# NOW LET'S TALK ABOUT *NOW*:
## IDENTIFYING CUE PHRASES INTONATIONALLY

*Julia Hirschberg*
AT&T Bell Laboratories
Murray Hill, New Jersey 07974

*Diane Litman*
AT&T Bell Laboratories
Murray Hill, New Jersey 07974

## ABSTRACT

Cue phrases are words and phrases such as *now* and *by the way* which may be used to convey explicit information about the structure of a discourse. However, while cue phrases *may* convey discourse structure, each may also be used to different effect. The question of how speakers and hearers distinguish between such uses of cue phrases has not been addressed in discourse studies to date. Based on a study of *now* in natural recorded discourse, we propose that cue and non-cue usage can be distinguished intonationally, on the basis of *phrasing* and *accent*.

## 1. Introduction

*Cue phrases* are linguistic expressions -- such as *okay, but, now, anyway, by the way, in any case, that reminds me* -- which may, instead of making a 'semantic' contribution to an utterance (i.e., affecting its truth conditions), be used to convey explicit information about the structure of a discourse [4], [16], [5].[1] For example, *anyway* can indicate a topic return and *that reminds me* can signal a digression. The recognition and generation of cue phrases is of considerable interest to research in natural language processing. The structural information conveyed by these phrases is crucial to tasks such as anaphora resolution [6], [5], [16] and the identification of rhetorical relations among portions of a text or discourse [11], [8], [16]. It has also been claimed that the incorporation of cue phrases into natural language processing systems helps reduce the complexity of discourse processing [21], [4], [10].

Despite the recognized importance of cue phrases, many questions about how they are defined both individually and as a class -- and how they are to be represented, generated, and recognized -- remain to be examined. For example, in the general case, each lexical item that can serve as a 'cue phrase' also has an alternate interpretation.[2] While the 'cue' interpretation provides explicit information about the structure of a discourse, the 'non-cue' interpretation provides quite different information, such as conjunction (*but*) or adverbial modification (*anyway*). Distinguishing between these two uses is critical to the interpretation of discourse. In this paper, we address the problem of how this distinction might be made. We propose that, in speech, this distinction is made intonationally. We support our hypothesis by an analysis of cue and non-cue uses of the item *now* in recorded naturally occurring discourse.

In Section 2 we discuss the general problem of distinguishing between cue and non-cue usage and consider possible alternatives to our hypothesis. In Section 3 we present relevant aspects of the theory of English intonation assumed here for our analysis [13], [9]. Section 4 describes our data, presents the results of our analysis, and along with Section 5, discusses the implications of our results for the identification of cue phrases in general -- both in speech and in written text.

## 2. The Problem

Previous definitions of cue phrases as a class have been extensional and definitions of particular cue phrases procedural. For example, *now* signals a 'push' or 'pop' [5] of the attentional stack or 'further development' of a previous context [16]. Despite some recognition [5] that cue phrases are not always employed as cue phrases, no attempt has been made to discover how 'cue' uses of cue phrases are distinguished from 'non-cue' uses. When does *now*, for example, function as a discourse marker and when is it deictic?

Roughly, the non-cue or deictic use of *now* makes reference to a span of time which minimally includes the utterance time. This time span may include little more than moment of utterance, as in 1, or it may be of indeterminate length, as in 2.[3]

---

1. Previous literature has employed the terms 'clue word', 'discourse marker' or 'discourse particle' for these items [16], [4], [14], [18]. More recently Grosz and Sidner [5] have proposed the term *cue phrase* for these items, which we will adopt in this paper.

2. If 'non-lexical' items such as *uh* are classed as cue phrases, then this generalization may not hold for all cue phrases. However, even *uh* appears to have both 'cue' and 'non-cue' uses; i.e., it may signal a digression or interruption, or it may simply serve as a pause filler.

3. These and other examples are taken from a radio call-in program, Harry Gross's "Speaking of Your Money" [15]. The corpus will be described in more detail in Section 4.

**1.**

Fred: Yeah I think we'll look that up and possibly
uh after one of your breaks Harry.

Harry: OK we'll take one *now*. Just hang on Bill
and we'll be right back with you.

**2.**

Harry: You know I see more coupons *now* than I've
ever seen before and I'll bet you have too.

In contrast, the cue use of *now* signals a return to a previous topic, as in the two examples of *now* in 3, or introduces a subtopic, as in 4.

**3.**

Harry: Fred whatta you have to say about this IRA
problem?

Fred: Ok. You see *now* unfortunately Harry as
we alluded to earlier when there is a
distribution from an IRA that is taxable
...{discussion of caller's beneficiary status}...
*Now* the the five thousand that you're
alluding to uh of the --

**4.**

Doris: I have a couple quick questions about the
income tax. The first one is my husband is
retired and on social security and in '81 he
___ few odd jobs for a friend uh around the
property and uh he was reimbursed for that
to the tune of about $640. *Now* where
would he where would we put that on the
form?

While the distinction between cue and non-cue now seems fairly clear in the above examples, other cases are more difficult. Consider 5:

**5.**

Ethel: All right I have just retired from a position
that I've been in for forty some odd years. I
have -- I earned in 1981 about thirty
thousand dollars. *Now* I have a profit
sharing coming to me. My problem is shall I
take the ten year averaging...

From the transcription alone, either a cue or a non-cue interpretation is plausible. The caller might *have a profit sharing* due her at the moment of utterance (non-cue). Or, she might be using *now* to mark *profit sharing* as a subtopic (cue) -- leaving the time of the profit sharing unspecified.

How then do hearers distinguish cue from non-cue uses? One might propose that hearers use tense to delimit cases in which deictic *now* is possible. That is, it would seem

reasonable to propose that deictic *now* occurs only when the verb modified by *now* (or the main verb of the clause so modified) is temporally compatible -- i.e., *non-past*. For example, using the past tense in 1 -- *we took one now* -- seems distinctly odd. However, *we took one just now* is clearly felicitous. So, both cue and non-cue *now* are possible when the main verb is in the past tense. As examples 1- 3 above illustrate, both are also possible when the main verb is in the present tense. So, tense is clearly inadequate to distinguish between cue and non-cue uses of *now*.

Another possible diagnostic for non-cue *now* might be some notion of the general felicity of temporal reference in an utterance -- which might correspond to the felicity of substituting other temporal adverbials for *now*. For example, *we'll take one in an hour* would be felicitous in 1, as would *I see more coupons these days* in 2. Substituting other temporals for *now* in either example 3 (*Today the the five thousand that you're alluding to...*) or example 4 (*Monday where would he where would we put that on the form?*) would be infelicitous. However, this is only a necessary -- but not a sufficient -- test for deictic *now*. While a temporal adverbial *may* be substituted for *now* in 5 (e.g., *Today I have a profit sharing coming to me*), both cue and non-cue interpretations appear equally plausible from the transcription, as noted above. In fact, listeners have no hesitation in labeling this a cue *now*.

A third possibility is that hearers use surface order position to distinguish cue from non-cue uses. In fact, most systems that generate cue phrases assume a canonical (usually first) position within the clause [16], [21]. However, without intonational information, surface position may itself be unclear. Consider Example 6:

**6.**

Evelyn: I see. So in other words I will have to pay
the full amount of the uh of the tax *now*
what about Pennsylvania state tax? Can you
give me any information on that?

Although a cue reading is possible, most readers would assign *now* a non-cue interpretation if it is associated with the preceding clause, *I will have to pay the full amount of the...tax now* -- but a cue interpretation if it is associated with the succeeding clause, *Now what about Pennsylvania state tax?*. The actual recording of 6 clearly supports the latter interpretation: the strong intonational boundary between *tax* and *now* identifies the clausal boundary -- and, thus, indirectly, the surface position of *now* within its clause. Similarly, 7 would be ambiguous between a cue reading, *Well now, you've got another point*, and a deictic reading, *Well, now you've got another point* -- without intonational cues:

7.
  Fred: You stand up for your rights. Whatever you
    give to charity you claim.
  Linda:(laughs) I don't want the hassle of an of an
  Fred: Well *now* you've got another point and I
    think at at times the service counts on the
    fact that people don't want the hassle --
    and maybe we as Americans have to stand
    up a little bit more and claim what's due us.

Here it is clear from the recording that Fred intended the deictic use. Later, we will present evidence from our corpus that cue *now* can appear clause-finally, and non-cue *now*, clause-initially. So, surface position also appears inadequate to distinguish cue from non-cue *now*.

Finally, hearers might use syntactic information to discriminate between cue and non-cue usage. At least for *now*, this seems unlikely. Both cue and non-cue *now*'s are commonly classed as adverbials. So syntactic category does not differentiate. Furthermore, both can be attached at the sentence level. While non-cue *now* may also modify *VP*, it is difficult to imagine attaching cue now at that level -- since, by definition, it can make no 'semantic' contribution to either *S* or *VP*. However, this potential attachment distinction does *not* provide a means of distinguishing cue from non-cue *now* -- rather, attachment possibilities must be based on the prior cue/ non-cue distinction. So, syntactic structure provides no useful clues to the identification of cue versus non-cue usage in this case.

In summary, neither tense, nor the 'appropriateness' of temporal modification (or lack thereof), nor surface position, nor syntactic structure provides adequate information for distinguishing between cue and non-cue *now*. As we will show in the remainder of this paper, however, intonational features do provide such information.

## 3. Phrasing and Accent in English

The importance of intonational information to the communication of discourse structure has been recognized in a variety of studies [7], [20], [2], [17], [1]. However, just *which* intonational features are important and *how* they communicate discourse information is not well understood. Under-utilization of objective measures of intonational features in empirical research and the lack of a sufficiently explicit system for intonational description have made it difficult to compare and evaluate specific claims. For our study we have examined fundamental frequency (F0) contours produced using an autocorrelation pitch tracker developed by Mark Liberman. As a system of intonational description, we have adopted Pierrehumbert's [13] theory of English intonation.

In Pierrehumbert's system, intonational contours are described as sequences of low (L) and high (H) tones in the F0 (fundamental frequency) contour. A well-formed intermediate phrase consists of one or more pitch accents,

which are aligned with stressed syllables (with alignment indicated by *) on the basis of the metrical pattern of the text and signify intonational prominence, and a simple high (H) or low (L) tone that represents the **phrase accent**. The phrase accent controls the pitch between the last pitch accent of the current intermediate phrase and the beginning of the next -- or the end of the utterance. **Intonational phrases** are larger phonological units, composed of one of more intermediate phrases. At the end of an intonational phrase, a **boundary tone**, which may also be H or L and is indicated by '%', falls exactly at the phrase boundary. So, each intonational phrase ends with a phrase accent and a boundary tone.

A phrase's **tune**, or **melody**, has as its domain the intonational phrase. It is defined by the sequence of pitch accent(s), phrase accent(s), and boundary tone of that phrase. For example, an ordinary declarative pattern with a final fall is represented as H* L L% -- that is, a tune with H* pitch accent(s), a L phrase accent, and a L% boundary tone. Consider the pitch track in Figure 1 representing a simple intonational phrase composed of one intermediate phrase and with a typical declarative contour. (For ease of comparison of intonational features here, we present pitch contours of synthetic speech, produced with the Bell Labs Text-to-Speech System [12]. The analysis we will present in Section 4 is based upon recorded natural speech.)
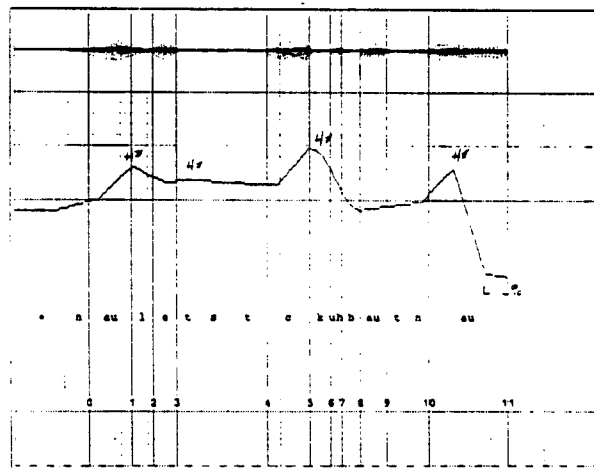


**Figure 1.** A Simple Declarative Contour

All the pitch accents in this phrase, including the **nuclear accent** -- the primary stressed syllable -- are high (H*). The phrase accent is L and the boundary tone is also low (L%).

A given sentence may be uttered with considerable variation in phrasing. For example, in Figure 1 *Now let's talk about 'now'* was produced as a single intonational phrase, whereas in Figure 2 *Now* is set off as a separate phrase.
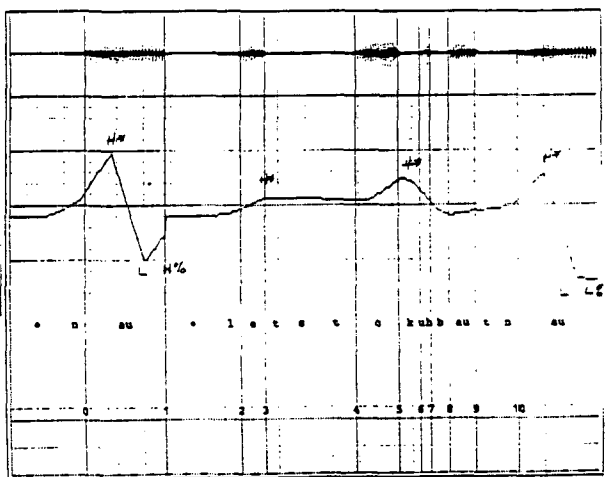
**Figure 2. Two Phrases**

The occurrence of phrase accents and boundary tones, together with other phrase-final characteristics such as pauses and syllable lengthening, enable us to identify intermediate and intonational phrases in natural as well as in synthetic speech.

**Pitch accents**, peaks or valleys in the F0 contour which fall on the stressed syllables of lexical items, make those items intonationally prominent. In Figure 3, the first instance of *now* has no pitch accent, while the second receives nuclear stress. (In our notation, the absence of a specified accent indicates that a word is **not** accented.)
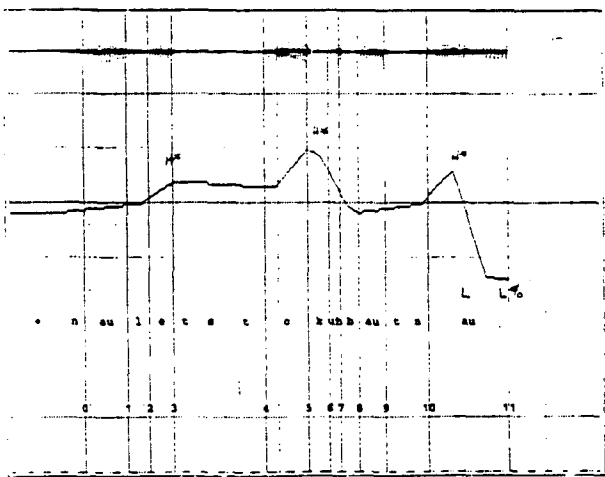


**Figure 3. Deaccenting 'Now'**

Contrast Figure 3 with Figure 1. In Figure 3, the first f0 peak occurs on *let's*; in Figure 1, the first peak occurred on *now*.

A pitch accent consists either of a single tone or an ordered pair of tones, such as L*+H. The tone aligned with the stressed syllable is indicated by a star (*); thus, in an L*+H accent, the low tone (L*) is aligned with the

stressed syllable. There are six pitch accents in English: two simple tones -- H and L -- and four complex ones -- L*+H, L+H*, H*+L, and H+L*. The most common accent, H*, comes out as a peak on the accented syllable (as, on *Now* in Figure 1). L* accents occur much lower in the pitch range than H* and are phonetically realized as local f0 minima. The accent on *Now* in Figure 4 is a L*.
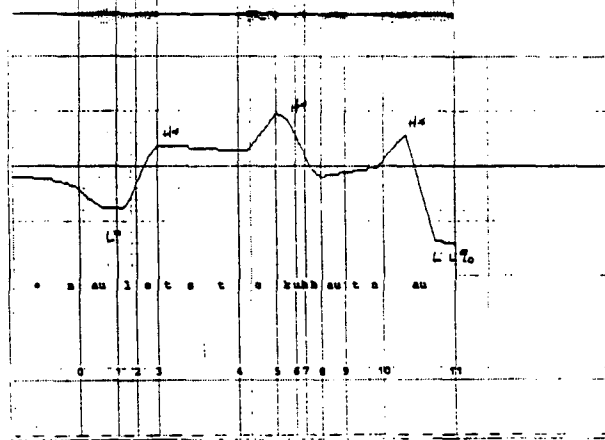


**Figure 4. Low Accent on 'Now'**

The other English accents have two tones. Figure 5 shows a version of the sentence in Figures 1-4 with a L+H* accent on the first instance of *now*.
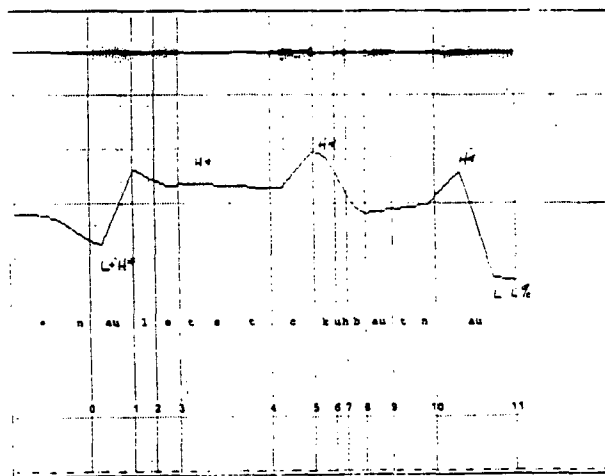


**Figure 5. An L+H* Accent**

Note that there is a peak on *now* (H*) -- as there was in Figure 1 -- but now a striking valley (L) occurs just before this peak.

While other intonational features, such as overall tune or pitch range,[4] may also provide information about cue phrase interpretation, so far we have found the most significant results by comparing accent and phrasing for cue and non-cue *now*.

166

## 4. Intonational Characteristics of Cue and Non-Cue *Now*

To investigate our hypothesis that cue and non-cue uses of linguistic expressions can be distinguished intonationally, we conducted a study of the cue phrase *now* in recorded natural speech. Our corpus consisted of recordings of four days of "The Harry Gross Show: Speaking of Your Money", recorded during the week of 1 February 1982 [15]. In this Philadelphia radio call-in program, Gross offers financial advice to callers; for the 3 February show, he was joined by an accountant friend, Fred Levy. The four shows provided approximately ten hours of conversation between expert(s) and callers.

We chose *now* to begin our study of cue phrases for several reasons. First, our corpus contained numerous instances of both cue and non-cue *now* (approximately 350 in all). In contrast, phrases such as *anyway*, *anyhow*, *therefore*, *moreover*, and *furthermore* appear fewer than ten times each. A second reason for our choice of *now* is that *now* often appears in conjunction with other cue phrases (as with *well* in 7, or *I see now*, *now another thing*, *ok now*, *right now*.) This allows us to study how adjacent cue phrases interact with one another. Third, *now* has a number of desirable phonetic characteristics. As it is monosyllabic, possible variation in stress patterns do not arise to complicate the analysis. Because it is completely voiced and introduces no segmental effects into the f0 contour, it is also easier to analyze pitch tracks reliably.

### 4.1 Sample One

Our first sample consisted of 48 occurrences of *now* -- all the instances from two sides of tapes of the show chosen at random.[5] The 48 tokens were produced by fifteen different speakers; 22.9% were produced by Harry Gross and 77.1% by other speakers.

We analyzed this data in the following way: First, three people (including the authors) determined by ear whether individual tokens were cue or non-cue. We then digitized and pitch-tracked the intonational phrase containing each token, plus (where same speaker) the preceding and succeeding intonational phrases. For this study we compared cue and non-cue uses along several dimensions: 1) We examined whether each instance of *now* was accented and, if so, noted the type of accent employed. 2) We identified differences in phrasing, including in particular whether or not *now* represented an entire intermediate or intonational phrase. 3) We noted where *now* occurred positionally in its intonational and its intermediate phrase,

whether first, not first but preceded only by other cue phrases, last, or none of these. 4) We looked at the type of intonational contour used over the phrase in which *now* occurred. 5) We noted when *now* occurred with (linearly adjacent to) other cue phrases. 6) We identified the position of the phrase containing now with respect to speaker turn. Of these, (1-3) turned out to distinguish between cue and non-cue now quite reliably. That is, accent type and phrasing distinguished between all 48 of the tokens in the sample.

Just over one-third of our sample (17) were determined to be non-cue and just under two-thirds (31) cue. The first striking difference between the two appeared in phrasing, as illustrated in Table 1: Of all the non-cue uses of *now*, none appeared as the only item in an intonational or intermediate phrase, while fully 42.0% of cue *now* represented entire intonational or intermediate phrases. (Of these 13 cue *now*'s, 8 were the only lexical item in a full intonational phrase.) A $\chi^2$ test of association between cue/non-cue status and phrasing shows significance at the .005 level $(\chi^2(1)=9.8)$.[6] So, this sample suggests that *now*'s which

| | INPHRASE | WHOLEPHRASE |
|---|---|---|
| NON-CUE | 17 | 0 |
| CUE | 18 | 13 |

**Table 1.** Phrasing for Cue and Non-Cue *Now*

are set apart as separate intermediate or intonational phrases are *very* likely to be cue nows.

Another clear distinction between cue and non-cue *now*'s in this sample emerged when we examined the position of *now* within its intermediate phrase. As Table 2 illustrates, all 31 cue *now*'s were 'first' (30 were absolutely first and

| | FIRST | LAST | OTHER |
|---|---|---|---|
| NON-CUE | 3 | 10 | 4 |
| CUE | 31 | 0 | 0 |

**Table 2.** Position within Intermediate Phrase

---

4. The pitch range of an intonational phrase is defined by its topline -- roughly, the highest peak in the f0 contour of the phrase -- and the speaker's baseline -- the lowest point the speaker realizes in normal speech, measured across all utterances. Since the baseline is rarely realized in an utterance, pitch ranges may be compared for a given speaker by comparing toplines.

5. Two instances were excluded from this sample since the phrasing was unavailable due to hesitation or interruption.

6. The $\chi^2$ test measures the degree of association between two variables by calculating the probability (*p*) that the disparity between expected and actual values in each cell is due to chance. The value of $\chi^2$ itself for (*n*) degrees of freedom (d.f.) is an overall measure of this disparity. The data show in Table 1 have $\chi^2 = 9.8$ for 1 d.f., p < .005. That is, there is less than a .5% probability that this apparent association is due to chance. Roughly, p < .01 or better is generally accepted as indicating 'statistical significance'; p > .01 becomes more controversial; p > .05 is generally considered *not* statistically significant; and p > .2 is good indication of a lack of discernible association between two variables. So, the data in Table 1, which are significant at the .001 level, appear very reliably associated.

one followed another cue phrase) in their phrase. Not only were these first in intermediate phrase -- they were also first in their (larger) intonational phrase. Only three non-cue *now*'s occupied a similar position (again, with one following a cue phrase). However, 10 non-cue *now*'s (58.8%) were *last* in their intermediate phrase -- and half of these were last in their intonational phrase. Again, the data show a very strong association ($\chi^2(2)=36.0$, p < .001). So, once intonational phrasing is determined, cue and non-cue now are generally distinguishable by position within the phrase, with cue *now*'s tending to come first in intonational phrase and non-cue *now*'s last (at least in intermediate phrase and often in intonational phrase as well).

Finally, cue and non-cue occurrences in this sample were distinguishable in terms of presence or absence of pitch accent -- and by type of pitch accent, where accented. Because of the large number of possible accent types, and since there are competing reasons to accent or deaccent items,[7] we might expect these findings to be less clear than those for phrasing. In fact, although their interpretation is more complicated, the results are equally striking.

The overall results of the 46 occurrences from this sample for which accent type could be precisely determined[8] are presented in Table 3:

|         | DEACCENTED | H*orCOMPLEX | L* |
|---------|------------|-------------|-----|
| NON-CUE | 2          | 15          | 0   |
| CUE     | 13         | 10          | 6   |

Table 3. Accenting of Cue and Non-Cue *Now*

Note first that large numbers of cue and non-cue tokens were uttered with a H* or complex accent (34.5% of cue and fully 88.2% of non-cue). The chief similarity here lies in the use of the H* accent type, with 9 cue uses and 8 non-cue (and 2 other non-cue tokens are either H* or complex). Note also that cue *now*'s were much more likely overall to be deaccented (44.8% vs. 13.3%). No non-cue *now* was uttered with a L* accent -- although 6 cue *now*'s were.

An even sharper distinction in accent type is found if we separate out those *now*'s which form entire intermediate or intonational phrases from the analysis. (Recall that these tokens are all cue uses. These *now*'s were always accented, since each such phrase must contain at least one pitch accent.) Of the 11 cue phrases representing entire phrases (and for which we can distinguish accent type precisely), 9 bore H* accents. This suggests that one similarity between cue and non-cue *now* -- the frequent H* accent

7. Such as, accenting to indicate contrastive stress or deaccenting to indicate an item is already salient in the discourse.

8. 2 cue *now*'s were either L* or H* with a compressed pitch range

-- might disappear if we limit our comparison to those *now*'s forming part of larger intonational phrases. In fact, such is the case, as illustrated in Table 4:

|         | DEACCENTED | H*orCOMPLEX | L* |
|---------|------------|-------------|-----|
| NON-CUE | 2          | 15          | 0   |
| CUE     | 13         | 0           | 5   |

Table 4. Accenting of *Now*'s in Larger Intonational Phrases

Again, these results are significant at the .001 level, $\chi^2(2)=28.1$. The great majority (88.2%) of non-cue *now*'s forming part of larger intonational phrases received a H* or complex pitch accent, while the majority (72.2%) of cue *now*'s forming part of larger intonational phrases were deaccented. Since all other cue *now*'s forming part of larger intonational phrases received a L* accent, only two *now*'s forming part of larger intonational phrases are *not* distinguishable in terms of accent type -- the two deaccented non-cue *now*'s. So, those cue *now*'s not distinguishable from non-cue by being set apart as separate intonational phrases *were* generally so distinguishable in terms of accenting. Since neither of the deaccented non-cue *now*'s appeared at the beginning of an intonational phrase -- as all cue *now*'s did -- all of the instances of *now* in our sample were in fact distinguishable as cue or non-cue in terms of their position in phrase, phrasal compostion, and accent.

We also examined whether cue and non-cue *now* patterned differently in terms of appearance with other cue phrases, with the following results:

|         | ALONE | WITHCUE |
|---------|-------|---------|
| NON-CUE | 9     | 8       |
| CUE     | 22    | 9       |

Table 5. Occurrence with Other Cue Phrases

Somewhat counter-intuitively, non-cue *now* tended to appear more frequently than cue *now* with other cue phrases -- although generally these other cue phrases were also used in their non-cue sense, e.g., *right now*. The cooccurrence is not, however, statistically significant ($\chi^2(1)=1.6$, p > .2). At any rate, the possibility that listeners identify cue *now* by its co-occurrence with other cue phrases receives no support from our data. Examination of the intonational contour used with phrases containing cue and non-cue *now*, and of the location of these phrases within speaker turn also produced no significant results.

So, we were able to hypothesize from this sample that cue and non-cue *now* are characterizable in the following ways:

Non-cue *now* forms part of larger intonational phrases and tends to be accented and to receive a H* or complex pitch accent. All non-cue uses in the sample did form part of larger intonational phrases and all but two -- which were deaccented -- were accented with a H* or complex accent. Cue *now* seems to form two classes: One class is generally set apart as a separate intermediate or intonational phrase. Something under half of our sample fell into this category. The other class, which constituted just over half of our sample, forms part of a larger intonational phrase and is either deaccented or uttered with a L* accent. Both classes share the property of appearing in initial intonational phrase position.

In summary, non-cue *now* is always distinct from cue *now* in our sample in terms of a combination of accent type, position in intonational phrase, and overall composition of the intermediate or intonational phrase. Thus we hypothesize that hearers might be able to distinguish between the two uses of *now* in three ways: by noting whether *now* formed a separate intermediate (or intonational) phrase, by locating *now* positionally within its intonational phrase, and by identifying the presence or absence of a pitch accent on *now* and the type of such accent where present. To test the validity of these hypotheses, we replicated our study with a second sample from the same corpus.

### 4.2 Sample Two

For our second sample, we examined the first 52 instances of *now* taken from another four randomly chosen sides of tapes.[9] This sample included tokens from fifteen speakers, with exactly half produced by the host and half by others.[10] This time, six people (including the authors) determined whether instances were cue or non-cue before we analyzed the intonational features. We next examined phrasing and accent used with these tokens to test the hypotheses derived from our first sample.

Again, just over one third of our sample (20) were determined to be non-cue and just under two-thirds (32) cue. The striking differences in phrasing noted between cue and non-cue *now* in sample one were again present in sample two: Again, around 40% (13) of cue *now*'s formed separate intermediate (8) or intonational (5) phrases; only one of the 20 non-cue *now*'s formed a separate intermediate phrase and none a separate intonational phrase. These results were significant at the .005 level -- again strong evidence of association between cue/non-cue status and phrasal composition. When we tested position of *now* within its intonational phrase in sample two, we again found that cue *now* generally began the intonational phrase: All but one cue *now* (this ended its phrase) began

---

9. We excluded 2 tokens from these tapes because of lack of available information about phrasing or accent and 5 others because our informants were unable to decide whether the *now* was cue or non-cue.

10. We speak to this issue below.

its phrase; again, most (60%) non-cue *now*'s came last in phrase, with two first. These results were significant at the .001 level.

Finally, our hypotheses about accent type were also borne out by our second study: The division of all cue and non-cue *now*'s by accent type appears even more pronounced in the second study: Of 20 non-cue *now*'s, 85% of non-cue were H* or complex and the rest deaccented; while of 31 cue *now*'s, 58.1% were deaccented, 19.4% H* or complex, and 22.6% L*. So, while non-cue *now*'s are almost identical to those in the first sample, cue *now*'s are more distinguished here from non-cue. When instances of *now* forming entire intermediate or intonational phrases are removed from the second sample, the accenting of cue and non-cue *now* is even more distinct: All cue *now*'s forming part of a larger phrase are deaccented, while only 15.8% of non-cue now are; the rest of the non-cue *now*'s receive a H* or complex accent (p < .001). So, our second sample confirmed our hypotheses that cue and non-cue *now* can be differentiated intonationally in terms of position within intonational phrase, composition of intermediate or intonational phrase, and choice of accent.

### 4.3 Speaker Independence

Although our second sample did confirm our initial hypotheses, the preponderance of tokens in both samples from one (professional) speaker might well be of concern. To test this, we compared characteristics of phrasing and accent for host and non-host data over the combined samples (n=100). The results showed no significant differences between host and caller tokens in terms of the hypotheses proposed from our first sample and confirmed by our second: First, host (n=37) and callers (n=63) produced cue and non-cue tokens in roughly similar proportions -- 40.5% non-cue for the host and 34.9% for his callers (p > .5). Similarly, there was no distinction between host and non-host data in terms of choice of accent type, or accenting vs. deaccenting (p > .1). Our hypothesis about the significance of position within intonational phrase holds for both host and non-host data with significance at the .001 level in each case. However, in tendency to set cue *now* apart as a separate intonational or intermediate phrase, there was an interesting distinction between host and caller: While callers tended to choose from among the two options for cue *now* in almost equal numbers (48.8% of their cue *now*'s are separate phrases), the host chose this option only 27.3% of the time. While analysis of data for callers and for all speakers shows that the relationship between cue use and separate phrase is significant at the .001 level, this relationship is not significant for the host data. However, although host and caller data differ in the proportion of occurrences of the two classes of cue *now* which emerge from our data as a whole, the existence of the classes themselves are confirmed. Where the host did *not* produce cue *now*'s set apart as separate intonational or intermediate phrases, he always produced cue *now*'s which were deaccented or accented with a L* accent. So, while individual speakers

may choose different strategies to realize cue *now*, they appear to choose from among the same limited number of options. In sum, the hypotheses proposed on the basis of our first sample are borne out by our analysis of the second -- and remain significant even when we eliminate the host from our sample.

### 4.4 Distinguishing Cue and Non-Cue Usage in Text

Our conclusion from this study that intonational features play a crucial role in the distinction between cue and non-cue usage in speech clearly poses problems for text. Do readers use strategies different from hearers to make this distinction, and, if so, what might they be? Are there perhaps orthographic correlates of the intonational features which we have found to be important in speech? As a first step toward resolving these questions, we examined the orthographic features of the transcripts of our corpus (which were prepared without particular consideration of intonational features) and made a preliminary examination of two sets of typescript interactions.

We examined transcriptions of all tokens of *now* in both our samples to determine whether phrasing was indicated orthographically.[11] Of all those instances of *now* (n=60) that were absolutely first in their intonational phrase, 56.7% (34) were preceded by punctuation -- a comma, dash, or end punctuation. 28.3% (17) were first in speaker turn, and thus othographically 'marked' by indication of speaker name. It should be noted that these units so distinguished were *not* necessarily syntactically well-formed units. So, in 85% (51) of cases, first position in intonational phrase was marked in the transcription orthographically. No *now*'s that were *not* absolutely first in their intonational phrase (in particular, none that were merely first in intermediate phrase) were so marked. Of those 23 *now*'s coming last in an intermediate or intonational phrase, however, only 60.9% (14) are immediately followed by a similar orthographic clue. Finally, of the 13 instances of *now* which formed separate intonational phrases, only 2 were so marked orthographically -- by being both preceded and followed by some punctuation. None of the *now*'s forming only complete intermediate phrases were so marked.

These findings suggest that only the intonational feature 'first in intonational phrase' has any clear orthographic correlate. However, since this feature does characterize 90.1% of the 63 cue *now*'s in our spoken data (merging both samples) -- and since 85.0% of these cue *now*'s are also orthographically marked for position as well (so that 80.1% of cue *now*'s can be orthographically distinguished) -- it seems that this correlation between intonation and orthography may be a useful one to pursue. It is also possible that a perusal of text, rather than transcribed speech, might indicate more orthographic clues to cue/non-cue disambiguation. We are currently examining two sets of

typescripts[12] of task-oriented text interactions.

### 5. Conclusions

Our study of the cue phrase *now* strongly suggests that speakers and hearers can distinguish between cue and non-cue uses of cue phrases intonationally, by making or noting differences in accent and phrasing. Cue and non-cue *now* in our samples are reliably distinguished in terms of whether *now* forms a separate intermediate or intonational phrase, whether it occurs first in its intonational phrase, and whether it is accented or not -- and, if accented, the type of accent it bears. In the absence of alternate known means of distinction between cue and non-cue use, we propose that speakers and hearers do differentiate intonationally. Our next step is to extend our study to other cue phrases, including *anyway*, *well*, *first*, and *right*. We also plan to examine the relationship between cue usage and pitch range manipulation [7], another indicator of discourse structure. The goal of our research is both to provide new sources of linguistic information for work in plan inference and discourse understanding, and to permit more sophisticated use of intonational variation in synthetic speech.

---

11. No instances of capitalization or other othographic marking of nuclear stress appear in any of the transcripts.

---

12. Ethel Schuster's transcripts of students being tutored in EMACS [19] and transcripts of people assembling a water pump [3]

# REFERENCES

1. Brazil, D., Coulthard, M., and Johns, C. *Discourse intonation and language teaching.* Longman, London, 1980.

2. Butterworth, B. Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research 4* (1975), 75-87.

3. Cohen, P., Fertig, S., and Starr, K. Dependencies of discourse structure on the modality of communication: telephone vs. teletype. In *Proceedings of the ACL,* ACL, Toronto, 1982, pp. 28-35.

4. Cohen, R. A computational theory of the function of clue words in argument understanding. In *Proceedings of COLING84,* COLING, Stanford, 1984, pp. 251-255.

5. Grosz, B. and Sidner, C. Attention, intentions, and the structure of discourse. *Computational Linguistics 12,* 3 (1986), 175-204.

6. Grosz, B.J. The Representation and use of focus in dialogue understanding. 151, SRI International, 1977. University of California at Berkeley PhD Thesis.

7. Hirschberg, J. and Pierrehumbert, J. The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting,* Association for Computational Linguistics, New York, 1986, pp. 136-144.

8. Hobbs, J. Coherence and coreference. *Cognitive Science 3,* 1 (1979), 67-90.

9. Liberman, M. and Pierrehumbert, J. Intonational invariants under changes in pitch range and length. In *Language sound structure,* M. Aronoff and R. Oehrle, Eds. MIT Press, Cambridge, 1984.

10. Litman, D. and Allen, J. A Plan recognition model for subdialogues in conversation. *Cognitive Science 11* (1987), 163-200.

11. Mann, W.C. and Thompson, S.A. Relational Propositions in Discourse. ISI/RR-83-115, ISI/USC, November 1983.

12. Olive, J.P. and Liberman, M.Y. Text to speech -- An overview. *Journal of the Acoustic Society of America, Suppl. 1 78,* Fall (1985), s6.

13. Pierrehumbert, J.B. The phonology and phonetics of English intonation. PhD Thesis, Massachusetts Institute of Technology, 1980.

14. Polanyi, L. and Scha, R. A Syntactic approach to discourse semantics. In *Proceedings of COLING84,* COLING, Stanford, 1984, pp. 413-419.

15. Pollack, M.E., Hirschberg, J., and Webber, B. User Participation in the Reasoning Processes of Expert Systems. MS-CIS-82-9, University of Pennsylvania, 1982. A shorter version appears in the AAAI Proceedings, 1982.

16. Reichman, R. *Getting computers to talk like you and me: discourse context, focus, and semantics.* MIT Press, Cambridge MA, 1985.

17. Schlegoff, E.A. The relevance of repair to syntax-for-conversation. In *Syntax and semantics,* 12: Discourse and syntax, T. Givon, Ed. Academic, New York, 1979, pp. 261-288.

18. Schourup, L. *Common discourse particles in English conversation.* Garland, New York, 1985.

19. Schuster, E. Explaining and Expounding. MS-CIS-82-49, University of Pennsylvania, 1982.

20. Silverman, K. Natural prosody for synthetic speech. PhD Thesis, Cambridge University, 1987.

21. Zukerman, I. and Pearl, J. Comprehension-driven generation of meta-technical utterances in math tutoring. In *Proceedings of the 5th National Conference,* AAAI86, Philadelphia, 1986, pp. 606-611.