

Prosody, Linguistic Diffusion and Conversational Inference

Author(s): John J. Gumperz and Hannah Kaltman

Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society (1980), pp. 44-65

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.

Prosody, Linguistic Diffusion and Conversational Inference

John J. Gumperz and Hannah Kaltman

Murray Emeneau is one of the first among American historical linguists to go beyond the mere listing of individual instances of areal diffusion and to provide detailed documentation of how and under what conditions areal borrowings can affect the grammatical system of a language. Prosody, the subject of this paper, has only begun to be incorporated into grammatical analysis and is rarely, if ever, studied in historical linguistics. What we would like to suggest is that prosodic systems, contrary to what has been assumed by some, are basically areally distributed and that studies of the interrelation of prosodic and grammatical signalling in conversation may provide important insights into pragmatic and interactive constraints on linguistic diffusion and change.

We use the term prosody to include: intonation, i.e., pitch levels on individual syllables and their combination into contours; changes in loudness; stress, a perceptual feature generally comprising variations in both pitch and loudness: variations in vowel length; phrasing, including utterance chunking and accelerations and decelerations within and across utterance chunks; and shifts in overall pitch register. These are perceptual confluences of variations in the three basic phonetic dimensions of frequency, amplitude and duration. Conversations are seen as made up of sequences of interactive exchanges in which both speakership and listenership in several channels contribute to the production of meaning.

Prosodic phenomena have been studied within the discipline of linguistics at the level of the word (Pike, 1945; Trager and Smith, 1951), the sentence (Halliday, 1967; Crystal, 1969) and the utterance (Brazill, 1975; Liberman and Sag, 1974,75) and have been examined as elements of syntactic (Berman and Szamosi, 1972) and pragmatic competence (Ladd, 1976; Liberman and Sag). Linguists and phoneticians, in spite of disagreements in detail, have reached a general consensus on which prosodic features have semantic or prosodic salience. Their work, which has revealed a great deal about the conventions of English prosodic usage and about the nature of the semantic information conveyed by prosody, forms the basis of our analysis. The basic question we ask, however, is one which has received relatively little attention so far: what sorts of information do speakers and listeners require prosody to provide in interpreting what is intended at any one point in a verbal exchange?

This concern with conversational inference raises problems

which are quite different from those ordinarily encountered in grammatical analysis. The data for grammatical analysis consists of sentence or clause units which have known boundaries. Furthermore, while such units may or may not derive from longer texts, the very fact that the analyst has singled them out for study assures that they receive the reader's attention. In everyday verbal interaction neither of these two conditions hold. Participants in a conversation, if they are to be heard at all, must actively create and maintain conversational involvement. That is, speakers have to induce potential listeners to attend and remain attentive until the message is complete, while listeners must signal their cooperation through choice and timing of verbal and non-verbal responses. This means that all participants must share a system for chunking the stream of talk into appropriate subunits, for signalling what is likely to happen, for managing turns of speaking, for distinguishing new information from qualifying remarks, and for negotiating the many subtle shifts in topic, focus or emphasis which characterize most types of everyday talk.

We assume that the maintenance of conversational cooperation requires judgments made simultaneously at several levels of generality through an inferential process which both interprets what has transpired and generates expectations about what is to come. Interpretation begins with informed guessing, based on knowledge of the extralinguistic situation and participants' goals and background. This yields general contextual expectations about what the activity is likely to be, its goals and possible outcomes, and what topics might be discussed. As soon as talk begins, however, these initial presuppositions are sharpened and modified by information signalled through both the form and the content of speech. We want to examine the role of prosodic signalling cues in this latter process of inference and to consider how prosody interacts with other modalities to signal thematic connections and contribute to interpretations of communicative intent.

Example One illustrates the kind of issues we intend to deal with. The passage comes from a videotape of an informal discussion among graduate students in anthropology, all of whom had participated in an advanced survey course covering various anthropological subdisciplines. The discussion revolves around the extent to which these subdisciplines were or continue to be related.

EXAMPLE 1

- T1. 1 A: ...and then you could see where / you could see more in
 2 depth where / how things are related // but 'I think
 3 that you absolutely have to see how / ..where the rela-
 4 tionships are //

interruptions of the rhythmic flow of talk. Speaker A sets the theme in line 2 with his claim that the subfields of the discipline are related and that the relationships become evident if you look for them. B questions A's claim and A counters by suggesting that current divisions did not exist in the early stages of the discipline whereupon B goes on to dispute the relevance of this historical argument.

What do A and B do that makes their utterances reasonable sequential contributions to continuing conversation? First of all, each turn has lexical and grammatical markers of cohesion such as have been discussed by linguists interested in discourse (Halliday and Hasan, 1976). For example T1 and T2 are both phrased in the first person. Turns T2, T3, T4 and T5 all start with "but." This use of "but" has three components: 1) negation, although not necessarily total negation; 2) that the speaker has something to add which changes matters -- "not only that, but"; and 3) a claim for the floor -- "but wait a minute." These four turns are all related in similar ways to the prevailing theme of the conversation. They both counter the previous turn and justify or add new support to the speaker's position. At the lexical level topical continuity is signalled by repeated use of key terms such as "related," "relationships," in turns T1, T2 and T6 and of phrases such as "did everything" in T3 and "its all spread out" which have related referents. Yet these and other cohesive markers do not simply exist. They are made salient and in a very real sense given their situated interpretation through their syntactic placement within the clause and through prosody.

The basic characteristics of English prosody that make this possible have been most clearly described by Halliday (1967). Halliday describes intonation in terms of three components, each the realization of separate "systems of meaningful choices" (Smeall, 1976) which he calls tonality, tonicity and tune.

Tonality provides the basic unit of analysis, the tone group or breath group, a smooth, continuous intonational contour, set off from other units by features of timing similar to what is called phrasing in musical performance. The tone group as Laver has pointed out, is the basic unit of speech planning or cognitive processing and thus corresponds to what has been called an information or idea unit (Chafe, 1980; Pawley and Syder, 1978). Linguists have observed that tone group boundaries tend to correspond to clause and sentence boundaries, and distinguish between minor tone groups which indicate that more is to come and major tone groups which signal finality (Trim, 1976). Where an expression is syntactically ambiguous prosody can function to provide information that is not otherwise available in the spoken sentence. For example:

My sister who lives in New York is very nice // (i.e.,
I have more than one sister, and the one who lives in
New York is very nice.)

My sister / who lives in New York / is very nice //
 (i.e., I have one sister, who both lives in New
 York and is very nice.)

Tone grouping here is a cue which distinguishes a restrictive relative clause from a non-restrictive relative clause. In line 10 of our passage tone grouping separates "right" from the preceding clause and gives it the force of a tag question. Crystal (1975) discusses the further possibility of tone grouping conveying additional affective or rhetorical connotation in cases such as:

I said sit down //

vs.

I said / sit / down // (e.g., I am very angry, or
 I think maybe you didn't hear me, or it is very im-
 portant to me that you sit down, etc.)

In lines 3, 5, and 6 of Example 1 the speakers' use of tone grouping followed by pausing conveys the effect of conscious reflection or active planning.

Tonicity in Halliday's work refers to the placement of the tonal nucleus, or tonic, within the tone group. For purposes of this placement, the tone group can be seen as divided into a number of sub units or 'feet' corresponding either to items of lexical content or to syntactic phrases, each of which can potentially carry a tonic accent on its most prominent syllable.

Crystal's formulation is that the norm is for the nucleus to be on the final lexical (as opposed to grammatical) item in the main NP-VP clause. Thus it is possible to find clear instances of tonicity disambiguating syntactic minimal pairs. This is in the case of utterances whose surface ambiguity is the result of lexical ambiguity. Such cases have been the principal focus for the analysis of intonation in generative-transformational analyses, since such sentences reflect two possible underlying deep structures, and they are also likely to have (but don't have to have) different nuclear placements:

George has ^ˈplans to leave // (he intends to leave)

George has plans to leave // (he has blueprints to
 deliver)

The final lexical item rule accounts for 80% of Crystal's data. In most of the remaining cases, the tonic is on a grammatical item which operates clearly as a member of a closed system, and tonicity then signals a contrast between the marked item and some other member or members of the system, for example:

I want it in the garden / near the fence / and not
behind anything //

Some cases do not involve systemic contrast, and can only be accounted for by considering the speaker's attitude. As for non-final tonic lexical items, Crystal has found five cases which can be "related unequivocally to a specific feature of syntactic structure." However, "the vast majority of cases of pre-final lexical tonic can be accounted for only by referring to some kind of semantic or lexical conditioning," such as lexical presupposition, or the distribution of new and given information.

Bolinger (1958,72) provides further insights into the semantics of intonational signalling in pointing out that longer tone groups in American English are characterized by two peaks or accents, and that the relative height of such peaks can signal attitude and distinctions between new and old information.

Tune, the pitch treatment of the nucleus, appears not to have any purely syntactic disambiguating function at the level of clause structure, but can indicate illocutionary force, as in:

John's going // (statement)
 John's going // (question)

This is a pragmatic distinction, rather than a purely syntactic one -- "John's going?" is likely to be an echo question, meaning "Did you say that John is going?", and not the literal question "Is John going?" In general, options of tune, and of total contour, have been found to have some correlation with the speech act functions of utterances. However, apart from the question-statement distinction, the only examples which show any similar regularity are highly formulaic.

These formulaic uses of intonation have received particular attention because contour supplies information that is not predictable from the lexical material and thus there appears to be a 'lexical' meaning intrinsic to the contour or to features of the contour. Examples which have been studied are introductions, as in:

John Mary Mary John (Rogers, 1978)

and warnings:

John ny (Ladd, 1978)

Attention has also been directed at the question of whether intonation is used to differentiate direct from indirect speech acts (literal vs. non-literal questions -- Liberman and Sag, 1975; neutral predications from imperatives -- Bolinger, 1972). But apart from this the literature on the semantics of prosody is relatively scant.

In addition to intonation proper, there are other features of prosodic signalling which are most frequently discussed under the heading of paralinguistic or expressive signs, but which

nevertheless are of considerable importance in conversational inference. These include pitch register shifts, i.e., lowering or raising of overall pitch level; increases or decreases in loudness; speech tempo. In contrast to intonational features which apply to syllables these latter features apply to longer stretches, usually whole tone groups or sequences of tone groups. They often serve as signs of emphasis or deemphasis or as attention getting devices, as in line 7 of our passage:

┌ but "ultimately it is

where the raise in pitch combines with tonicity to convey strong disagreement.

We can now return to a systematic analysis of the conversation in Example One and examine the signalling load of these clause-level devices in longer stretches of discourse. We must note that any conception of prosodic meaning derives in large part from method of analysis, which in the case of linguistics has been that of looking for minimal contrasts at the single-utterance level. In this way, one can separate cases where prosodic phenomena are predictable from other aspects of linguistic structure from cases where they are not. A large proportion of prosodic usage thus becomes predictable from grammar. In conversation the situation is clearly more complex.

Note for example the placement of the tonic in the three tone groups in T1 containing the verb "see," in lines 1,2, and 3. In its first occurrence "see" stands alone, in the second it is a constituent of the phrase "see more in depth," and in the third the phrase is "have to see." It is evident that simple binary distinctions such as those between marked and unmarked or contrastive-noncontrastive do not fully account for how A's intent is signalled here by shifts in tonic placement. In the first tone group one might argue that the tonic placement is unmarked (and that "where" is simply the beginning of an abandoned tone group). The fact that "see" is not accented in the second and third tone groups might be explained by an argument similar to that made in Halliday's discussion of anaphora (1967b). But we need other considerations than mere contrastivity to explain what the stress on the elements which are stressed accomplishes. Halliday argues that since anaphora reflect known information they are inherently unstressed, and that whenever they are stressed they are consequently contrastive. On this basis we would expect either that the entire phrase "see more in depth" would be unstressed, or that the major stress would fall on "depth." Instead the stress falls on "more," which clearly does not contrast here with its opposite, "less." What is signalled here by this stress on "more," along with the minor tone group boundary again after "where," is that this second tone group, being a recycling of the first, is therefore a correction or a qualification of what has been said before. This is accomplished, in a way which would not have been effective had "depth"

received major stress, due to the evaluative function of "more." Similarly, the emphatic stress on "have to," by highlighting modality, signals that the listener must look at the situation under discussion in a certain way if he is to understand the speaker's argument. This is quite different from a simple contrastive statement about the existence or nonexistence of an obligation. Thus tonic placement, when viewed in these relational terms, allows us to trace speakers' thought processes, and their strategies in developing a theme.

Similar arguments apply to the use of tune on "related" and "relationships" in lines 2,3-4, and 6. The first "related" is introduced with an especially emphatic fall-rise (marked phonetically both by contour and by slowed rhythm and vowel elongation). "Relationships" in lines 3-4 has the less emphatic high fall, which serves both to keep the topic in focus and to continue the line of argument. The final "related" in line 6 continues to carry the tonic, thus maintaining the theme, but the pitch movement is reversed, highlighting the question. The topic of relationships is firmly established in the transition from T1 to T2, in which B chooses the topic, places it in focussed position, and addresses it with a question. B could have responded to the issue of whether it is important to see the relationships rather than questioning their existence. His response is framed around material which is implicit in A's utterance; the subject "it" in line 6 only makes sense on the basis of the presupposition that A has indirectly asserted that "it's" all related. A's reply in line 7 has the syntactic form of answers in such pairs as "Is it related?", "Yes, it is," where the main verb is elided. Here, syntactic knowledge enables us to fill in the elided verb; in addition, the low falling tune on "is" copies the falls in A's previous turn and therefore prosodically emphasizes that A is reasserting her claim in the face of B's question. In all these cases, tonic placement and tune, along with syntax and lexicon, guide the listener in inferring relationships among utterances and supplying nonlexicalized information.

The expression "they did everythin'" in lines 9-10 illustrates another aspect of prosodic signalling. Particularly in informal conversation, prosody is a factor which allows participants to use a minimum of lexical specificity to tie together parts of an argument. Here, it is clear that the speaker is saying something like that for earlier anthropologists the subfields were unified (that is, related). The very general expressions "to do everything" and "everything" have specific and obvious referents in this conversation ("doing" anthropology and doing "everything" contained in its current subdisciplines). Obviously these referents do not inhere in the words "do everything" but must be inferred. The phrase is highlighted syntactically by being the comment in a topic-comment structure, and prosodically by the heavy final, falling, emphasis on "everything." Given that nothing has intervened in the conversation, the prevailing theme is still that of

relationships, thus the relevance of "everything" -- a single term grouping 'related' referents -- and what people "do" as a means of discussing what "is."

In lines 11-14 the stress on a series of related words ("then," "now"; "now," "came") emphasizes the thematic link between turns T4 and T5, and marks out the four steps of an argument. When B first stresses "then" in line 11 he singles out one aspect of the preceding discussion, the temporal. The stressed "then" signals that this first phrase of B's will be balanced by forthcoming material; this opposition is signalled when it occurs in the second phrase, both by the parallel prosodic structure, with nuclear stress on the temporal adverb "now." A then repeats the internally-parallel structure in her T5, and uses this structure to reverse the argument. That is, she starts with a phrase describing and stressing "now," so that "now" becomes the first rather than the concluding member of a pair. Her final counter implicitly refers to "then," although it does not repeat it, both because her parallel intention has been signalled, and because of the past tense on the stressed verb "came."

In this sequence of turns, prosody signals rhetorical structure by highlighting key thematic items. The argument is stated in very abstract terms, but involves a fairly intricate set of inferences concerning concrete facts and situations. Let us look more closely at the first half of T4, "that's then." "That" in this passage has what Traugott (1979) calls its 'discourse meaning' rather than its lexical meaning. We can go further and say that in this usage it acquires a sort of double discourse meaning. The referent of "that" is 'what you have said.' This expression points to both 1) the act of having said what you said, and 2) the propositional content of what you said. The implication carried by "that" is that there will be a 'this,' a what I am going to say; i.e., that the speaker is going to say something further and that what s/he says will be related in a certain way to what has already been said.

"Then" also signals multiple levels of meaning. First of all, it is anaphoric within the discourse to "back in nineteen hundred." Then, given this temporal frame in the text, "then" also takes on a somewhat formulaic connotation: back in the old days. Thus "back in nineteen hundred" is paraphrased by "then," a structurally simple reference which also adds further connotations, connotations which are signalled in part by the prosodic prominence of the word. The entire expression, "that's then" signals proleptically, i.e., signals information about what the speaker will say next. As we have observed, "that" implies a "this." It is similar with "then": while "back in nineteen hundred" could have been responded to by talk about any time period, "then" can only be contrasted with "now."

B's utterance is an instance of a class of phrases which could be considered as providing a formulaic discourse strategy for some speakers: statements with a "that's..., that's not..." construction, where the contrasted objects of the two clauses are

opposing pairs of either general adverbs or pronouns. The proleptic signalling value of "that" plus an element which is one of a contrasting pair can be so strongly conventionalized that hearers can be expected to be able to fill in the completion; therefore in some contexts the second clause may be optional.

To summarize, we have been arguing through this example that interpretation at the level of conversation is a function of an inferential process which has as its input syntactic, lexical, and prosodic knowledge, and that judgments of intent are based on speakers' ability to relate the information received in these channels. To understand how this inferential process works, we can go back once more to the linguists' discussion of marked and unmarked tonic placement. Recall Crystal's example:

I want it in the garden / near the fence / and not
behind anything //

Broadly speaking one might argue that in uttering this sentence, a speaker violates prosodic rules such as: "the tonic falls on the last lexical item in the tone group" or "anaphora, prepositions, conjunctions, etc. are normally unstressed." But what is at issue here is the definition of terms such as 'rule' and 'normal.' The sentence as spoken is neither ungrammatical nor necessarily inappropriate. The speaker is selecting among available options and the resulting inferences are similar to those conveyed by pragmatic processes like topicalization or passivization.

This does not mean that notions like grammaticality and appropriateness are not relevant to the interpretation of prosodic signalling. Tone grouping, tonic placement and tune are clearly grammatically constrained. But the study of conversational inference provides a different perspective on what these constraints are and what they signal than does sentence by sentence analysis. Thus an utterance like

I want / it in the garden

with a falling tune and a tone group boundary after "want" would clearly be odd, but an articulation like:

I / want / it / in / the / garden

would be perceived as expected, only if the activity called for listing. In an informal conversation like Example 1 it would be odd unless a shift in activity were signalled in previous talk.

Perhaps the best way to describe this interpretation process would be as an instance of Gricean implicature. A shared interpretation therefore relies on a shared understanding of a limited or closed set of options that can be chosen, and of how these are constrained by the nature of the linguistic and extralinguistic context.

In a sentence by sentence analysis, it is the total available corpus of utterances or the analyst's impressions of what sounds or is judged by native speaker to be acceptable that determines an assigned interpretation. In the analysis of conversational inference the frame, the subset of available option, and the interpretations they convey are all determined by preceding talk. The first step in such an inferential process is the participants' perception of what is highlighted, what previous bits of information it is most likely related to, and how it is related. It is on the basis of such perceptions that participants call on their knowledge of the semantic range of the terms used, of interclausal syntactic relations, and of conventional discourse strategies, to arrive at an interpretation of communicative intent and to trace the line of argument. Shared assumptions about how tonality, tonicity, and tune interact with grammar and lexicon to suggest relationships are thus a precondition for shared interpretation and for the maintenance of conversational involvement.

We opened this paper by suggesting that prosodic conventions are differentially distributed. In other words, that the rules and assumptions governing the interaction of prosody, grammar, and lexicon, and the consequent signalling functions of prosody in discourse may vary across subgroups of speakers of English. To illustrate this point we will draw on both conversational and elicited data, collected in England and in the U.S., from bilinguals who are native speakers of North Indian languages such as Hindi, Urdu, Punjabi, or Gujerati, but who regularly communicate in English both with others of similar backgrounds and with native English speakers in the surrounding community. We have pointed out elsewhere (Aulakh et al., 1978) that the differences in question are not simple instances of interference such as one finds in second language learning situations. The individuals concerned are not isolated speakers of English trying to learn a new set of language skills. Rather they have their own systematic conventions of using English which are highly effective within their own communities, and which are likely to persist as long as they remain effective and are reinforced by ingroup use.

Our initial observations on what these differences are are based on close analysis of natural conversations, and especially of breakdowns in conversational cooperation as indicated by disruption of conversational synchrony (Erickson and Schultz, 1980).

Our theoretical conceptions of signal-guided inference and our observations of natural conversation suggested that such breakdowns were the result of differences involving two sorts of communicative effect: (1) the use of prosody in signalling normal information flow, and the differentiation of 'normal' from contrastive and expressive intentions; (2) the use of prosody to signal various kinds of intra and intersentential relations such as subordination, or utterance finality or non-finality. We therefore constructed a list of single sentences which in terms

of their grammatical structure and the nature of their contents embodied this range of communicative intention: the list starts with simple existential propositions, goes on to compound constructions, complex constructions, parallel contrastive statements, and expressive items (e.g., "That's fantastic! Really beautiful!") (the list is given as an appendix). The list of items was read by two main informants, both women, one British and one Punjabi. Two observations regarding representativeness and naturalness of this data: Firstly, informants of the same backgrounds as these subjects, who knew the speakers personally, confirmed in each case that the recording represented a 'normal' style of speech; that is, neither of the subjects spoke with an obvious 'reading' intonation. Secondly, particularly in the case of the Indian speaker, as our hypotheses developed we supplemented the initial data with samples from other speakers.

Starting with the simple sentences, we can make several observations which carry over onto the more complex ones and acquire additional consequences. In English, the grammatical 'basic sentence' is a single clause, consisting of a subject noun phrase (NP) and a predicate verb phrase (VP), each of which can contain optional modifying elements, and optional qualifying phrases in the predicate, either adverbials or prepositional phrases. This level of syntactic organization is directly reflected in Western English prosody, where the 'basic' prosodic unit is also the clause. Thus in the case of a simple, i.e., one-clause sentence: 1) the sentence will comprise one tone group; 2) the tone group will have a smooth, unified contour; 3) the tone group will have two or more most-prominent syllables, corresponding to peaks of information; one will be the nucleus; 4) the contour will end in a distinct fall or rise.

Indian English systematically contrasts with Western English in its prosodic treatment of simple sentences in all but the first of these characterizations; that is, 1) the sentence will probably be spoken as a single tone group (that is there will be no pauses), but 2) there will be no 'unified contour'; rather, there will be two or more subunits separated by fairly abrupt changes in pitch; 3) there will be no clearly-defined nucleus; 4) the pitch change on final syllables will be much narrower; frequently they will be held high and level. For example,

- 2) W.E.: This is a book
 I.E.: This () a book (the "is" was not pronounced by the subject)
- 3) W.E.: John is reading a book
 I.E.: John is reading a book

Transcription conventions here: the lines above the utterance are perceived intonational contour, the subscripts indicate stressed syllables, (underlining, to be used in later examples,

indicates more emphatic or heavily stressed items.) In the Western English versions, the sentence as a whole has a steadily falling 'envelope' contour, with small peaks corresponding to stressed syllables. In the Indian English examples, in contrast, there is no real tonal contour in this sense. Instead each sentence is divided into several 'sense units' or prosodic pieces. Each of these has relatively level pitch on the central information-carrying items. There are sharp boundaries between the pieces. In these two examples, these are achieved by the occurrence of a sharp fall after a level syllable, occurring on unstressed items ("is" or "a"), with the pitch then rising somewhat more gradually to become level again on the next stressed item. Thus there is a succession of level tones, each of which is higher in pitch than the immediately preceding environment.

Furthermore, in the Indian English versions there are no nuclear syllables. Two factors are at work here. First of all, there is the breakup into sense units just described; thus, at least intonationally almost every content word is highlighted. Secondly, the distribution of stress in Indian English differs significantly from that of native Western speakers. Paradoxically, to Western ears, Indian English can sound either full of stress, and staccato, or droning and monotonous. This is because on the one hand Indian English speakers rarely reduce syllables and pronounce almost all consonants with a higher degree of articulation than native speakers, thus in one sense employing a great deal of stress; on the other hand, no syllables are stressed significantly more than any others.

Thus two of the striking features of Indian English in contrast to Western English are: the subdivision of utterances into small chunks; the rhythmic marking by stress of several words, with no one syllable made tonally prominent. There is some evidence that both of these phenomena have a basis in the languages of North India. In "A Reading Transcription for Hindi," W.E. Jones says, "...each syntactic piece -- of one or more words in length -- is spoken with a sub-contour" (1971b). A 'syntactic piece' consists of each NP or VP in a sentence, and adjectival or adverbial phrases may also be separate syntactic pieces. Hindi also has a large number of particles and verbal auxiliaries, and there are syntactic rules which incorporate these into either the NP or VP to make syntactic pieces. Prosodically, each subcontour rises in a sequence of level pitches, or consists of a single rise if a piece is monosyllabic. It should be apparent that this description exactly fits our description of the Indian English treatment of simple sentences.

As for stress, and the concomitant question of the existence of nuclear syllables, there are three important factors to note. First of all, it has been noted that, while there is word-level stress in Hindi, the differences in intensity between stressed and unstressed syllables are much less than for English. Secondly, as in Indian English, reduced syllables are almost non-existent in Hindi (Ohala, 1977; Jones, 1971a). Thirdly, the

intonational correlates of perceptual stress differ between English and Hindi. In Hindi, a stressed syllable is either high or rising in pitch, and the following stressed syllable falls. In English, no such simple statement can be made, but in general a stressed syllable will have a markedly greater degree of pitch movement than surrounding syllables -- thus frequently stressed syllables will contain a change in direction of pitch movement (i.e., a rise-fall), or else will fall or rise more sharply than preceding and following syllables.

Thus, to summarize, in comparison to Western English, Indian English bases its prosodic conventions on (a) different syllable-level phonology; (b) a different level of syntactic breakdown; (c) different phonological means for making prosodic distinctions and relations.

One result of these differences shows up in simple statements with a contrastive focus. In workshops in communication skills with Indian English speakers we have used the following exercise, which points up these differences and their automatic, habitual nature:

Instructor: A (student's name), what's your phone number:

A: 834-9578

I: 835-9578?

A: No, 834-9578

When Western English speakers take the role of A, they automatically stress the "4" the second time they repeat the number, by giving it increased pitch and loudness. This pattern in effect signals, "You got something wrong, this is the digit that you had wrong, and this is what it should have been." The Indian students, on the other hand, tended to repeat the phone number exactly as it had been said the first time. If they made any change, it was to shift up in pitch register, starting on the corrected digit and continuing in high register to the end. When the different Western English strategy was pointed out to them, and they tried to duplicate it, they were unable to highlight the single digit. Instead, they combined stress in the form of increased loudness with the raising of pitch register, and again continued both to the end of the string: "eight three FOUR NINE FIVE SEVEN EIGHT." Sometimes they anticipated, and began the stress on the numeral preceding the one which should have been marked.

This example also illustrates one general level difference in the allocation of signalling function among the various channels which make up prosody. In Western English broad pitch register movements tend to be used at boundaries to indicate shifts in

the nature of the speech activity, or in degree of intimacy, or as attention-getting devices. In Indian English these shifts seem to carry some of the 'grammatical' load which in Western English is carried by tonicity and tune, that is, the marking of points of information structure and flow.

This picture extends to longer sentences. In Western English, the choices deployed in the placement of stress and pitch changes signal the relevance of items of information in the light of what has gone before -- unexpected or new or qualifying information will be focussed. Tonal nuclei tie together semantically related items, given the focus and communicative intent of the whole. We build on tonal wholes at the clause level to build relations in connected sequences, where we can develop prosodic parallels and contrasts. For example, in a compound coordinated sentence:

4) Do you ¹want a cup of tea / or do you ¹want a cup of
 coffee //

The two clauses are prosodically similar in several respects. Each clause has a head on "want" and a tonic on the object -- the main verb and the main noun are in focus. While each clause is smooth both rhythmically and intonationally, there are two sub-parts in the phrasing: "do you want" and "a cup of X." In each of these rhythmic groups the accent is on the last syllable. Thus "do you want" signals the speaker's focus on the addressee's intent and the illocutionary questioning function. "A cup of X" contextualizes the utterance as being involved with the selection of items to fill the slot indicated by "want," and signals that there has been a change in that item in the second clause. The two semantically critical dimensions here -- concern with desire, objects possibly desired -- are those which receive prosodic prominence.

The two halves of the sentence also contrast with each other in several respects. "Tea" rises, while "coffee" falls. This contrast closes the list -- nothing else will be offered. In addition, the second contour is lowered as a whole with respect to the first. The effect is almost a metaphor for the two-option choice presented in the utterance: 'on the one hand, and on the other' -- the 'hands' being 'pointed to' by the two pitch registers.

In Indian English, as we have said, each 'sense unit' is in some way made distinct -- with this example we'll see better how this can be done, and how the prosodic reflection of semantic integrity is different.

4) ¹Do you ¹want / a / ¹cup of ¹tea / or / ¹do you ¹want /
 a / ¹cup of ¹coffee

Like (4), here each clause is phrased in two parts. Analyzing

these as poetic feet, in the Western English example we had "do you want" and "(a) cup of tea" as two anapests, with the accents on the last syllables. Here in the Indian English version we have these phrases as two dactyls with the accents on "do" and "cup." The two feet are much more independent here than in the Western English version -- there is a complete rhythmic break, and "a" and "or" each are a separate tone group. There is a sharp downward pitch movement in addition to stress on "do" and "cup" each time these words occur, followed by a slight rise extended over the next, unstressed, syllable. The final words, the ones we saw accented in the (4) are not stressed here. However, they are set off prosodically -- each is spoken at a level pitch, which is the highest level reached in its segment. The parallel units, "do you want" and "cup of X," are prosodically parallel as in (4). The contrast between the two halves is signalled by the second clause as a whole being at a slightly lower pitch. Two changes occur between the first and the second clauses of this utterance: a change in the object noun, and a change in the degree of closure to the utterance. In (4) these were signalled by tonicity, accenting the noun, and tune, contrasting a rise with a fall. To Western ears, the emphases in (4) are 'misplaced,' so that the utterance appears not to signal these changes.

In contrastive constructions, Western English uses tonicity, tune, and stress placement together: the nuclei are placed on the two elements which frame a contrasting pair, and these are contrasted by having different shape and/or differing relative heights. The items focussed by tonicity need not be the words which carry the change; if not, this word or words will also be highlighted. Thus for example:

5)WE: If you don't give me that cigarette / I will have to
 "buy a cigarette //
 or If you 'don't give me that cigarette / ...

6)WE: If you 'take this course / you won't have to take the
 "other course //

In Indian English we have:

5')IE: If you don't give me [┌]that 'cigarette / I will have to
 buy [┌]a 'cigarette //

6')IE: If you take [┌]this 'course / you won't have to take the
 [┌]other 'course //

In (5') the contrasting phrases "don't give" and "have to buy" are not picked out, as they are in (5), by stress. The main emphasis seems to be on "cigarette" both times. In comparing (6) and (6'), "this course" and the "other course" get identical prosodic

treatment in (6'). In (6), in the first clause, the qualifier is not stressed, and is, contrastively, in the second.

In complex sentences the Indian English pattern continues -- either stress or raised pitch or both on each sense unit. The interrelation of these sense units gets more intricate -- they can be simple subject-predicate sentences, noun phrases, verb phrases, adverbials, conjunctions. We won't go into predicting how an utterance will be broken up, but we'll give examples of breakups at each of these levels and consider the contrast with Western English signalling strategies in these examples, and follow with some generalizations. Subject-predicate phrases as sense units:

7) I think / he said ...

8) I heard him say / she thought ...

In these cases the S-P unit in each tone group gets parallel prosodic treatment, with the emphasis on the subject. In contrast, in Western English the distributive nature of the main verb in such embeddings is prosodically reflected. The main verb will get the tonic. Succeeding subordinated S's will either continue in the same contour, or, if it gets triplicate as in (8) a new contour will start with emphasis on the third verb, but at a lower pitch level -- there will be an overall envelope effect.

These differences have more pronounced effect when an embedded sentence is relativized. Here in Indian English a tone group boundary will be placed before, i.e., setting off, the embedded sentence:

9) What he said was that / he wouldn't come //

An example of an isolated VP:

10) When he [↑]came home / the book was gone //

Both these examples sound contrastive to Westerners. In (9) the first tone group ends abruptly on "that," sounding like, "and not this." In (10) the emphasis is created by the pitch rise starting on "came," which could be implying contrast with some other possibility such as "left." A similar example:

11) [↑]Did you see anything / [↑]when you came home //

The pitch rise emphasizes "did," sounding like a contrast with "didn't."

There appears to be a logic at work in all these examples which is different from the contrastive interpretation which Westerners would derive from the Western unmarked emphatic patterns. Here relationships are signalled by picking out items, and the content of the relationship picked out derives from

lexical content -- an aspect of the literal meaning of a word or phrase is being pointed to and which that is is dependent on the verbal context. Thus in an S with a "when" introducing a relative modifying clause, the "when" will be highlighted to point to the relativizing function, as in (11). In a case like (10), on the other hand, where the relativized clause as a whole is topicalizing (not modifying), if "when" were stressed, the effect would be to point again to an element of grammatical function, and with "when" at the head of the utterance, this would make it a question. This is what is happening with the stress on "did" in (11) which points to the question function. (These phenomena can be related to the functional particles in Indian languages, see Aulakh et al.) The converse of this is what happens when an Indian English speaker is using extra-emphatic stress to signal a contrastive intention. In these cases, Westerners hear the intended emphasis, but the dimension of content being pointed to is different, so that a Westerner will derive a different reading from what is intended (Mishra, 1980).

There are also some seemingly formulaic illocutionary and/or expressive usages which distinguish Indian and Western English. For example, in Western English tag questions will be set off by pitch register from the question they follow -- usually they'll be lower, sometimes higher (we're talking about the start of the tag, not whether they rise or fall -- distinguishing confirmation-requests, semi-imperatives, etc.) In Indian English tag questions continue at the same level reached in the main question. Direct questions frequently are said with pitch steadily rising over the utterance:

- 12) Did he leave his key at home?
 13) Where did he leave his key?

Our findings on Indian English concern the origin of contextualization conventions in the structure of these speakers' native languages. However, we will advance the hypothesis that the differences in question are not language-based in the sense this word usually has but are distributed in accordance with what anthropologists call cultural or linguistic areas (Emeneau, 1964). The balance in signalling loads among channels reflects the need for verbal communication to do all the following: satisfy cognitive constraints of attention, memory, and information processes, reflect participants' shared perceptions of their physical and mental experiences, and be intelligible as an activity with a discernable function. This balance, and the roles of each of the various relational mechanisms, is the outcome of the history of a communicative tradition, in the sense of the nature of the communication for the accomplishment of which language has been shaped: the pragmatic presuppositions involved in norms for verbal communication in its social context, reflecting status

relations and paradigms of face-to-face or of distanced communication.

Thus contextualization conventions reflect long-established economic, political, and religious networks where speakers of historically distinct languages have been in contact for hundreds of years. Our data indicate that the contextualization conventions of Indian English will hold for many native speakers of Indo-Aryan languages such as Hindi, Urdu, and Punjabi, as well as for speakers of the generically unrelated Dravidian languages such as Telugu. We also have initial evidence from English conversations collected in Hong Kong and Southeast Asia that the prosodic conventions of Southeast Asian English show some similarity to those of Indian English (see Young, 1980). These statements about areal distribution are of course very preliminary and need to be tested through empirical investigation. However, if it is true the prosodic conventions we have described for speakers of North Indian languages speaking English result in part from the mapping of prosodic conventions of the speakers' native languages onto English, and if social conditions are such that these mappings are maintained and institutionalized over time, then the conversational behavior of Asian bilinguals can be seen as an area where language-based grammatical systems meet with areally-based prosodic and discourse systems. The study of the interaction of these systems in conversation can provide important insight into processes of semantic change, the grammaticalization of lexical items, and other basic historical processes of linguistic change.

* * * * *

Work on this paper was supported by NIMH grant MH26831. We are grateful to Gurinder Aulakh, Richard Bonte, Arpita Mishra, Mary C. O'Connor, and Celia Roberts for assistance in the analysis.

REFERENCES

- Aulakh, Gurinder, J.J. Gumperz, and H. Kaltman, 1978. Thematic Structure and Progression in Indian English Discourse. Manuscript.
- Berman, A. and M. Szamosi, 1972. Observations on sentential stress. Language 48:304-25.
- Bolinger, Dwight L., 1958. A theory of pitch accent in English. Word, Vol 14, pp. 109-49.
- _____, 1972. Relative height. In D. Bolinger, ed, Intonation. Harmondsworth, Middlesex: Penguin.
- Brazil, David, 1975. Discourse Intonation. Monograph, English Language Research.
- Chafe, W., 1980. The deployment of consciousness in the production of a narrative. In W. Chafe, ed., The Pear Stories. Norwood: Ablex.
- Crystal, David, 1969. Prosodic Systems and Intonation in English.

- Cambridge: Cambridge University Press.
- _____, 1975. The English Tone of Voice. London: Edward Arnold.
- Emeneau, Murray, 1964. India as a Linguistic Area. In D. Hymes, ed., Language in Culture and Society. N.Y.: Harper and Row.
- Erickson, F. and J.J. Schultz, 1980. Talking to 'The Man': Organization of Communication in School Counseling Interviews. N.Y.: Academic Press.
- Halliday, M.A.K., 1967. Intonation and Grammar in British English. The Hague: Mouton.
- _____, 1967. Notes on transitivity and theme in English, part 2. Journal of Linguistics 3:2, 199-244.
- _____, and R. Hasan, 1976. Cohesion in English. London: Longmans
- Jones, W.E., 1971. Syllables and word-stress in Hindi. J.I.P.A., pp. 74-78.
- _____, 1971. A reading transcription for Hindi. J.I.P.A., pp. 88-97.
- Ladd, D. Robert, 1976. Intonation, Main Clause Phenomena, and Point of View. Manuscript.
- _____, 1978. Stylized intonation. Language 54:517-540.
- Laver, R., 1972. Voice quality and indexical information. In R. Laver and J. Hutchison, eds., Face to Face Interaction. Middlesex: Penguin.
- Liberman, M. and I.A. Sag, 1974. Prosodic form and discourse function. In CLS 10, Chicago Linguistic Society.
- Mishra, Arpita, 1980. Discovering Connections. In Proceedings of the Berkeley Linguistic Society VI,
- O'Connor, J.D. and G.F. Arnold, 1961. The Intonation of Colloquial English. London: Longmans.
- Ohala, Manjari, 1977. Stress in Hindi. In L. Hyman, ed., Stressfest. Southern California Occasional Papers in Linguistics.
- Pawley, A. and F. Syder, 1978. Sentence formulation in spontaneous speech: the one-clause-at-a-time hypothesis. Typescript.
- Pike, Kenneth L., 1945. The Intonation of American English. Ann Arbor: University of Michigan Press.
- Sag, I.A. and Liberman, M., 1975. The intonational disambiguation of indirect speech acts. In CLS 11, Chicago Linguistic Society.
- Smeall, Christopher, 1976. Intonation. Special field exam, U.C. Berkeley.
- Trager, George L. and H.L. Smith, 1951. An outline of English structure. Studies in Linguistics, Occasional Paper No. 1.
- Traugott, E.C., 1979. From referential to discourse meaning. Lecture presented to the Colloquium of the Berkeley Sociolinguistics Group.
- Trim, J., 1976. English Intonation. Department of Linguistics, Cambridge University: Cambridge University Typescript.
- Young, Linda, 1980. Inscrutability revisited. In Proceedings of the Berkeley Linguistic Society VI,

Partial List of Elicited Sentences

1. This is a book
4. He is an engineer
5. John is reading a book
8. That big table is standing in the corner over there
10. That chair is the biggest in the room
13. Do you want a cup of tea or do you want a cup of coffee
14. He went home sat down and had a cup of coffee
17. When he came back the book was gone
18. There was nothing in the house when he came home
19. Did you see anything when you came home
20. I think he said he was going home
21. I heard him say she thought it was alright
22. It's very hot today isn't it
23. He left his key at home didn't he
24. Where did he leave his key
25. Did he leave his key at home
27. All he said was that he was not going to come back
29. I don't know if he wants to come along or if he would rather stay.
31. I'm telling you now and I wouldn't tell you again I will not go
33. If you don't give me that cigarette I will have to buy a cigarette
34. If I can't read that book I will have to read another book
37. Whatever they can do I can do also
39. If you are busy at the moment you don't have to come You can do what you want
40. I wasn't driving too fast I was going slowly Why do you ask
41. Mister Smith knows John He can tell you all about him
47. He didn't eat the food He cooked the food He wasn't supposed to eat the food
48. John there's a telephone call for you
49. That's fantastic Really beautiful
53. I wanted to send you the book but I have only just received it myself and I haven't had time to look at it
56. Yesterday you said it was five pounds Today you say it costs six pounds Why is it so much more
57. It was marked down for the sale and when the sale was over we put it back to the regular price
59. I understand your problem Why can't you understand my problem
61. Without looking into it how can you know it's all right
66. Please don't bother to come with me I know the way out

