

**EVALUATING DESCRIPTIONS OF INTONATION**  
**A COMPARISON OF**  
**DISCOURSE INTONATION AND CRYSTAL'S DESCRIPTION**

**RICHARD T CAULDWELL**

**English for Overseas Students Unit**

**School of English**

**University**

**December 1993**

**Recompiled April 2014**

## Introduction

The purpose of this paper is to compare Discourse Intonation (DI) (Brazil 1978, 1985) and Crystal's (CRYS) description of the systems of intonation (Crystal, 1969, 1975; Crystal & Davy, 1975). The comparison reported below was conducted in order to justify my non-adoption of an attitudinal approach to intonation for my PhD research. For the PhD, which involved analysis of poetry readings, introductions and interview responses of Philip Larkin, I adopted DI. I chose to compare DI's performance with CRYS because the latter is widely known and has been used in the analysis of poetry readings (e.g. Byers, 19XX). Most importantly CRYS has published tapes with accompanying transcriptions to illustrate his description at work (Crystal & Davy, 1975). The existence of tapes and transcriptions makes it possible to compare how CRYS transcriptions relate to DI transcriptions.

The adoption of a description of intonation involves a standardisation process which requires new transcribers to attune themselves to perceive and notate those moments in speech which that description finds significant. The transcribers' perceptual capacity is sensitized to notice those features of speech that the description has deemed important and to ignore other features that are not deemed important. The standardisation process requires considerable personal investment in hours of ear-training and learning how to interpret transcriptions. It is thus difficult for transcribers who are accustomed to working with the conventions of one description to change to using the conventions of another description. It is not the purpose of this paper to persuade adherents of other approaches to adopt DI: its purpose is to inform users of DI how it compares in relation to CRYS in the transcription of a piece of data.

Attitudinal approaches offer something very attractive. They propose that there are links between a given intonational feature, such as 'rising intonation' and attitude such as 'scorn'. This claim, if true would seem to make it possible to identify, merely by looking at a transcription, any attitudes that the reader of poetry may be conveying.

There are a number of descriptions of intonation which view attitude as the most important component of meaning. I could have chosen other 'attitudinal' descriptions, notable among them being O'Connor and Arnold (1973). However the advantage of using Crystal's description is that he has published tapes and transcriptions of spontaneous spoken texts (Crystal & Davy, 1975) – it is thus possible to teach oneself how to operate the descriptive apparatus on one's data. O'Connor and Arnold have published a tape but it consists solely of the fabricated examples which illustrate their textbook<sup>1</sup>. An additional advantage of adopting Crystals description is that with it he

---

<sup>1</sup> This does not mean that it is impossible to use O'Connor and Arnold's transcription scheme to represent all intonational phenomena. Indeed Knowles (1991) and Wichmann (1987) do so. It does mean however that it is not possible to train oneself to use it on spontaneous data.

performed a number of experiments which attempted to validate his hypothesis about the attitudinal approach.

## 1. Principles

There are four important publications relating to Crystal's description of the systems of intonation:

- Prosodic Systems and Intonation in English (Crystal 1969, henceforth PSIE)
- The English Tone of Voice (Crystal 1975, henceforth ETV)
- Investigating English Style (Crystal & Davy, 1969, henceforth IES)
- Advanced Conversational English (Crystal & Davy 1975, henceforth ACE)

PSIE and ACE are particularly important. PSIE contains the fullest statement of CRYIS, and ACE is accompanied by a tape and thus provides the necessary data for transcriber training.

CRYIS views the systems of intonation as being a sub-set of the prosodic systems of English, and that the prosodic systems are part of a more general phenomenon – non segmental phonation.

Table 1 (cf. page 5) shows that there are seven prosodic systems<sup>2</sup>; the labels available in each system vary in number from five (*voice qualifications*) to fourteen (*pitch range*). This compares with a total of four systems in DI, where the choices range from two (*prominence*) to five (*tone*).

CRYIS uses the terms 'system' to refer to a phonetic parameter or cline, (e.g. *loudness*) for which he presents a series of labels. These labels enabled him to characterise degrees of gradience along the cline: for example for (part of) the 'loudness' parameter: *piano, pianissimo, forte, fortissimo*.

CRYIS determines the number of labels for each system by making an appeal to *linguistic significance*. He asks whether a feature's use

can distinguish two otherwise identical utterances, so that linguistically untrained native speakers would consistently maintain the two utterance as being in some sense 'different' in meaning. (PSIE p. 19).

The advantages of having the labels-on-a-cline approach is that there is not theoretical limit to the number of terms and distinctions one can introduce. It is thus possible to adjust the number of terms to cope with all the fine phonetic phenomena of a particular piece of data, and thus produce a very finely graded transcription of a recording. The

---

<sup>2</sup> Tension is intermediate between being a prosodic system and a paralinguistic system

major disadvantage of the labels-on-a-cline approach is that it does not differentiate between gradience and system (in the DI sense).

CRYS's view is that prosodic phenomena have three functions: *accentual*, *attitudinal* and *grammatical*. The accentual function is 'governed by non-linguistic, situational factors and is unpredictable from the grammar' (PSIE, p. 263) – and it is mostly brought into play by the choice of tonicity (cf. 1.2 below).

CRYS takes the view that the attitudinal function is an ever-present function:

all cases where intonation is primarily of grammatical importance are also of attitudinal relevance, but not all cases of attitudinal function (i.e. all utterances) display a grammatical function (PSIE) p. 289-90)

But six years later, by the time of writing ETV (1975), he views most intonational phenomena as having no meaning:

the vast majority of tones in connected speech carry no meaning – that is, they communicate no new information, because their occurrence is syntactically predictable. (ETV, p. 34)

CRYS views intonation as a phenomenon which

has a very clear centre of pitch contrast and a periphery of reinforcing (and occasionally contradicting) contrasts of a different order. (PSIE p. 196).

Thus of the prosodic systems listed in Table 1, 'pitch movement' and 'pitch range' are central, with 'loudness' and 'pause' being next in importance.

Table 5.1 Summary of the systems of Non-segmental phonation.

NON-SEGMENTAL PHONATION										
← More linguistic					Less Linguistic →					
Prosodic systems					Paralinguistic systems					
Tone	Pitch-range	Pause	Loudness	Tempo	Rhythmic-ality	Tension	Voice Qualifiers	Voice Qualifications	Voice Quality	Vocal Reflexes
eg <i>falling</i> <i>rising</i>	eg <i>booster</i> <i>low drop</i>	eg <i>unit</i> <i>brief</i>	eg <i>stressed</i> <i>unstressed</i> <i>piano</i> <i>forte</i>	eg <i>clipped</i> <i>drawled</i> <i>allegro</i> <i>lento</i>	eg <i>rhythmic</i> <i>arhythmic</i> <i>spiky</i> <i>glissando</i>	eg <i>stared</i> <i>lax</i> <i>tense</i> <i>precise</i>	eg <i>whisper</i> <i>breathly</i> <i>husky</i> <i>creak</i>	eg <i>laugh</i> <i>giggle</i> <i>tremulous</i> <i>sob</i> <i>cry</i>		eg <i>cough</i> <i>sneeze</i>

Table 5.1 Summary of the systems of Non-segmental phonation.

[TABLE 1 AROUND HERE]

## 1.1 Tone units

CRYS's tone unit has a maximum of four structural parts: the *pre-head*, the *head*, the *nucleus*, and the *tail*, of which the first three have importance in expressing differences in meaning. The following tone unit, on which CRY5 and DI agree (as far as their different descriptions allow) shows CRY5's notation conventions for the tone unit:

CRYS	<i>Pre-head</i>	<i>Head</i>	<i>Nucleus</i>	<i>Tail</i>
	that there was	so many	ENtrances	'all round
DI	that there was	SO many EN	trances all round	
	<i>Proclitic segment</i>	<i>Tonic segment</i>	<i>Enclitic segment</i>	

ACE, p. 21

CRYS's 'pre-head' is roughly equivalent to DI's proclitic segment; CRY5's head is roughly equivalent to DI's tonic segment, except that DI includes the tonic in the tonic segment, whereas CRY5 gives the nucleus its own structural position after the head.

Unlike HALL and DI, CRY5 defines the phonetic criteria for the tone-unit boundary. He claims that there are three factors in normal (meaning 'not too hurried' PSIE p. 205) speech which enable one to recognise a tone unit boundary:

- the presence of a preceding nuclear tone
- a change in pitch direction to the onset syllable of the head of the following tone unit
- juncture features such as a pause.

## 1.2 Tonicity<sup>3</sup>

CRYS regards tonicity as having either an accentual or a grammatical function 'the former governed by non-linguistic, situational factors, and is unpredictable from the grammar.' (PSIE, p. 263). For example in

'I thought she was wearing a very nice dress'

tonicity is not grammatically predictable, therefore it is accentual:

the placement of the nucleus is a variable which functions independently of grammatical considerations; it is up to the speaker which word he emphasises most strongly, and he is bound only by general situational restraints, to defy which might produce nonsense, but not ungrammaticality (PSIE p. 263)

---

<sup>3</sup> Tonicity is HALL's term, it is not one that either CRY5 or DI adopt, but it provides a convenient shorthand, so I use it here.

CRYS identifies eleven cases (which he says represent a selection, not a complete list) in which tonicity has a grammatical function. His examples<sup>4</sup> include 'Was she wearing a green dress or a red one?' on which he comments:

[this] ... has non-optional tonicity: the item 'red' must be nuclear because of the grammar of the co-ordinate construction (p. 263)

It could be argued (from the DI perspective) with equal justification that it is the context implied by the wording of the example which makes it likely that the nuclear tone will go on *red*. The context is one in which there is an existential paradigm consisting of two members, a 'green dress' and a 'red dress', and that colour is the only means of differentiating them, and therefore conversational conditions are such that it is necessary for the speakers to establish which dress was worn. It is only possible to imagine one context in which this question is likely to be asked: this is what makes the nuclear tone on red 'compulsory', not the fact that it is an alternative question.

CRYS discusses the accentual function further by analyzing the form classes on which the tonic falls. He states that his finding agree 'very largely' with that reported by Quirk et al. (1964) who reported that 58% of tonics fall on nominal groups and that 22% fall on verbal groups (cited in PSIE p. 266). This is, clearly, a suggestion that the accentual and the grammatical motivations for tonicity are linked.

By 1975 CRYS is noting that tonicity may also be conditioned by attitude. In discussing the following example 'what are you doing it FOR/(context e.g. impatience)' (ETV, p. 24) he comments

one has to assume that it is the attitude of the speaker which has conditioned the tonicity

CRYS's ideas about the motivation of tonicity thus vary: it can be accentual, grammatical or attitudinal depending on the example. His explanation of the accentual reasons for tonicity is interesting in that he opens the door to context, but ultimately he is loathe to let go of the concept of 'normal intonation':

There would ... seem to be a strong pressure for lexical words to retain some prominence even when tonicity is marked on grammatical items (PSIE, p. 267)

---

<sup>4</sup> The other ten cases of grammatically conditioned tonicity include restrictive and non-restrictive clauses, end-placed vocative and apposition, etc.

### 1.3 Tone choice

CRYS recognises two types of tones: nuclear tones, which occur in the nuclear position in the tone unit, and non-nuclear tones which can occur in the head of the tone unit.

#### Nuclear tones

The number of tones that CRYs recognises varies from publication to publication. PSIE (p. 225) lists a total of twelve, ACE (p. 17) lists six, but the transcripts contain notation for a seventh, the rise + fall. The comparison below assumes that CRYs's description has seven tones.

CRYS	Label	DI
\	Falling	p
/	Rising	r+
–	Level	o
∨	Falling-rising	r
∧	Rising-falling	p+
\ + /	Fall + rise	
/ + \	Rise + fall	

Table 2 Crystal's nuclear tones (as used in ACE) with DI equivalents

Table 2 shows CRYs's tones aligned with their DI equivalents. The seven tones listed in the table are largely compatible with HALL's, except that CRYs has a level tone.

CRYs also allows for non-nuclear kinetic tones (PSIE, p. 221) or 'glides' occurring before the nuclear tone, which are not in themselves nuclei. These he sees as instances of 'glissando' (a label from the rhythmicality parameter). He admits that it is possible to mistake such glides for the first elements of compound tones, and he gives five criteria for distinguishing between instances of glissando and the first element of compound tones, one of which is that pauses may occur between the last element of a glissando sequence of syllables 'whereas this rarely happens in a compound tone' (PSIE, p. 222).

We can observe the effects of a hidden glissando if we examine the longest tone unit in the CRYs transcription. It consists of fifteen words and seventeen syllables (counting *er* as one syllable and *every* as two syllables) and is rendered thus in ACE:

| | | mean you ↑ don't – er | every 'now and a'gain the ↑ team 'builds a 'new STAND |

This was transcribed as three tone units in DI:

```
01 || ... i mean you DON'T... ||
02 || p EVery now and again ||
03 || p the team BUILDS a new STAND ||
```

DI has a falling tone beginning on the first syllable of *every* which continues over the words *now and again*; CRYs interprets this as non-nuclear glided, an instance of



glissando. However, in line with a number of simplifications which took place between PSIE and ACE, the glissando is not marked.

For CRY5, there are thus two possible interpretations for change in pitch direction occurring before a pause: either (a) nucleus + tone unit boundary, or (b) glissando elements + pre-nuclear pause. This leads to problems for the transcriber.

#### 1.4 Stress

PSIE is careful to provide separate notations for loudness and movements upwards in pitch. In practice, speakers rarely make syllables louder without also making a concomitant change in pitch. This fact is recognised in ACE where the two parameters are at some points conflated as the list of symbols for stress in ACE shows:

| = first prominent syllable of the tone-unit

' = the next syllable is stressed

↑ = the next syllable is stressed and also steps up in pitch

'' = extra strong stress

CRY5 uses the term 'prominent' in a different sense from DI. For CRY5 this means the first stressed syllable in the tone unit, that is the syllable that marks the beginning of the head. For CRY5 no particular meanings are attached to the occurrence of stress, although in the experiment reported below he notes a link between:

(a) strong stressed syllables and the attitudinal labels 'angry', 'disapproving', 'vexed', 'impatient', 'irritated'

(b) high unstressed syllables and the attitudinal labels 'angry', 'bored', 'questioning', 'conspiratorial', 'impatient', 'satisfied'

For more on this see 4 below.

#### 1.5 Pause

CRY5 distinguishes between silent and voiced pauses noting that there is a formal, but not necessarily functional difference between the two. For voiced pauses he proposes using phonetic symbols, but in ACE such notation is not used: filled pauses are represented orthographically as *er*, *erm*, etc. For silent pauses he proposes four labels:

Brief pause     ` (raised period)

Unit pause     -

Double pause --

Treble pause ---

A 'unit pause' is defined as 'the interval of an individual's rhythm cycle from one prominent syllable (arsis) to the next, within a stable tempo' (PSIE, p. 171). Crystal's definition regards pause length to be relative to the speakers' speed of utterance: the unit pause being analogous to one rhythmic beat, and the double and treble pauses being two and three times the length of a unit pause.

A brief pause is defined as 'a silence perceivable shorter than (and usually approximately half as long) a unit length' (PSIE p. 171). The length of pause that is a junctural feature of a tone unit boundary 'is never as long as brief' and is not given a separate notation (PSIE p. 171).

## **1.6 Secondary systems**

In PSIE CRY5 separates the two parameters of pitch movement and pitch range. That parameter known as the 'simple pitch range system' (syllables) has seven labels: an unmarked norm (for which there is no notation) which assumes that in any sequence of two syllables, the second one will be at a slightly lower level; a label to indicate where a second syllable stays at the same level as the first (continuance); three labels for jumps upwards (booster, high booster, and extra high booster) and two labels for a jump down in pitch (drop and low drop). These seven labels are reduced to two in ACE: only the declination line (no notation) and the booster (a jump up in pitch) remain.

## **2. Comparison**

The methodology for the investigation follows closely that of the investigation into Halliday's approach in Cauldwell (1993). I made a DI transcription of the tape of the 'Talking about Football' text (ACE pp. 19-23) without referring to Crystal's transcription. Because of the quality of the tape, I was not able to hear all the utterances printed in the text of the discussion. These unheard utterances – numbering 37 Crystal tone units in all – have been left out of the investigation.

In what follows, I briefly state the differences between the CRY5 and DI transcriptions of this text, and then make broad-brush statements (generalisations) of what the consequences would be if these differences were true for all transcriptions made in these descriptions. These generalisations are preceded by the word 'Thus' and are given in bold.

## 2.1 Tone units

The DI transcription of the Football text has 88 more tone units than CRY5: the DI transcription contains 310 tone units, the CRY5 transcription contains 222 tone units. Thus

- **a DI transcription will have 40% more tone units than a CRY5 transcription (if the difference reported here were to be true of all other transcriptions)**

There is, from the CRY5 point of view, substantial agreement between the two systems on where the tone unit boundaries lie: where CRY5 has a tone-unit boundary, there is a 95% chance that DI will also have one. DI however places a boundary wherever there is a pause, whereas CRY5 allows pauses to occur inside tone units.

The DI transcription includes 26 incomplete tone units – that is tone units without a tone – which CRY5 prefers to accommodate within a tone unit<sup>5</sup>.

For example, where (below) DI has the first tone unit in the text as incomplete, CRY5 transcribes the same stretch as another head within the tone unit:

DI

```
01 || ... well WHAT'S the ... ||  
02 || p what's the FAILure with the FOOTball ||
```

CRY5

| well | what's the . | what's the failure with the FOOTBALL |

There are 26 incomplete tone units of this type. Thus of the total of 88 DI's extra tone units, there remain 62 (88-26=62) to account for.

CRY5 allows for compound tones within a single tone unit:

| I went to 'Stamford ↑ \ BRIDGE last year / ONCE |

```
01 || p I went to stamford BRIDGE ||  
02 || r+ LAST year ONCE ||
```

In this example there is a falling tone which starts on BRIDGE, and a rising tone which starts on ONCE. For CRY5, this is a single compound tone; for DI the tones on BRIDGE and ONCE are separate choices. CRY5 transcription contains 24 such compound tones in the data. This accounts for 50 (together with the 26 incomplete tone units) of the differential of 88.

There thus remain 38 (of the original differential of 88) DI tone units to be accounted for: these extra tone units occur where DI transcribes a tone where CRY5 has a stress (of

---

<sup>5</sup> CRY5's own transcription contains seven incomplete tone units which are not accommodated inside complete tone units.

one kind or another). DI transcription has more tones and hence more tone units than a CRY5 transcription.

## 2.2 Tonicity

There are only three occasions in the data where CRY5 has a tone unit boundary and DI has none. Thus

- **99% of the time, where CRY5 has a tone unit boundary, DI will have one**

There are only 17 occasions where CRY5 has a tone and DI not mark a tone. Thus

- **93% of the time, where CRY5 has a tone, DI will also have a tone**

However, because DI has more tones and tone units than CRY5, the degree of compatibility is not as high when viewed from the DI perspective. Thus

- **71% of the time, where DI has a tone, CRY5 will also have a tone<sup>6</sup>**

There are thus 29% of DI tones which CRY5 does not transcribe as nuclear. This is a significant proportion of non-agreement, which we come to below. It is worth mentioning now that in by far the majority of these cases involve DI have tones on syllables that CRY5 marks as stressed. What is it that DI hears on these stressed syllables, that CRY5 does not represent in transcription?

---

<sup>6</sup> In terms of raw numbers, this means that of the 284 tones that DI captures, 201 (70.77%) occur on syllables which Crystal transcribes as nuclear.

## 2.3 Tone choice

In this section I investigate how far DI and Crystal agree in assigning tone choice.<sup>7</sup>

	% Matches	% Non-matches
p	81	19
r+	76	24
r	67	33
o	11	89
Ave	67	33

**Table 3 Percentage of matches between ACE tones and DI tones**

The last row of Table 3 shows that of the total number of DI tones, 67% match with CRY5 tones. Thus

- **67% of the time, where DI has a tone, CRY5 will have a matching tone**

The following are examples of agreement for each tone:

CRY5 | what's the 'failure with the | ↑ \ FOOTBALL |

DI // what's the FAILure with FOOTball //

(ACE, p. 19 TU 1)

CRY5 | have a 'nice | ↑ -- MEAL |

DI // o HAVE a nice MEAL //

(ACE, p. 20 TU 53)

CRY5 | of his V NORMAL |

---

<sup>7</sup> In order to make the comparison, certain adjustments have to be made to ensure that the counts are directly comparable. For DI I have removed from the calculation the 12 instance of uncertain tone and the 26 incomplete tone units (which of course have no tones): there are thus 272 (310-26-12=272) DI tones to account for. For CRY5 I have counted the compound tones (\+/ & /+\) as pairs of simple tones: this brings the CRY5 tone count up from 222 to 246. But from this total of 246 it is necessary to deduct the number of tone units with no nuclear tone – 7 – making 239 Crystal tones in all. The following count as matches: (a) where p tone coincides with a \ (fall) or the first part of \+/ (fall + rise) or the second half of a /+\ (rise + fall); (b) where r+ tone coincides with a / or the second part of a \+/ or the first half of a /+\; (c) where r tone coincides with either \/ or \+/\

DI // r of his NORMal //

(ACE, p. 20 TU 43)

CRYS | all 'ninety /TWO |

DI // r+ ALL ninety TWO //

The figure of 67% represents the average number of matches for all instances of the four tones occurring in the data (p+ tone does not occur). Table 3 shows that the percentage of matches varies from a high of 81% (for p tone/fall) to a low of 11% (for level tones).

	Other tones	Stressed Syllables	Unstressed Syllables
p	2	22	3
r+	4	7	1
r	4	8	0
o	6	23	10
Totals	16	60	14

**Table 4 Co-occurrence of non-matches with other tones, stressed and unstressed syllables**

Table 4 shows whether the non-matches occurred with (a) tones other than those with which they were expected to match; (b) on syllables that CRY S has as stressed (c) on syllables that CRY S has as unstressed. The last row in the table shows that 16 cases in which CRY S and DI agreed there was a tone, but disagreed as to which tone it was. This is a level of 6% disagreement and is the level one would expect from the auditory method of transcription.

There are 60 cases where DI has a tone, and CRY S has a stressed syllable. By far the majority of them (45) concern p tones and o tones. This amounts to 22% of the DI tone total, and therefore requires some explaining. There are 14 cases where DI has a tone where CRY S has an unstressed syllable, although this is a small proportion of the the overall numbers, 10 of these cases occur where DI has o-tone. It is clear from this that an explanation is needed of non-matches of p tone and o tone. Before I embark on the explanation of why this is so, I wish to compare DI's and CRY S's treatment of what CRY S terms stress, pause and pitch range.

## 2.4 Stress

words	syllables	% stresses	% prominences
1338	1037	41	31

**Table 5 Percentage of Stresses and Prominences as a proportion of syllables**

Table 5 shows that there were 1037 transcribable words consisting of 1338 syllables in the football text<sup>8</sup>. DI has fewer prominences – 31% of all syllables – than CRY5 has stresses – 41% of all syllables. These percentages are very similar to those for HALL comparison. It would thus seem to be safe to assert that for both scripted and non-scripted data:

- **DI's proportion of prominences will be 10 % less than CRY5's proportion of stresses**

## 2.5 Pause

In the data CRY5 has 109 pauses. Of these, 68 co-occur at tone unit boundaries, and 41 occur inside tone units. Of the 41 which occur inside CRY5 tone units, 40 co-occur with DI tone unit boundaries. The remaining pause was not heard by me to be a pause.

## 2.6 Pitch range: Key and Termination

As mentioned above, CRY5's ACE transcriptions indicate whether or not there is a step up in pitch by use of an up arrow [↑]. The absence of an arrow can mean one of three things: (a) the syllables are following the declination line (a gradual lowering of pitch on each successive syllable) (b) the syllable is at the same pitch as the preceding one (c) there is a step down bigger than would be expected from following the declination line.<sup>9</sup> To determine the extent to which this notation is compatible with DI's system of key and termination, I looked in detail at what happened to choices of termination on p tone, the most common tone in the data.

What we might expect is as follows:

- DI's high termination would co-occur with CRY5's booster
- DI's mid and low termination would co-occur with the absence of a booster (i.e. with the declination line)

---

<sup>8</sup> Those words that I was not able to catch from the tape when doing the DI transcription are not included in the word count.

<sup>9</sup> PSIE had a notation for all these (and more) possibilities, see Table 1 above.

	% High Termination	% Mid Termination	% Low Termination
Booster	44	49	7
No Booster	16	66	18

**Table 6 Percentage matches between Boosters and Termination**

(Table 6 excludes from the calculation cases where DI has a p tone and CRYs has no tone).

Table 6 shows the 44% of the time, a 'booster' will co-occur with high termination, and 56% of the time it will not. The absence of a booster co-occurs in 84% of cases (66% mid and 18% low) with non-high termination choices. The table shows the number of non-matches: 41 of the total of 117 cases, that is 35% are non-matches. This degree of non-matching is a result of CRYs and DI attending to different features of the stream of speech to determine where significant steps in pitch level take place: like HALL, CRYs attends to the immediately preceding syllable: DI attends to the preceding choice of key on a previous onset prominence. This difference can result in completely opposing views of which significant choices have been made: this difference is seen at its most extreme on the four occasions where CRYs has a booster where DI has low termination.

| where I |went to see CHELSEA | play | ↑ LEEDS |

| - and | Leeds 'played SHOCKINGLY | - | worst game the | ↑ ever PLAYED |

(ACE pp. 22-23 TUs 234-238)

there is no way of knowing from looking at this transcription that both *Chelsea* and *Leeds* are both are both mid-termination choices (the presence of a booster implies a high termination on *Leeds*). The transcription implies that *Chelsea*, *shockingly*, and *played* are at the same pitch level. In fact both *shockingly* and *played* have big jumps downwards, and thus realise low termination in the DI transcription.

// p i went to see CHELsea //

// p play LEEDS //

// p and LEEDS played ↓ SHOCKingly //

// p WORST game they ever ↓ PLAYED //

These differences are not attributable to transcriber error; they are due to the conflict between CRYs's description which attends to differences between successive syllables and DI which attends to relationships between prominences.



Part of the reason for the difference is that DI accepts, as realisations of p tone, falls which are preceded by a slight jump up in pitch. This has the consequence that for a high termination to be selected in DI, the speaker has to produce a jump in pitch which is notably higher than those produced by the little, almost automatic, jumps which precede a fall. CRY5 (in PSIE) regards the small pre-fall jumps as a phenomenon to be separately notated, and does not distinguish, as DI claims to do, between the automatic jumps (necessary concomitants of the fall) and the jumps made as a deliberate choice.

CRY5's notation originally (PSIE) contained 17 labels for upward and downward movement in pitch range. In ACE this is reduced to two: there is only the declination line (un-notated) by which to indicate the "downward" relationship between syllables. In ACE therefore the absence of notation preceding a syllable means that either the following syllable is on the same level, or it follows the declination line downwards. Thus from looking at a CRY5 transcription it is impossible to tell whether a syllable stays at the same level as the preceding syllable, whether it is marginally but non-significantly lower than it, or whether it is significantly lower than it.

### 3. Discussion

It is now time to return to discussion of the discrepancy in the numbers of tones between DI and CRY5 which was postponed from 2.3 above

#### 3.1 Non-nuclear glides and p tones

There are 23 falling "tones" which ACE does not represent in the transcription.

##### Falling "tones" and non-nuclear glides

It would appear that, as with HALL, CRY5 interprets as "stress" certain moments in the data which DI interprets as "stress plus fall". As pitch movement is one of the phonetic correlates, there are bound to be cases of legitimate disagreement between descriptions on whether or not the pitch movement is sufficient to indicate stress alone, or is great enough to signal tone as well. It would seem that in an example such as the following, the upward jump to a stressed syllable on *all* followed by the declination line is in DI's terms enough to warrant a transcription of a p tone:

CRY5 | - - and he | goes to | ↑ all the 'matches AWAY |

DI // p he GOES to ALL the matches // r+ aWAY //

(ACE, p. 19-20 TU 33)

A proportion of the disagreements can be attributed to such cases. Others, though, can be attributed to the pressure to make tone unit boundaries match grammatical boundaries.

PSIE's concept of the non-nuclear glide allows the transcription to include pre-nuclear falls. The 'notation' for such falls consists of notes in the margin indicating the stretch of syllables over which these glides occur. This method of signaling non-nuclear falls is not used in ACE.

In the following example the syllable *one* though it has a fall (a p tone in DI terms) is not marked as nuclear because this would force the CRYSTranscriber to place a tone unit boundary in the middle of the noun group *only one modern ground in England*:

CRYSTranscription: || think he 'said there was | ↑ only 'one 'modern GRO / UND in ↑ ENG \ LAND |  
 DI transcription: // p i think he said there ONLY ONE // r MODern ground // p in ENGLand //

(ACE, p. 21 TU 131)

In order to preserve the clause-element tone unit link, CRYSTranscription seems to ignore "tones" that a DI transcription would capture.

### 3.2 o tones and pauses

Like DI but unlike HALL, CRYSTranscription allows for the existence of a level tone. However, CRYSTranscription has only 10 occurrences of level tone against DI's count of 44. They agree on the location in only five cases. Of the non-matches, six occur on other tones, ten on syllables CRYSTranscription has as unstressed and 23 on syllables CRYSTranscription marks as stressed.

In the following example, the two descriptions agree that the tonic/nuclear tone on *surroundings* is level but DI has three other level tones:

CRYSTranscription: |. in .| very . | ↑ plushy SURROUNDINGS |  
 DI transcription: // o IN // o VErY // o PLUSHy // o surROUNDings //

(ACE, p. 20 TU 54)

In the CRYSTranscription transcription the first syllables in *very* and *plushy* are both marked as stressed but *in* has no such marking and is thus unstressed. But (to a DI transcriber's ears) there is no perceptible difference between the realisations of *in* and of the first syllables of *very* and *plushy*.

In the football recording, the words on which CRYSTranscription does recognise level tones are all 'lexical':

AMERICA	west BROMWICH
BOB	MEAL
SURROUNDINGS	WARM

WAY	get OUT
KNOW	CIGARS

In contrast, less than 50% of DI's total of o tones (19 out of 40) occur on lexical words. The majority of o tones occur on grammatical words of 'filled pauses': 9 cases on *er* or *erm*; three on *about* two on *and* and there are single occurrences on words such as *who*, *this*, *in*, etc.

It seems that the ACE transcription disregards "level tones" on grammatical words and on filled pauses, which, had they appeared on content words would have made that word an instance of level tone. One possible reason is that the meanings ascribed to level tone –*absence of emotional commitment, sarcasm, irony, boredom* – (see below) – are difficult to apply to non-lexical words and non-words such as *er* and *erm*.

For DI, wherever there is a pause, there has to be a tone unit boundary; for CRY5, a pause may occur either at a tone unit boundary or inside tone-units. Of the 41 pauses that ACE has inside tone units, 37 are associated with (from the DI perspective) incomplete tone units, or with o tones. In nearly two-thirds of these cases (21) of pause, there is a o tone in the immediate environment; and in the remaining 16 cases there are incomplete tone units either one or no prominence.

The pauses, incomplete tone units and zero tones are symptomatic of re-starts, or a searching for appropriate words. In DI's terms, there is a focus on encoding the language (oblique orientation) rather than on communicating with a hearer. CRY5 places these pauses inside tone units because he is interested in the relationships between grammar and prosodic phenomena – he takes a linguist's perspective. In contrast, DI, because it attempts to take the speaker's perspective, places these encoding difficulties in separate tone units.

#### Tone unit boundaries

In the discussion above I mentioned that if a description allows for a tone unit to have either one or two tones, it is essential to provide a definition of a tone unit boundary in order to prevent ambiguity of coding. PSIE does provide such a definition but unfortunately, the definition is not sufficiently tight to eliminate ambiguities. For 'normal ... not too hurried' speech (PSIE, p. 205) boundaries occur where there is (a) a preceding nuclear tone (b) a change in pitch direction to the onset syllable of the head of the following tone unit (c) junctural features such as a pause. The problem with this definition is that all three of these criteria can occur (either alone or together) inside the tone unit. Non-nuclear tones, and pause inside tone units, and steps up in pitch, can all

occur inside a tone unit. Thus, as noted above, a transcriber can treat the occurrence of a tone followed by a pause as either:

- a nuclear tone + tone unit boundary
- non-nuclear tone + a pause inside the tone unit

The PSIE discussions on the difference between nuclear and non-nuclear tones is essentially an argument for preserving the relationship between intonation and elements of grammar, which introduces briefly the idea that a tone unit may include subordinate and superordinate tones (p. 270) ‘which are not equally prominent’ (p. 221). This is potentially confusing because PSIE provides no statement of what happens when they are ‘equally prominent’. Additionally, arguments about ‘more and less prominent’ are arguments about gradience, and hence it is possible for judgments to go either way depending on which phonetic feature the transcriber chooses to regard as most significant.

Where the transcriber hears “tones” which do not complete with expectations for the normal relationship between intonation and grammar, it is possible to ignore them. They can be regarded (but in the case of ACE not coded) as cases of non-nuclear tones. In ACE they occur as instances of ‘stress’ followed by the declination line; but it is impossible to distinguish them from (a) other stretches which do not have non-nuclear tones (not all ‘stresses’ are nuclear tones) or from (b) stress followed by ‘continuance’.

#### 4 Testing the attitudinal hypothesis

PSIE (chapter 7) describes an experiment which sought (a) to identify the links between attitude and prosodic systems and (b) to find out how the different prosodic systems combined in conveying such attitudes. There were three components to the experiment:

- six undergraduate students who were native speakers of English
- twenty attitudinal labels

excited	dismayed	haughty	puzzled
amused	matter of fact	precise	disapproving
pleased	bored	questioning	vexed
apologetic	conspiratorial	worried	impatient
satisfied	grim	irritated	angry

- three sample sentences

- (1) Michael Johnson was in the office again this morning
- (2) This is the third time he's been in to see me in a week
- (3) There's that little black dog in the garden again.

Each of the subjects was given a piece of paper with one of the sentence written on it. There were then supplied with the labels – one by one, in an order which kept labels similar in meaning apart – and asked to record that sentence twenty different ways in the following stages:

Stage 1 – They were given sentence one

Stage 2 – They were given the first of the twenty labels

Stage 3 – They were asked to “imagine themselves in an appropriate situation for such a tone of voice” (p. 297)

Stage 4 – They recorded the sentence

Stage 5 – They then made a judgment on a 5 point scale of whether it was easy or difficult to read in that tone of voice.

The participants then repeated stages 2 to 5 for the other 19 labels, then followed the same procedure for the other two sentences.

The results were then analyzed from two perspectives: the extent to which an attitudinal label attracted particular prosodic contrasts; the extent to which specific prosodic contrasts co-occurred with an attitudinal label. The following are among the conclusions drawn from the experiment:

1. Certain ‘non-segmental contrasts ... enter into the definition of *many* labels and have an extremely important attitudinal role to play’ (PSIE, p. 306, my emphasis).
2. A ‘given non-segmental feature rarely has a single semantic field’ (PSIE, p. 306)
3. PSIE states that ‘specification by pitch-movement alone ... clearly will not suffice’ (PSIE, p. 307)

These conclusions amount to an admission that there are no causal links between attitudes and prosodic systems. Despite the assertion in conclusion 1 about the fact that many prosodic contrasts have an ‘extremely important attitudinal role’ it is clear that the experiment failed to pin down any relationship between any single prosodic contrast and an attitudinal label. Even more destructive for the attitudinal hypothesis in conclusion 3: pitch movement, a system which is central to intonation is not on its own a correlate of attitude.

Two follow-up experiments reported in PSIE provided more evidence for rejection of the attitude hypothesis. In the first, 20 utterance which were ‘felt to be good example of performances’ (p. 307) of each of the labels were presented to the same subject in the

first experiment who were then asked to match the labels with the utterances. CRY5 describes the results thus:

No student obtained better than a 60 per cent correct identification, and *no two students made the same set of identifications and misidentifications* (PSIE pp. 307-308, emphasis in the original)

In the second of the two follow-up experiments, a few weeks later, the subjects were given the same utterances and asked to assign labels to them without being given the list of labels used in the original experiment. PSIE reports that 'correct identification dropped to 20 per cent...' and to make matters worse '...nearly 100 labels were introduced ... ' which the subjects tended to use '... in pairs or threes' (PSIE, p.308).

PSIE thus demonstrates the futility of a reliance on the stability of attitudinal labels. PSIE ends with a refusal to be pessimistic about these results, and calls for further research.

## 5. Contours and Context

CRY5 would seem to be in an awkward position: he has made clear his commitment to the primacy of the attitudinal function (cf. 1 above) but the results indicate that the attitudinal view is untenable. He speculates:

This suggests that it may be more useful to talk of contours having 'primary' functions in specific situations and situations having 'primary contours' (PSIE, p. 306)

It is difficult to see precisely what this means, but what it seems to mean is that context seems to be the dominating factor in the conveying of intonational meaning. What he seems to mean by 'situations having primary contours' is that speakers/hearers have expectations about what intonational meanings are appropriate in a specific context. Indeed context has played an important role in the experiment: the six subjects 'were asked to imagine themselves in an appropriate situation .. for a tone of voice, and to use the sentence accordingly.'

Despite the evidence in favour of rejection of that attitudinal hypothesis, ETV ( XX years later) presents statements of the links between tone (pitch movement and pitch range) and attitudinal labels, which is shown in Table 7.

Attitudinal meanings		
Nucleus Type	Environment	Meanings
Level	final tonic of S	<i>absence of emotional commitment</i> which may be interpreted as sarcasm, irony, boredom
	non-final tonic	implication of routineness
Rising + Low drop	final tonic of S	<i>personal inconclusiveness</i> non-committal, unaggressive, polite, respectful <i>social openness</i> casual, friendly persuasive, warning, grim ( <u>with appropriate kinesics</u> )
	non-final tonic	attitudinally neutral
Falling + low drop	final tonic in S	attitudinally neutral
	non-final tonic in S	<i>personal definitiveness</i> abrupt, insistent <i>unsociability</i> cool, irritated, rude
Falling + high booster	in any position	<i>definite emotional commitment</i> emphasis, surprise, warmth, <u>depending on kinesics</u>
Rising + high booster	in any position	<i>definite emotional enquiry</i> query, puzzlement, surprise
Falling-rising	in any position	<i>uncertain outcome</i> doubt, hesitation, suspicion, threatening
Rising-falling	in any position	<i>definite outcome</i> impressed, satisfied, smug, or the reverse, <u>depending on kinesic accompaniment</u>

Table 7 Tones and their attitudinal meanings (derived from ETV, p. 38)

Table 7 shows us that ETV contains three adaptations which go some way towards accommodating the findings of the experiments in PSIE.

First is the admission that that context plays a part in expressing the different meanings: the phrase ‘with appropriate kinesics’ or an equivalent appears three times in the table. Second each attitudinal gloss begins with a ‘general meaning’ (in italics in the table). This is a description of what CRYS earlier alludes to (PSIE, p. 285) as ‘an area of semantic relevance’.

The third change is the abandonment of the strict parametric approach. Prosodic contrasts from the pitch range and tone parameters are taken together as conveying certain meanings. For example the combination of falling nucleus plus high booster is claimed to convey the meaning of *definite emotional commitment*.

Despite the attempt to accommodate the findings of the experiment in this table, problems remain which seriously undermine the attempt to make causal links between intonation and attitude. These problems have to do with labels and context.

## 5.2 Labels

As PSIE (pp. 295-7) points out, there are many problems associated with the use of labels: they are difficult to interpret precisely, it is difficult to verify their meaning; and people mean different things by these labels on different occasions.. The use of such labels is thus non-systematic. The set of general labels from Table 7 have been abstracted and re-displayed in Table 8 below, and they are accompanied by examples of their specific interpretations.

	General label	Specific Meanings
1a	absence of emotional commitment	sarcasm, irony, boredom
1b	definite emotional commitment	emphasis, surprise, warmth
1c	definite emotional enquiry	query puzzlement surprise
2a	personal inconclusiveness	non-committal unaggressive, polite respectful
2b	personal definitiveness	abrupt, insistent
3a	social openness	casual open warning grim
3b	unsociability	cool, irritated, rude
4a	uncertain outcome	doubt, hesitation, suspicion, threatening
4b	definite outcome	impressed satisfied, smug

**Table 8 General labels and specific meanings from ETV**

The first thing to say (re 1a) is that it is highly questionable that *sarcasm* and *irony* appear under the heading *absence of emotional commitment* they could with equal, if not more, justification be classed as examples of *presence of emotional commitment*. The labels *cool* (3b) *casual* (3a) and *non-committal* (2a) would be better examples of *absence of emotional commitment* but they are placed in other categories which are more 'emotionally committed'.



The second thing to say is that the label *definite emotional enquiry/query* (1c) seems to be an example of confusing the attitudinal and grammatical functions of intonation which CRY5 elsewhere has exhorted us not to do. *Query* is not in itself an attitude or an emotion that it may be accompanied by one or several emotions. The same can also be argued with the terms *emphasis* and *warning*.

The situation with the relationship between attitudinal labels is thus so fluid – it is impossible to get stable points of reference – that such lists can only be said to be impressionistic. They are statements of possible groupings of meanings whose psychological and linguistic status Crystal had himself demonstrated to be insubstantial in the two follow-up experiments to his major experiment described in PSIE.

CRY5 is a committed attitudinalist, and believes that further research will eventually identify links between attitude and intonation. But his results do lead to a very clear conclusion which is that there are no systematic links between the prosodic systems of English and the conveying of attitudinal meanings. CRY5 acknowledges the results which point towards this conclusion without every admitting the conclusion itself.

### 5.3 Context

We have seen above that CRY5 speculates that context may have a part to play in the understanding of attitudinal meaning. Context makes its appearance in the tabulated information in three ways: first, in the statements of the grammatical environment; second in the concept of kinesis, ; third in the reference to ‘interpretation’. My discussion will focus on ‘kinesics’.

The use of the term ‘kinesics’ implies a situation in which both hearer and speaker can see each other, and thus have available signaling systems other than utterance for communication. It is CRY5’s belief that ‘kinesic accompaniment’ can reverse the meanings that would otherwise have been conveyed by the accompanying intonation: for him rising-falling nuclear types have the general meaning *definite outcome* and mean *impressed satisfied*, or *smug* ‘or the reverse, depending on kinesic accompaniment’ (ETV, p. 38).

Basically, what CRY5 is saying is that contextual factors such as kinesics are essential to the interpretation of attitudinal meanings: they enable the hearer to decide which of the meanings *casual*, *friendly*, *persuasive*, *warning*, *grim* are intended by rising nucleus with a low drop; it lets the hearer know which meaning (*smug* or its opposite?) the rise fall nucleus is conveying.

This amounts to saying ‘whenever you hear me use a rise-fall, it can mean any one of *impressed*, *satisfied*, or *smug* (but you might think of other labels to describe what is being conveyed) or *unimpressed*, *unsatisfied* or *unsmug* (or other labels) depending on how I’m looking at you, and what my hands are doing’. This is tantamount to saying that “a rise-fall can mean anything”. Once you have got to this stage, one has to

acknowledge that there is no relationship at all between attitude and nucleus types other than the fact they can co-occur.

Context, then, explodes all the hypotheses about intonation and attitude contained in the Table 7 above. This is because, once one admits that a nuclear type can convey an attitude or its opposite (depending on context), one has to admit the possibility that a *rise-fall* might co-occur with meanings which are partial opposites or partial synonyms (i.e. any other label).

Appeals to the reader's expectations of a context are very useful (but dangerous) PSIE uses the notion of context in the experiment in which subjects had to imagine themselves in a context in which the attitudinal labels might be used. ETV uses an appeal to context to help readers of place, and therefore 'hear' examples in the way he want you to hear them.

[what are you doing it FOR] (context e.g. IMPATIENCE) (ETV, p. 24)

In this example, the parenthesis helps the reader imagine a stereotypical situation in which the utterance might be heard to be spoken 'lovingly'.<sup>10</sup>

However, CRY5 seeks to downplay the importance of context. He argues that there are five possible meanings of context:

1. The co-occurring syntactic and lexical pattern of utterance (e.g. sentence final and non-final)
2. Preceding and subsequent syntactic and lexical patterns (e.g. a sequence of parallel sentence structures)
3. Preceding and subsequent intonational patterns
4. Relevant co-occurring and preceding semiotic behaviour, especially facial expression
5. Co-occurring and preceding observable alterations in situation

This list moves from lexis and syntax (1 & 2) through intonation (3) to what CRY5 elsewhere refers to as kinesics (4), and ends (5) with the danger item – situation. CRY5 argues that sense 5 is largely irrelevant to the semantic analysis of intonation, because by the time one has run through the first four elements there is no longer any need to refer to the fifth. He is thus attributing considerable theoretical significance to the order of items. Implicit in this order is a view of speech processing where the speakers are deemed to work first on the lexical and syntactic 'context' (1 & 2) then on intonational patterns, and so on down the list.

---

<sup>10</sup> This example is used in a discussion of whether there are impossible intonational renderings of particular sentences: the conclusion is that 'it is always possible to think up some context which makes the sentence appear quite normal'. (ETV, p. 29).

If one takes the perspectives of speakers and hearers however, it is perfectly possible to argue that the reverse order – i.e. 1 situation, 2 semiotic behaviour, etc. – is the one that has the greater psychological priority in speech processing. In fact much research into schema theory etc. would confirm this latter view.

In order to demonstrate how the use of items 1 to 4 on the list render item 5 unnecessary, CRYS takes the following exchange as an example:

A | you've got something in your ^ HAIR |

(said in a jocular tone, thinking that it is no more than a fallen leaf)

B | \ HAVE I |

(ETV, p. 31-32)

CRYS glosses this exchange as follows:

The low falling tone here, where one might have expected a livelier reply (such as a high rise) to suit SA's jocular mood, provides a contrast that indicates B's displeasure – let us call it “offended”

CRYS states that ‘sense 1 is relevant , for with a less *abrupt* syntax and lexis there would be less likelihood of an offensive interpretation’ (ETV, p. 32, my emphasis). This is trickery – CRYS uses the word *abrupt* to describe the syntax: there is nothing inherently abrupt in stating ‘Have I’ in any context. Whether or not an utterance is abrupt is a matter for the participants’ perceptions, and it relates to their expectations of what is appropriate for the context in which they are speaking. CRYS’s choice of the word ‘abrupt’ means ‘short and rude’, but if one were to substitute the word *succinct* for ‘abrupt’ then his argument becomes weaker.

CRYS then argues that once senses 1-4 are covered, sense 5 is unnecessary. However, prior to beginning his descriptions of the importance of the first four senses of context has provided a characterisation of context in terms of sense 5, the very sense he hopes to dispense with. The words in parentheses in the exchange ‘said in a jocular tone, thinking that it is no more than a fallen leaf’ constitute an explanation of the context in sense five – ‘Co-occurring and preceding observable alterations in situation’.

## 6. Where does this leave Attitudinal meanings?

What then is the relationship between the conveying of attitudes and intonation? There is no doubt that attitudes are conveyed by the way people speak, and that intonational contrasts do co-occur with the attitudes what are being conveyed. The point to grasp is that there is no causal relationship between the occurrence of intonational choice or a set of intonational choices (or prosodic choices) and the conveying of an attitude.

It is easier to demonstrate the fact that there is no causal link between intonational choice and attitude than it is to explain how attitudinal meanings are conveyed. Any assertion that is made about the relationship between tone choice and attitudinal meaning can be refuted by using either of the following tests:

- place the utterance in a new context and use the same tone choices – different meanings will be conveyed
- retain the same context as in the original example, and change the tones, the attitudinal meaning will remain the same

By way of demonstration, consider the following dialogue adapted from Bradford (1988) in which an angry boss (A) asks a quivering secretary (B) where a young employee, John, is.

A: // p WHERE'S ↑ JOHN //

B: // p he's GONE HOME //

A: // r+ GONE HOME // p ↑ SUREly not //

B: // r he HAS //

In this context it is tempting to attribute the meaning 'fear', 'subservience', or 'cowering insistence' to the r tone in the last turn. If we use the first test, and change the context to one in which A and B are hosts at a large party at which John has been misbehaving and his departure unexpected, the r tone on the last turn could be said to be communicating 'delight', or 'unanticipated relief'. The different meanings are clearly attributable to the context, not the tone choice.

Other assertions might be that the p tone with high termination on *surely not* conveys 'surprise', and that the r tone on *has* conveys 'insistence'. If we use the second of the two tests we will find that these meanings would still be conveyed even if the intonational choices were changed:

A: // p GONE HOME // r ↑ SUREly not //

B: // p he HAS //

Whatever the intonational choices, the meanings 'surprise' and 'insistence' are aspects of the context: in a context in which someone is expecting someone to be present, but that person is absent, 'surprise' is naturally present whatever tone is used. Equally, when B reasserts the fact of John's absence by saying 'he has' the very fact that B reiterates can be interpreted as 'insistence' whatever tone is used.

It is perfectly possible for attitude to be conveyed by non-prosodic means. 'Disappointment' or 'shame' or 'fear' can be conveyed by looks: 'His face said it all'. However it is possible to convey an attitude by means of the voice alone, as over the

telephone or on audio tape. A speaker conveys an attitude in a context, to a particular hearer (or set of hearers) who is aware of, or who is alerted to, the feelings that the speaker is likely to have at the particular moment of speaking. The presence of the attitudinal meaning is revealed/highlighted by the speaker's use of non-normal prosodic choices.

What are these 'non-normal prosodic choices'? There is a set of parameters which includes such things as volume, register, tenseness, degree of phonetic realisation of intonational choices which vary continually whenever a person speaks. Although there are 'norms' for each speaker, there is no single norm for a single speaker, but multiple norms which vary with context. A teacher's 'norms' of speech for addressing a class will be different from the norms for speaking with friends in a home, which will again differ from the norms for talking while tramping upstairs, or while jogging.

The reasons for the differences between the norms can be thus physiological (breathiness due to jogging) or sociological (pupils as opposed to friends).<sup>11</sup> Hearers are subliminally aware of these norms and are quick to recognise deviations from them. While jogging, speaker may breath heavily and this may well result in changes in length of tone units, breathiness, and volume: but this will be perceived as normal for them, for this activity, thus a hearer is unlikely to attribute attitudinal meanings to the breathiness, short tone units, and changes in volume. I however, a healthy person in the middle of a relaxed conversation began to exhibit increased breathiness, or greater variations in volume, hearers would seek to attribute a feeling to her which is relevant to what she is saying.

Thus by the term 'non-normal' I am referring to a hearer perception of what is non-normal for that particular speaker at the moment of utterance. Normality in this sense is entirely relative to a particular speaker in a particular context. Equally, judgements of normality are made by hearers in the same context and are related to their previous experience of the hearer's voice.

The particular way in which the norm is deviated from may predispose hearers to perceive the meaning a certain way. Additional creak might predispose hearers to describe the local meaning a 'gruff' or another hearer 'grumpy' additional 'breathiness' might in certain context be taken to indicate 'excitement' or 'fear'. There is an element in this of the hearers' putting themselves in the place of the speaker, and asking the question 'what would I mean in this context if I were to deviate from my norm. in a similar way?'

---

<sup>11</sup> It is often possible for me to identify which or our friends, relatives, or acquaintances my wife is talking to over the phone by the norm she adopts for her second utterance (the first is normally 'Hello') 'Oh how are you?'

When a deviation from the relevant norm occurs, the hearer (again the process is normally subliminal) seeks to attribute a meaning, relevant to the context, for the deviation from the norm.

## 7. Conclusion

As the experiment described in PSIE indicates, there is no generalisable cause between the prosodic variation by which the deviation from the norm is achieved and the meaning the hearer attributes to the deviation. Although the same meanings may be perceived to be present in similar contexts (where the speaker and hearers remain the same), they will certainly not be generalisable to all speakers in all contexts. The relationship between attitudinal meanings and prosodic choices is one of co-occurrence not causation – in principle, any prosodic choice may co-occur with any meaning.

## References

- Bradford, B. (1988). *Intonation in context*. Cambridge: Cambridge University Press.
- Brazil, D. (1975) *Discourse intonation*. Discourse analysis monographs no. 1. English Language Research: University of Birmingham.
- Brazil, D. (1978) *Discourse intonation II*. Discourse analysis monographs no. 2. English Language Research: University of Birmingham.
- Brazil, D. (1997) *The Communicative value of intonation in English*. Cambridge: Cambridge University Press.
- Byers, P. B. (1979). A formula for poetic intonation. *Poetics*, 8, 367-380.
- Byers, P. B. (1980). Intonation, prediction and the sound of poetry. *Language & Style*, 13/1, 3-14.
- Cauldwell, R. (1993). Evaluating descriptions of intonation: A comparison of Discourse Intonation and Systemic intonation. Unpublished typescript.
- Crystal, D. (1969a). Intonation and grammar in British English. [Review of M. A. K. Halliday (1967)]. *Language*, 45/2, 378-393.
- Crystal, D. (1969b). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Crystal, D. (1975). *The English tone of voice*. London: Edward Arnold.
- Crystal, D., & Davy, D. (1975). *Advanced conversational English*. Harlow: Longman.
- Knowles, G. (1987). *Patterns of spoken English*. London: Longman.
- O'Connor, J. D., & Arnold, G. F. (1973). *Intonation of colloquial English*. [2nd Edition]. London: Longman.
- Crystal, D. & Quirk, R. (1964). *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton.

Quirk, R. Duckworth, A.P., Svartvik, J. Rusiecki, P.P. L. & Colin, A.J.T. (1964) Studies in the correspondence of prosodic to grammatical features in English. *Proc. Ling IX*, pp. 679-91

Wichmann, A. (1987). *Stylistic variation in prosody: a preliminary study of the Spoken English Corpus*. Unpublished master's dissertation, University of Lancaster.