists who take the traditional (British) contour approach to intonational analysis. Such approaches have never produced, so far as I know, a definition of the contour, in spite of the widespread use of the term and the concept. If the phrase as described here continues to stand up under scrutiny, it could prove to be an important entity for intonational analysis, since important adjustments seem to be made at phrase boundaries. Pitch norms may be reestablished at phrase boundaries if they have wandered too far up or down. This is possible because the material following a hyper-closure is not heard as relating to preceding material as part of the same unit. A fall this size apparently receives the perceptual interpretation "break". Schematic representations of the three contour types follow.

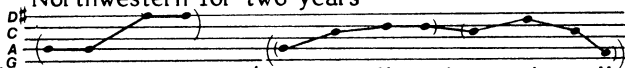


7. Tones in sequence. We will try to apply the basic concepts outlined above to a few brief texts to see how they contribute to discourse structure and cohesion. While doing this, we will compare the effectiveness of these descriptive concepts with a few others proposed by Crystal and Quirk in 1964. Theirs was some of the earliest work in this area. C & Q noted that tone units in sequence often form distinctive patterns and proposed that "even if such relationships did not force themselves on our attention through configurations in substance itself, one might set up the existence of higher-order patterns (above tone units) as a hypothesis in view of the hierarchic structure of prosodic features that it is found necessary to postulate for other

aspects of speech data" (p. 52). They insisted that subordination be defined solely on prosodic grounds, identifying what they considered the primary characteristic of the subordinate tone-unit: "its pitch contour, while having a complete and independent shape within itself, falls broadly within the total contour presented in the superordinate tone-unit" (*ibid.*). Such units could either precede or follow the superordinate nucleus, singly or in combination.

Certain similarities between this view and the hierarchical approach being investigated here are at once obvious: both subscribe to the notion that contour sizes can be used to imply subordination between segmental elements; both attempt to analyze prosodic data on its own terms. In fact neither succeeds fully in this, but I think the effort should be made as far as possible. There are also differences in the two views -- the greatest being the overall conception of how pervasive relational pitch gestures are within linguistic constructs. There are some less substantial but still important differences in the descriptive devices proposed by the two approaches as well. These will be considered conjointly with a few lines from an old-time medicine show pitchman, Doc Bloodgood:

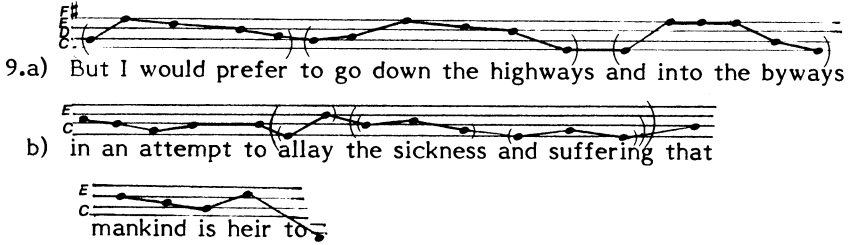
3.a) actually I'm really not a doctor; I did attend
Northwestern for two years



b) I'm not licensed b/ I'm not allowed to make calls c/

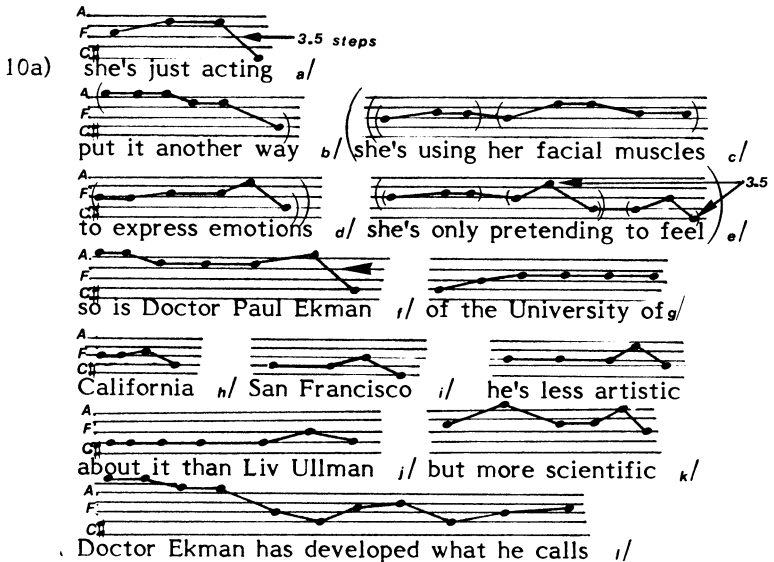
In (b) we have clauses in sequence -- temporally and semantically. The first receives a rising or open contour, the second a closed contour. The open contour on (b) signals an incomplete constituent and the speaker's wish to have the following matter considered part of a unit with it. By my reckoning, the contour on (c) implies subordination to (b). Both begin on the same note, A, but whereas "licensed" rises to D# (and stays two syllables, giving the peak extra prominence), "make" rises only to D for one syllable. This interpretation is in keeping with the semantics, which propose a cause-effect relationship between (b) and (c): I'm not allowed to make calls because I'm not licensed. The unitary status proposed for (b-c) may derive from the fact that both elaborate on the earlier statement "I'm really not a doctor": 'Here is the reason I can't claim to be one and the consequences of not being one.'

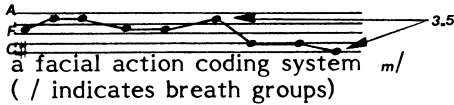
Crystal and Quirk's stipulation that contours be directionally similar to be considered pairs will not permit the interpretation of (b) and (c) as subordinate relatives. The contour on (b) will almost certainly be considered a rise and that on (c) a fall. This is unfortunate, because the first part of (c) -- "I'm not allowed" -- does actually meet the directional requirement by matching the rise in (b). It is also appropriately reduced, creating an even stronger case for subordinative status. But it cannot be admitted. Similarly, in the following example, also from Bloodgood,



the words "highways" and "byways", obviously semantic coordinates, are interpreted intonationally as coordinates by being raised and lowered to exactly the same notes. The set of potential coordinates following -- "sickness" and "suffering" -- are given different intonational treatment. The second of these words is both lower and narrower in contour than the first, indicating subordination as in example 8. This subordination cannot be admitted using C & Q's criteria either, if by having the subordinate contour fall within the superordinate they would disallow the low extension on "suffering". I think the contour similarity is in this case creating a subordinate relationship which is better accounted for by the matching contour shapes, descending peaks, and return to B on "suffering", which creates a unit of the sequence "allay the sickness and suffering".

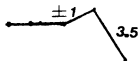
8. An analysis. We will see how the concepts of open and closed contours and of tone-matching are applied to semantic interpretation in the following example sentences. The speaker, Charles Osgood, is introducing a psychologist whose work is to be reviewed on Walter Cronkite's "Universe". Osgood's voice is heard while views are shown of Liv Ullman in "Scenes from a Marriage":





The highest pitches (A and B) are used for "put it" in b), "so is" and "Ekman" in f), "more scientific" in j), and "Doctor" in k). Of these all are sentence onsets except "more scientific" — which soon reveals itself as one of Osgood's main points. Note that the high reached on "more scientific" is equal to the onset high for the preceding two sentences, which places it on a high local norm. It is higher, however, than its own sentence's onset. Onset (i) is lower than the others in this speech sector by 2.5 steps — an unusually large variation. Its content, however, was clearly meant as a humorous aside by Osgood, as it accompanied footage of Ekman making faces one after the other. On (j) he returns to the discourse theme, and the high norm at A is readopted. Onset (a) is at a low F and reflects the status of that sentence as a completion on a preceding set of remarks. The onset in (b) ranks as a new cluster marker due to its rise from the preceding onset. In sum the speaker has used his high pitches to divide and punctuate the speech string into sentence-like units and larger.

Not only are the onsets of sentences signaled intonationally; terminations also signaled — perhaps more dependably than onsets. Sector (a) ends in a fall from G to C, a 3.5-step fall across the short span of two syllables. This, as we have mentioned, was the preferred method of closing sentences for most of our speakers. The most typical contour is much the same as that in (a) — a flat body rising about 1 step and then falling 3.5 steps, forming the pitch equivalent of this figure:



Occasionally the falls were spread over several syllables, although this was less common. An example of a gradual falling termination is available in (e). Whether the fall was abrupt or spread it fell a characteristic distance, 3.5 steps, a large majority of the time. This falling interval was used so regularly in conjunction with statement terminations that I now regard it informally as a morpheme. It is likely, since the interval formed by 3.5 steps is a half-octave and is thus harmonious and easily learned, that this pitch interval has been adopted in English as a terminal marker. In this sample it is used on (a), (e), (f), (j), and (l): every sentence. (To count intervals allow .5 steps for each line and .5 for each space on the graphs.) Recognition of a learned interval in terminal environments could account for the often expressed intuition that English statements have a "distinctive kind of falling pitch" (Chafe, 1979).

Not only are the onsets and terminations regular in their behaviors in this passage; rise and fall levels on individual contours are extremely methodical. Note that the short introductory "Put it another way" is given its own contour which ends on D. The two contours in c), both open, await closure which occurs at the end of d) — on the same note as the preceding closed contour. Sector e) includes 3 contours,

the second of which closes again on D, just prior to the terminal contour. Thus all phrase boundaries in this sentence are made on a single note with the exception of the terminal closure, which is lower.

A glance at the rise levels shows them to be no less regular. After the initial contour, every phrase contains peaks at the same high G. The resemblance between rise patterns in the sentence interior units is striking. On each of the sectors (1) "she's using her facial muscles", (2) "to express emotions", and (3) "she's only pretending", there is an onset at E and a subsequent rise to F and then to G. The similar number of syllables within each of these sectors in combination with the pitch patterns gives a strong rhythmic integration to the whole. On each side of this set of structural triplets is a slightly aberrant contour; the first marks the onset and the last the termination of the encompassing unit. Since d) and the first phrase of e) begin and end on precisely the same notes, they are classified coordinate structures — elements given equal significance as subparts of the sentence. The two contours in c), however, differ slightly. They use the same onset notes and together form the same F-G rise pattern, but the fall, is short of closure. This prevents their being construed as completed units while keeping up the rhythmic organization of the sentence's elements. The unit of which they are a part ends in D), and it is that entire element (c-d) which is held in equivalence with the final closed contour and termination elements. The speaker's strategy has been to grant equal status to the two clauses but to dilute the strength of this division with a triplet rising pattern.

Finally, having considered that phrases are internally structured by contours, we can see the entire process being reiterated at a higher level: the sentence is internally structured by the phrase units which in this case are presented as equals. And at levels above the sentence the process occurs again. Paragraphs contain clusters which contain individual sentences. The cycle is bounded ultimately by the width of speaking ranges which can no longer accommodate wider intervals.

Summary. We have seen examples of both coordination and subordination created through the use of a simple iconic device. Equal pitch height has been applied in equivalent environments such as sentence or contour onsets to imply equivalent significance. Unequal pitch height in equivalent environments signals subordination. Subordination can also result from conjoining similar contour shapes and lowering or compressing the pitch range on one. Structural parameters which have been found relevant to the constituent-building process include relations between absolute pitches in critical environments (at pivot points, for example) and interval relations between concatenated pitches.

It appears from all this that incredible burdens are being placed on speakers' and listeners' tonal memories. In order to control pitch heights used for sentence onsets, for example, one would have to recall precisely the pitch on which the last onset occurred. I doubt, however, that they are being recalled without the aid of some more local reminders. We probably cannot recall absolute pitches well enough to

reproduce them over more than about one or two intonational phrases at a time. It appears to be one of the functions of phrase boundaries -- as distinct from simple contours -- that they release the memory from its recall obligations on absolute pitch levels used for contour perception. Contours, once processed, may be stored as configurations with a few salient attributes. Speakers probably do not control pitch levels by recalling absolute pitches of former contours so much as by recalling critical aspects of the immediately preceding contour and setting them into relation with comparable aspects of the current contour. Critical features that may be recalled in this way are onset levels, peak levels, and fall levels. There is no doubt also that some part of this "tonal" memory is reconstructed through semantic cues provided by the segmental information.

We have attempted to represent an outline of the pitch network that binds units into discourse coherence. Since our data were taken primarily from monologue, in some ways a speech situation that privileged speakers (to plenty of time, freedom from interruptions, etc.), it will be interesting to learn whether the analysis techniques generated from this data will transfer to conversation.

Bibliography

- Bing, J. 1979. Aspects of English prosody. Bloomington: Indiana University Linguistics Club.
- Bolinger, D. 1972. Accent is predictable (if you're a mind-reader). *Language* 48.633-645.
- _____. 1982. Nondeclaratives from an intonational standpoint. *Parasession on nondeclaratives*. Chicago: Chicago Linguistic Society.
- _____. 1983. Intonation and gesture. *American Speech* 58.156-174.
- Brazil, D., Coulthard, M. and C. Johns. 1980. Discourse intonation and language teaching. London: Longman Group Ltd.
- Chafe, W. 1977. The recall and verbalization of past experience. *Current issues in linguistic theory*, ed. by R. Cole. Bloomington: Indiana University Press.
- _____. 1979. The flow of thought and the flow of language. *Discourse and syntax*, ed. by T. Givón. New York: Academic Press.
- Collier, R. 1975. Perceptual and linguistic tolerance in intonation. *International Review of Applied Linguistics* 13.293.308.
- Cooper, W. and J. Sorensen. 1981. *Fundamental frequency in sentence production*. Heidelberg: Springer.
- Crystal, D. and R. Quirk. 1964. *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton.
- Fox, A. 1973. Tone sequences in English. *Archivum Linguisticum* 4.17-26.
- Ladd, D. R. 1983. Phonological features of intonation peaks. *Language* 59.721-759.
- Liberman, M. 1978. *The intonational system of English*. Bloomington: Indiana University Linguistics Club.
- 't Hart, Johan and A. Cohen. 1973. Intonation by rule: a perceptual quest. *Journal of Phonetics* 1. 309-21.
- Trim, J.L.M. 1959. Major and minor tone groups. *Le Maitre Phonetique* 112. 26-29.