The Utterance, and Other Basic Units for Second Language Discourse Analysis

GRAHAM CROOKES

University of Hawaii—Manoa

The selection of a base unit is an important decision in the process of discourse analysis. A number of different units form the bases of discourse analysis systems designed for dealing with structural characteristics of second language discourse. This paper reviews the more prominent of such units and provides arguments in favour of the selection of one in particular—the utterance.

INTRODUCTION

Discourse analysis has been an accepted part of the methodology of second language (SL) research for some time (see, for example, Chaudron 1988: 40–5; Hatch 1978; Hatch and Long 1980; Larsen-Freeman 1980), and its procedures are becoming increasingly familiar and even to some extent standardized. As is well-known, an important preliminary stage in the discourse analysis of speech is the identification of units relevant to the investigation within the body of text to be analysed, or the complete separation of the corpus into those basic units. In carrying out this stage of the analysis, the varying needs of second language discourse analysis have caused investigators to make use of all of the traditional grammatical units of analysis (morpheme, word, clause, etc.), as well as other structural or interactional features of the discourse (for example, turns), in addition to a variety of other units defined in terms of their functions (for example, moves). These units have formed the bases of a number of different analytic systems, which can mainly be classified as either structural or functional (Chaudron 1988), developed to address differing research objectives.

In carrying out structural discourse analyses of oral text, researchers have been confronted with the fact that most grammars are based on a unit that is not defined for speech, but is based on the written mode of language—that is, the sentence. Various different units have been applied to replace the sentence, but there has been little comparative discussion concerning basic units of analysis considered as options from which the discourse analyst must select, or concerning desirable criteria for selection. In the following survey, the most common of these items will be defined, and their origins and function discussed: the T-unit (and related variants), the turn, the utterance, and (because of its prominence in first language analysis), the tone unit. Then, the relationships which exist between them will be described as far as possible. The final section of the paper discusses these units from the viewpoint of basic criteria for the selection and evaluation of discourse analysis systems, and argues for the preferability of the utterance on these grounds.

Applied Linguistics, Vol. 11, No. 2 © Oxford University Press 1990

BASIC UNITS OF ANALYSIS

T-unit and variations

The *T-unit* is 'one main clause plus whatever subordinate clauses happen to be attached or embedded within it' (Hunt 1966: 735). Thus *Mary hit John* is one T-unit, and *Mary hit John, but she is my best friend* is two T-units. Originally designed for the assessment of syntactic development in children's first language (L1) writing, measures based on the T-unit have been applied to the development of SL learners' written English (Larsen-Freeman 1978; Larsen-Freeman and Strom 1977; Scott and Tucker 1974) and to the analysis of oral second language material, in English (Larsen-Freeman 1983) and Japanese (Harrington 1986).

The *c-unit* (communication unit: Loban 1966) is closely related to the T-unit, but has the advantage that isolated phrases not accompanied by a verb, but which have a communicative value, can be coded. Such phrases typically appear in answer to a question:

Q: Where's my hat?
A: On the table.

In this case, the answer is not a T-unit, and could not appear in an analysis using T-units, but it is a c-unit. Chaudron (1988: 78) observes that these two units are in practice roughly equivalent, and Kroll (1977: 85) regards them as identical.

The *idea unit* (Kroll 1977) is less well-known than the T-unit (for analyses using it see, for example, Chafe 1980; Danielewicz 1984). As defined by Kroll, it is

> a chunk of information which is viewed by the speaker/writer cohesively as it is given a surface form . . . related . . . to psychological reality for the encoder. (Kroll 1977: 85)

Kroll developed the unit because she felt that existing structurally relevant measures (T-unit and c-unit) were too specific to the written modality, were grammatical rather than psychological units, and failed to reflect the communicative nature of the corpus she was working with (oral monologues and written versions of the same). Her precise definition is somewhat lengthy:

> (1) a subject and verb counted as one idea unit together with (when present) a (a) direct object, (b) prepositional phrase, (c) adverbial element, or (d) mark of subordination
> (2) full relative clauses counted as one idea unit when the relative pronoun was present
> (3) phrases which occurred in sentence initial position followed by a comma or which were set off from the sentence with commas were counted as separate idea units
> (4) verbs whose structure requires or allows a verbal element as object were counted with both verbal elements as one idea unit
> (5) reduced clauses in which a subordinator was followed by a non-finite verb element were counted as one idea unit
> (6) post-nominal *-ing* phrases used as modifiers counted as one idea unit

> (7) other types of elements counted as idea units were (a) absolutes, (b) appositives, and (c) verbals. (Kroll 1977: 90)

An example of the segmentation of written text into idea units is as follows (Kroll 1977: 91):

> Sue roared all the harder./ She claimed I looked funny,/ clinging there,/ screaming. [4 idea units]

Turn

A widely used discourse analysis unit is the *turn*, which is commonly defined as one or more streams of speech bounded by speech of another, usually an interlocutor. Thus, in the following (created) example, A has two turns, B one:

A: Are you going home?
B: Sure, I'll be leaving in ten minutes.
A: Great.

Although the definition just mentioned is standard, it has unfortunately also been used to refer to the utterance (Crystal 1969: 277; Sinclair and Coulthard 1975). The potential for confusion makes it particularly important to be as clear as possible what is meant when the term is used (cf. van Lier 1988: 100–3). As a phenomenon in discourse, the turn was the center of attention in influential early work on conversation analysis, a specialized area of work on discourse (see, for example, Sacks, Schegloff, and Jefferson 1974), from where it has spread to mainstream discourse analysis (cf. Levinson 1983; van Lier 1988).

Tone unit/tone group

Although the *tone unit* (sometimes called tone group) has rarely, if ever, been used for the investigation of SL discourse, it cannot be left out of the present discussion. SL researchers have regularly borrowed and adapted analytical tools developed by L1 investigators, and the tone unit is important in the analysis of spoken English as a first language. It has been utilized in the most prominent line of grammatically-oriented research into spoken English, in the work associated with Quirk, Crystal, and their associates (particularly studies involving the Lund corpus—for example, Quirk *et al.* 1985), in the related work on language sampling and assessment (Crystal, Fletcher, and Garman 1976), and in research on other varieties of English (Oreström 1985).

Crystal and Quirk refer to the tone unit as 'the most striking prosodic unit in English speech'. They continue:

> For us, the tone-unit is a stretch of speech . . . in which there is a climax of pitch prominence which takes the form of 'nuclear' pitch movement of—in the case of level tones—pitch sustention . . . the nucleus is generally realised on a single syllable, though the pitch movement or sustention may be continued on one or more further syllables which constitute the 'tail' of the tone-unit . . . Little more need be said here . . . because the basic typology does not differ for the most part from that described in considerable

> detail in the standard works that have treated English intonation in terms of contours rather than phonemic levels. (Crystal and Quirk 1964: 50–1)

In addition to pitch-related means of identification, Crystal (1969: 206) identifies a second major characteristic as

> the presence of junctural features at the end of every tone-unit. This usually takes the form of a very slight pause, but there are frequently accompanying segmental phonetic modifications (variations in length, aspiration, etc.) which reinforce this.

An example of a stream of speech separated into tone units is the following:

> but I personally / have never found / a boffin or a statistician / who worked things out in theory which ever came out in practice. (after Crystal 1969: 260)

A parity between this unit and the apparently similar 'tone unit' of Halliday (1967) and the 'intonation contour' of Pike (1945) is assumed by some researchers (for example, Brown, Currie, and Kenworthy 1980: 40; see also Crystal and Quirk 1964: 51, fn. 7). Although these may have originally been very close, Coulthard (1985: 100ff.), having identified the antecedents of the tone unit in the work of Halliday (e.g. 1963), states that recent developments (see Brazil 1985; Brazil, Coulthard , and Johns 1980) have weakened the tone unit's original close connection to the grammatical clause (as previously existed in the Hallidayan model). Since Brazil sees intonation, and the selection of different sub-units within the basic intonation contour, as a way a speaker has to provide meaning additional to that conveyed lexically, Brazil has emphasized the functional (as opposed to structural) aspect of the tone unit. This has enabled him to fit it more closely with the prominent functional discourse analysis system of Sinclair and Coulthard (1975). A similar change is to be found in the work of Kreckel (1981), who has adopted a purely semantic interpretation of the tone unit as a unit of speech containing or equivalent to one 'message' (though confusingly, she elsewhere refers to it as 'define[d] as a continuously spoken clause' (Kreckel 1982: 56)).

Utterance

Use of the *utterance*, (defined intonationally) as a unit into which the stream of speech could be separated, probably antedates work based on the other units discussed here. Cooper and Sorensen (1981) state that

> [l]inguists traditionally relied on their perception of F_0 [fundamental frequency] fall-rise patterns in conjunction with the perception of pauses to demarcate the syntactic structure of utterances. (Cooper and Sorensen 1981: 4)

and Lea observes that

> for decades, linguists have claimed that intonation indicates the immediate constituent structure of English sentences (Jones 1909, 1932; Pike 1945; Wells 1947). (Lea 1973: 18)

The definition of the utterance implied here was utilized (perhaps somewhat loosely) in early L1 acquisition research. Scollon (1974: 206, 1976: 152) cites

use of 'terminal intonation contour' to segment speech, by Brown and Bellugi-Klima (1964), and he used it in his own dissertation work (an L1 acquisition case study). Scollon's remarks concerning the utterance bear detailed consideration, principally because of their influence on L2 work, but also because of their tentative and exploratory nature.[1] Scollon states (1974: 219) that his criteria for identifying (and defining) utterances were developed and applied as a way of making explicit the intuitions he initially relied on in the process of analysis, but he makes it clear that from the start of his investigation, he responded to both intonation contours and pauses in segmenting the speech of his subject. On classifying the intonation patterns of his subject's speech, he found that more than half of them corresponded to Lieberman's 'unmarked breath group' (see below). His initial criterion for identification of utterances was 'bounded by pauses' (1974: 41), so the discovery that most intonation patterns corresponded to unmarked breath groups emerged from his interpretation and analysis of the recorded speech, rather than from an initial criterion for segmentation. To analyse the recordings he had made of later stages in his child subject's speech, he revised his initial working definition of utterance (1974: 239), making more definite the two criteria mentioned so far, and including two aspects of a semantic criterion. The semantic criterion was added because Scollon sought a way of determining whether two-morpheme productions were holophrases or separate utterances. This he obtained through stating the criterion that the morphemes be 'independent' (i.e. capable of being used productively and separately by his subject). 'Independence' in this sense can be taken as implying separate semantic units. In addition, he refers (1974: 220) to an 'appropriate semantic connection' between elements of speech, as a final factor which he used for deciding on utterance boundaries.

Some SL researchers (for example, Long 1980; Sato 1985) derived an operational definition of the utterance from Scollon (1976: 153) and other L1 acquisition research (Ochs and Schieffelin 1979). Sato (1985: 83–4) defines it as 'a stream of speech under a single intonation contour bounded by pauses'. A subsequent modification was that of Crookes and Rulon (1985: 9):

> an utterance [is] defined as a stream of speech with at least one of the following characteristics:
> (1) under one intonation contour,
> (2) bounded by pauses, and
> (3) constituting a single semantic unit.

This did not mention how long a pause time would be indicative of an utterance boundary, however. For Scollon (1974: 206), a pause of less than 0.6 seconds did not indicate an utterance boundary, one between 1 and 8 seconds did. In general practice, an appropriate time duration which may be taken as indicating an utterance boundary will depend on the nature of the corpus being analysed. Chaudron (1988: 45) provides a modified definition, incorporating (variable) pause time explicitly.

With reference to the third part of the definition mentioned above, additional

support for using a semantic criterion in identifying utterances may be found in the work of Brown, Currie, and Kenworthy (1980) (assuming that speech unified by topic is equivalent to semantically-related speech). Brown *et al.* found systematic co-occurrences between intonational patterns and topic shifts. Because of theoretical disagreements concerning topic identification (they cite Grimes 1975; Li 1976; Van Dijk 1977; and see also Brown and Yule 1983; Keenan and Schieffelin 1976; Van Oosten 1984), their position is hedged, but they find that 'the strongest indication that a speaker is changing direction within the overall topic is generally intonation' (Brown *et al.* 1980: 26).

It is noteworthy that clinical linguistics, an area of research usually quite separate from first or second language acquisition studies, also makes use of the utterance as a base unit, defined in a surprisingly similar fashion to that mentioned above. In the course of an exposition of the development of a speech sampling system, Shewan (1988) defines an utterance as:

> a complete thought, usually expressed in a connected grouping of words, which is separated from other utterances on the basis of content, intonation contour, and/or pausing. (i) *Content*. A change in content is used as one criterion for segmenting utterances ... (ii) *Intonation Contour*. A falling intonation contour signals the end of an utterance. A rising intonation signals the end of an utterance if it is a question ... (iii) *Pauses*. Pauses are used in conjunction with the above two criteria to segment utterances. (Shewan 1988: 124)

Despite the fact that Shewan reviews an extensive body of literature in her paper, none of it overlaps with any of that discussed in this section—i.e. this definition of the utterance appears to have been independently developed.[2]

To summarize the definitional work on the utterance: this unit is specified by way of intonational, pausal, and semantic criteria. It has been developed from a widely used but not carefully defined unit in linguistics, through its application to child L1 acquisition work, to its use and increasingly careful specification in L2 learning research, and has independent support from its use in a similarly defined form in recent clinical linguistics research.

POSSIBLE RELATIONSHIPS BETWEEN UNITS

(1) *Turn to other units*

This is a straightforward connection. Basically, the turn is a potentially superordinate category, which can contain one or more of the other units. Sato (1985: 84) provides an example in which the first speaker's turn consists of two utterances, and the second speaker's has just one:

> Speaker A: You said you were hungry/ Are you going to eat this or not?//
> Speaker B: You know I hate raw fish/

With regard to the turn/T-unit connection, since a T-unit is defined grammatically, a turn can contain zero T-units (though it could not be 'empty' if analysed in terms of the other units discussed here). An example of SL discourse analysis which deals in both turn and utterance is Long *et al.* (1984), where turn-complexity is operationalized as number of utterances per turn. Duff (1986) uses both c-unit and turn in analysing SL discourse, and Oreström's (1983) study of L1 conversation relates turn to tone-group.

(2) *T-unit to tone unit*

The relationship between the tone group and the T-unit (as well as between tone unit and close relatives of the T-unit such as the c-unit and idea unit) is conceptually somewhat distant, insofar as the T-unit has an explicit written and grammatical formulation, whereas the tone group derives principally from the nature of oral (English) language. Some indication of the extent to which a tone unit analysis would coincide with a T-unit analysis is indicated by Crystal's remark (1969: 262) that, 'there is a high degree of coincidence of tone-unit boundary with junction between elements of clause structure' which, for his data, constituted 84 per cent for beginning-points and 79 per cent for end-points. The elements of clause structure to which he is referring are, however, such items as subject or complement. It seems likely, therefore, that a T-unit would often contain, without overlap, more than one tone unit.

(3) *T-unit to utterance*

As with the turn, the utterance is potentially a superordinate category to the T-unit: an utterance can usually be rendered as one or more T-units (though it may contain none). But it is a little strange to reduce a unit of oral language to one defined (or conceived) in terms of written language. A simple measure of utterance complexity (T-units per utterance) might thus be obtained, but perhaps an equally good one would be S-nodes per utterance. This at least avoids the extensive reduction of raw data necessitated by a T-unit analysis (see below).[3]

(4) *Tone unit to utterance*

If Halliday's (1963) position, that the unmarked tone unit is coextensive with grammatical clause, is accepted, this suggests that any relationship between these two units will only be direct in discourse where utterances are mostly uni-clausal. In fact, tone units are often shorter than utterances (as defined following, for example, Chaudron 1988), and Crystal (1969: 256) mentions an average figure of five words per tone unit. Considering the matter of a list may clarify this. First, it must be noted that a tone-unit

> will have one peak of prominence ... after this nuclear tone there will be a tone-unit boundary ... indicated by ... a perceivable pitch-change, either stepping down or stepping up. (Crystal 1969: 205)

and

> [r]elatively few tone-units are of a single syllable ... most of these being restricted to response utterances, lists, and the like. (Crystal 1969: 208)

In other words, once the nuclear tone of an utterance has been identified, a pitch-change implies a new tone unit. Thus the sequence of pitches usually found in a list of items spoken in English would give a sequence of tone units. Under the definition of utterance mentioned above, they would probably not give a sequence of utterances. Although there are certainly pitch changes in the elements of a list spoken as part of a normal conversation in English, the list as a whole normally has a single intonation contour, and so would probably be transcribed as a single utterance by a discourse analyst using the Chaudron definition.

CRITERIA FOR DISCOURSE ANALYSIS UNIT SELECTION

There are at least two stages where a discourse analysis system can fail to meet standards in the process of being applied to a body of recordings and language transcripts. The first of these is the segmentation of the transcribed stream of speech into units, the second is the process of classification of units thus identified. At both points, the system can be judged concerning its validity and reliability, the two basic criteria for any measuring instrument. A measure is reliable if it gives the same reading on the same item on different occasions; and it is valid if it measures what it is intended to measure (cf. Hammersley 1987).

(1) *Reliability*

Though reliability does not imply validity, an unreliable discourse analysis system cannot be valid (Frick and Semmel 1978: 158; Nunnally 1978: 192; cf. Chaudron forthcoming). The reliability of a discourse analysis system is usually indicated in terms of an index of interrater agreement (cf. Chaudron, Crookes, and Long 1988), and herein may lie a problem for the tone unit. Although Crystal has cited a fairly high degree of interrater reliability with regard to identification of tone unit boundaries (84.8 per cent, Crystal 1969: 203), Brown *et al.* (1980) found the system unusable for analysis of their corpus of speech. Despite all being trained linguists, Brown and colleagues were unable to identify tone unit boundaries at a satisfactory level of agreement, particularly when 'the speaker was trying to work out what he wanted to say as he was saying it' (1980: 41). Difficulties caused by this kind of speech may have been exacerbated by the variety of English of the corpus—Edinburgh Scottish English—because, according to Brown *et al.*, various aspects of the 'neutral intonation pattern' (1980: 19–20) in this variety are not as salient as in 'RP'. Instead of using the tone unit, Brown *et al.* defined a basic unit which depends on pauses and topic shifts.

However, even confining the use of the tone unit to standard (British) English may not be sufficient. Recent developments of the tone unit are oriented towards an analysis of how intonation encodes meaning or function, and so Brazil, Coulthard, and Johns (1980: 45–6) are able to make the following statement:

> Whatever description of intonation one uses to analyze recorded speech, there are inevitable difficulties in deciding where to put the tone unit boundaries in a small number of cases . . . One significant advantage of our description . . . is that it gives us a principled reason for saying that tone unit boundaries are not in fact of great importance.

This seems to be equivalent to saying that an analysis using the tone unit can be conducted even if analysts (such as Brown *et al.*) cannot agree on the position of the boundaries between units. If this is the case, it may be satisfactory for some functional analyses, but it seems likely that a structural analysis using tone units will encounter difficulties.

Figures for reliability analyses are increasingly reported in discourse analysis studies. It is noteworthy that reliability problems have not been reported for the other principal units discussed here (but see note 5, below); studies which report such data on the other units under consideration provide figures at or above conventionally acceptable levels.

(2) *Validity*

The aspect of validity which most critically applies to measuring instruments such as a discourse analysis system is 'correspondence' (Brinberg and McGrath 1982, 1985). In the present case, it is necessary that the discourse analysis system correspond to, or be appropriate for the aspect of discourse, of interest in a given investigation—a property also referred to as 'instrument validity' (Brinberg and McGrath 1982: 14). For example, with regard to the preliminary stage of discourse analysis (obtaining a physical record), even phonological transcriptions reflect varying degrees of truth, as indicated by the terms 'broad' and 'narrow transcription'. Indeed, it has been argued (Scollon 1974, 1976) that early child L1 acquisition studies were weakened by their failure to use an appropriately narrow transcription system. By using a broad system which did not correspond to the nature of the speech signal, such studies simply lost data which would otherwise have been interpretable as the investigations proceeded. (See also Ochs 1979, for further detailed discussion.)

Just as serious is the criticism levelled by Kroll (1977) at the T-unit and c-unit. Referring to O'Donnell (1974), one of the first studies to use the T-unit for spoken discourse, she points out that in order to use the T-unit for this purpose, it is necessary to ignore speaker disfluencies (or 'mazes'). In the case of the c-unit, she notes that its originator (Loban 1966) was only able to perform c-unit analysis by selecting out and separately tabulating the problems such disfluencies constituted for the analytic unit. 'Loban's data bypasses the mazes', she remarks (Kroll 1977: 86).[4]

In general, structural investigations of SL discourse are concerned with the results of the psychological processes of language production. It may be argued, therefore, that the demand for instrument validity is particularly served if the basic unit of a discourse analysis system corresponds to, or directly reflects such processes. *Ceteris paribus*, such a unit would be the preferred one for analyses of such processes and their results, such as degree of complexity of speech.

On these grounds, the turn may be eliminated from consideration. Since its boundaries are determined by the processes of speaker interaction, it does not

reflect the psychological processes of an individual's speech production alone, but is additionally influenced by the many social variables which determine the flow of multi-party discourse. (It also becomes meaningless when monologue is considered.) The T-unit, and related forms, are defined principally in syntactic terms, and their connection with speech production mechanisms must therefore also be indirect.

The tone unit has already been identified as having some possible reliability problems. Its validity, in terms of a connection to psychological speech processes, is somewhat unclear. Some early work in psycholinguistics identified a unit known as the 'phonemic clause' (Boomer 1965) as the fundamental unit of speech processing, and at that time Laver (1970) argued that it was identical with the tone unit. There has, however, been little, if any, work in discourse analysis making use of the phonemic clause, and Boomer's research has recently been heavily criticized on methodological grounds, such as selective use of data (O'Connell and Slaymaker 1984). Consequently, it is difficult to be certain concerning the validity of the tone unit even if the equation of tone unit and phonemic clause is accurate.[5] Finally, almost no work has been done to establish the applicability of this unit outside of English (unlike the utterance). Given some of the problems of other units mentioned so far, the utterance may be more satisfactory when reliability and particularly validity are considered, for the following reasons.

The pause and the intonation contour in speech have been held by a variety of authors to be indicative of the psychological processes involved in the creation of an utterance (see, for example, Petrie 1987), and it has already been observed that most discussions of the utterance use 'intonation contour' as one defining criterion. Across languages, the most common intonation contour is that traced out by the fundamental frequency (F_0) from a usually high initial setting at the beginning of the stream of speech, to a low final setting which is widely recognized as constituting the end of a speaker's stream of speech (Cooper and Sorenson 1981; Lieberman 1984). One group of authorities further describe this pattern as a gradual decrease of F_0 over time, and see it as appearing typically in declarative statements. They refer to it as *declination* (see Abramson 1962; Bolinger 1964; Chiba 1935; Cohen and t'Hart 1967; Fonagy 1971; Hadding-Koch 1961; Jones 1909; Maeda 1976; O'Shaughnessy 1976). Cooper and Sorenson (1981: 20) state that

> F_0 declination . . . appear[s] reliably in typical conversation . . . as well as in unpracticed reading of a lengthy passage of text.

Declination in declarative utterances is generally observed across languages (Bolinger 1964; Vaissiere 1983). Cooper and Sorensen (1981: 28) refer to it as 'the most salient global attribute' of declarative speech and also observe that

> declination provides one of the most fertile grounds for studying the nature of the speaker's central commands in speech production, including the representation of syntactic units as well as the manner in which such units are processed during the speaker's planning and execution of speech. (1981: 3)

Cooper and Sorenson (1981) see physiological factors as the cause of declination (citing the work of Collier 1975, and Maeda 1976, as supporting this position). They note that

> [i]n the case of declination, the gradual falling pattern of F_0 throughout the course of a main clause may be attributed to the combined weight of a number of factors, including a gradual decrease in subglottal pressure and a gradual slackening of the speech musculature, including the vocal folds. (op. cit. 1981: 160)

The coincidence of the ends of intonation contours with the ends of syntactically defined units is at least circumstantial evidence implicating the utterance (defined intonationally) with the psychological processes involved in creating syntactic units. What other reasons could there be which would cause intonation to reflect cognition?

Lieberman (1967, 1970, 1976, 1984) and Cooper and Sorensen (1981), although representing two distinct lines of research, all link intonation with both physiological and psychological processes. Most linguists would grant that many elementary properties of speech have physiological constraints at bottom, and indeed, experimental work on adults (for example, Lieberman 1967) and observation of children (Lieberman 1984) support the position that the fluent speaker co-ordinates air pressure and volume in the lungs with duration of phonation in producing coherent speech. Similarly, Cooper and Sorensen (1981) found that F_0 declines faster over shorter sentences. They found higher values for the first F_0 peak in longer sentences, suggesting

> that the speaker employs a look-ahead mechanism which computes the approximate overall upcoming utterances length. (op. cit. 1981: 38)

Cooper and Sorensen (1981), and Lieberman (for example, 1984; see also Berkovits 1984; Boyce and Menn 1979, for related studies) agree that there is a complex interaction between psychological speech production and planning processes, and the physiological mechanisms which render thought into acoustic energy. Both of these two separate lines of work support the position that an intonation contour is a physiologically conditioned but meaning-carrying indicator of the unity of propositional and syntactic form which originated it. Accepting the grammatically defined sentence as 'the minimal unit of speech for a complete semantic interpretation', Lieberman (1984) suggests that under normal conditions there is a legitimate equating of utterance and sentence. In brief, his position is that

> the articulatory maneuvers that we use to produce the intonation, or 'melody', that can group a series of words together demonstrate the psychological reality of the sentence. (Lieberman 1984: 98)

This unit he refers to as the 'breath-group' (Lieberman 1984: 118), which seems to be equivalent to the utterance as it has come to be used in SL discourse analysis.

In his most recent work in this area, Lieberman (Lieberman *et al.* 1985; cf. Liberman and Pierrehumbert 1984; Umeda 1982) disagrees with Cooper and

Sorensen concerning the actual form of the basic intonation contour, as it manifests itself in spontaneous speech. The experimental work of Lieberman *et al.* testing Maeda's (1976) position on declination appears to show that declination is less pronounced in spontaneous speech than in single sentences and read text. The data of Lieberman *et al.* suggest that although a high F_0 appears at the beginning of an utterance, and there is a marked, sharp fall at the end, the intervening section of the intonation contour may be flat, rather than declining all the way through. This is actually not a problem for the definition of the utterance being used in SL discourse analysis, of course.

Accepting that the utterance or 'unmarked breath group' is closely tied to psycholinguistic production processes does not mean it must be assumed that there is always a single, unitary internal unit, which is first computed and then output as a single piece of the physical signal (though this was an equation made in early speech production studies, for example, Goldman-Eisler 1958). Phonation occurs in parallel with speech planning, which takes place on several levels (semantic, syntactic, lexical, discoursal, etc.). Furthermore, while a clause may be the syntactic unit of an utterance (or breath group) on one occasion, a set of clauses may be covered by one breath group or utterance on another, and a single word at yet another time (cf. Dechert 1980). As Cooper and Sorensen observe, while declination

> seems to be programmed over a domain encompassing a major syntactic constituent, [a] speaker may select a domain of coding that is more superordinate than normal when the constituents contained in this domain are short in terms of their number of words or syllables, when speaking fast or when the material contained in the domain is highly cohesive on semantic and/or pragmatic grounds . . . with resetting [i.e., beginning a new intonation contour] typically accompanying only the boundary between two sentences. (Cooper and Sorensen 1981: 161)

CONCLUSION

This paper has surveyed a widely scattered body of literature concerning important discourse analytic units. In the course of it, it has been argued that the utterance has a strong claim for attention in structural discourse analysis, perhaps to the point of displacing other units as a default choice on grounds of reliabilty and validity. It has not been claimed that the utterance is the one and only production unit of speech (particularly as some have said that the search for a single such unit is futile, for example, Dechert 1980), though its links to language production processes are important. The basic position taken is that a combination of practical and theoretical considerations make it a better prospect for SL discourse analytic purposes than the other candidates, in the light of the evidence presently at hand.

It is in the nature of scientific research that progress is not linear, and it is quite common for major conceptual advances to be made from a less than solid methodological base. Eventually, however, there comes a need for some regrouping, to strengthen lines of communication. In the present case, the issues

involved are complex, and their empirical investigation, though not impossible, has proven quite difficult, so we should not expect the process of clarification to be accomplished all at once. The primary objective here has simply been to aid future SL discourse analysis by setting out some of the fundamental issues, bringing together some diverse and important lines of work, in the hope that further work in this area will show a greater degree of coherence and cumulativeness than has hitherto been the case.

(Received October 1988)

NOTES

[1] It is unusual, and instructive, to see how discourse analytic categories can be developed through careful analysis of protocols rather than being defined *a priori*.

[2] Some recent L1 acquisition work has used early research in this area. Siegel (1963) gives a detailed operational definition of 'vocal response units', which has recently been re-used as a definition of utterance by Rondal, Ghiotto, Bredart, and Bachelet (1987).

[3] The presence of an S-node is taken as indicated by a tensed (or untensed) verb (within whatever base unit of analysis is being used)—see, for example, Long (1980). S-node-based measures have been used in a number of SL studies, without, however, much explicit discussion of their foundation.

[4] Questions have also been raised concerning the validity of the T-unit and measures based on it as an index of SL development. The question as to whether it is adequate in this regard is a question subsidiary to whether it is valid in an absolute sense for the analysis of speech. Those considering a T-unit analysis as the basis for assessing development should consult Gaies (1980), Kameen (1983), Harrington (1986), Larsen-Freeman (1983), and Vann (1979).

[5] And again, if the equating is correct, O'Connell and Slaymaker (1984: 281) report that 'researchers at the University of Giessen have found it impossible to identify phonemic clauses reliably and have turned to other units'.

REFERENCES

Abramson, A. S. 1962. 'The vowels and tones of standard Thai: acoustical measurements and experiments.' *International Journal of American Linguistics* 28.

Berkovits, R. 1984. 'Duration and fundamental frequency in sentence-final intonation.' *Journal of Phonetics* 12: 255–65.

Bolinger, D. 1964. 'Intonation as a universal.' *Proceedings at Linguistics IX*. Hague: Mouton.

Boomer, D. S. 1965. 'Hestitation and grammatical encoding.' *Language and Speech* 8: 148–58.

Boyce, S. and **L. Menn.** 1979. 'Peaks vary, endpoints don't: implications for intonation theory' in C. Chiavello *et al.* (eds.): *Proceedings of the Fifth Annual Berkeley Linguistics Society*. Berkeley, Calif.: University of California at Berkeley Linguistics Society.

Brazil, D. C. 1985. *The Communicative Value of Intonation*. Birmingham, UK: English Language Research.

Brazil, D. C., M. Coulthard, and **C. Johns.** 1980. *Discourse Intonation and Language Teaching*. London: Longman.

Brinberg, D. and **J. E. McGrath.** 1982. 'A network of validity concepts within the

research process' in D. Brinberg and L. Kidder (eds.): *Forms of Validity in Research*. (*New Directions for Methodology of Social and Behavioral Science, 12*.) San Francisco, Calif.: Jossey-Bass.

Brinberg, D. and **J. E. McGrath.** 1985. *Validity and the Research Process*. Beverly Hills, Calif.: Sage.

Brown, G., K. L. Currie, and **J. Kenworthy.** 1980. *Questions of Intonation*. Baltimore, Md.: University Park Press.

Brown, G. and **G. Yule.** 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, R. and **U. Bellugi-Klima.** 1964. 'Three processes in the child's acquisition of syntax.' *Harvard Educational Review* 34: 133–51.

Chafe, W. L. 1980. 'The deployment of consciousness in the production of a narrative' in W. L. Chafe (ed.): *The Pear Stories*. Norwood, NJ: Ablex.

Chaudron, C. 1988. *Second Language Classrooms: Research on Teaching and Learning*. New York: Cambridge University Press.

Chaudron, C. (forthcoming). 'Validation in second language classroom research and the evaluation of teaching methods' in R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, and H. Swain (eds.): *Foreign Language Pedagogy: a Commemorative Volume for Claus Færch*. Clevedon: Multilingual Matters.

Chaudron, C., G. Crookes, and **M. H. Long.** 1988. *Reliability and Validity in Second Language Research.* Technical Report No. 8. Center for Second Language Classroom Research, Social Science Research Institute, University of Hawaii, Honolulu.

Chiba, T. 1935. *A Study of Accent: Research into its Nature and Scope in the Light of Experimental Phonetics*. Tokyo: Phonetic Society of Japan.

Cohen, A. and **J. t'Hart.** 1967. 'On the anatomy of intonation.' *Lingua* 19: 177–92.

Collier, R. 1975. 'Physiological correlates of intonational patterns.' *Journal of the Acoustical Society of America* 58: 247–55.

Cooper, W. E. and **J. R. Sorensen.** 1981. *Fundamental Frequency in Sentence Production*. New York: Springer-Verlag.

Coulthard, M. 1985. *An Introduction to Discourse Analysis* (2nd edn.). London: Longman.

Crookes, G. V. and **K. Rulon.** 1985. *Incorporation of Corrective Feedback in Native Speaker/Non-native Speaker Conversation*. Technical Report No. 3. Center for Second Language Classroom Research, Social Science Research Institute, University of Hawaii, Honolulu.

Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Crystal, D. and **R. Quirk.** 1964. *Systems of Prosodic and Paralinguistic Features in English*. The Hague: Mouton.

Crystal D., P. Fletcher, and **M. Garman.** 1976. *The Grammatical Analysis of Language Disability: a Procedure for Assessment and Remediation*. London: Edward Arnold.

Danielewicz, J. M. 1984. 'The interaction between text and context: a study of how adults and children use spoken and written language in four contexts' in A. D. Pellegrini and T. D. Yawkey (eds.): *The Development of Oral and Written Language in Social Contexts*. Norwood, NJ: Ablex.

Dechert, H. W. 1980. 'Pauses and intonation as indicators of verbal planning in second-language speech productions' in H. W. Dechert and M. Raupach (eds.): *Temporal Variables in Speech*. The Hague: Mouton.

Duff, P. A. 1986. 'Another look at interlanguage talk: taking task to task' in R. R. Day (ed.): *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, Mass.: Newbury House.

Fonagy, I. 1971. 'Synthèse de l'ironie.' *Phonetica* 23: 42–51.

Frick, T. and **M. I. Semmel.** 1978. 'Observer agreement and reliabilities of classroom observational measures.' *Review of Educational Research* 48: 157–84.

Gaies, S. J. 1980. 'T-unit analysis in second language research: applications, problems and limitations.' *TESOL Quarterly* 14/1: 53–60.

Goldman-Eisler, F. 1958. 'Speech production and the predictability of words in context.' *Quarterly Journal of Experimental Psychology* 10: 96–106.

Grimes, J. E. 1975. *The Thread of Discourse*. The Hague: Mouton.

Hadding-Koch, K. 1961. *Acoustico-phonetic Studies in the Intonation of Southern Swedish*. Lund: Gleerup.

Halliday, M. A. K. 1963. 'The tones of English.' *Archivum Linguisticum* 15: 1–28.

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

Hammersley, M. 1987. 'Some notes on the terms "validity" and "reliability".' *British Educational Research Journal* 13/1: 73–82.

Harrington, M. 1986. 'The T-unit as a measure of JSL oral proficiency.' *Bulletin of the ICU Summer Institute in Linguistics (Tokyo)* 19: 49–56.

Hatch, E. 1978. 'Discourse analysis and second language acquisition' in E. Hatch (ed.): *Second Language Acquisition: a Book of Readings*. Rowley, Mass.: Newbury House.

Hatch, E. and **M. H. Long.** 1980. 'Discourse analysis, what's that?' in D. Larsen-Freeman (ed.).

Hunt, K. W. 1966. 'Recent measures in syntactic development.' *Elementary English* 43: 732–39.

Jones, D. 1909. *Intonation Curves*. Berlin: Teubner.

Jones, D. 1932. *An Outline of English Phonetics* (3rd edn.). New York: Dutton.

Kameen, P. T. 1983. 'Syntactic skill and ESL writing quality' in A. Freedman, I. Pringle, and J. Yalden (eds.): *Learning to Write: First Language/Second Language*. New York: Longman.

Keenan, E. O. and **B. Schieffelin.** 1976. 'Topic as a discourse notion: a study of topic in the conversation of children and adults' in C. N. Li (ed.): *Subject and Topic*. New York: Academic Press.

Kingdon, R. 1958. *The Groundwork of English Intonation*. London: Longman.

Kreckel, M. 1981. *Communicative Acts and Shared Knowledge in Natural Discourse*. London: Academic Press.

Kreckel, M. 1982. 'Communicative acts and shared knowledge: a conceptual framework and its empirical application.' *Semiotica* 40: 45–88.

Kroll, B. 1977. 'Combining ideas in written and spoken English: a look at subordination and coordination' in E. Ochs Keenan and T. L. Bennett (eds.): *Discourse Across Time and Space. Southern California Occasional Papers in Linguistics 5*. Los Angeles, Calif.: Department of Linguistics, University of Southern California.

Larsen-Freeman, D. 1978. 'Evidence of the need for a second language acquisition index of development' in W. Ritchie (ed.): *Second Language Acquisition Research*. New York: Academic Press.

Larsen-Freeman, D. (ed.). 1980. *Discourse Analysis in Second Language Research*. Rowley, Mass.: Newbury House.

Larsen-Freeman, D. 1983. 'Assessing global second language proficiency' in H. W.

Seliger and M. H. Long (eds.): *Classroom Oriented Research in Second Language Acquisition*. Rowley, Mass.: Newbury House.

Larsen-Freeman, D. and **V. Strom.** 1977. 'The construction of a second language acquisition index of development.' *Language Learning* 27/1: 123–34.

Laver, J. 1970. 'The production of speech' in J. Lyons (ed.): *New Horizons in Linguistics*. Harmondsworth: Penguin.

Lea, W. A. 1973. 'Segmental and suprasegmental influence on fundamental frequency contours' in L. M. Hyman (ed.): *Consonant Types and Tone*. Los Angeles, Calif.: U.S.C. Press.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Li, C. N. (ed.). 1976. *Subject and Topic*. New York: Academic Press.

Liberman, M. 1979. *The Intonational System of English*. New York: Garland.

Liberman, M. and **J. Pierrehumbert.** 1984. 'Intonational invariance under changes in pitch range and length' in M. Aronoff and R. T. Oehrle (eds.): *Language Sound Structure*. Cambridge, Mass.: MIT Press.

Lieberman, P. 1967. *Intonation, Perception, and Language*. Cambridge, Mass.: MIT Press.

Lieberman, P. 1970. 'Towards a unified phonetic theory.' *Linguistic Inquiry* 1: 307–22.

Lieberman, P. 1976. 'Phonetic features and physiology: a reappraisal.' *Journal of Phonetics* 4: 91–112.

Lieberman, P. 1984. *The Biology and Evolution of Language*. Cambridge, Mass.: Harvard University Press.

Lieberman, P., W. Katz, A. Jongman, R. Zimmerman, and **M. Miller.** 1985. 'Measures of the sentence intonation of read and spontaneous speech in American English.' *Journal of the Acoustical Society of America* 77/2: 649–57.

Loban, W. 1966. *Language Ability: Grades Seven, Eight, and Nine*. Washington, DC: Government Printing Office.

Long, M. H. 1980. 'Input, Interaction, and Second Language Acquisition.' Ph.D. dissertation, UCLA.

Long, M. H., C. A. Brock, G. Crookes, C. Deicke, L. Potter, and **S. Zhang.** 1984. *The Effect of Teachers' Questioning Patterns and Wait-time on Pupil Participation in Public High School Classes in Hawaii for Students of Limited English Proficiency.* Technical Report No. 1. Center for Second Language Classroom Research, Social Science Research Institute, University of Hawaii, Honolulu.

Maeda, S. 1976. 'A Characterization of American English Intonation.' Unpublished Ph.D. dissertation, MIT.

Nunnally, J. 1978. *Psychometric Theory* (2nd edn.). New York: McGraw-Hill.

Ochs, E. 1979. 'Transcription as theory' in E. Ochs and B. B. Schieffelin (eds.).

Ochs, E. and **B. B. Schieffelin.** (eds.). 1979. *Developmental Pragmatics*. Orlando, Fla.: Academic Press.

O'Connell, D. C. and **F. L. Slaymaker.** 1984. 'Evidence for the phonemic clause as an encoding unit.' *Language and Communication* 4/4: 273–83.

O'Donnell, R. C. 1974. 'Syntactic differences between speech and writing.' *American Speech* 49: 102–10.

Orestrom, B. 1983. *Turn-taking in English Conversation*. Lund: Gleerup.

Orestrom, B. 1985. *A Corpus of Shetland English*. Stockholm: Almqvist and Wiksell.

O'Shaughnessy, D. 1976. 'Modelling Fundamental Frequency and its Relationship to Syntax, Semantics, and Phonetics.' Ph.D. thesis, MIT.

Petrie, H. 1987. 'The psycholinguistics of speaking' in J. Lyons, R. Coates, M. Deuchar, and G. Gazdar (eds.): *New Horizons in Linguistics 2*. London: Penguin.

Pike, K. L. 1945. *The Intonation of American English*. Ann Arbor, Mich.: University of Michigan Press.

Quirk, R., S. Greenbaum, G. Leech, and **J. Svartvik.** 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rondal, J. A., M. Ghiotto, S. Bredart, and **J.-F. Bachelet.** 1987. 'Age-relation, reliability and grammatical validity of measures of utterance length.' *Journal of Child Language* 14: 433–46.

Sacks, H., E. Schegloff, and **C. Jefferson.** 1974. 'A simplest systematics for the organization of turn-taking in conversation.' *Language* 50/4: 696–735.

Sato, C. J. 1985. 'The Syntax of Conversation in Interlanguage Development.' Ph.D. dissertation, UCLA.

Scollon, R. T. 1974. 'One Child's Language from One to Two: The Origins of Construction.' Ph.D. dissertation, University of Hawaii.

Scollon, R. T. 1976. *Conversations with a One-year Old: A Case Study of the Developmental Functions of Syntax*. Honolulu: University of Hawaii Press.

Scott, M. and **G. Tucker.** 1974. 'Error analysis and English language strategies of Arab students.' *Language Learning* 24: 69–99.

Shewan, C. M. 1988. 'The Shewan Spontaneous Language Analysis (SSLA) system for aphasic adults: description, reliability, and validity.' *Journal of Communication Disorders* 21: 103–38.

Siegel, G. M. 1963. 'Adult verbal behavior in "play therapy" sessions with retarded children.' *Journal of Speech and Hearing Disorders, Monograph Supplement 10*.

Sinclair, J. McH. and **R. M. Coulthard.** 1975. *Towards an Analysis of Discourse*. Oxford: Oxford University Press.

Umeda, N. 1982. '"F_0 declination" is situation dependent.' *Journal of Phonetics* 10: 279–90.

Vaissiere, J. 1983. 'Language-independent prosodic features' in A. Cutler and D. R. Land (eds.): *Prosody: Models and Measurements*. New York: Springer-Verlag.

Van Dijk, T. A. 1977. *Text and Context*. London: Longman.

van Lier, L. 1988. *The Classroom and the Language Learner*. New York: Longman.

Van Oosten, J. 1984. 'A unified view of topic.' *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, 372–85.

Vann, R. J. 1979. 'Oral and written syntactic relationships in second language learning' in C. A. Yorio, K. Perkins, and J. Schachter (eds.): *On TESOL '79*. Washington, DC: TESOL.

Wells, R. S. 1947. 'Immediate constituents.' *Language* 23: 81–117.