DAFYDD GIBBON

Intonation as an Adaptive Process

1. Aims and scope of this study

This paper describes a procedural model for English intonation, and applies it to complex dialogue data. In the first section, descriptive (§ 1.1.) and heuristic (§ 1.2.) assumptions are discussed. A set of descriptive categories for English intonation is presented in § 2, followed by an analysis of complex intonation data in § 3 and an outline of a procedural syntax for English intonation in § 4. In § 5, some extensions of the notion of 'process' to include 'adaptation to context' are proposed, and in § 6 the main properties of the theory are briefly evaluated in terms of a dynamic speaker-hearer model.

The main points developed in the paper concern

(1) the metalocutionary hypothesis (§ 1.2.);
(2) the status of discourse tokens (§ 1.2., § 3);
(3) the articulatory bases of intonation and their relation to perceptual patterns (§ 2, § 5);
(4) the pulse accent theory (§ 2);
(5) iterative and recursive processes in discourse phonology, organised in prosodic frames (§ 4);
(6) the adaptation of intonation processes to context, described as a feedback system (§ 5);
(7) consequences of this procedural approach with regard to the methodological problem of 'fuzziness' in intonation description (§ 2.3., § 6).

A major aim of this approach is to provide conceptual bridges between phonetic, structural and functional aspects of intonation, and for this reason concepts are adapted from a fairly wide selection of disparate fields, from the physiology of speech to discourse analysis. Emphasis is laid on analysing 'data in use' rather than intuitively constructed 'data'.

1.1. Intonation form and function

The view is taken here that the forms of intonation can initially be described at a linguistic level *sui generis,* and that there is an 'interpretation function' of a quasi-semantic kind mapping intonation patterns on to locutionary structures at various levels of description. This is a classic view of the status of intonation, and is in one form or another characteristic of

most major authors from Pike (1945) through Jassem (1952) and Crystal (1969). I have called this view, in an explicit form, the 'metalocutionary hypothesis', since intonations mark, *inter alia*, structural properties of locutions and their functions in dialogue, acting as a 'suprasegmental' or, semantically speaking, 'metalocutionary' system (Gibbon, 1980).

In addition to the classical component of this view, the following hypothesis is maintained: the metalocutionary interpretation function presupposes a level-selection or domain-switching function which selects the relevant level of locutionary structuring for intonational marking. This level-selection function defines stylistic or functional variation in intonation, and is sensitive to the discourse context. Locutionary structure is more complex than the intonational means available for indexing it, and includes at least the following kinds of structure as values of the level-selection function and thus as domains for prosodic indexing:

(1) Word and syllable structure.
(2) Phrase structure: relations between syncategories and categories, specifiers and heads of phrases (e.g. in an $\bar{X}$ syntax), early vs. late linear position (each paired with weaker and stronger accents, respectively); constituent boundary indexing.
(3) Semantic structure: operator/operand scope, as with negation, quantification, degree and focus adverbs (stronger accent on the operand).
(4) Topical (semantic frame) structure: anaphora/contrast, other given/new relations in topic development (also weak/strong).
(5) Speaker attitudes of modal (knowledge, belief, obligation) and appraisive types (emphasis, pejoration, amelioration).
(6) Discourse organisation: turn-taking processes, speech act sequencing, indexing of completion/noncompletion of dialogue constituents.
(7) Discourse type (style, genre, register) indexing: both via the selection function itself and via specific genre markers (e.g. the 'chroma' feature of § 2.1., § 3).

For instance, in read-aloud citations, level (2) will tend to dominate; in spoken narrative, (4) is more relevant, while in small-talk, (6) may be dominant. A case of domain selection between levels (2) and (5) or (6) is discussed by Bing (this volume), Boves & al. (this volume).

This descriptive background is stated here *in nuce* (cf. also Gibbon, 1982) since, although the body of the paper is more concerned with the formal organisation of intonation processes, it presupposes this background.

1.2. Intonation and data complexity

The research strategy underlying the work described in this paper involves, primarily, finding methods of systematically describing complex intonation data, in particular discourse tokens paired with context specifications. The strategy proceeds from an informal 'interpretation' of forms

and functions of such data in the direction of a more precise formal description; an important feature of the strategy is that it is necessary to proceed along a broad front, taking form and function description along roughly at the same pace.

It has proved heuristically useful to distinguish between two data types and to try and combine use of these types in the descriptive cycle. These two types centre around concepts such as *paradigmatic, static, structural, classificatory* on the one hand, and *syntagmatic, dynamic, procedural, raw* on the other. Some of these words may seem odd partners, and there is systematic overlap between the groups, but the data types, which I shall refer to here as 'P-data' and 'S-data', are real enough in practice. P-data are edited productions of sentences or sentence constituents, preplanned, read as citation forms (i.e. as unimaginatively as possible), and paired with paradigmatic metalinguistic judgments of constituency type ('sentencehood', 'intonation group') or more complex equivalences ('ambiguity', 'substitutability', 'paraphrase'); these subsentential paradigmatic judgments are often supplemented by sentential functional, syntagmatic judgments on anaphora, subjecthood, 'focus-of'. S-data are discourse tokens with context specifications ('context' in the sense of 'cotext' as well as the more general sense), from a variety of registers along scales such as 'most spontaneous' to 'most rehearsed'; the context specifications are syntagmatic metaliguistic judgments on function in discourse, and paradigmatic judgments on constituent types are regarded as derived from these (cf. also Gibbon, 1976a: 44f.).

In developing precise descriptions of selected points, P-data are indispensable but they embody certain heuristic assumptions:

(1) The style/register of pre-planned utterances read and judged as citation forms is representative of all speech and allows access to the underlying 'language'.
(2) The 'clear case' technique based on straightforward P-data judgments allows quasi-inductive 'predictions' to be made about less clear cases on the basis of an exact theory, providing a suitable strategy for discovering the underlying 'language'.

Both of these assumptions are, however, questionable. To avoid the restrictions which they bring, a therory, whether developed using P-data or S-data or both, must be tested continually on S-data in order to falsify existing descriptions by extending the data base. One valuable contrary technique to (2) is, in fact, the description of easily isolated, conspicuous, but by no means 'clear' cases; continued interest in 'call contours' (cf. Gibbon, 1976: Ch. 4; Ladd, 1978), at first sight a peripheral issue, is a case in point. This technique is, no doubt for psychological rather than logical reasons, widespread but unsung in all branches of linguistics. A reason for using S-data, counter to (1), is to account for 'functional' properties of speech: 'functionalistic' need not mean 'seeking teleological explanations', it can

also mean a descriptive approach with the sense of 'embedded in a contextual matrix'. In this sense, 'function' is no less a formal or structural notion than that of grammatical function. In a process-oriented approach, S-data therefore have more than a mere alibi role in linguistic heuristics.

Within the present research strategy, S-data have also made in desirable to incorporate into the description as coherent an account of the process properties of speech as possible; these include the temporal organisation and the functional or stylistic variation of speech. If intonation is conceived as the temporal organisation of pitch in speech, not an over-controversial conception, then P-data at once become highly suspect where intonation description is concerned. Likewise, some of the very functions of intonation are to mark styles or speech registers, again restricting the utility of P-data.

In view of S-data judgments on temporal organisation and functional variation, distinctions such as language/speech, competence/performance, programme/execution seem too simple to be of much relevance. Where a 'real-time job' (e.g. speaking) is involved, it seems reasonable to assume that any programme will take a real-time environment into consideration. P-data, however, may convey an impression of a-temporality, owing to their pre-edited nature. This study is a step in the direction of a process-oriented explanation.

2. *Descriptive categories for English intonation*

The descriptive categories required for the analysis of English intonation are, fundamentally, dynamic relations between articulations and percepts; a fully described category would be a pair ⟨articulation, percept⟩. The term 'percept' refers to the user's perception both of his own productions and of incoming singals; in the first case they include proprioceptive as well as auditory percepts (cf. also § 5 and § 6). The articulations are thoracic and laryngeal gestures, and the percepts are essentially pitch height and trajectory gestalt impressions; temporal patterning of intensity plays a related role. Presumably, the relations containing these pairs are learned during the earliest stages of speech acquisition. It should be noted that the relation between the phonological features outlined here (e.g. Cricothyroid pulse) and the physiological reality (e.g. the plurality of factors affecting cricothyroid muscle operation) is not a simple one; the categories are no less abstract than the segmentally relevant articulatory categories (e.g. Coronal, Voicing, & c.), and are mnemonically convenient abbreviations for complex activites (cf. Hardcastle, 1975).

The relationship between articulations and percepts is syntagmatic, part of a process sequence within a feedback cycle, not a static 'correlation'. As

an expository aid, however, indications of 'correlations' will be provided, using the symbol '≙'.

Articulatory categories will be in the foreground of attention, since they allow a simpler overall description. Three kinds of descriptive category will be distinguished: primitives (§ 2.1.), complex processes (§ 2.2.), and strategies (§ 2.3.).

2.1. Primitives

The primitives are local (short-scope; roughly speaking: word-oriented) and global (long-scope; roughly speaking: phrase/sentence/text-oriented).

The local primitives are accent pulses and their modifications (some modifications being reflexes of global features determined by constructions or strategies).

(1) Accent pulses. Accents are conceived as laryngeal pulses. The larynx is regarded as a complex elastic body whose steady state with respect to vocal cord tension is altered by 'stretching' ('Cricothyroid') and 'compressing' ('Sternohyoid') pulses. After a pulse, the larynx returns (other things being equal) to its pre-pulse state. The auditory reflexes of these modifications are a pitch change in one direction on the leading flank of the pulse, followed by a change in the reverse direction after the pulse. The articulatory 'correlates' are conceptually simpler than the auditory patterning, and are presupposed by the latter. They are more abstract, however, from an empirical point of view. Aspects of the pulse theory of accent are shown pictorially in Figure 1.

The two phonological accent types are (note that the names are not intended to imply a simple association with single muscles):

i. Cricothyroid pulse (≙ upward pitch movement), denoted by '↑';
ii. Sternohyoid pulse (≙ downward pitch movement), denoted by '↓'.

The ↑ pulse can be felt with a finger as a narrowing of the intercartilage ridge at the front of the larynx, the ↓ pulse sometimes as a lowering of the entire larynx. Not all short-range pitch movements are functions of laryngeal pulses. Pitch perturbations of supraglottal origin, the use of subglottal mechanisms, and other factors, affect vocal cord stretching and compression. Bolinger's Accents A and C could be partially explicated in these terms as ↑ and ↓, respectively; the ensuing fall or rise is a natural auditory reflex of the post-pulse relaxation predicted for unmodified cases by the laryngeal pulse theory. Other more complex forms of accent are due to pulse modifications. Similarities with Lieberman's theory (1967) are evident, at least for the ↑ pulse.

(2) Accent modifications. The central accent modification is pulse *amplitude* (≙ degree of pitch change, ≙ Δf_0 (i.e. change in fundamental frequency), ≙ prominence), which realises accentual gradation. A second set of accent modifications affects pulse *timing* relative to syllable structure:

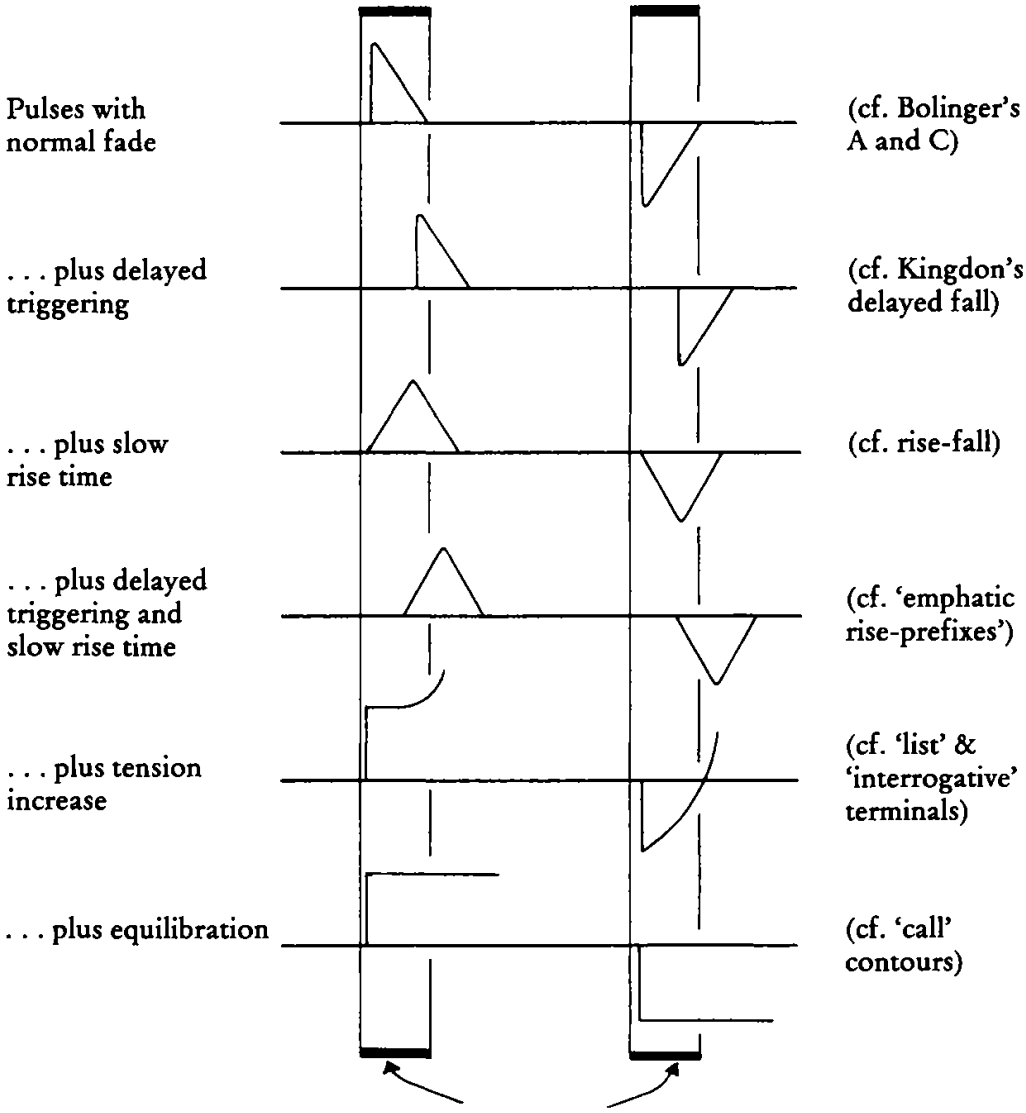

Figure 1: Schematic outline of pulse types and modifications (baseline modifications not shown); comments apply mainly to rise pulses.

i. pulse leading flank:
   a. delayed triggering (syllable-initial, post-syllable-onset, post-syllable-nucleus, post-syllable), as in Kingdon's delayed rise-fall, Bolinger's variant of Accent A (characteristic of some dialect and style variation and some types of emphasis, interacting here with syllable onset timing features which produce fortition, e.g. stronger aspiration);
   b. slow rise time involving syllable lengthening (accounting for some rise-fall tones, Armstrong & Ward's emphatic rise-prefix, Kingdon's 'homosyllabic prehead');

ii. pulse trailing flank:
   a. fade (the neutral case – cf. remarks above);
   b. cricothyroid tension increase (≙ rising pitch, ≠ pulse);
   c. laryngeal equilibration (≙ level tone, cf. 'chroma', Gibbon, 1976a, b; 'stylization', Chao, 1956; Ladd, 1978; Brazil's *o* tone, 1978 & this volume).

It seems plausible to postulate a more abstract category of accentual pulse, with pulse polarity as a modification, e.g. with ↑ as unmarked and ↓ as inverted, i.e. marked pulse. Evidence for this may be seen in Bing's discussion of equivalences between Bolinger's A and C accents (1980), as well as in the peakline-baseline complementation feature discussed below under (2) v.

The global primitives can be divided into boundary features and baseline features: they are properties of complex processes (cf. § 2.2., and the γ-frames and π-frames of § 4.).

(1) Boundary features. The boundary features mark the beginning and end of laryngeal pulse trains:

i. Initial boundary: a complex function of Anacrusis/Prehead and Head/Onset height specification, where the function of Head height (Crystal, 1971; cf. also Trim, 1964: 378; Brazil's 'key'; Lehiste's 'paragraph' boundary marking, 1975; Brown & al.'s 'initial stressed peaks', 1980: 136; cf. also Ozga, 1980) appears to be primary, strategically modifiable by Prehead height (cf. Bing's initial boundary tone, 1980). In articulatory terms, Head height appears to be contextually determined to some extent by subglottal conditions.
ii. Final boundary, with several different feature types:
   a. Lengthening of accented syllable and other timing factors involving delay, such as 'terminal pause';
   b. Pulse mirroring: a sequence of [α pulse] ∧ [−α pulse] pulse (the use of square brackets and '−α' is that of generative phonology; '∧' means concatenation), with the second pulse before the soonest segment which is next lowest in sonority in the subsequent unaccented stretch (cf. Liberman's similar explanation of tone distribution in 'call contours', 1978). This also allows a generalisation over call contour phenomena and some other contour types. If the specification is [α = ↑], then the following inversion, [−α], or ↓, can be explained functionally as accelerating the natural post-pulse decay which already receives perceptual prominence by virtue of the syllable lengthening function (cf. also § 2.3., the Principle of Sequential Pitch Contrast).
   c. Cricothyroid tensing (≙ pitch rise).
   d. Larynx equilibration (≙ pitch sustention).

The occurrence of pulse mirroring, tensing and equilibration in both local and global functions may be seen as an indication of a functional am-

bivalence of 'level' assigament which is characteristic of prosodic patterning (cf. § 2.3.).

(2) Baseline features. Throughout the history of English intonation description, indications are to be found that two kinds of 'baseline' are involved: a 'baseline proper', a virtual pitch value toward which unaccented syllables tend (depending on their relative proximity to a pulse), and a line along which auditory pitch peaks tend to be aligned (cf. Bolinger's 'tangent', or Body/Head contours). I shall call the former *baseline* and the latter *peakline.* The most similar recent phonological uses of this distinction are by Brown & al. (1980) and Bing (1980). Given a baseline and a peakline, a number of global relational properties of prosodic patterning can be defined over ⟨peakline, baseline⟩ pairs:

i. Bandwidth ('span', 'range') is a simple peakline-baseline relation defined as pitch difference ($\triangleq \Delta f_0$).

ii. Peakline slope, [α Δpeakline] (i.e. type of change in peakline: up, down, level), with downdrift as the neutral case and a rise (cf. Palmer's 'scandent head') or more complex tangents as a marked case. Whether peakline slope induces baseline slope (or vice versa) is an open question. Three main kinds of 'level' peakline slope may be distinguished:
   a. low, as in relaxed conversational reporting style, often taken as the standard American English pattern but also characteristic of analogous styles in other dialects of West Germanic languages;
   b. sustained mid, as in one speech genre of overt narration (as in E(*q*) of § 3); this feature is similar to the laryngeal equilibration noted twice above (cf. § 2.3.; and iv, below);
   c. high pitch, as in 'uncontrolled', excited speech.

iii. Peakline-baseline convergence, [peakline $\rightarrow_p$ baseline], again of two types:
   a. an unmarked type, the 'trailing off' which is presumably allied to neutral downslope and may not need to be postulated separately from this;
   b. sustained convergence, as in mid-discourse 'narrative tags' (cf. § 3, § 4).

iv. Peakline-baseline 'identity', [peakline $=_p$ baseline]: this is strict 'chroma' or 'stylization', which occurs in its most extreme form in chants and song (Chao, 1956), and is entirely dependent on larynx equilibration supported by negatively phased auditory feedback (cf. § 5, § 6).

v. Peakline-baseline complementation, in which relative peakline-baseline heights are inverted; this is an interesting theoretical possibility with considerable descriptive potential. Figure 2 illustrates a recorded example, in which the upper frame is a train of ↓ pulses (with 'upward' relaxation indicated by rising lines), and the lower consists of ↑

pulses (with 'downward' relaxation indicated by falling lines). This feature provides a device for describing certain features of style and dialect variation in a unified fashion, and distinguishes some global uses of ↑/↓ alternation from the local uses already noted, with cases such as the following (where $\alpha\langle_p\beta$ means 'α is lower in pitch than β' and $\alpha\rangle_p\beta$ means 'α is higher in pitch than β'):

a. Edinburgh dialect: [peakline $\rangle_p$ baseline], Glasgow dialect: [peakline $\langle_p$ baseline] (Brown & al., 1980: 19);
b. North German dialects: [peakline $\rangle_p$ baseline], South German dialects: [peakline $\langle_p$ baseline] (an oversimplification, but cf. v. Essen's "Zickzackmelodie", 1956);
c. In standard English dialects, child-adult and adult-child speech, with [peakline $\langle_p$ baseline], often functionally alternating with [peakline $\rangle_p$ baseline] (cf. Figure 1).
d. As a stylistic discourse strategy in German (cf. Gibbon & Selting, 1983).

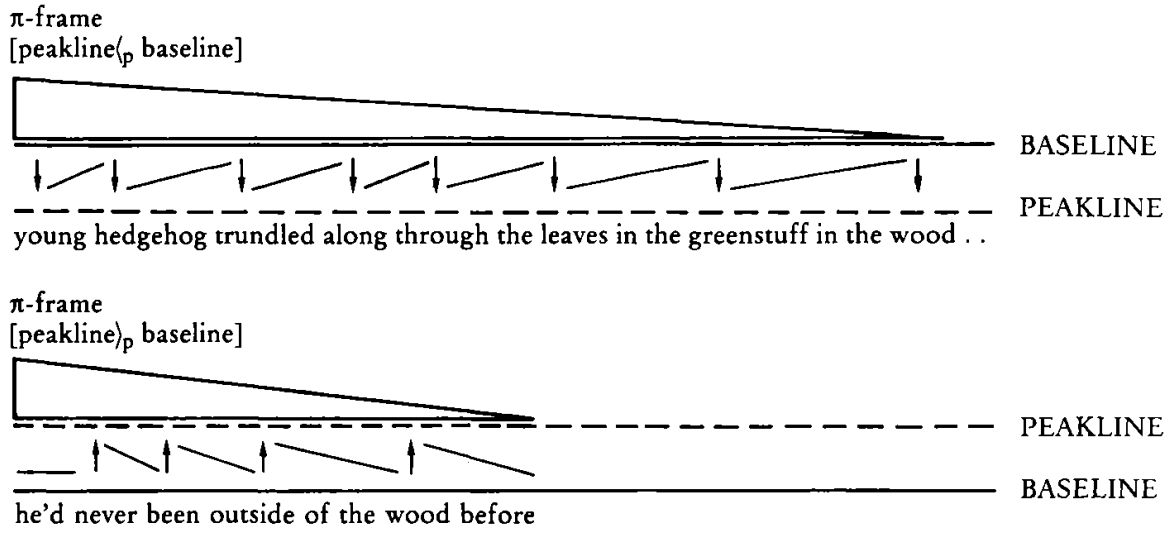

Figure 2: Use of different peakline-baseline specifications to mark stages in the development of an adult-to-child narrative (from a BBC Children's Hour broadcast).

2.2. Complex processes

The problem of representing intonation patterns in formal terms is still unsolved, though a wide range of suggestions has appeared in recent years. In the present theory, the line taken is different again, and uses outline augmented transition networks to define construction types as processes. In order to clarify the distinction between two fundamental posi-

tions with regard to notational systems, it is expedient to distinguish between *categorial* and *procedural* explications of intonation structures.

For some aspects of intonation structure, a representation of dependency relations (e.g. head/satellite) is required (cf. § 1.1., especially accentuation). For this purpose a temporally uninterpreted phrase structure or categorial syntax would be appropriate. For instance, if α is an Intonation Group, and β is a Nucleus, then α/β is the category of Head (= Pretonic); if γ is a Head (= Onset), then $\gamma\backslash(\alpha/\beta)$ is Kingdon's category of Body, i.e. the sequence of accented and unaccented syllables between Head (= Onset) and Nucleus. Semantic interpretations (e.g. in terms of 'topic' and 'comment' or more structured explications of such notions) can then be built constructivistically on these categories; cf. remarks above in § 1.1. on semantic interpretation. The integer notation of Chomsky & Halle (1968) and particularly the $\langle s, w \rangle$ notation of autosegmental and metrical phonological theories can possibly be re-analysed as approximations to categorial syntaxes for intonation.

However, this kind of syntax seems likely to be too rigid for much of the prosodic patterning found in English. There are other formal properties of intonation constructions for which a procedurally interpreted autonomous rewrite grammar or a transition network are more suggestive. The latter approach will be taken here. A distinction will be made between fixed, iterative and recursive processes. The thesis is put forward here that all three process types are required to describe the prosodic patterning of discourse. They are not necessarily all required for the description of P-data.

Fixed processes are items consisting of fixed sequences of global and local specifications. They are found in the stereotyped patterns associated with pragmatic idioms (greeting, interjections, exclamations, & c.), and in

Levels of analysis	Structures postulated at each level
1. Major accentual pattern, 'bar':	(head – (break(s))) – nucleus
2. Accentual sequencing:	$a_{|}$–. . .–$a_{i|}$–. . .–a_{i_n}–. . .–a_m
3. Syllabic sequencing:	$s_{|}$–. . .–s_g–. . .–$s_{j|}$–. . .–s_{j_o}–. . .–s_k–. . .–s_p

Figure 3: Symbolisation of levels of analysis in partial explication of Klinghardt & Klemm (1920), from Gibbon, 1976: 106 (the bracketed plural – '(s)' – is used to denote an optionally iterative category).

some theories they are postulated as the only type, enumerable in an intonation lexicon (Liberman, 1978). They will not be considered here.

Iterative processes have frequently been postulated for the description of accent sequences in more or less explicit fashion for many decades. In a previous critical survey of research I represented one iterative property of 'Tone Groups' in the diagrammatic form of Figure 3; in Fox (this volume) and Halliday (1967), this iterative property is made explicit. Simple transition networks formulated as finite state machines may be used to describe this property of intonation patterning. Since 1978 I have been using transition networks for instruction purposes to synthesise 'stylisations' of pitch accent sequences in English (in a sense approximating to the usage of 't Hart) with a microcomputer (cf. Gibbon, 1981b); networks and tones are interactively definable by the operator. Several approaches have used similar notions (Reich, 1969; 't Hart & Cohen, 1976; Gibbon, 1981a; Pierrehumbert, 1980). It is evident that a formal explication of rhythm, among other things, will have to rely heavily on an interative principle at some point, whether this is understood, in real terms, as a cerebral 'clock' frequency or as a perceptual scanning and pattern-recognition principle (cf. Neisser, 1967).

In a previous study I have shown that in German at least two levels of iterative process can be identified in complex expository discourse (Gibbon, 1981a). At the lowest prominence level, there are sequences such as r, r *r*, rr *r*, . . . with various explainable 'exceptions', where 'r' symbolizes a ↓ accent and '*r*' symbolises a more prominent ↓ accent (i. e. with higher pulse amplitude or following cricothyroid tensing). At the next prominence lev-

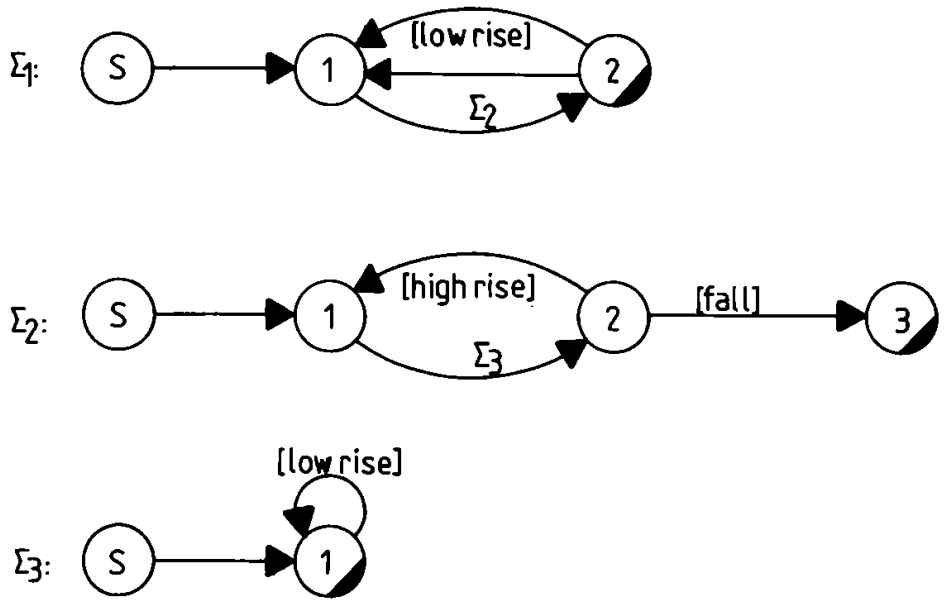

Figure 4: Outlines of augmented transition networks for the German data from Gibbon (1981).

el, sequences such as *f, rf, rrf,* ... (with intervening lower prominence pulses) occur, where *r* has the same meaning and *f* denotes a ↑ pulse with post-pulse syllable lengthening and, optionally, leading flank modifications. This structure is generated by the network represented in Figure 4 (cf. also Figure 5).

Each of these levels is obviously right-regular. However, the similarity between the levels suggests a limited degree of *recursion* as well as *iteration.* This suggestion makes for 'flatter' structures than binary trees with ⟨s, w⟩ pairs, where this distinction is not made; a result which also has intuitive appeal. Complex prosodic patterning in S-data therefore points towards a necessity for combining the two kinds of 'control flow'. Figure 5 shows a simplified network for some aspects of the data discussed above (A), and a generalisation of this network (B) which uses recursion, taking an initial maximum or 'nuclear' prominence assignment (e.g. the integer 1) as input and then choosing between pulse output or calling itself with reduced prominence value. On completion of interation, a terminal pulse with the currently specified value of β added to a minimum pulse value α is generated. A complex analysis and more highly structured model is discussed in § 4; for this model, two types of complex process dominating accent sequences are postulated: γ-frames and π-frames, concerned with the global specification of baseline and boundary features; respectively.

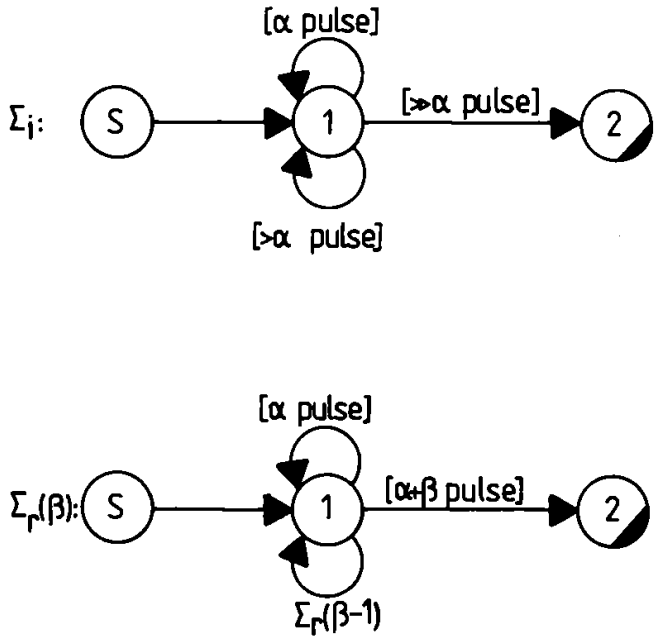

Figure 5: A (top). Iterative network equivalent to Σ_2 and Σ_3 of Figure 3. Higher values of α are interpreted here as higher pulse prominence (a function of pulse amplitude and polarity).
B (bottom). Generalisation of the above by including recursion. The current value of β, a pulse prominence increment value, is passed to the network as a parameter on each call. Lower values of β are interpreted as higher pulse prominence.

2.3. Strategies

The preceding discussion has shown several points at which similar processes appear to be in operation at different levels of structure. One relatively trivial set of examples encompasses 'relaxation' after a ↑ pulse and the overall downslope or downdrift of the peakline; in this set, various articulatory mechanisms contribute towards producing an auditory impression of falling pitch. A perceptually or functionally oriented priciple applying here might be a *principle of cooperation of means.* Another case could be the use of laryngeal and subglottal gestures with timing modifications to intensify a paticular auditory effect such as a pitch rise. On the other hand, the use of a low Prehead to modify a high Head, producing a complex initial boundary tone, or of wide pitch bandwidth (peakline minus baseline), or of alternating ↑ and ↓ pulses, or of ↑↓ pulse pairs to accelerate a pitch fall at terminal boundaries, points to the operation of another perceptual principle: a *principle of sequential pitch contrast.* Finally – contrasting with the last-named principle – in level tones, in sustained-peakline speech, and in stereotypic speech involving [peakline $=_p$ baseline] specifications, a *principle of equilibration* appears to be in operation, with the two laryngeal functions operating antagonistically and under the stabilising influence of negatively phased auditory feedback: a pitch drop trigger the ↑ mechanism, and a pitch rise triggers the ↓ mechanism (and/or the respective subglottal mechanisms). The effect is to counteract accentual overshoot (for ↑) or undershoot (for ↓), as well as to work against relaxation and downdrift effects. In a previous study of 'call contours' (1976a: Chap. 4.3.) I called this functional feature of discourse phonology 'chroma', taking the term from Bachem (1950); see also § 5 below.

It is not clear how the notion of strategy is to be integrated into the overall description; the ideas behind the notion are simple, however. First, strategies are intended to represent the complementary functional status of perceptual as against articulatory processes within a dynamic feedback cycle. Second, they represent an attempt to generalise over processes which are phonologically different in structure but related in function. Third, if considered *without* a hypothesis on underlying articulatory processes, they are likely to lead to indeterminacy and uncontrolled variety in descriptions, since arbitrary decisions have then to be made in chossing between perceptual models. That there is indeed little clarity in this field is perhaps informal support for the necessity of a dynamic production-perception relation such as that postulated here.

3. Intonation in dialogue structure

The metalocutionary hypothesis outlined in § 1.1. is part of theory of intonation semantics pertaining to properties of locutions which are indexed

by intonational processes. The present section is an attempt to combine this approach to intonation semantics with the phonological intonation theory developed in § 2, by modelling selected aspects of the intonation of a discourse token, and sketching their relationship with – in this case – semantic aspects of dialogue progression (mainly levels (4) and (6) of § 1.1.). The semantic mapping itself is not explicitly modelled.

The discourse from which the present examples are taken is on a commercially available recording, which readers should consult (Crystal & Davy, 1975, Extract 4; the accompanying book contains a tonetic transcription). It is an extract from an informal tale-swapping session among acquaintances, in which overall semantic coherence is provided by a single superordinate frame: pigs, experiences with pigs, and pejorative appraisal of these experiences. The extract consists of three anecdotes, each with its own distinct subframe, with bridging dialogue; only the first will be considered in detail. The anecdotes and bridging portions constitute two kinds of discourse:

(1) Narrative within a 'stable' semantic subframe, with a single, long, undisputed narrative turn.
(2) Subframe negotiation between anecdotes with 'unstable' turntaking patterns which include considerable turn-overlap.

This patterning suggests that an appropriate model for such discourse will have to include a self-regulating, context-testing principle to account for the alternation of 'stable' and 'unstable' group interaction states (cf. also § 6). The first of these state types will be examined here, using Anecdote 1. The 'stable' semantic subframe on which it is based will be described, followed by a partial account of its intonational marking in terms of the categories outlined in § 2.

3.1. A 'stable' narrative contribution to dialogue

In a previous study (1981a: 93ff.) I have described how intonation patterns mark the more or less cooperative development of topical structures. Development occurs in several stages or substages within a previously negotiated semantic frame or subframe. Depending on discourse type, frames and subframes may hold for one turn or a whole sequence, and although the turn-taking structure of the present example is quite different, the principle of adherence to a negotiated semantic frame or subframe still holds. To provide a context for the present discussion, the anecdote is reproduced here in a purely orthographic transcription; for more detail, readers should consult the tapes and the published tonetic transcription.

A. Oh, and one pig died, because it ate too much.
B. Oh really?
A. Oh it was revolting. Oh, they were terrible, the pigs.
C. Oh . . .
A. They made a dreadful row in the morning when it was feeding time.

And one pig, it was erm a young pig, about *that* size, you know, middling, anderm it was dead, and it was lying there. I'd never seen a dead pig before. Absolutely stiff.

B. Di- the children saw it, did they?

A. Oh they were engrossed, you know.

C. Oh yes!

A. It was marvellous, erm they thought this was wonderful. And erm they asked why it was dead, and er the farmer apparently didn't want his wife to know, because he'd overfed them before and she'd been furious – and of course he was trying to keep it from her. But all the kids were agog about this dead pig, and he was telling them not to tell the farmer's wife

D. Yeah.

A. and all this . . . So this pig was absolutely dead, so they put it on – they have a sort of smouldering heap that smoulders all the time – so they went to burn the pig, and all the kids *(laughter)* were hanging over the gate watching this pig. And they were very taken that the pig had died because it had eaten too much, you know.

D. What a marvellous death!

B. A moral in that somewhere!

This anecdote elaborates a semantic frame with the structure "(*p* because *q*) & *r*", where:

p stands for "one pig died",

q stands for "it ate too much",

r stands for an appraisal (perhaps to be glossed as 'fascinated pejoration') of "*p* because *q*" or its constituents.

Topical development within this frame is as follows:

Stage 1 (statement of the frame): "one pig died because it ate too much".

Stage 2 (elaboration of the frame as a function of the elaboration of its constituents, in the following order):

$E_1(p)$: "and one pig it was erm a young pig . . ."

$E(r)$: "oh they were engrossed you know . . ."

$E(q)$: "they asked why it was dead . . ."

$E_2(p)$: "so this pig was absolutely dead . . .".

Stage 3 (restatement of the frame):

r: "and they were very taken that"

p: "the pig had died" because

q: "it had eaten too much".

This development is marked in several ways. First, it is very noticeable that turn-taking is highly restricted, limited to a clarification and a few interjections; this contrasts strongly with the following section of the dialogue, in which a new semantic frame is negotiated. Second, topical development in the anecdote is marked by conjunction patterns. Of particular interest are the largely asyndetic sequences associated with *r*, coordination

as a mark of improvisation in E(*q*) – cf. *anderm, because, and, and of course* (≈ *therefore), but, and* – and, most notably, the "restart" signals, after breaking off an entagled construction, in the form of three occurrences of *so.* Third, developments in topical structure are marked by different dinds of prosodic frame; these structures will be informally discussed in the next subsection and analysed in § 4.

3.2. Preliminary observations on pitch patterning

The intonation patterns in Anecdote 1 offer a number of interesting features which are not captured well by a linear, in-text tonetic transcirpiton. The main points to notice for present purposes are the following:

(1) Pause organisation is largely independent of the organisation of long pitch stretches;

(2) There are significant relations between long pitch patterns and elements of discourse structure, such as high pitch Heads coupled with downslope (cf. the beginning and end of the anecdote), pitch lowering relative to neighbouring sequences (e.g. *oh they were terrible the pigs*), or the tendency towards pitch sustention at a mid level (following *didn*'t want his wife to know) both on level and on longer stretches within E(*q*).

(3) A number of 'special effects' such as higher pitch, rise-fall tones, mark lexically appraisive items such as *oh, terrible, engrossed, marvellous.*

One problem associated with in-text notations is that they embody a large number of implicit judgments about the phonetic properties of the environments of accentual sequences as well as of particular syllables, and possibly also judgments about properties of accompanying locutions. The following three excerpts show several different kinds of 'onset', whose differences do not come out in such transcriptions, and whose similarities are not at all evident on either perceptual or distributional grounds (| = 'onset', ↑ = 'booster', not pulse accent in the sense of § 2; cf. Crystal & Davy, 1975):

- |and erm —— it was ↑DÈAD (a pause separates the 'onset' from the rest of the tone group);
- erm they |thought this was ↑WÓNDERFUL ('thought' is part of a very rapid mid-to-low sequence —— a long prehead with *wonderful* as onset and nucleus?);
- and er· the |farmer · ap↑parently·↑didn't want his ↑wife to KNŌW (first two accented syllables rhythmically separate from the rest – perhaps three onsets, the first two as 'orphans' – i.e. not part of a completed tone group – and the third, *didn't,* higher in pitch than *par,* as onset of the completed tone group?).

On the other hand, the perceptual and distributional similarities between onsets and boosters, as well as their relationship with pausal structure (cf. the third example) suggest that rigid application of full tone group structure to all occurrences of accent sequences in tokens is wrong,

and that a less Procrustean approach which allows a linear development with the possibility of uncompleted sequences, on the one hand, and includes non-accentual criteria (such as overall pitch height and pitch bandwidth) for various different kinds of prosodic grouping on the other, does more justice to the facts. A systematic description along these lines is attempted in § 4, where sets of representations of quasi-autonomous levels of organisation (following the descriptive strategy of Jassem & Gibbon, 1980, Fig. 3) are used, rather than in-text, single-level representation strategies; the latter are useful for many purposes, but often unrevealing.

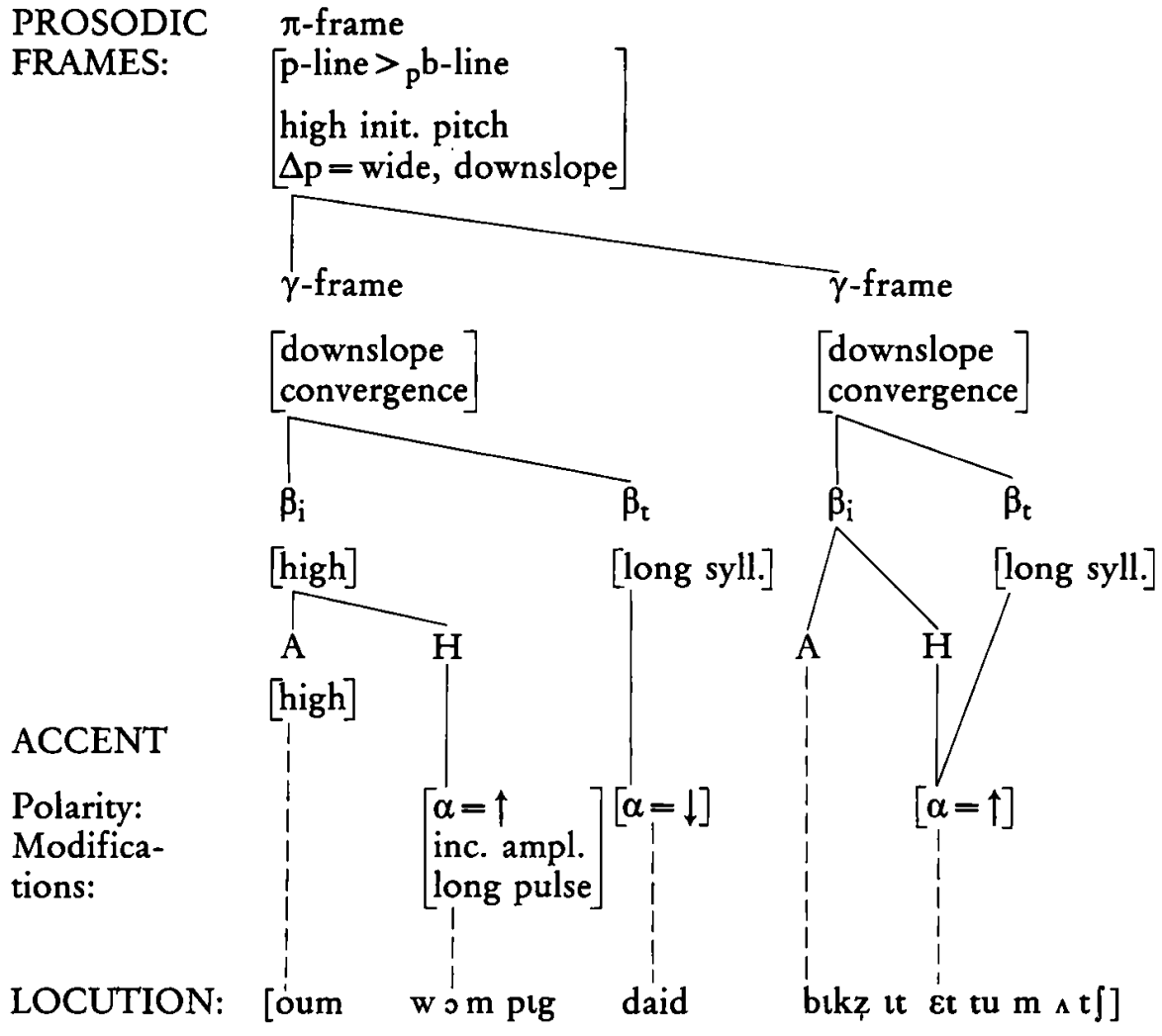

Figure 6: Prosodic frames and selected specifications of selected primitives for "*p* because *q*" *(one pig died because it ate too much)* in the Crystal & Davy corpus extract (see text).

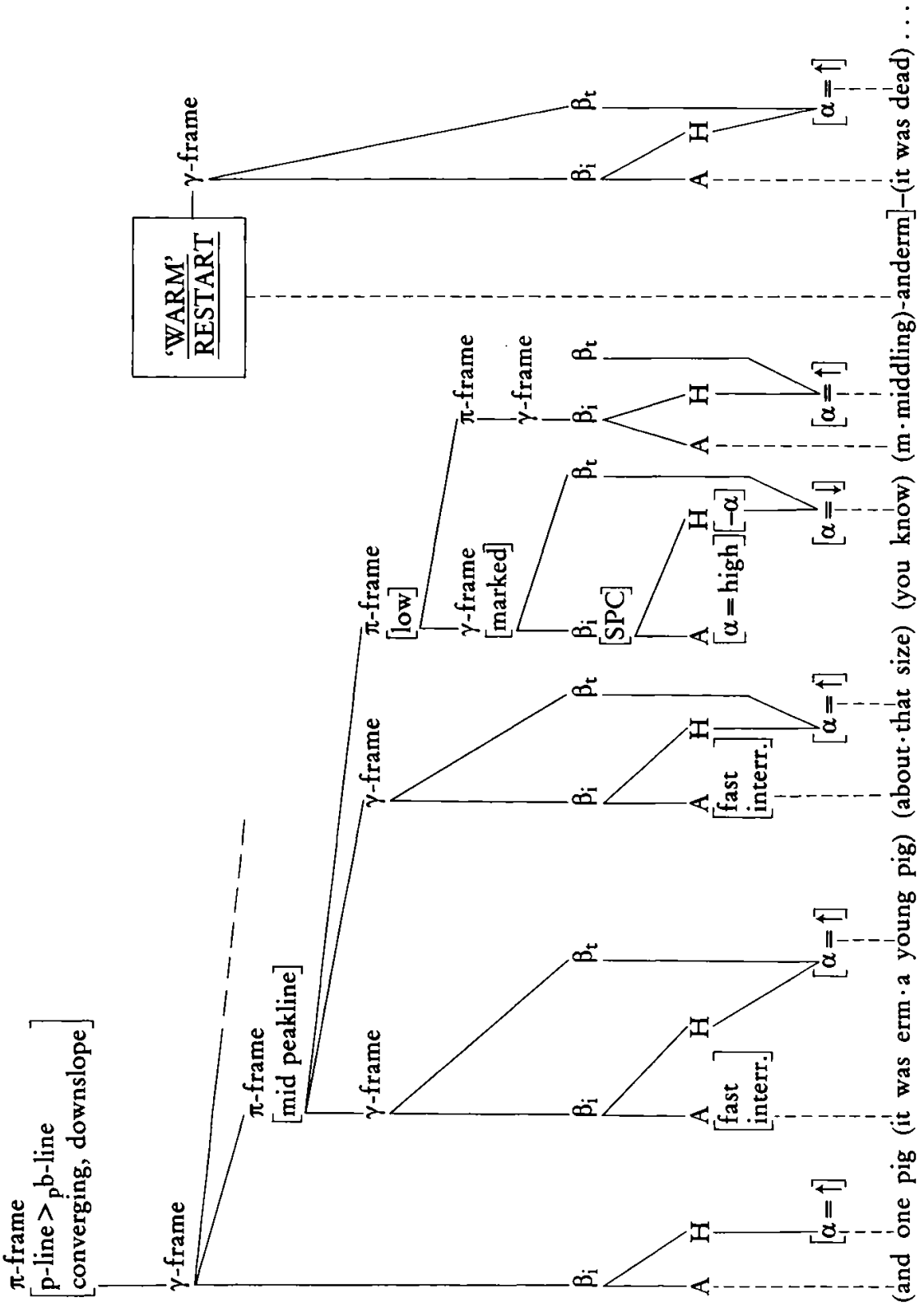

Figure 7: Prosodic frames and selected specifications for E1 (*p*) from the Crystal & Davy corpus extract, showing 'warm' restart ('warm' since existing top-level π-frame specifications are used). Mapping of prosodic categories on to the locution is symbolised by parentheses and dotted lines; forced closure of open γ-frames after the restart is symbolised by a large square bracket.

4. *Prosodic frames in discourse*

At the risk of being accused of reading too much into a limited set of examples, I shall restrict attention to "*p* because *q*" in Stage 1 and $E_1(p)$ of Stage 2 in the dialogue concerned. It would be scarcely possible to cover more detail in the space of a single article. The prosodic structures of the relevant extracts are represented in Figures 6 and 7. The main aspects of an appropriate syntax for such structures are shown in Figure 7.

The nodes π and γ are interpreted as organisational *frames;* the processes which operate within these frames have access to features denoted by the 'multilabels' (Eikmeyer, 1980) which are attached to nodes in the figures.

The overall picture conveyed by the figures is relatively complex; however, at each level organisation is relatively simple, as Figure 8 shows. The principles underlying the interrelationships between levels are also simple, though the combinatory possibilities are large, leading to complex structures in actual use. The apparent complexity of specific structures is perhaps an artefact of the analysis, a kind of 'optical illusion' to which all analytic description is prone. Conditions on mapping between specifications at different levels of embedding cannot be discussed here.

π-frames

Observable features associated with π-frames are:

(1) long-scope baseline feature specification, in particular [peakline $\rangle_p$ baseline], downslope, peakline-baseline convergence.

(2) High initial peakline value (ultimately manifested at β_i); cf. the 'initial boundary features' of § 2.1. above. By this criterion, π corresponds approximately to Fox's "paratone group" (1973) or Brazil's "pitch sequence" (1978).

(3) In Figure 6, peakline modifications can be observed in two cases; in each case, the category with altered peakline is marked as a π-frame.

γ-frames

Observable criteria for γ-frames are chiefly boundary criteria (see below), though the global π-frame specifications (e.g. downslope) are also accessed in the internal organisation of γ-frames; these also show internal downslope (to a lesser extent than the longer γ-frame sequences within π-frames). In different terminology, the higher-level specifications 'percolate' down to the lower level. The superordinate π-frame peakline may be determined heuristically from the heights of Heads of successive γ-frames.

Recursion and iteration

The π-frames are recursive with a peakline change at each recursion; this

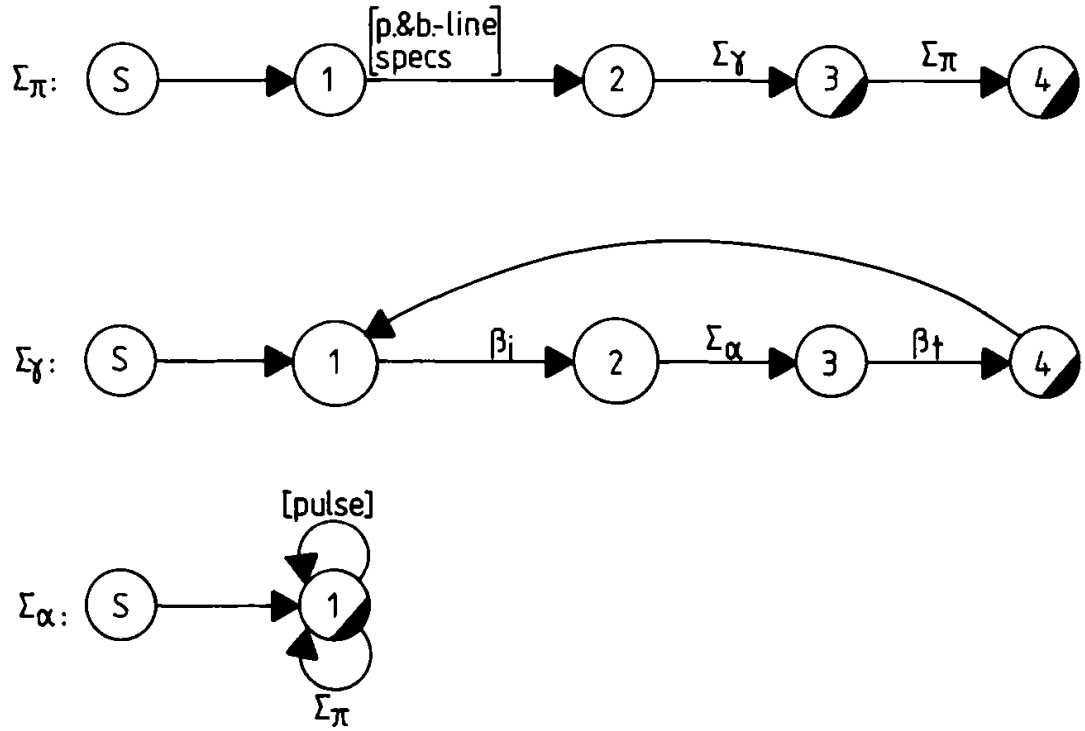

Figure 8: Networks for description of structures represented in Figures 5 and 6. Note the two kinds of recursion of π-frames: right-branching, in, Σ_π and centre-embedding, in Σ_α (cf. "narrative tag" and parenthesis marking, respectively). Application to other data types will necessitate specifying more end states, *inter alia*. Details of ATN tests and actions based on § 2 are not included; a full phonetic interpretation requires such specifications.

is an empirical claim based on observation of peakline (especially Head height) patterning. Three cases of π-frame recursion are shown in Figure 6 (the recursion analysis is also reflected in the parallel organisation of locutionary structure at these points: discontinuities with embedding also occur). Exit from recursion is either regular (perhaps here in the parenthetic *you know*?), with return to the previous peakline, or by 'purging', i. e. an inelegant or 'brute force' return to the highest level, involving reinitialization to the original peakline value, not a continuation of the immediate preinterruption value (cf. *anderm* in Figure 6, also marked by slow timing). A similar event occurs at the locutionary level: a kind of 'left dislocation' with an anaphoric ('resumptive') subject *it* which may be interpreted procedurally in the same way. A further non-prosodic signal of a similar process is the use of *so*, already noted in § 3.1.

The γ-frames are not recursive, but iterative within π-frames. The π/γ-frame distinction is obscured in the customary p-data, the speech style of reading sentences aloud, since π-frames and γ-frames tend to be co-extensive in such data. Within γ-frames, pulses occur in a lower level iterative control structure. In previous intonation studies concerned with these questions this distinction between recursive and iterative control flow at different levels has not been made. The networks shown in Figure 8 re-

present the levels concerned. The highest level system, $\Sigma\pi$, first defines the global specification and then calls the next system, Σ_γ. On return to the highest level, recursion to Σ_π itself is possible, yielding 'narrative tags' or similar marking stuctures. The Σ_γ system generates initial boundary specifications, transfers control to the accentual rhythm system, Σ_α, and on return generates terminal boundary specifications; it has the option of iterating (looping). The lowest level system, Σ_α, iteratively generates accentual pulses or, alternatively, recurs to Σ_π to generate parenthesis markings. Accent iteration explicates the notion of 'subjective rhythm'.

Boundaries

The boundaries of a γ-frame are designated β_i and β_t. The initial boundary of a γ-frame is, in the illustrations, marked by a $\uparrow$ pulse of high amplitude; if the γ-frame is, in turn, π-frame-initial, then the pulse is higher still in amplitude.

A criterion for the initial boundary of a γ-frame is, in a large number of cases, Anacrusis ('A' in the figures), i.e. a sequence of unstressed syllables spoken more rapidly than elsewhere and not in the general rhythmic pattern (Jassem & Gibbon, 1980; Jassem, this volume). Anacrusis is taken to mark the initial unstressed segment of a γ-frame before the Head (the first accented syllable of a γ-frame). The initial boundary is complex, not correlating with a single boundary tone, but consisting of *both* Anacrusis (where present) *and* Head. The unmarked β_i has a mid-pitch Anacrusis and a $\uparrow$ pulse, ($\triangleq$ high Head, Onset). In one fairly common marked case (cf. the fixed parenthetic expression *you know* in $E_1(p)$, Anacrusis is high, followed by a $\downarrow$ pulse, ($\triangleq$ low Head), showing the *Principle of Sequential Pitch Contrast.* Anacrusis is an important category, both heuristically and theoretically. Its phonetic properties, as a fast, arhythmic series of unstressed syllables (Jassem) and separated by a pitch level change from the preceding group (Crystal), signal the start of a new γ-frame. Three special properties of Anacrusis are illustrated in the present corpus extract:

(1) It operates together with the category Head to form a complex boundary signal β_i; the neutral case is lax, i.e. baseline-level Anacrusis with $\uparrow$ Head. Tense ($\triangleq$ high pitch) Anacrusis also occurs, either with $\uparrow$, forming an overall tense β_i, or with $\downarrow$(as in *you know*), creating a conspicuous boundary effect by means of the *Principle of Sequential Pitch Contrast* (§ 2.3.).

(2) It is not necessarily separated from the preceding γ-frame by a pause; in several cases, it follows the preceding β_t extremely rapidly (marked as [fast] in the figures).

(3) It is also independent of, and may be interrupted by, minor hesitation stretches, filled or unfilled. A criterion for continuity of the Anacrusis in such cases is continutity of pitch contour near baseline level. It has al-

ready been noted that the hesitation type *anderm* has different discourse status.

The γ-frame terminal boundary β_t has different properties from β_i- the traditional terminology of Nucleus and Tail is adopted here for exposition, but not used in the figures. The neutral case of a β_t is simply a ↑ pulse followed by a fade at all levels: π, γ and pulse level. The fade constitutes an unmarked Tail. Marked cases are of several kinds:

(1) An [α Accent]^[–α Accent] sequence on Nucleus and Tail, respectively (≙ steep falls and steep rises).

(2) [tense] or [tensing] Tail.

(3) Tail with laryngeal equilibration (= 'chroma', 'stylization'). Note that (1) and (3) may combine. If so, double-fall or double-rise 'call contours' are produced; (3) alone results in a level contour. In the former case, the [–α Accent] contributes to the *Sequential Pitch Contrast* 'conspiracy' which in the neutral case applies automatically through the cumulative effect of 'fade' at all levels.

Nucleus/Tonic

Recent investigations (Brown & al., 1980) have cast doubt on the validity of the tradtitional notion of 'Nucleus' or 'Tonic', with results which indicate that impressions of 'nuclear prominence' are a function of lexico-contextual information (including contrastive contexts) and of sequence-final position *per se*, and are not elementary prosodic judgments. Investigations of Head status have tended to point in the other direction, which may mean that the so-called Nucleus is not as well-defined a category as Head in prosodic frames – despite traditional descriptive priorities. The γ-frames are frequently interrupted and re-started in 'real-time speech'; a well-defined initial boundary is, perphaps for this reason alone, a more important point of orientation than a final boundary. When γ-frames are concluded after a recursion, either the peakline before the recursion is recovered, or a new Head is established. In either case, the Head is the determining category, not the Nucleus.

Accentual pulses

The π and γ systems illustrated in the figures have a variety of interesting properties. The pulse accent sequences themselves belong to a system of rhythm generation which includes secondary accents. The pulses are adapted to the current context in three main ways:

(1) position in the peakline-baseline vector defined in the dominating π-frame or frames, and by the boundaries β_i and β_t in the dominating γ-frame(s);

(2) semantic conditions determined by the current locutionary constituent, contrastive contexts being associated with increased pulse amplitude (= higher/lower deviation from the baseline) and syllable lengthening,

and appraisive contexts with longer leading flank rise time together with syllable lengthening, yielding certain kinds of 'emphatic' 'rise prefix' including delayed rises, rise-falls, 'homosyllabic preheads';

(3) in other styles such as reading aloud, syntactic structure (which provides constraints of detail in other styles, too) may be the main available context, in which case pulse amplitude is mainly determined by locutionary specifier-head and other operator-operand relations, approximately as postulated in conventional P-data-based 'sentence stress' descriptions; cf. also Brazil, this volume).

Although the modifications to accents occur locally, they are determined by global context specifications which are accessible to pulse routines; the pulses themselves do not necessarily have to be specified even for polarity, if the global [baseline $\rangle/\langle_p$ peakline] feature is accessed. Local specifications such as (1), (2) and (3) above can override the global specifications, but local specifications are marked specifications. This applies in particular to pulses in non-Head and non-Nucleus environments, which do not have any direct affiliation to frame categories in the figures.

These conditions on pulse accent sequences may be restated in the following terms: this approach implies that there is no actual accent *type* Head or Nucleus; Head and Nucleus are *positions* in γ-frames, i.e. functional terms analoguous to 'subject' and 'object', etc., in sentence syntax; they are associated with 'syncategorematic' position-marking systems β_i and β_t, whose phonetic correlates (boundary tones, pulse modifications) may be said to have a theoretical status within γ-frames analogous to case inflexions within sentences.

5. *Adaptation as context sensitivity*

Towards the end of § 4 the 'context-sensitivity' of pulse modification was discussed, and described in terms of three kinds of context: position in π-frames and γ-frames, relation to locutionary context, and relation to other aspects of context. Pulse modification was said to be based on local look-up of global values. Some of the context values involved in such 'prosodic handshaking' will be dicussed in this section and followed by a more general interpretation in terms of a feedback cycle in § 6.

(1) The association of peakline-baseline features with π-frames has already been discussed. Although these were represented as being assigned to the current π-frame, they may be even more general 'genre' or 'style' features, such as the [peakline $\rangle_p$ baseline] feature in adult-child speech styles. Another genre-marking feature of this kind is [peakline $=_p$ baseline], a manifestation of the larynx equilibration strategy used in some narrative styles with repetitive character (cf. Ladd, 1978, on 'stylization'). This contour flattening effect occurs in the discourse discussed in § 3, du-

ring E(*q*) of Stage 2; it appears auditorily as a mid, narrow bandwidth pitch contour and in the use of level tones on *dead, know, before, wife.* The context specification table necessary for feature assignments of this kind must contain information on the structure of the current discourse participant and channel configuration; a suggestion concerning the context information involved in such 'call contours' is given in Gibbon (1976), and a more elaborate contextual framework of the kind required is used in a different descriptive application in Gibbon (1981).

(2) The instances of π-frame recursion described in § 4 appear to be triggered by a context-lookup connected with assessments made by the speaker about the amount of information the addressee already has (cf. the parenthetical *you know*) or the amount of background information he still needs in order to be able to grasp the main point of the narrative (cf. the 'narrative tags' such as those associated with the first recursion in Figure 6 or the part of sequence E(*q*) in Stage 2 which ends in *. . . and all this*). The narrative tags involve mid pitch-range peakline-baseline convergence and an auditory impression of fading intensity over at least one π-frame.

(3) The preceding two instances involved context look-up at the π-frame level. Anecdote 1 in this discourse does not show as much variation at π-frame level as Anecdote 2 (not shown), in which far more occurrences of pulses and associated sequential pitch contrast mechanisms ('upward' laxing, 'final' rising pitch) occur; it also includes more low level iterative structures similar to those mentioned in § 2 for a German dialogue: complete sequences of accents in γ-frames end in Nuclei; the γ-frames themselves are in sequences within a π-frame, and the last γ-frame in such sequences ends in a more prominent Nucleus. In particular, the occurrence of 'rise-tags', i.e. upward relaxation after a final pitch-lowering mechanism, as marks of subordination to higher-level structures (cf. Gibbon, 1980a: 86, and Bing, this volume), are more common in Anecdote 2.

(4) Local adaptation to lexical properties of the current locutionary constitutent is shown in a number of cases in the dialogue; in particular, appraisive expressions such as *oh, revolting, terrible, engrossed, marvellous, wonderful, furious* are associated with increased pulse amplitude and/or pulse leading flank delay features which typically yield auditory rise-fall contours.

(5) The 'canonical' Armstrong & Ward Tune 1 type of contour, with its typical pronounced downslope from a high pulse amplitude at Head position, is also used in adaptation to discourse structure. It occurs, with minor variations, in the initial and final stages of the anecdotes, the explicit statements of the semantic frame on which the anecdote is built; cf. the following (in the Crystal & Davy transcription):

Anecdote 1: oh and |one 'pig DÌED| because it ÀTE too 'much| . . . the ↑pig had ↑DÌED| be|cause it had ÈATEN too 'much| you |KNŌW|

Anecdote 3: al|though a ↑friend of ÒURS| who · was |so · ↑passionately ↑FÒND of PÍGS| . . .
and would |lean 'over and ↑TÀLK to them| |FÒNDLY| |WÒULDN'T he|

This contour also occurs in mid-narrative, possibly as the unmarked contour with no locally effective contextual adaptation. If this is the case, there may be an argument for accepting such contours 'normal intonation' judgments on P-data; in any case, it is hardly surprising that this tune should have something approaching genre-marking status for formal reading aloud of written texts, e.g. in news bulletins (where each sentence may be a statement of a new semantic frame) or in P-data (isolated sentences or standardised brief dialogue exchanges). It has close analogues in other West Germanic languages, such as the contours described by v. Essen (1956) for German, or the 'hat contour' of 't Hart for Dutch.

6. *On the descritpive potential of a process orientation*

A large number of open questions remain. One of these is whether the pulses themselves have a reasonably simple relationship with physiological reality, or whether they are best regarded as abstract phonological features. Internal descriptive evidence for their validity is given by their status within a detailed and consistent descriptive system; external evidence from the physiological phonetic literature indicates that the suggestion may not be far from the truth (cf. the summary in Ohala, 1978). Another point is the empirical validity of the distinction between recursive and iterative solutions to the descriptive problems, in particular the validity of the distinction between π-frames and γ-frames. However, whatever the most plausible models for formal and substantive properties of intonation might be, it will almost certainly be necessary to turn to procedurally oriented models if ways out of some of the old impasses are to be found. One way of solving the hoary 'levels vs. contours' controversy is to find a third category which is neither the one nor the other; this has been done here for part of the English intonation system with the pulse accent thesis. A similar solution to the problem of baseline definition is the distinction between *baseline* and *peakline,* which offers descriptive advantages in several areas.

The claim that a descriptively useful distinction can be made between iterative and recursive properties of intonation processes was put to the test in § 4 as a means of explaining pitch bandwith narrowing and pitch height 'restart' patterns.

The strategies discussed in § 2.2.1, if they have any reality in speech processes, help to explain why intonation analysis on a perceptual basis appears over-simple and obscures distinctions between the different procedural factors which contribute to a percieved or measured pitch func-

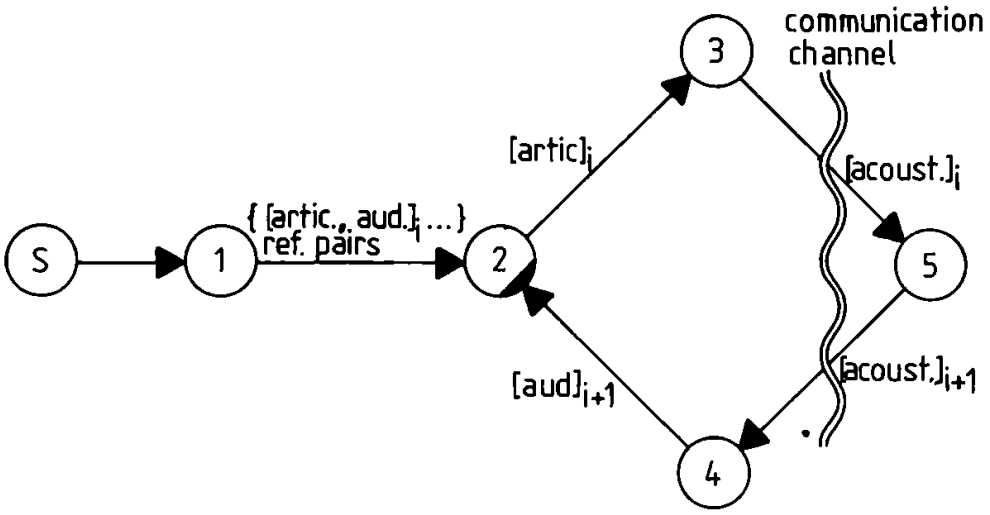

Figure 9: Outline of a feedback network as required in the implementation of each level of the systems shown in Figure 7.

tion. On the other hand, the strategy concept also tends to make any intonation hypothesis based on articulatory features difficult to falsify or confirm. A further implication of the auditorily based overriding strategies, in particular that of larynx equilibration under the stabilising influence of negatively phased auditory feedback, is that they control a functional adaptation cycle which is sensitive to situational and locutionary context, but is also operative within the phonological processes of intonation as a dynamic interäction of articulatory and auditory factors. This interactive cycle is outlined in the network of Figure 9. The indices symbolise stages in temporal operation. The full cycle has a planing stage, an execution stage, as well as checking and (as required) correcting stages. In this aspect of the present approach lies the motivation for suggesting in § 2 above that the fundamental units of intonation organisation, and possibly for all phonological organisation, are (articulation, percept) pairs. An advantage of this conception is that it systematically rejects a 'speaker/hearer' idealisation for discourse phonology, preferring a dynamic combination of speaker and hearer perspectives.

To summarise: a process-oriented model of intonation for English has been proposed, which incorporates a recursive category, the π-frame, an iterative category, the γ-frame, and a conception of accents as articulatory pulses. These were applied to the description of discourse data, and formulated as a system of transition networks which take contextual feedback at phonological and functional levels into account. Much is still speculative, there are gaps in the theory, and new kinds of empirical evidence will be needed in order to test the claims. But if it suggests ways of combining the various lines of analysis required in the study of discourse phonology, even if it does not to turn out to be viable in detail, the theory will have served a useful purpose.

Bibliography

Abercrombie, D. et al. (eds.) (1964). *In Honour of Daniel Jones.* London: Longman.

Bachem, A. (1950). Tone height and tone chroma as two different pitch qualities. *Acta Psychologica 7*, pp. 80–88.

Bing, J. M. (1980). Aspects of English Prosody. Bloomington: IULC. (Diss. Amberst 1979).

Bing, J. M. (this volume). A discourse domain identified by intonation.

Brazil, D. (1975). *Discourse Intonation. Discourse Analysis Monographs* No. 1. English Language Research, Birmingham University.

Brazil, D. (1978). *Discourse Intonation II. Discourse Analysis Monographs* No. 2. English Language Research, Birmingham University.

Brazil, D. (this volume). The intonation of sentences read aloud.

Brown, G. et al. (1980). *Questions of Intonation.* London, Croom Helm.

Chao, Y. R. (1956). Tone, intonation, singsong, chanting, recitative, tonal composition and atonal composition in Chinese. In Halle, M., Lunt, H. and McLean, H. (eds.), pp. 52–59.

Chitoran, D. (ed.) (1976). *Second international conference of English contrastive projects.* Bucharest, U. P.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English.* New York, Harper & Row.

Crystal, D. (1969). *Prosodic Systems and Intonation in English.* Cambridge, C.U.P.

Crystal, D. (1971). Relative and absolute in intonation analysis. *Journal of the International Phonetic Association* 1: 17–28.

Crystal, D. and Davy, D. (1975). *Advanced Conversational English.* London: Longman.

Dechert, H. W. and Rauppach, M., eds. (1980). *Temporal Variables in Speech: Studies in Honour of Frieda Goldmann-Eisler.* The Hague, Mouton.

Eikmeyer, H.-J. (1980). *Transformationsgrammatiken mit Multilabels. Definition und Anwendungsmöglichkeiten.* Hamburg, Buske.

v. Essen, O. (1956). *Grundzüge der hochdeutschen Satzintonation.* Ratingen, Henn.

Fox, A. (1973). Tone sequences in English. *Archivum Linguisticum* 17–26.

Fromkin, V. A., ed. (1978). *Tone. A Linguistic Survey.* New York, Academic Press.

Gibbon, D. (1976a). *Perspectives of Intonation Analysis.* Bern, Lang.

Gibbon, D. (1976b). Performatory categories in contrastive intonation analysis. In Chitoran, D. (ed.) pp. 145–156.

Gibbon, D. (1981a). A New Look at Intonation Syntax and Semantics. In James, A. and Westney, P. (eds.) pp. 71–98.

Gibbon, D. (1981b). *Forschungsprojekt Propädeutik der Phonetik. Erster Teilbericht; Digitale Frequenzanzeige mit einem Mikrocomputer zur Visualisierung der Sprachmelodie.* Fachhochschule Köln, Fachbereich Sprachen, März 1981.

Gibbon, D. (1981c). Idiomaticity and functional variation: A case study of international amateur radio talk *Language in Society* 10: 21–42.

Gibbon, D. (1982). Intonation in Context. An Essay on Metalocutionary Deixis. In Rauh, G. (ed.), pp. 195–218.

Gibbon, D. & M. Selting (1983). Intonation und die Strukturierung eines Diskurses. *Zeitschrift für Literaturwissenschaft und Linguistik,* LILI 49.

Halle, M., Lunt, H. and McLean, H., eds. (1956). *For Roman Jakobson.* The Hague, Mouton.

Halliday, M. A. K. (1967). *Intonation and Grammar in British English.* The Hague, Mouton.

t'Hart, J. and Collier, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics* 3: 235–255.

James, A. and Westney, P., eds. (1981). New Linguistic Impulses in Foreign Language Teaching. Tübingen, Narr.

Jassem, W. (1952). *Intonation of Conversational English (Educated Southern British), as: Travaux de la Société des Sciences et des Lettres de Wrodaw,* Series A, No. 45, Wrocław.

Jassem, W. and Gibbon, D. (1980). Re-defining English stress. *Journal of the International Phonetic Association* 10: pp. 2–16.
Jassem, W., Hill, D. R., Witten, I. H. (this volume). Isochrony in English speech: Its statistical validity and linguistic relevance.
Klinghardt, H. and Klemm, G. (1920). *Übungen im englischen Tonfall.* Cöthen.
Ladd, D. R. (1978). *The Structure of Intonational Meaning.* Diss., Cornell University.
Lehiste, I. (1975). The phonetic structure of paragraphs. In Nooteboom, S. G. and Cohen, A. (eds.), pp. 195–203.
Liberman, M. Y. (1978). *The Intonational System of English.* Bloomington, IULC.
Lieberman, P. (1967). *Intonation, Perception, and Language.* Cambridge, Mass., M.I.T. Press.
Neisser, U. (1967). *Cognitive Psychology.* Englewood Cliffs, N. J., Prentice-Hall.
Nooteboom, S. G. and Cohen, A. (eds.) (1975). *Structure and Process in Speech Perception.* Heidelberg/New York, Springer.
Ohala, J. J. (1978). Production of Tone. In Fromkin, V. A. (ed.), pp. 5–40.
Ozga, J. (1980). 'A functional analysis of some pause and pitch step-up combinations', In Dechert, H. W. and Rauppach, M. (eds.), pp. 221–225.
Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation.* Diss., M.I.T.
Pike, K. L. (1945). *The Intonation of American English.* Ann Arbor, The University of Michigan Press.
Rauh, G., ed. (1982). *Essays on Deixis.* Tübingen, Narr.
Reich, P. A. (1969). The finiteness of Natural Language *Language* 45: 831–843.
Trim, J. L. M. (1964). Tonetic stress-marks for German. In Abercrombie, D. et al. (eds.), pp. 374–383.