STRUCTURAL DISCOURSE ANALYSIS UNITS FOR SECOND LANGUAGE RESEARCH

GRAHAM CROOKES
University of Hawai'i at Manoa

DISCOURSE ANALYSIS has been an accepted part of the methodology of second language (SL) research for some time (see, e.g., Chaudron, 1988, pp. 40-45; Hatch, 1978; Hatch & 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 analyzed, 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 (e.g., turns), in addition to a variety of other units defined in terms of their functions (e.g., 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, nor of 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

---

1 I am grateful to Robert Bley-Vroman, Craig Chaudron, Charlene Sato and Mike Long for their comments on earlier versions of this paper.

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.

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, p. 735). Thus *Mary hit John* is one T-unit, and *Mary, who is my best friend, hit John* 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 & Strom, 1977; Scott & 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, which do have a communicative value, can be coded. Such phrases typically appear in answer to a question:


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, p. 78) observes that these two units are in practice roughly equivalent, and Kroll (1977, p. 85) regards them as identical.

Idea unit

The idea unit (Kroll, 1977) is less well-known than the T-unit (for analyses using it see, e.g., Chafe, 1980; Danielewicz, 1984). As defined by Kroll (1977, p. 85), 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 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. (1977, p. 90)

An example of the segmentation of written text into idea units is as follows (Kroll, 1977, p. 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, p. 277; Sinclair & 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, pp. 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 e.g., Sacks, Schegloff & 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 can not 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—that associated with the work of Quirk, Crystal, and their associates, particularly studies involving the Lund corpus (e.g., Quirk, Greenbaum, Leech & Svartvik, 1985), the related work on language sampling and assessment (Crystal, Fletcher & Garman, 1976), as well as other varieties of English (Oreström, 1985).

Crystal & Quirk (1964, p. 50) 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. (pp. 50–51)

In addition to pitch-related means of identification, Crystal (1969, p. 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. (p. 206)

An example of a stream of speech separated into tone units is the following (after Crystal, 1969, p. 260):

but I personally / have never found / a boffin or a statistician / who worked things out in theory which ever came out in practice /

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 (e.g., Brown, Currie & Kenworthy, 1980, p. 40; see also Crystal & Quirk, 1964, p. 51 fn. 7). Although these may have originally been very close, Coulthard (1985, p. 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 & 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, he 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 & 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”; 1982, p. 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 & Sorensen (1981, p. 4) state that

[l]inguists traditionally relied on their perception of Fo [fundamental frequency] fall-rise patterns in conjunction with the perception of pauses to demarcate the syntactic structure of utterances and Lea (1973, p. 18) observes that

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

The definition of the utterance implied here was utilized (perhaps somewhat loosely) in early L1 acquisition research. Scollon (1974, p. 206, 1976, p. 152) cites use of “terminal intonation contour” to segment speech by Brown & Bellugi-Klima (1964), and used it in his own dissertation work (a 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. Scollon states (1974, p. 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” (p. 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 an initial criterion for segmentation. To analyze the recordings he had made of later stages in his child subject’s speech, he revised his initial working

2 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.
definition of utterance (p. 239), making more definite the two criteria mentioned so far, and including two aspects of a semantic criterion. The semantic criterion was added because Scallon 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 (p. 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 (e.g., Long, 1980; Sato, 1985) derived an operational definition of the utterance from Scallon (1976, esp. p. 153) and other L1 acquisition research (Ochs & Schieffelin, 1979). Sato (1985, pp. 83-4) defines it as "a stream of speech under a single intonation contour bounded by pauses". A subsequent modification was that of Crookes & Rulon (1985, p. 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 Scallon (1974, p. 206), a pause of less than 0.6 seconds did not indicate an utterance boundary, one between one and eight 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 analyzed. Chaudron (1988, p. 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 et al. (1980) (assuming that speech unified by topic is equivalent to semantically-related speech). Brown et al. found systematic cooccurrences 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 & Yule, 1983;
Keenan & 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, p. 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, p. 124) 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.

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.3

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.

---

3 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 reused as a definition of utterance by Rondal, Ghiotto, Bredart & Bachelet (1987).
Possible relationships between units

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, p. 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 analyzed 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 analyzing SL discourse, and Oreström's (1983) study of L1 conversation relates turn to tone-group.

T-unit to tone unit

The relationship between the tone unit 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 unit 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, p. 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% for beginning-points and 79% for end-points. The elements of clause structure to which he is referring are, however, such items as subject, or complement, and as a result, only 46% of the clauses in Crystal's data are co-extensive with tone units.
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).

Tone unit to utterance

Any relationship between these two units will probably only be close in discourse where utterances are mostly uni-clausal. In fact, tone units are often shorter than utterances (as defined following e.g., Chaudron, 1988), and Crystal (1969, p. 256) mentions an average figure of five words per tone unit. An additional factor which may contribute to the discrepancy between tone unit and utterance is the occasional existence of very short tone units, such as the elements of lists. 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, p. 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, p. 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

4 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, e.g., Long (1980). S-node-based measures have been used in a number of SL studies, without, however, much explicit discussion of their foundation.
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 exhibits an overall pattern of change in intonation,
with an initial peak and a gradual decline in fundamental frequency (see
Liberman & Pierrehumbert, 1982, p. 171 for a specific example). Consequently,
it 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 as it is 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).

Reliability

Though reliability does not imply validity, an unreliable discourse
analysis system cannot be valid (Frick & Semmel, 1978, p. 158; Nunnally, 1978,
p. 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 & Long, in preparation), 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%, Crystal,
1969, p. 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, p. 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, pp. 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 & Johns (1980, pp. 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 inter-rater 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 6,
below)—studies which report such data on the other units under consideration
provide figures at or above conventionally acceptable levels.
Validity

The aspect of validity which most critically applies to measuring instruments such as a discourse analysis system is "correspondence" (Brinberg & 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 & McGrath, 1982, p. 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, p. 86).5

In general, structural investigations of SL spoken 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.

5 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 subsidiary question 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).
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 & 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.6 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 e.g., 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 (Fo) 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 & Sorenson, 1981; Lieberman, 1984). One group of authorities further describe this pattern as a gradual decrease of Fo over time, and see it as

---

6 And again, if the equating is correct, O’Connell & Slaughter (1984. p. 281) report that “researchers at the University of Giessen have found it impossible to identify phonemic clauses reliably and have turned to other units”.


Fo 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 & Sorensen refer to it as “the most salient global attribute” of declarative speech (1981, p. 28), 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, p. 3)

Cooper & 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 Fo 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. (p. 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. In addition, there may be other reasons which would cause intonation to reflect cognition.

---

7 A dissenting view is that of Liberman & Pierrehumbert (1982), who recognize the existence of declination, but argue that it can be predicted simply by positing underlying "step accents" for English, dispensing with any physiological mechanism.
Lieberman (1967, 1970, 1976, 1984) and Cooper & Sorensen (1981), although representing two distinct lines of research, both 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 (e.g., Lieberman, 1967) and observation of children (Lieberman, 1984) support the position that the fluent speaker coordinates air pressure and volume in the lungs with duration of phonation in producing coherent speech. Similarly, Cooper & Sorensen (1981) found that Fo declines faster over shorter sentences. They found higher values for the first Fo peak in longer sentences, suggesting

that the speaker employs a look-ahead mechanism which computes the approximate overall upcoming utterances length. (p. 38)

Cooper & Sorensen (1981) and Lieberman (e.g., 1984; see also Berkovits, 1984; Boyce & 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. Initially accepting the grammatically defined sentence as "the minimal unit of speech for a complete semantic interpretation", Lieberman (1984, p. 98) suggests that under normal conditions there is a legitimate equating of utterance and sentence. In brief, his position is

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. (1984, p. 98)

This unit he refers to as the "breath-group" (Lieberman, 1984, p. 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, Katz, Jongman, Zimmerman, & Miller, 1985; cf. Liberman & Pierrehumbert, 1984; Umeda, 1982) disagrees with Cooper & 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 Fo 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, e.g., 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 & 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 term 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. (1981, p. 161)
Conclusion

This paper has surveyed a widely scattered body of literature concerning important discourse analytic units. 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 reliability 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—e.g., 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 1, 1988.

Author’s address for correspondence:
Graham Crookes
Department of English as a Second Language
1890 East–West Road
University of Hawai‘i at Manoa
Honolulu, HI 96822
DISCOURSE ANALYSIS UNITS

References


