

Intonation in Context. An Essay on Metalinguistic Deixis

Dafydd Gibbon

1. Context, content and method

The meanings of intonation patterns are indexical, or context-dependent. But they are not only indexical in a limited semiotic sense (as symptoms of speaker states) or in a generalised semantic sense (as any dependence on utterance context): they are deictic, in a strict sense to be described in this essay. Intonation deixis primarily involves stretches of speech itself rather than non-verbal objects or events, and is metalinguistic in a quasi-Tarskian sense which will be referred to here as 'metalinguistic'.

The more general methodological theses underlying this approach include the following conception of the role of context descriptions:

- (1) Context descriptions have a dual role
 - i. as heuristic aids in delimiting the boundaries of linguistic descriptions;
 - ii. as objects of functionalistic linguistic descriptions.
- (2) All linguistic methodologies utilise informal context descriptions in the first sense in defining their own descriptive content.

The second consideration also applies, for instance, to empirically parsimonious approaches such as generative linguistics, where the heuristic notion of context is restricted to the native speaker and those judgments on utterances which are not specific to either production or perception, or to discourse sequencing (but cf. § 5.1. and the notion of "natural response"). I use the term "utterance" rather than "sentence" advisedly: any linguistic data base, however inexplicitly acknowledged, involves actual judgments of speech utterances or written tokens at some stage, while sentences are abstract products of the description and represent hypotheses about some properties of observable utterances.

Intonation is not a completely autonomous semiotic system, of course, since it clearly co-functions with other speech systems in focussing, defocussing, locution integration ('cohesion' marking), discourse strategy control (turn-taking, calling, ritual speech). Its autonomy is distributional at a formal level; in addition, the autonomy assumption has considerable heuristic value.

In the area of accentuation, the autonomy hypothesis makes it possible to distinguish between such spatio-temporally realised properties of utterances as pitch accents, which may be described in an independent intonation syntax, and a more abstract notion of 'stress' as a structural property of the dominant sister constituent of locutionary words or phrases (cf. Jassem & Gibbon, 1980; also § 4.2. below). Depending on the semantic interpretation of a locution in a given context, an intonation assignment function may synchronise pitch accentuation with a given 'stressed' (= 'normally accentable') syllable, or leave it unaccented. For a discussion of 'normality' in such contexts, cf. Gibbon, 1976: 28 ff.) The apparent perceptual prominence (a cover category for accent, stress and other judgments) of syllables is a function of the overall reconstruction of concepts from immediate physical cues plus stored knowledge about structural 'stress', and is not a simple datum. In this area of speech, the human tendency to perceive in functional terms rather than in more form or sense oriented (though still cognitively reconstituted) terms has often led to fallacious conclusions about so-called 'stress patterns'. This tendency can fortunately be reduced by perceptual training and the use of instrumental methods.

An intonation syntax contains a rhythm generator, a system of accent specifications, a long-scope pitch strategy system, and an 'interrupt' control system for coping with parentheses and *ad hoc* (ostensive and appraisive) contexts for accentuation. It has already been noted that procedural formulations such as these reflect concern with temporal aspects of intonation use.

It has been suggested that the basic syntax of English accentuation, as of other West Germanic languages, is right-linear; that is, it may be generated by a right-linear finite state grammar, finite state machine or transition net-work (Reich, 1969; t Hart & Cohen, 1973; Pierrehumbert, 1980). This appears to be true for restricted data types such as read-aloud sentences and, with reservations to be noted below, for some other data types, too. The main reservation for fluent speech is that at least two levels of structure (independently of the structural distinction between primary and secondary 'stress' in words) can be identified in complex discourse, which are beyond the scope of a finite state machine (cf. Fox, 1973; Gibbon, 1981a).

Similarly, the ideal speaker-listener modelled by such descriptions is, for an ontological realist, a hypothesis about some properties of the real people who make the concrete judgments in spatio-temporal contexts (e.g. "Do you get 'Here John SAW Mary?' - "No."), which are explicated in grammars.

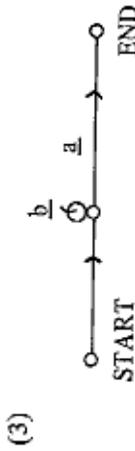
Formal aspects of the intonational sign system of English are described in § 2, a deictically relevant notion of context is developed in § 3 and applied to intonation in § 4, and in § 5 a selection of descriptive problems is treated, with locutions (text) figuring as part of the context for meta-locutionary deixis. The orientation of this essay is perhaps more theoretical than purely descriptive; for this reason, the examples given are largely constructed and illustrative, rather than observed, transcribed spontaneous utterances with greater evidential value. Procedural formulations in this approach reflect a concern with the role of time in the use of intonation.

2. Intonation

By intonation I mean the the temporal organisation and use of the phonetic parameters known as 'phonation rate' in the articulatory domain, 'fundamental frequency', 'f₀' in the acoustic domain, and 'pitch' in the auditory domain (Lehiste, 1970). This is a broader definition than is usually given, covering both long and short scope aspects of pitch patterning (e.g. pitch accent and utterance-long pitch gestalt, respectively). The organisation of accentuation will be in the centre of attention in this paper. Intonation, including accentuation, must be distinguished from other prosodic and from the paralinguistic features of speech (Crystal, 1969) which contribute to the naive native speaker's judgments of 'manner of speaking', 'tone of voice' (cf. 'gentle/throaty/whining' voices) and which are described by a wide range of *verba dicendi* with adverbs of manner in narrative writing. Intonation is partially synchronised with other features from this complex, as well as with gestures from the kinesic domain. In a full description, these items would also have to be considered.

Intonation patterning is better described as a quasi-autonomous semiotic subsystem within language than as an appendage to locutionary syntax. Intonation-locution relations vary greatly with speech register: data derived from the reading aloud of highly edited written language, particularly isolated sentences, tends to be more highly structured and oriented toward locutionary syntax than does fluent spontaneous discourse, which foregrounds serial projections of semantic relations (cf. § 5.1.).

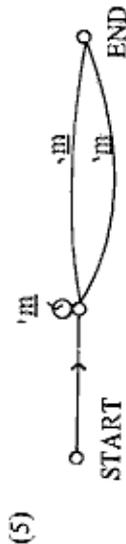
The right linear structure of accentuation may be shown by means of a state diagramme (reduced here to a minimum):



Here *a* is a terminal, *b* a non-terminal accent. Accent shaping rules assign phonetic values to the output of the machine with respect to the immediate context, which is (a) the current state of the machine, and thus also (b) the immediately preceding output. Illustrations of such patterns in tonetic notation would be (Lewis, 1977: 36):

- (4) *Diner* 'Waiter, there's a fly in my soup.
Waiter 'Then per'haps you'd pré'fer a-red wine, sir.

(The tonetic signs mean: high-low pitch accent [']; mid-low pitch accent [·]; high pitch accent [ˈ]. The comma denotes a break.) A network such as (5) will generate these patterns:



(The dummy sign *m* is used to identify pitch accent height). An example of multiple level structuring from the same source is the following (*cop. cit.*: 37):

- (6) That's 'not the 'question, v really. 'She has all these dreadful 'allergies.

(Accent specifications are analogous to those given above.) The main patterns are accounted for by a network like (5), with the high accents as non-terminal, the falling pitch accents as terminal symbols. In addition, a higher-level relation between accents ending in a rise and accents ending in a fall is required, and an interrupt modification for the appraisive context 'dreadful', which enters into this higher-level relation, already anticipated by the non-terminal fall-rise tag which follows the first right-branching sequence. The higher-level relation has the same kind of right-linear structure as the lower-level rhythmic component, though a tendency to binary structuring exists here (cf. also § 5.1., on topical structure development). This indicates that a more abstract transition network generating calls to the lower-level routines plus higher-level output is required. Although the strings could be generated by a single lower-level machine, the problem of systematic sem-

antic interpretation would remain (§ 4); for the same reason it is not descriptively adequate to account for tags, which also enter into a higher-level structuring, by a linear extension of a finite state machine (Pierrehumbert, 1980: 98 f.).

Certain important features of intonation as a whole have been omitted from this account, such as long-scope discourse-oriented variation in pitch height (cf. Brazil, 1975, 1977). Enough has been described for the purposes of this essay, however. The questions which now remain to be examined in later sections concern contextual semantics or pragmatics: what kind of semantic interpretation do these intonational patterns receive? In particular, what makes their constituents into deictic terms? The first step will be to give an appropriate account of context and metalinguistic deixis.

3. Context description

The notion of context developed here is fairly rich, mainly because of an interest in the second role of context descriptions (§ 1).

First, it is necessary to include temporal progression in the context description. Prosodic features can hardly be discussed without referring to a temporal vector: rhythm, relative duration of syllables and words, directionality (the temporal correlate of non-commutability of constituents). Further, the pairing of intonation patterns with locutions involves synchronising partially autonomous semiotic subsystems in language (Gibbon, 1976: 62 f.). The 'shifter' characteristics of locutionary deictic terms also require consideration of a temporal vector, especially if context-change in ostensive functions within discourse is concerned:

- (7) This is one — and this, and this . . .

Second, person deixis and the 'vocative' uses of intonation (§ 5.4) require replacement of ideal speaker-listeners with at least a speaker-plus-listener distinction. And once this distinction has been made, it is only a short step to the hypostatization of a speaker-listener relation in the form of Jakobson's 'contact' or channel (1960). The channel not only co-defines the participant roles which are relevant for deixis, but also spatial relations such as 'proximal' and 'distal', and temporal constraints on the use of speech, such as signal transience or storage (cf. speech vs. writing).

Figure 1 outlines an indexical co-ordinate system for deictically relevant context descriptions (for detailed justification cf. Gibbon, 1981b). It consists of two parts: a primitive indexical structure, and a set of levels of analysis (interpretation/description/representation).

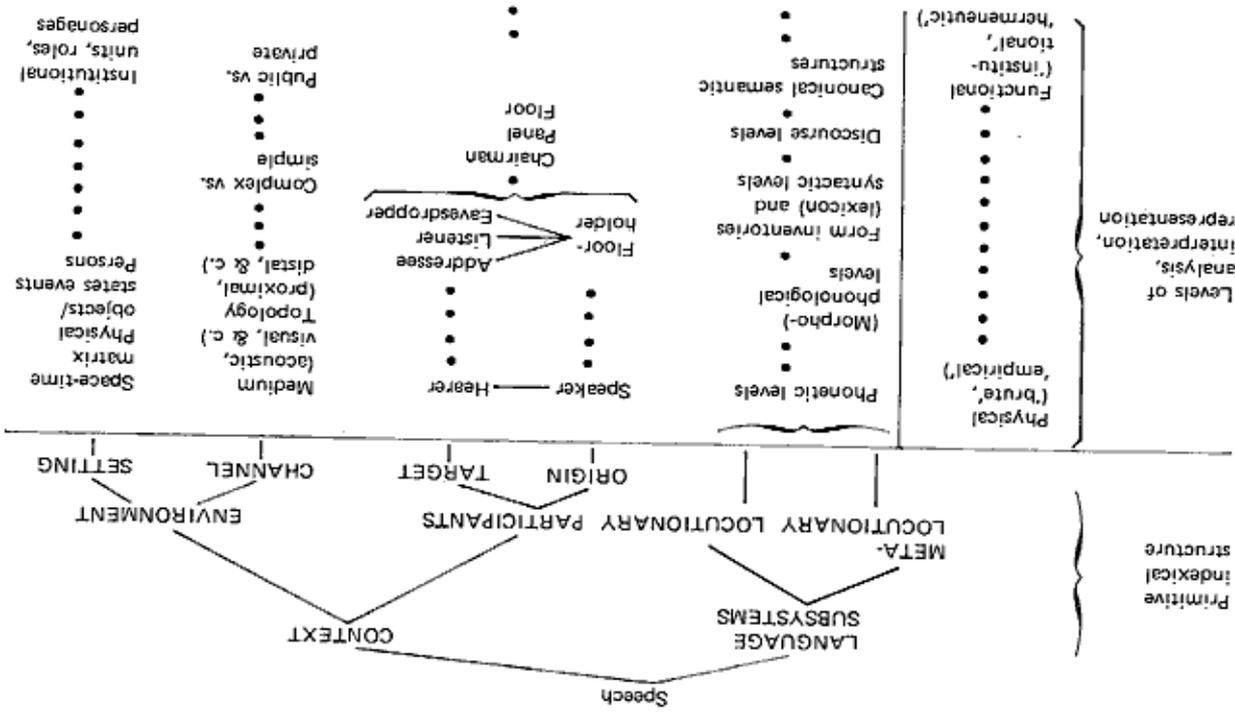


Figure 1: Outline of an indexical co-ordinate system for context descriptions, with (a) primitive indexical structure; (b) selection of parameters constituting levels of analysis relative to the primitive indexical structure.

The levels of analysis define individuals or 'objects' of different logical and descriptive sorts; the individuals (e.g. units, categories, relations) defined at each level are functions of the criteria of analysis used. Briefly, they range from physically or sensorily defined individuals (objects, events, states) to hermeneutically defined individuals like institutional 'personages', 'actions', 'roles'. The individuals at each level are related to each other in terms of the relations provided by the primitive indexical structure.

The primitive indexical structure constitutes a deictically relevant, spatio-temporally interpretable context. This holds initially at the physical levels of analysis and, by virtue of congruences between the different levels, at the functional levels, too. At a physical level of analysis, for instance, the participants in a dialogue are *transmitters* and *receivers* of noises. At more abstract level relevant for interpreting the personal pronoun system and for discourse organisation they are *speakers* and *addressees*. At a more obviously functional level they are *floorholders* and *audience*. The levels have congruence relations in 'normal' use, but may be used non-congruently, as when a heckler shifts from *addressee* to *speaker* to make his comment without, in general, altering his global audience status. Local discourse interruption similar to this occurs frequently in discourse loops (§ 5.3.), including echo questions and repetitions, without disturbing the global discourse role configuration.

The indexical co-ordinate system is multidimensional. The same, or comparable individuals at some level of analysis can enter into a variety of primitive indexical relations, as when a person may participate or be part of the setting or (as indirect speech shows) even part of a complex channel (the messenger function). Speech itself may be the vehicle of communication, part of the setting, or both, if a metalanguage is being used. Cross-classification of individuals with respect to indexical relations is thus an essential property of the system, which thereby allows more flexible and explicit context descriptions than lists of indexical co-ordinates (Lewis, 1972: 173 ff.) or lists of linguistically contextual 'features' (Rauh, 1978: 178 ff.), and appears as a complex system of 'context properties' analogous to those of Cresswell (1973: 114 ff.).

In procedural terms, the system may also be interpreted quite easily in a quasi-cybernetic framework defining persons interacting with their environments (including other persons) via a channel. Quasi-cybernetic models have frequently been suggested in linguistics and psycholinguistics, from Hockett (1955) through Miller, Galanter & Pribram (1960) and many others to the present day (cf. Tracy, this volume; also Gibbon, 1981b, for a detailed linguistic application). A system of this kind is of particular relevance for the explanation of discourse structures such as loops (§ 5.3.).

In a realist ontology, each participant can be thought to possess sets of linguistic and contextual schemata analogous to the indexical co-ordinate system. Against such a background, the role of deictic terms is to provide orientation points for the spatial matching and temporal synchronisation of speaker and addressee schema pairs. This applies in particular to the use of deictic terms in ostensive definitions and other demonstrative speech acts where, in psychological terms, synchronised locution-accent-gesture combination have a 'metaprocedural' attention-drawing function within discourse (Tracy, this volume; cf. also § 5.3. below).

Lyons defines deixis as follows (1977: 637):

By deixis is meant the location and identification of persons, objects, events, processes and activities being talked about, or referred to, in relation to the spatiotemporal context created and sustained by the act of utterance and the participation in it, typically, of a single speaker and at least one addressee.

The indexical co-ordinate system satisfies the requirements of this definition. It should be noted, however, that intonation (like other deictic terms) simply provides orientation points for co-ordinate matching and synchronisation which enable objects, events, states to be 'placed' and 'timed', relative to the known positions or times of other individuals such as the participants. Contrary to Lyons' suggestion, it does not appear to be the job of deictic terms to identify, re-identify (or — a prerequisite for identification? — describe) individuals (Rauh, 1978: 30). This is fundamental for the metalocutionary functions of intonation (§ 4.2).

There is a further level of deictic functioning of intonation in discourse which is often neglected: the vocative, initiatory, perhaps 'pre-deictic' procedures of constituting a contact or *channel* (cf. Fig. 1) in order to build up a spatio-temporally realised primitive indexical structure, as in calls (§ 5.4.). This level is shared by locutionary systems and may perhaps be considered one of the most primitive kinds of deictic functioning.

4. Contextual semantics for intonation

Three main deictic areas were noted at the end of § 3, in which intonation is involved:

- (8) i. provision of spatiotemporal orientation points for spatial matching and temporal synchronisation of the indexical co-ordinate systems of speaker and addressee (simple deictic function);

- ii. 'metaprocedural' attention-drawing to some of these orientation points (demonstrative function);
- iii. 'pre-deictic' constitution of indexical co-ordinate systems or their repair (the 'vocative' function).

All three functions are metalocutionary in a sense to be made more explicit below.

4.1. 'Fregean' interpretation and the metalocutionary hypothesis

The present approach describes meaning assignment to intonations constitutively, with the meanings of complex forms as functions of the meanings of the constituent forms. This 'Fregean' approach (Cresswell, 1973: 75) contrasts with a lexical or idiom-based approach to the assignment of meanings to intonations, as when holistic judgments such as 'ironic', 'reserved', 'eager' are used to construct a kind of intonation lexicon. This is not to say that there are not distinctly idiomatic aspects of intonation meaning; this is intuitively evident. But these unanalysed, paradigmatic meaning lists are as derivative here as idiom meanings would be in locutionary semantics. 'Ironic' intonation, for instance, would not stand up to serious definitions of irony: the role of wider than expected pitch bandwidth or slower rhythm and lengthening of accented syllables is simply to provide a prosodic pointer to the locus of an ironical incongruence between a locution and its context. The same device can index hyperbole or other-rhetorical devices under a different locutionary semantic interpretation relative to a different context.

The idea that the constituents of intonations index and mirror the constituents of locutions in a hierarchical fashion consonant with Frege's Principle constitutes the *metalocutionary hypothesis*. This hypothesis is foreshadowed in Praguean views of intonation semiotics, clearly formulated by Jakobson (1939 (1962: 288)):

Der Satz ist eine Sinneseinheit, die dem Wort übergeordnet ist, und jedes lautliche Mittel, welches ihre Abgrenzung, Einteilung oder die Hierarchie ihrer Bestandteile anzeigt, ist gleichfalls ein autonomes Zeichen.

The hypothesis may be understood syntactically or semantically. Since intonation forms are signs, like locutionary forms, a strict semiotic definition would see intonation-locution relations as part of syntax, i.e. as relations between signs. Sequences of pitch accents under an overall pitch contour are then analogous to standard treatments of punctuation in written language — the points, brackets, indentations, paging, type faces, numbering which index structures in writing. Accents 'punctuate' the temporal

vector of discourse in a quite literal sense and thereby provide synchronisation cues for participants. Like written punctuation, intonation forms may therefore be seen simply as 'auxiliary symbols' in an overall syntax of speech. But neither a simple pre-Tarskian semiotic definition of syntax nor the mere naming of a sign-class as 'auxiliary symbols' constitutes an explanation of the specific syncategorematic roles of intonation. It is with the aim of focussing on these specific roles that the alternative semantic metalocutionary hypothesis is presented.

One reason why the syntactic explanation is unsatisfactory lies in the complexity of speech patterning and its underdetermination by phonetic cues. Each level of complex structuring in speech competes with others for the relatively meagre means of intonational structure indexing which are available; the problem may be illustrated as in Figure 2. In view of this one-many relation, one component of an intonation interpretation function has to be a selection function which switches between levels; the various levels are, metaphorically speaking, 'possible worlds' for the intonational 'language'. No attempt will be made to pursue this metaphor in terms of a formal explication.

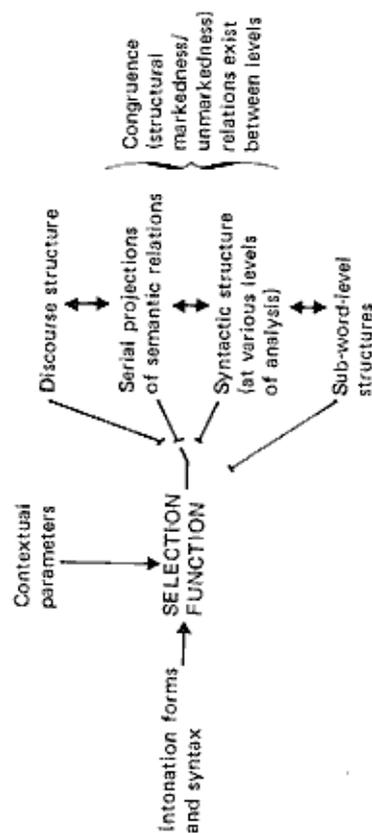


Figure 2: Outline of intonation meaning assignment with a selection function for possible metalocutionary deictic domains

The level selection function defines prosodic speech styles: informal speech concentrates on discourse structure (i.e. internal speech act structure, cohesion in speech act sequences, turn-taking patterns and cohesion between turns), while in more complex monologue-like contributions to dialogue, serial projections of semantic relations (§ 5.1.) increase in importance. Syn-

tactic structures are often more clearly marked in formal styles of presentation such as reading aloud, where the written text provides the only source of cues for metalocutionary indexing. At the other extreme, in stereotyped and ritualised speech there may be little marking which is relevant to locutionary meanings, but increased rhythmic effects and other kinds of prosodic stylisation (§ 5.4.), which are related to word-internal and low-level phrase structure, and have aesthetic, ludic or other phatic functions.

The level selection component of intonation meaning assignment is 'super-valuative' in that it uses top-down context-determined disambiguation to resolve intonation meanings; a Fregean purist would perhaps not accept this as strictly Fregean.

Using these principles of describing intonational meaning, the main aspects of deictic metalocutionary uses of intonation will be characterised in § 4.4., prior to a description of a selection of specific location-indexing processes in § 5.

4.2. Three metalocutionary deictic functions

Apart from the level selection function, the intonation interpretation function contains a strictly semantic component consisting, in turn, of three specific functions: position, relative prominence, and integration or closure, which I shall label A, B1 and B2, respectively. The first applies to single accents, and the second and third are relational, applying to accent sequence (minimally — and perhaps most frequently — to pairs).

Function A, the position function, assigns temporal positions to accents and constituents of locutions relative to some context. In this function, pitch accents are often said to 'focus' these constituents. This they do ambiguously (cf. Chomsky, 1971), since (as in gestural ostension used with or without locutionary deictic terms) a given position can be the locus of the head of a whole hierarchy of denotata, as in the following, where α denotes the focussed constituent:

$$(9) \dots \text{I saw a man in a red shirt} \\ \left(\begin{array}{c} \underline{\alpha} \\ \underline{\alpha} \\ \underline{\alpha} \end{array} \right) \\ \left(\dots \left(\begin{array}{c} \underline{\alpha} \\ \underline{\alpha} \\ \underline{\alpha} \end{array} \right) \right)$$

Note that there is no fixed relation between accents and specific locutionary categories. This relation is mediated by the level selection function (§ 4.1.) and varies with context. Like other deictic terms, the accent simply 'places' and 'times' items of particular relevance without actually identifying these

items or motivating their relevance (§ 3). The task of describing these denotata belongs to other levels of analysis, just as the task of describing the denotata of locutions is extrinsic to locutionary description.

Function B1, relative prominence, assigns a sequence of temporal points to a series of pitch accents and a series of sister constituents at some locutionary level determined by the level selection function, whereby the main accent of the pitch accent sequence is paired with the head of a dependency relation (cf. Esser, 1975) and the others are synchronised with the dependent sisters. In the unmarked case, the main accent is the rightmost, i.e. the terminal accent in the sense of the right-linear syntax outlined in § 2, and (in English) the head of the dependency relation is the rightmost operand of some operator such as negation, affirmation, or, more generally, any non-specifier in the sense of the X' theory of phrase structure (cf. Jackendoff, 1977). The factors involved in characterising the denotata of main, or 'nuclear', accents have been discussed in many places during the past ten years (the *PLANS to leave vs. plans to LEAVE* controversy — cf. Fuchs, 1980, and § 5.3. below).

In addition to being deictic, and therefore indexical in a narrow sense, this function is also iconic by virtue of its formal similarity to the selected locutionary structure (cf. the dual role of *so* with a gesture involving *two hands, fingers, & c.* in relational ostension: *It was 'so big*). In function B1, the prominence relations between pitch accents mirror the dependency structures in locutions. Like function A, this function does not identify or characterise the items which it indexes: it simply synchronises them with contextual co-ordinates, i.e. it 'locates' them in a temporal sense. Whether the relations concerned are subject-verb-object, modifier-head, hyponym-hyponym, given-new, question-reply is a matter for co-determination by the available locutionary structures and contextual co-ordinates via the level selection function. Whether B1 is a locally determined function, as implied by the syntactic local-determination hypothesis (Pierrehumbert, 1980) or whether it is implemented at a number of hierarchical levels, implying a structure more complex than a single-level transition network, is an empirical question. Evidence from spontaneous speech points toward the more complex structure.

The function B2, integration or closure, is constituted by the relative prominence function, B1, and long-scope pitch height and overall gestalt assignment strategies of the kinds outlined by Brazil (1975, 1977). The tendency of discourse contributions in English to start high and finish low, or to contain narrower bandwidth and rising accents in the earlier sections and wider bandwidth and falling accents in later sections, contribute to

B2. In the *dreadful allergies* example given in § 2, the appraisive interrupt on *dreadful* modifies the expected non-terminal accent, producing a cadence effect: the prominent rising-from-low emphasises that the contribution is still not closed and then enters into a conspicuous rising-falling long-scope pattern which integrates the adjective and noun into a highly prominent closing sequence.

Like the other two functions, and other deictic terms, B2 does not identify or characterise the locutionary structures which it indexes: it simply positions them in time. The constituent involved is co-determined by the locution (e.g. the lexical appraisive property ascribed to *dreadful*, possibly also to *allergies*) and context. Relevant constituents may range from individual words in a slow, compository, deliberative oral style, to a complete sprint or horse-race commentary given in a single overall 'excited' intonation contour.

5. Selected metalocutionary deictic interpretations

The compositional factors involved in intonation interpretation, in which function B2 presupposes prior application of B1, and the B functions presuppose prior application of A, underlie the discussion of specific interpretations in the present section. The perspective to be followed here, however, is the more general one outlined in § 3, in which the major distinctions are between simple deictic, metaprocedural and pre-deictic functions.

5.1. Simple metalocutionary deixis

The basic area of simple metalocutionary deixis is that which is most directly concerned with the three assignment factors A (position), B1 (relative prominence), and B2 (closure), together with the level-selection function. Previous work has centred on the contribution of sentence and word structure to unmarked intonation assignment; these points have been relatively well explored, for instance, within generative phonology (cf. Chomsky, Halle & Lukoff, 1956; Chomsky & Halle, 1968; Bresnan, 1971; Liberman, 1975 for the main developments, and the account in Jassem & Gibbon, 1980: 6 ff.). It was assumed for a time that the surface phrase structure of locutions was adequate for defining a notion of normal stress to be realised in normal contexts by normal accentuation (cf. the discussion of B1, above). Bresnan showed that lexical co-occurrence constants are involved, which she described using the cyclic transformational system then current. Others (cf. Bolinger,

1972; Allerton & Cruttenden, 1979; Ladd, 1978; Fuchs, 1980) showed that lexical and contextual semantic factors are involved. These factors, with the level selection function, amount to what the Prague School termed 'functional sentence perspective' or 'communicative dynamism' as applied to intonation (cf. esp. Daneš, 1960). The contribution of locution syntax to this complex presumably interacts intimately with processing constraints relating right-branching tendencies in locution syntax (cf. Kimball, 1973) and the right-linear character of accentual patterning, as reflected in the traditional idea that intonation nuclei occur on the rightmost lexical item.

In spontaneous speech, other locutionary levels are more obviously involved than locution syntax, in particular the topic development structure in which paradigmatic semantic relations are projected on to the serial syntagmatic vector of discourse (Gibson, 1981: 93). The main relations concerned have been detailed by van Dijk (1977: 106) as the following: general/particular, whole/part, set/subset, set/element, including/included, large/small, outside/inside, possessor/possessed, non-symmetrical relations which determine a "normal ordering of state descriptions" (ibid.). The first five are varieties of hyperonym-hyponym relation, and their serial projections are characteristic of additive development in topical structure:

- (10) 'Jack's bought a 'car --- 'Yes it's a 'Morgan

However, the identical prosodic forms may be used with hyponym-cokyponym relations rather than those listed by van Dijk to index contrastive (adversative) development of topical structure:

- (11) 'Jack's bought an 'M'G --- 'No it's a 'Morgan

Contrast, as this example shows, is a function of a locution in context, not of intonation.

"Normal ordering" constraints of a closely related kind are found in the syntagmatic projections of negation-affirmation in dialogue which constitute agreement and disagreement strategies, as more explicit versions of these examples show:

- (12) 'Jack's bought a 'car --- 'Yes he ,has it's a 'Morgan

- (13) 'Jack's bought an 'M'G --- 'No he ,hasn't it's a 'Morgan

The following example shows two disagreement stages:

- (14) . . . 'Jack hasn't bought an 'M'G --- 'Yes he ,has . . .

It will be noted that the disagreement strategy, whether negative or affirmative in locutionary syntax, is indexed in these examples as an unenclosed se-

quence by means of a rising accent; the constraint to regard these sequences as unenclosed is strongest in the case of an explicit negative, as in the case of (13) and the first half of (14); the second half of (14) might easily have been a falling accent, indicating closure. There is a simple explanation for these facts in the pragmatic character of affirmative and negative claims: a negative claim asserts nothing directly, it simply delimits the range of possible affirmative claims and thereby creates an expectation that a preceding affirmative claim had been taken from a wrongly delimited set of possible affirmative claims, and that a subsequent affirmative claim will be the correct one. In this pragmatic sense, negative claims are, so to speak, 'open at both ends', a property regularly indexed by intonation in discourse.

In connexion with the preceding examples, it may be objected that negatively marked constituents of a disagreement strategy may be indexed as closed. The preceding remarks were only concerned with unmarked or congruent relations between different levels of language; where apparent exceptions are found, this amounts to a disagreement in our metalanguage pointing to wrong delimitation of the initial claim. An adequate re-delimitation of the above claim would involve the 'competing levels' hypothesis of § 4.1.; in children's disagreement play (or childish disagreements among adults!), repetitive sequences may occur, in which the level selection function determines a lower locutionary level as the appropriate source of data for a stereotypic style:

- (15) . . . No it 'isn't --- Yes it 'is --- No it 'isn't . . .

Superordinate discourse strategies may also trigger the closure index, another instance of competing levels:

- (16) . . . 'No it 'isn't and 'please lets talk about something 'else . . .

Syntagmatic relations of the unmarked or congruent kinds discussed above underlie the unanalysed notion of 'natural response' used by Chomsky (1971), Jackendoff (1972) and others. Consideration of the syntagmatic character of the judgments involved, rather than only their consequences for sentence paradigms, shows that contrastive accent is a functional, not a phonetic fact, even though other prosodic facts such as anaphoric de-stressing may appear to prove the contrary at first glance. The expression 'contrastive intonation assignment' is therefore not to be parsed as ((contrastive intonation) assignment), but rather as (contrastive intonation assignment).

5.3. Counter-ostension: de-stressing

A special area of simple metalinguistic deixis concerns 'de-stressing', that is, the isolation of certain locutionary stretches from synchronisations with accent sequences. It affects anaphoric, repeated or 'given' stretches of some serial projection of semantic relations into a topical structure in the sense of the preceding section. Anaphora means, for present purposes, not simply the repetition of verbal forms or extensional co-reference of nouns and pronouns, but also the repetition of intensional constructs: non-restructured sections of cohesive topical structure sequences. Repetition is to be understood semantically, for instance as the mention (or pronominalisation) of a genus, some species of which, represented by a co-hyponym of some other expression in the structure, is accentually indexed:

- (17) Jack's bought a 'car --- 'Yes a 'blue car in fact
 (18) Jack's bought a 'car --- 'Yes a 'blue one in fact

Anaphora itself is neither a necessary nor a sufficient condition for de-stressing, just as (in discourse contexts) it is neither a necessary nor a sufficient condition for pronominalisation, deletion and other related facts. It is not a necessary condition, since de-stressing may also be involved in parentheses and tags; nor is it a sufficient condition, since there are counter-conditions under which the de-stressing mechanism is blocked -- a case of competing levels again. In particular, superordinate discourse strategies in which the topic constituent in binary topical structure sequences is accentually indexed provide counter-conditions:

- (19) If 'Jack , bought the cake then he can 'eat it
 (20) If 'Jack , bought the cake then 'Jack can 'eat the cake
 (21) 'Jack saw ,Sue and 'she saw 'him
 (22) 'Jack saw ,Sue but 'she didn't see 'him
 (23) 'Jack saw.Sue but she didn't see 'him

Note the independent functioning of pronominalisation and de-stressing in such discourse contexts.

A different case is the following, where an anaphoric item is the locus of a non-closure index:

- (24) 'Jack 'criticised 'Mary but she ignored ∨ him

Postnuclear accentual indexing is subject to additional discourse constraints which justify the non-closure indexing at a higher discourse level: here, the justification could be that a parallel source of information (speaker environment or previous knowledge) is available and suggests a possible com-

pletion. Such completions are often glossed verbally in the teaching handbooks in ways such as the following:

- (25) (. . . because she hates him anyway / . . . because his criticisms are always irrelevant / . . . because they always disagree)

The number of possible completions is unlimited, though their type is limited by possible projections of semantic relations in topical structures. The descriptive point to be made is a structural one: the location is indexed as incomplete and simply switched off; rhetorically, the effect is one of inuendo. Partial or dummy completions may be made, with a wide range of rhetorical effects, using interjections, tags (. . . isn't it, . . . of course, . . . you know), 'knowing looks', and the like. A subset of contextual indicators of this kind have been analysed in detail by Bublitz (1978).

Certain special cases of anaphora result in a pseudo-contrastive use of accent on adjacent non-anaphoric items, which Ladd has termed 'default accent' (cf. Ladd, 1978; Fuchs, 1980; Ronat, to appear). Default accent contexts are perhaps to be explained by the observation that they involve a reversal of van Dijk's "normal ordering" from genus to species:

- (26) Hairy Jake is a 'Bartok fan.
 Really? I thought he didn't 'listen to music.

In isolation, this accent placement might suggest contrastive use of *listen* (vs. *play, compose, & c.*). This is not the case in this context; *Bartok-music* are in hyponym-hyperonym (species-genus) order and therefore anaphoric on deductive grounds: the genus covers the previously mentioned species. In Ladd's account the accent therefore defaults to the preceding non-anaphoric lexical item. Other kinds of 'hidden' anaphora are also possible, the most well-known kinds involving appraisive substitutes:

- (27) Theodore disliked Eddie because Eddie had dared to criticise the fool.

Again, the rhetorical effect of this de-stressing is inuendo or irony. The details of the semantic interpretation of de-stressed locations are the subject of much ongoing research.

5.3. Metaprocedural deixis: appraisal, parentheses, discourse loops

Intonation, particularly in its accentual aspects, may be used to index the various explicit or implicit attention-drawing control structures used in discourse. Attention-drawing is one of the fundamental aspects of a pragmatic interpretation of any accentual pattern, but there are a number

of structures which involve a conspicuous interruption of the regular level-synchronisation processes for special purposes. Three distinct types of *ad hoc* 'interrupt' of this kind will be briefly outlined here.

The first, appraisal indexing, has already been noted in connexion with the *-dreadful* *allegies* example of § 2. This kind of accentuation is traditionally known as emphatic, rather than contrastive, and it appears to be a scalar factor, with prominence varying from low to high relative to the whole of the rest of the discourse constituent, not in relation to sister constituents as with the neutral accentuation strategy. Unlike contrastive accent (i.e. contrastive accentual functions), appraisal indexing is frequently coupled with phonetic modifications such as syllable termination delay (lengthening) and accent leading flank delay which results in a conspicuous rise-fall accentuation:

(28) It was a 'lovely party

The scalar, vague or fuzzy, 'more-or-less' rather than 'all-or-none', character of these modifications has been discussed by Bolinger (1961) and, in lexical areas, by representatives of generative semantics. Where the scope of the appraisal function is a whole phrase or sentence, the features associated with *ad hoc* appraisive accentual enhancement may combine in sequence, resulting in intonational *glissando* (Crystal, 1969: 164) another 'more-or-less', scalar feature which may vary from quite small pitch movements within a rhythmically marked but otherwise unexceptional accentual sequence to a highly conspicuous sequence of accentual pitch movements which might, on more atomistic criteria, be interpreted as independent 'nuclei'. Note the recorded data in Crystal & Davy (1975), for instance Extract 1, 11. 125 ff., or Extract 4, about pigs, 1. 58 (with the locution *they were horrible filthy storing things weren't they*) and 1.71:

(28) these ^dirty shuffling 'monsters in 'acres of 'mud

The parenthesis-indexing function of intonation, which may involve a slight increase in tempo, lowering of overall pitch, and narrowing of the bandwidth of accents, constitutes attention-drawing of a very different kind: to background information, to explicit statements of the speaker's opinion, involving *verba dicendi* and *verba cognoscendi*, non-restrictive relative clauses, appositions, tags, and the various 'adjectives' (Kingdon, 1958). The processing of parenthetical interrupts appears to differ from regular syntactic processing within locutions themselves, as the differences between relative and non-relative clause syntax, and the oddities of complex sentences with verbs of speaking and knowing in English; with the latter, both parenthetical and quasi-catenative uses (*Who did you say John claimed Mary thought Hal invited?*) have a distinctly stereotypic character.

The third kind of control structure which may be intonationally indexed is the discourse loop, of which two kinds will be contrasted here. Each of these involves a kind of ostension, one on the object-description level and the other on a metalinguistic level. They are illustrated by the following sequences.

(29) *Object ostension:*

A: I 'saw something —
 B: 'What — Ostensive loop call }
 A: 'That —
 B: 'Oh — Return from loop } Loop

(30) *Metalinguistic ('uptake') ostension:*

A: I 'saw something —
 B: 'What — Uptake loop call }
 A: I 'saw something —
 B: 'Oh — Return from loop } Loop

These structures have the character of procedural loops, since they may be iterated *ad lib* until the ostensive function has been fulfilled.

The area of metaprocedural deixis is, by its very nature, more heterogeneous than the interpretation of more orthodox structures; their relatively *ad hoc* relations to other aspects of discourse structure can hardly be said to be a unifying feature of this class of functions.

5.4. Vocatives, intonation and pre-deictic procedures

The grammatical category of Vocative is, in inflected languages, treated as a case within the noun paradigm. It is evident that on distributional grounds, its role differs from that of other cases; in fact, it has a distinctly parenthetical character, and proper names, used in this way, are subject to similar insertion constraints as other parenthetical expressions. Vocatives are used as independent discourse constituents with the deictic function of addressee denotation, and belong to a higher level of discourse processing than other nouns or noun phrases in the locutions into which they are inserted. Their parenthetical character relates them closely to the metaprocedural strategies outlined in the preceding subsection: they contribute in specific ways to the channel-sustaining control and attention-steering levels at which accentual indexing also operates. Vocatives are a special case of person deixis, and, like other deictic terms, are used in a variety of different deictically determined speech acts, here calls, farewells, and tag-like appellative inserts which mark the speaker's concentration on the addressee

and hold or recapture the addressee's attention. The functions are 'phatic', in the sense of Jakobson's (1960) modification of Malinowski's (1923) term.

Intonation contours, too, are involved in this special kind of person deixis with phatic function, as, for instance, in the functionally named 'call contours' or 'vocative intonations'. These have been analysed in considerable detail phonetically and functionally over the past forty years (cf. the critical outlines and specific analyses in Gibbon, 1976: 274 ff. and Ladd, 1978a).

Just as there are many uses of Vocatives in different kinds of discourse context, there are also many uses of specific intonations in discourse contexts (cf. the 'competing levels' hypothesis of § 4.1.); in other words, there is no one formally or functionally definable call contour or vocative intonation. In some contexts, the 'chanter' falling or rising double level contour is used; these contours are characteristically used in teleglossic ('distal') channel configurations, where speaker and addressee are at some distance; a further restrictive context in which they are used is that of stereotypic speech, whether play, ritual, or aesthetic acts (cf. Pike, 1945; Chao, 1956). These contexts are highly marked for addressee oriented or group oriented activity in which the main function is phatic, i.e. the initiation, sustention or termination of the group activity itself, rather than some other more specific communicative function.

Rising or falling-rising terminal accents are used to signal resumption of communication rather than completely fresh initiation:

(31) ✓ Bill take a look at `this

This turns out to be a particular case of the B2 function of marking open and closed stretches in the development of discourse. In tag position, the intonation of Vocatives is determined by regular discourse considerations, too:

(32) I'll `tell you something John

(33) You `know something, John . . .

In the first member of this pair, the indirect metaprocudural imperative not to interrupt the current discourse role configuration is intonationally indexed (the 'de-stressed' proper name), while in the second, the Vocative bears a rise-tag index which simply suggests non-closure, i.e. that a continuation is to follow.

The special metaprocudural of these uses of intonation patterns is reflected in the existence of whistled speech surrogates which mirror them as acoustic icons: there exist falling double level whistled signals used to call attention

at a distance, (rising-)falling-rising signals used for a similar purpose but at closer range, and 'interjective' whistled signals used as icons of the appraisal-indexing rising-falling accentuation discussed in the preceding subsection (the so-called 'wolf whistle'). These whistled surrogates for specific brief discourse contributions apparently vary from one language (or 'speech area' - 'Sprechbund' where languages which are unrelated or only distantly related by locutionary criteria share highly similar prosodic and paralinguistic structures and strategies) to another as the corresponding pitch patterns vary.

These special uses of Vocatives and of intonation are pre-deictic rather than deictic in the conventional sense; their pre-deictic character can be clarified by pointing out that they are concerned with those highly specialised discourse strategies which serve to establish the 'preparatory conditions' for more conventional speech acts, in the sense of Searle (1969): they help to establish 'normal input and output conditions', to sustain these, and even to create explicit terminations of discourse transactions when input and output to a previously established channel are no longer required.

6. A temporary conclusion

The general picture which has emerged in this essay is that the concept of deixis provides a meeting point for intonational and locutionary semantics. Although the similarities between locutionary and intonational deictic 'terms' should not be exaggerated, they are identifiable without much difficulty, and promise a more unified treatment of the meaning of intonation and locutionary meanings than has previously been possible. In particular, pitch accents have been shown to have various ostensive functions, both simple and relational, with respect to the temporal vector of discourse, and even a specialised form of person deixis is shared by intonational and locutionary forms. The indexical 'language' of intonation is far simpler than the 'language' of the locutionary levels, and its semantic domain is more restricted; it is a metalanguage in a quasi-Tarskian sense, which deals not in truth values but in segments of temporal vectors, using intonational gestures to index the temporal positions of discourse constituents in a given discourse context and to synchronise the development of discourse schemata by the participants in discourse.

The simplicity of the intonational 'language' and the complexity of the locutionary 'object language' which it indexes, made it necessary to postulate a system of discourse control at which a level selection function operates; the level selection function is, from a purely descriptive point of view, a way of describing a kind of stylistic or functional

variation in intonation, referred to in the body of the essay as the 'competing levels' hypothesis. Process-oriented notions like this point to the possibility of developing an account of intonational and locutionary form and meaning along reasonably unified lines; they are a commonplace in phonetics and have frequently been used in functional descriptions in a highly suggestive fashion. Since all these levels, superficially so different, have to be synchronised in speech, why should they not be accessible along structurally similar paths and chartable by similar concepts, models and theories?

Postscript

I wish to thank Rosemarie Tracy for the constant criticism and encouragement which made it possible for this paper to see the light of day, and to the editor of this volume for resigning herself to what turned out not to be inevitable after all.

References

- Allerton, D.J. and Cruttenden, A. (1979), Three reasons for accenting a definite subject. *Journal of Linguistics* 15, 49-53.
- Bolinger, D.L. (1961), Generality, Gradience and the All-or-None. The Hague, Mouton.
- Bolinger, D.L. (1972), Accent is predictable (if you're a mind-reader). *Language* 48, 633-644.
- Brazil, D. (1975), *Discourse Intonation*. Discourse Analysis Monographs No. 1. English Language Research, Birmingham University.
- Brazil, D. (1977), *Discourse Intonation II*. Discourse Analysis Monographs No. 2. English Language Research, University of Birmingham.
- Bresnan, J. (1971), Sentence stress and syntactic transformations. *Language* 47, 257-281.
- Brettschneider, G. and Lehmann, Chr., eds. (1980), *Wege zur Universalienforschung*. Tübingen, Narr.
- Bublitz, W. (1978), *Ausdrucksweisen der Sprechereinstellung im Deutschen und Englischen*. Tübingen, Niemeyer.
- 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. (1956), 52-59.
- Chomsky, N. (1971), Deep structure, surface structure and semantic interpretation, in Steinberg, D.D. & Jakobovits, L., eds. (1971), 183-216.
- Chomsky, N. and Halle, M. (1968), *The Sound Pattern of English*. New York, Harper & Row.

- Chomsky, N., Halle, M. and Lukoff, F. (1956), On accent and juncture in English, in Halle, M., Lunt, H. and McLean, H., eds. (1956), 65-80.
- Cresswell, M. (1973), *Logics and Languages*. London, Methuen.
- Crystal, D. (1969), *Prosodic Systems and Intonation in English*. Cambridge, C.U.P.
- Crystal, D. and Davy, D. (1975), *Advanced Conversational English*. London, Longman.
- Danes, F. (1960), Sentence intonation from a functional point of view. *Word* 16, 34-64.
- Davidson, D. and Harman, G., eds. (1972), *Semantics of Natural Languages*. Dordrecht, Reidel.
- Dijk, T.A. van (1977), *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. London, Longman.
- Esser, J. (1975), *Intonationszeichen im Englischen*. Tübingen, Niemeyer.
- Fox, A. (1973), Tone sequences in English. *Archivum Linguisticum* 4 (new series), 17-26.
- Fuchs, A. (1980), Accented subjects in 'all-new' utterances, in Brettschneider, G. and Lehmann, Chr., eds. (1980), 449-461.
- Gibbon, D. (1976), *Perspectives of Intonation Analysis*. Bern, Lang.
- Gibbon, D. (1981a), A new look at intonation syntax and semantics, in James, A. and Westney, P., eds. (1981), 71-98.
- Gibbon, D. (1981b), Idiomaticity and functional variation: A case study of international amateur radio talk. *Language in Society* 10, 21-42.
- Gibbon, D. and Richter, H. eds. (to appear), *Pattern, Process and Function in Discourse Phonology*. Berlin, De Gruyter.
- Halle, M., Lunt, H. and McLean, H., eds. (1956), *For Roman Jakobson*. The Hague, Mouton.
- Hart, J. and Cohen, A. (1973), Intonation by rule - a perceptual quest. *Journal of Phonetics* 1, 309-327.
- Hockett, C.F. (1955), *A Manual of Phonology*. International Journal of American Linguistics Memoir 11, Baltimore.
- Jackendoff, R.S. (1972), *Semantic Interpretation in Generative Grammar*. Cambridge (Mass.), M.I.T. Press.
- Jackendoff, R.S. (1977), *X' Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monographs, No. 2. Cambridge (Mass.), M.I.T. Press.
- Jakobson, R. (1939), Zur Struktur des Phonems, in Jakobson, R., *Selected Writings*, Volume 1. The Hague, Mouton, 280-310.
- Jakobson, R. (1960), Closing statement: Linguistics and poetics, in Sebeok, T., ed. (1960), 350-377.
- James, A. and Westney, P. (eds.), *New Linguistic Impulses in Foreign Language Teaching*. Tübingen, Narr.
- Jassem, W. and Gibbon, D. (1980), Re-defining English stress. *Journal of the International Phonetic Association* 10, 2-16.
- Kimball, J. (1973), Seven principles of surface structure parsing in natural languages. *Cognition* 2, 15-47.
- Kingdon, R. (1958), *The Groundwork of English Intonation*. London, Longmans.
- Ladd, D.L. (1978b), *The Structure of Intonational Meaning*. Diss., Cornell University.

- Ladd, D.R. (1978a), Stylized intonation. *Language* 54, 517-540.
- Lehiste, I. (1970), *Suprasegmentals*. Cambridge (Mass.), M.I.T. Press.
- Lewis, D. (1972), General semantics, in Davidson, D. and Harman, G., eds. (1972), 169-218.
- Lewis, J.W. (1977), *People Speaking. Phonetic Readings in Current English*. London, O.U.P.
- Lieberman, M.Y. (1975), *The Intonational System of English*. Diss., M.I.T.
- Lyons, J. (1977), *Semantics, Volume 2*. Cambridge, C.U.P.
- Malinowski, B. (1923), The problem of meaning in primitive languages, in Ogden, C.K. and Richards, I.A. (1923), Supplement 1, 296-336.
- Miller, G., Galanter, E. & Pribram, K. (1960), *Plans and the Structure of Behavior*. New York, Holt.
- Ogden, C.K. and Richards, I.A. (1923), *The Meaning of Meaning. A Study of the Influence of Language upon Thought and of the Science of Symbolism*. London, Routledge & Kegan Paul.
- 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, University of Michigan Press.
- Rauh, G. (1978), Linguistische Beschreibung deiktischer Komplexität in narrativen Texten. Tübingen, Narr.
- Reich, P.A. (1969), The finiteness of natural language. *Language* 45, 831-843.
- Ronat, M. (to appear), Logical form and prosodic islands, in Gibbon, D. & Richter, H., eds. (to appear).
- Searle, J.R. (1969), *Speech Acts. An Essay in the Philosophy of Language*. Cambridge, C.U.P.
- Sebeok, T., ed. (1960), *Style in Language*. Cambridge (Mass.), M.I.T. Press.
- Steinberg, D.D. and Jakobovits, L., eds. (1971), *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge, C.U.P.
- Tracy, R. (this volume), Cognitive processes and the acquisition of deixis.