5

Intonation Units

Consciousness is like vision. The similarities are probably not accidental, since the eye is anatomically an extension of the brain, and since for most of us vision is so fundamental a part of conscious experience. One way in which consciousness and vision are alike is in the very limited amount of information each can focus on at one time. There is foveal vision and focal consciousness. Surrounding this small area of maximum acuity lies, on the one hand, peripheral vision and, on the other hand, peripheral consciousness, both of which not only provide a context for the current focus but also suggest opportunities for its next moves. Beyond peripheral consciousness lies a vast treasury of information, some of which will at some time be brought into focal or peripheral consciousness, but all of which lies unattended at the moment. Consciousness and vision are alike in one other way as well. Both are in constant motion, the eye with its brief fixations, the mind with its continual shifting from one focus to the next. Both vision and consciousness exist in a state of constant restlessness.

Activation States

According to whether some idea is in the focal, peripheral, or unconscious state, we can speak of it as active, semiactive, or inactive. Thought and language involve continual changes in these activation states. Our concern in part 2 is with the effects such changes have on language, and on what language can tell us about the nature of the changes.

It might not be too misleading to associate active and inactive information with short-term and long-term memory respectively. I do not use those terms here, partly because of their possible implication that memory is a place. In the long run it may be less fruitful to speak of something being *in* memory or retrieving something from memory than to view these phenomena in terms of activation. Western psychology may have been misled by the fact that in European languages the process of remembering is reified in *memory* as a noun. One of the endearing qualities of Frederic Bartlett (1932) is the fact that his book was titled not *Memory,* but *Remembering.*

However that may be, although psychology has exploited the notions of

short-term and long-term memory, it has not so obviously allowed for a semiactive state. Such a state has been recognized indirectly in the notion of *context,* as when items in short-term memory are thought to be influenced by their surroundings., It seems to have been recognized more directly in James's use of words like *'Psychic overtone, suffusion, or fringe,* to designate the influence of a faint brain-process upon our thought, as it makes it aware of relations and objects but dimly perceived" (James 1890, 1: 258). Bruce Mangan has recently stimulated renewed interest in James's distinction, and has drawn an analogy to the menu-bar on some computer screens that "functions to indicate the existence of ... information that can be *potentially* called to the screen in detail—just as the fringe radically summarizes information that can be called into focal attention" (Mangan 1993, p. 98). I believe, however, that a more apt analogy is to vision, as when Bernard Baars observes that "we would be missing something important if we only dealt with focal consciousness, just as we would miss something vital in human vision if we studied only fovea! sight. Some of the most remarkable capacities of the visual system reside in the periphery and the same may be true of conscious experience in general" (Baars 1993, p. 135). The same analogy to vision was drawn seven hundred years ago by the theologian-philosopher Duns Scotus, who asserted that

> for every single perfect and distinct intellection existing in the intellect, there can be many indistinct and imperfect intellections existing there. This is evident from the example of vision, the field of which extends as a conical pyramid at the lower base of which one point is seen distinctly, and yet within that same base many things are seen imperfectly and indistinctly; but of these several visions, only one is perfect, namely, that upon which the axis of the pyramid falls. If this is possible in one of the senses, all the more so is it possible in the intellect. (Wolter 1986, p. 173; cf. Brett 1965, p. 295, and Mangan 1993, p. 89)

It is interesting too that James related "fringe" consciousness to (discourse) topics in a manner that will occupy us in chapter 10.

Speakers realize, of course, that one or more other minds are involved in the communicative use of language. As they speak, they not only take account of the changing activation states of information in their own minds, but also attempt to appreciate parallel changes that are taking place in the minds of their listeners. Language is very much dependent on a speaker's beliefs about activation states in other minds. Such beliefs themselves constitute an important part of a speaker's ongoing, changing knowledge, and language is adjusted to accord with them. Beliefs about other minds have various sources. To a considerable extent they are based on previous linguistic interaction—on things said within the same discourse, but also things remembered from previous talk.: Others are derived from nonlinguistic interaction, from shared experiences, and from shared cultures. Whatever the sources may be, conversation could not function as it does unless speakers took account of activation states in minds beyond their own.

Before we look more closely at the interplay between the three activation states and language, it is worth noting that the number three is probably too small. Although it is convenient in this work to deal only with the distinction between active, semiactive, and inactive information, a fuller understanding must almost certainly allow for further divisions of this continuum.

At one end of the continuum may be located what has been called *echoic* memory, the ability to shift one's consciousness of sound from the semiactive to the active state during the first few seconds after it has ceased to be present in the air (e.g., Neisser 1967, Glucksberg and Cowan 1970). Sound remains briefly available to active consciousness even if it failed to enter that state while it was physically present. This ability is clearly observable through introspection, as when we are able to retrieve something that was said to us, even though we may have been reading a newspaper when it was actually said and failed to focus active consciousness on it then. It is an ability that has a clear relevance to language, for it allows us to process sound sequences as wholes, not just "from left to right" as the sound enters our ears. It means that there is no real difference in the way we process *the gray house* and *la maison grise,* because in both cases the phrase is available to consciousness in its entirety. This ability compensates, in a small but important way, for the evanescence of sound, making it briefly scannable as a whole in the way a visual scene can be scanned. In this chapter we will meet, a unit of mental and linguistic processing—the intonation unit—that seems to be of exactly the right size to be processed in its entirety with the help of echoic memory, a fact suggesting that this ability functions crucially as a support for language. Indeed, it would not be far-fetched to speculate that echoic memory evolved as a necessary component of the evolution of language.

In a different part of the continuum, it is likely that the semiactive-inactive distinction includes more than just that simple dichotomy. In this book..any information that is neither fully active nor demonstrably semiactive will be called inactive. But there are reasons, even linguistic ones, to suspect that inactive information may be stored at either a shallower or a deeper level, the passage from the former to the latter being influenced by sleep, time, and the relative salience of the information. As one manifestation of this shallow-deep distinction, the ability to recall the temporal sequencing of events—the knowledge that one event happened before another—may be retained at the shallower level but lost at the

deeper one. This shallow-deep distinction and the manner in which it may influence the use of temporal adverbs and other linguistic features were explored in Chafe (1973). Although that work relied on constructed examples, it may nevertheless be suggestive of reasons for dividing the inactive category into at least a shallow and deep component.

Finally, we need to allow for the possibility that the three or more activation states are less categorical than they are depicted here—that they have fuzzy boundaries. However that may be, the effect of these states on language *is* categorical, and it is their linguistic effects that will concern us. Most of this chapter is devoted to the movement of ideas into and out of the fully active state. In later chapters, and especially chapters 7, 9, and 10, we will come to appreciate the relevance of semiactive information as well.

The Study of Prosody

The term *prosody* as used here embraces a variety of perceptual and physical properties of sound, including pitch, loudness, timing, voice quality, and the presence or absence of vocalization itself. In spite of ever-increasing research, the significant features and functions of prosody are still wide open to further exploration. Until roughly the second half of this century it was necessary to rely on the perceptual abilities of skilled investigators for prosodic observations. After World War II the sound spectrograph made it possible to observe the physical nature of pitch, loudness, and timing with considerable accuracy, but the labor involved in pitch measurements was arduous and time-consuming. More recently it has become much easier to make visual displays that open new worlds of observational possibilities. One way of dealing with this bonanza has been to approach prosody from the perspective of a phonetician, using displays of frequency, intensity, and duration as the primary data for understanding what language does with these aspects of sound. A well-known line of current research, for example, focuses on fundamental frequency for its primary data (e.g., Pierrehumbert 1980, Pierrehumbert and Beckman 1988), subsequently attempting to understand the semantic and discourse phenomena with which this one aspect of prosody is associated (e.g., Pierrehumbert and Hirschberg 1990).

The approach followed in this book developed out of a different tradition, in which the sounds of languages are transcribed in terms of perceived phenomena judged to express significant aspects of function and meaning. For this approach the breakthrough provided by current technology has been the ability to relate perceptual and physical observations. We perceive sounds in ways that do not fully correspond to their acoustic properties, but access to the latter can provide helpful insights into and correctives to our perceptual observations. As we try now to develop a better understanding of the flow of consciousness and language, prosody will be found to contribute in ways that cannot he ignored for spoken language or even, perhaps surprisingly, for written.

Intonation Units and Their Delimitation

Anyone who listens objectively to speech will quickly notice that it is not produced in a continuous, uninterrupted flow but in spurts. This quality of language is, among other things, a biological necessity. Because speech sounds are produced by expelling air from the lungs, the air must be periodically replaced if the speaker is to remain alive. It is remarkable that language and this obvious physiological requirement have evolved together in such a way that we are able to speak for long periods of time without getting out of breath. Eric Lenneberg once called attention to the fact that "breathing undergoes peculiar changes during speech. What is astonishing about this is that man can tolerate these modifications for an apparently unlimited period of time without experiencing respiratory distress. . . I believe it is fair to say that we are endowed with special physiological adaptations which enable us to sustain speech driven by expired air" (Lenneberg 1967, pp. 80-81; see also Goldman Eisler 1968 on the relation between speech pauses and breathing). The need to breathe would alone produce the spurtlike quality of speech, but if one examines the linguistic and psychological nature of the spurts, it becomes clear that more is involved. Breathing would require nothing more than *an* interruption of vocalization at regular intervals. One finds, in fact, that this physiological requirement operates in happy synchrony with some basic functional segmentations of discourse.

These functionally relevant segments are not delimited by pauses alone, since pauses may occur within them, and although they are often separated by pauses, that is not always the case. From now on I will refer to these segments of language as *intonation units.* Various other names have been used for units of a similar, though not in all ways identical kind: *tone unit,* for example, by various British linguists, or *intonation group* (Cruttenden 1986), or *intonation(al) phrase* (Bing 1985; Pierrehumbert and Beckman 1988, where evidently it is the *intermediate phrase* that corresponds to the intonation unit here). There is also a correspondence between the intonation unit and what Dell Hymes (1981) calls a *line* (as in a line of poetry). Because these various other terms do not always delimit a unit that coincides consistently with the intonation unit as it is understood here, the use of a distinctive term is justified. Intonation units,

Figure 5.1 Acoustic Properties of Example (1)

for example, need not be limited to only one primary accent, as is arbitrarily required of such units in the British tradition (but cf. Ladd 1986).

Researchers are always pleased when the phenomena they are studying allow them to identify units. Units can be counted and their distributions analyzed, and they can provide handles on things that would otherwise he obscure. Unless all of us have been deceiving ourselves badly, language does make use of units of various kinds—vowels, consonants, and syllables, for example, or words and sentences, and now intonation units. It would be convenient if linguistic units could be identified unambiguously from phonetic properties: if, for example, phonemes could be recognized from spectrograms, or intonation units from tracings of pitch. For good or bad, however, the physical manifestations of psychologically relevant units are always going to be messy and inconsistent. If one breaks eggs into a frying pan, it may or may not be easy to tell where one egg leaves off and another begins. It may be similarly easy or difficult to read off the boundaries of intonation units directly from displays of acoustic data.

- The features that characterize intonation units may involve any or all of the following: changes in fundamental frequency (perceived as pitch), changes in duration (perceived as the shortening or lengthening of syllables or words), changes in intensity (perceived as loudness), alternations of vocalization with silence (perceived as pausing), changes in voice quality of various kinds, and sometimes changes of turn. Figure 5.1 shows (above) the wave form and (below) the fundamental frequency of a well-

defined intonation unit whose boundaries are confirmed in various of these ways. The relatively narrow transcription in (1) below attempts to capture certain acoustic properties that are relevant in the discussion that follows. Later, for practical reasons, I will fall back on a broader transcription system that will represent only those features that bear most directly on the topics discussed.

(1) .. and so the **háll** is real 1ó= ng%.
... (.36) [next intonation unit]

Preceding the vocalization is a very brief pause of about .07 second. Pauses of .10 second or less are transcribed simply with two dots. Following the vocalization and before the next intonation unit is a lodger pause of .36 second, transcribed with three dots followed in parentheses by a measurement of the pause length (an accuracy to hundredths of a second is more than adequate). By convention, boundary pauses are shown at the beginning of each intonation unit. Among other things, then, (1) is set off by pauses.

One of the major cues to intonation unit boundaries is change in duration, captured in part by the notion of "anacrusis" (Cruttenden 1986, pp. 24, 39). Example (1) begins with a sequence of three rapid syllables *(and so the)* occupying roughly .10 second each, shown with smaller type. The transcription system employed in the rest of this work does not mark accelerated syllables in this way, but they will nevertheless play a role in the determination of intonation unit boundaries. After the first three words there are two words *(hall* and *real,* separated by a rapid *is)* whose duration lies in the range from about .20 to .30 second, a normal length for one-syllable words. The intonation unit ends with a word of extended length *(long)* occupying .43 second, the lengthening shown with an equals sign after the vowel. This pattern of acceleration-deceleration, proceeding from reduced-length syllables up to about .15 second, through normal-length syllables from about .15 second to about .35 second, to extended-length syllables longer than .35 second, is characteristic of many intonation units and may in some instances be the primary evidence lot their delimitation. (Obviously these figures need to be adjusted for slower and faster speaking rates.)

When it comes to pitch, it happens that (1) coincides with a "declination unit" (Schuetze-Coburn, Shapley, and Weber 1991). There are three words with noticeably high pitch *(hall, real,* and *long),* each lower than the preceding (maxima of 299 hertz, 211 hertz, and 192 hertz respectively). As Schuetze-Coburn et al. show, such declination units often extend over several intonation units, but at least their beginnings and endings provide evidence for many intonation unit boundaries.

A more consistently present indicator is a terminal pitch contour of

some kind at the end of each intonation unit. A variety of contours are observable in natural speech, where further study of their properties and functions is much needed. Example **(1)** ends with the familiar falling pitch contour associated with the end of a declarative sentence or a question-word question, transcribed here with a period. The terminal contours that are distinguished in transcriptions in this work include this falling pitch, a high rising pitch of the type associated with a yes-no question (transcribed with a question mark), and any other, nonterminal pitch contour (transcribed with a comma). These distinctions are adequate for our immediate purposes, but a better transcription system would replace the comma with markings of various more specific contours.

A particularly common change in voice quality is creaky voice (laryngealization or "fry"). It is conspicuous here at the end of the lengthened word *long*, where it is indicated with the percent sign. Intonation units often end and sometimes begin with creaky voice, which thus provides still another clue to their delimitation. Creaky voice may obscure acoustic displays of falling pitch contours, as is the case at the end of figure 5.1.

In summary, the identification of (I) as a coherent intonation unit is supported by a convergence of (a) the pauses preceding and following it, (h) the pattern of acceleration-deceleration, (c) the overall decline in pitch level, (d) the falling pitch contour at the end, and (e) the creaky voice at the end. These and other features are discussed and exemplified in more detail in Chafe (1992c).

Prominences

Besides perceiving speech as segmented into intonation units, we perceive certain elements within an intonation unit as more prominent than others. The acoustic correlates of prominence are also complex and variable. There are degrees of prominence, and there are several ways in which prominence may be realized. Here I arbitrarily use the term *accent* for prominences that are realized as pitch deviations from a mid or neutral baseline, usually a higher pitch but occasionally a lower one, represent such pitch deviations with accent marks, regardless of whether they rise above or fall below the baseline. When one of these accented elements is also either loud or lengthened or both, I say that it has a *primary* accent and show it with an acute accent mark. **A** pitch deviation alone, without accompanying loudness or lengthening, is said to characterize a *secondary* accent, shown with a grave accent mark, Of course an element may be either loud or lengthened without a pitch deviation; in such cases I say only that it is loud or lengthened, but not that it is accented.

As an illustration of prominences, we can look again at the intonation unit cited in (1), repeated here (see again fig. 5.1).

(2) .. and so the **háll** is rèal ló=ng%.

Three of these words—hall, *real*, and *long*—*are* accented, all showing heightened pitch (with downstep). However, not only is *hall* higher pitched than the other two, it is also significantly louder, as indicated with boldface type. To anticipate the discussion in chapter 6, the exaggerated prominence of this word expresses its contrastiveness. The idea of this hall was introduced eight intonation units earlier, but in (2) the hall is contrasted with' the living room, the bedroom, and the bathroom, all introduced in the meantime. It is not unusual for contrastive elements to show exaggerated pitch deviation as well as exaggerated volume.

In the predicate of (2) the heaviest load is carried by the word *long* which is both high pitched (before the fall) and lengthened. The intensifer *real* is high pitched but neither loud nor lengthened, and thus is said to carry a secondary accent. We find, then, three different manifestations of prominence in this intonation unit: the high-pitched and loud *hall*, *the* high-pitched and lengthened *long*, and the word *real* with high pitch only.

In this work, the prosodic features that are marked consistently in transcriptions include (a) pauses (marked by sequences of dots and sometimes, when relevant, by measured pause times); (b) terminal contours (marked with periods, question marks, and commas); and (c) accents (marked with acute and grave accent marks). Noted only occasionally, when relevant to the discussion, are changes in overall pitch level, accelerations and decelerations, and voice quality. All these features, however, enter into the segmentation of discourse into intonation units, indicated throughout by the placement of each such unit in a separate line.

Intonation Unit Sequences

The following conversational excerpt illustrates a few of the complications typical of intonation unit sequences. The notations (A), (B), and (C) identify different speakers. The preceding talk had been about a fatal accident that involved an elephant.

(3) a(A) ... (0.4) Have the .. animals,
b(A) (0.1) ever attacked anyone in a car?
c(B) (1.2) Well I
d(B) well i heard of an elephant,
e(B) that sat dówn on a VW one time.
f(B) ... (0.9) There's a gir

g(B) Did you éver hear thát?
h(C) (0.1) No,
i(B) ... (0.3) Some élephants and these
j(B) ... (0.1) they
k(B) ... (0.7) there
1(B) these gáls were in a Vólkswagen,
m(B) .. (0.4) and uh,
n(B) ... (0.3) they uh kept hónkin' the hórn,
o(B) (0.2) hóotin' the hóoter,
p(B) ... (0.6) and uh,
q(B) ... (0.4) and the .. élephant was in fróntof em,
r(B) so = he jùst proceeded to sìt dòwn on the VW
s(B) ... (0.3) But thèy .. had .. managed to get óut first.

Noteworthy is the fact that (3)c, f, i, j, and k were truncated intonation units that never arrived at their terminal contours. One can also note that there was no pause separating the truncation in (3)c from the beginning of (3)d, or the truncation in (3)k from the beginning of (3)1. There was also no break between (3)q and (3)r, a segmentation dictated by a terminal pitch contour at the end of (3)q, a resetting of the pitch baseline at the beginning of (3)r, and a durational phenomenon the reverse of that observed in (2): the last five words of (3)q were accelerated, whereas the ifrst word of (3)r was lengthened.

Segmenting speech into intonation units and identifying primary and secondary accents are skills that can only be learned with instruction and practice. (Useful guides are Cruttenden 1986, pp. 35-45 and Du Bois et al. 1992; see also Du Bois et al. 1993.) But nothing can substitute for hands-on practice with recordings of real language under the guidance of an experienced transcriber. Unfortunately these abilities, like the ability to record phonetic dictation of any kind, cannot be learned from a hook. In a better world they would be as important a part of the training of a linguist as the ability to transcribe vowels and consonants.

The Function of Intonation Units

In spite of problematic cases, intonation units emerge from the stream of speech with a high degree of satisfying consistency, not just in English, but in all languages I have been able to observe and in fact in all styles of speaking, whether conversation, storytelling, oration, the performance of rituals, or even (or especially) reading aloud. That fact suggests that they play an important functional role in the production and comprehension of language. As we consider what that role might he, we can return to the notion of activation states. It is intuitively satisfying to suppose that each intonation unit verbalizes the information active in the speaker's mind at its onset. Let us hypothesize that an intonation unit verbalizes the speaker's focus of consciousness at that moment.

At the onset of an intonation unit, according to this view, often but not always following a pause, a certain small amount of information is active in the mind of the speaker. Typically, some of that information will have become active during the pause, though other parts of it are likely to have been activated previously. In chapter 9, we will meet a strong constraint on how much information can be newly activated at one time. It may be that all of the *information* to be verbalized in the upcoming intonation unit is active for the speaker at this onset point, but disfluencies show that people sometimes revise their choice of *wording* while an intonation unit is already in progress, as illustrated in (4):

(4) ... Her she has an enlarged heart.

Evidently the speaker began to say *her heart*, but, for reasons considered in chapter 7, she quickly shifted to a different wording.

During these successive activations the minds of the speaker and the listener are necessarily out of phase. At the completion of an intonation unit the speaker must intend that a reasonable facsimile of his or her focus of consciousness will have become active in one or more other minds. It is through this dynamic process of successive activations, first for the speaker and then, through the utterance of an intonation unit, for the listener, that language is able to provide an imperfect bridge between one mind and another.

Types of Intonation Units

Viewed a little more closely, intonation units fall into several types. While many ways of categorizing them can be imagined, the following breakdown into three major types is useful because certain aspects of an analysis can be directed at one of these types to the exclusion of the others. We have already noticed that some intonation units are truncated or *fragmentary*. The successful units can be subcategorized into those that convey *substantive* ideas of events, states, or referents and those that have *regulatory* functions in *the* sense of regulating interaction or information flow. The distinction between substantive, regulatory, and fragmentary intonation units is illustrated in (5), which provided the context for (4):

(5) a(A) Well, (regulatory)
b(A) isn't she healthy? (substantive)

c(B) ....Mhm, (regulatory)
d(A) …I mean she (fragmentary)
e(A) I knów she has (fragmentary)
f(C) More or léss. (substantive)
g(A) .. She has [something with her] gállbladder, (substantive)
h(B) [gállbladder and, (substantive)
i(B) …héart tròuble and, (substantive)
j(B) [back problems.] (substantive)
k(A) [She has héart] tròuble, (substantive)
1(C) ... Her she has an enlàrged héart. (substantive)

In a finer analysis, regulatory units can be subdivided further. Some regulate the development of the discourse, as when (5)a prepares for the contextually relevant question in (5)b. Others have to do with interaction between the participants, as when (5)c responds to the question in (5)b. Still others express the speaker's mental processes (as in expressions like *oh* or *let me see),* or judgment of the validity of the information being conveyed (like *maybe* or *I think).* Thus, regulatory units serve at least the following functions, whose boundaries are less categorical than this listing suggests:

textual (e.g., *and then, well)*
interactional (e.g., *mhm, you know)*
cognitive (e.g., *let me see, oh)*
validational (e.g., *maybe, I think)*

Regulatory intonation units coincide to a large extent with the devices that have been discussed under the label *discourse markers* (Schiffrin 1987), which often constitute intonation units in themselves, though they may also be expressed as parts of larger units.

The Size of Intonation Units

A certain insight into consciousness and linguistic processing can be gained just from examining the size of intonation units. The simplest and most obvious measure is the number of words an intonation unit contains. Regulatory and substantive units differ significantly in this respect and for that reason are best measured separately. Fragmentary units can be ignored, since one can only guess how long such a unit would have been if it had been completed. To begin with *regulatory* intonation units, their mean length in the measured sample is 1.36 words, with a modal length of one. Thus the regulation of discourse flow, whether it functions textually, interactionally, cognitively, or validationally, is accomplished in very short segments of speech:

(6) So,
Yeah.
Hm.
Sort of,

The mean length of *substantive* intonation units in the measured sample is 4.84, with a modal length of four. Apparently a focus of consciousness is typically expressed with four words of English.[1] It is important to realize that this figure is valid for English only; languages that pack more information into a word show fewer words per intonation unit, as discussed in chapter 12.

The word, it should be noted, is not a wholly satisfactory measure of information. Aside from the fact that different languages include different amounts of information in their words, both words and the morphemes of which they are composed express a variety of different types of information. In (7), for example, there is a sense in which the words *gal* and *Volkswagen* are more informative than the other words:

(7) these gals were in a Volkswagen,

Furthermore, there are many instances in which unitary ideas are expressed by sequences of words. In (8), for example, the two-word sequence *heart trouble* conveys one idea:

(8) She has heart trouble,

It would thus be a mistake to assume that each word counts equally, or that the number of separate ideas verbalized in an intonation unit bears any simple relation to the number of words. Despite these reservations, it is a striking fact that the number of words in an intonation unit remains within a narrow range for any one language, reflecting in a gross way a strong constraint on the capacity of active consciousness.

Intonation Units and Clauses

Many substantive intonation units have the grammatical form of single clauses. Many others are parts of clauses, but the mean proportion of single-clause substantive intonation units in the measured sample is about

1. In earlier discussions of this topic I identified a modal length of five words. The discrepancy can be traced primarily to a more careful identification of intonation unit boundaries in more recent work, where more subtle criteria have increased the number of boundaries and thus reduced the unit size.

60 percent. It appears that speakers aim at verbalizing a focus of consciousness in the format of a clause, although for reasons explored in chapters 7 and 9 they are often forced to spread the clause across several intonation units.

A clausal intonation unit may assert the idea of an event or state. For example, (9) and (10) verbalize ideas of events, (11) the idea of a state:

(9) ... and these gàls were tàking píctures.
(10) but then your back gets sway bàck.
(11) .. She has something with her gállblàdder,

In general, a state involves a situation or property that exists for a certain period without significant change, whereas an event typically involves a change during a perceptible interval of time. It is helpful to think of an event as something that *happens—either* something someone does (an action), as in (9), or something that happens to someone or something (a change of state), as in (10). A state, on the other hand, rather than happening, simply exists for a greater or lesser period of time, as in (11).

If we think of a typical substantive intonation unit as having the form of a clause, and if we think of a clause as verbalizing the idea of an event or state, we can conclude that each such idea is active, or occupies a focus of consciousness, for only a brief time, each being replaced by another idea at roughly one- to two-second intervals. Event and state ideas, in other words, are highly *transient* in active consciousness. They are constantly being replaced by other event and state ideas.

It can also be observed that each event or state idea is, by and large, activated only once within a particular discourse. This is not to say that the same idea cannot be reactivated; we will shortly notice ways in which that can happen. But transient and nonrepeated activation is the rule. It seems that the mind does not usually dwell on an event or state idea for more than a second or two. Any sample of ordinary speech will show a constant progression from one such idea to the next, of the sort illustrated by the sequence in (12):

(12) a(A) ... Cause I had a ... a thíck patch of bárley there, (state)
b(B) mhm, (regulatory)
c(A) about the sìze of the .. kìitchen and líving room, (state)
d (A) and I went òver ít, (event)
e(A) .. and then, (regulatory)
f(A) ... when I got dóne, (event)
g(A) I had a little bit léft, (state)
h(A) .. so I tùrned aróund, (event)

i(A) and I went and sprày ed it twíce. (event)
j(A) .. and it's just as yèllow as …. can bé. (state)

This movement reflects our mental experience: "Thought is in constant change" (James 1890, 1:229). To some extent the continual replacement of event and state ideas reflects the world we live in, which is usually itself in flux. Even when that is not the case, however, consciousness continues to move from one such idea to another, and it seems impossible to keep it still. This restlessness forces us to keep sampling in small chunks the information available to us.

But there is another kind of idea that is more persistent. Each of these event or state ideas contains within it other, included ideas that can be said to be *participants* in the events or states. These participants are typically the ideas of people, objects, or abstractions, for which the term *referents* is appropriate. The state idea in (12)a includes as participants a referent verbalized as *I (the* idea of the speaker) and another verbalized *as a thick patch of barley*. The event idea in (12)d includes as participants these same two referents, verbalized this time as *I* and *it*, as also in (12)i. With a few exceptions such as *raining* and *being cold* (of *the* weather), things do not happen and states do not exist without the inclusion of referents who perform them, are affected by them, or participate in them in other ways. In English, referents are typically factored out from the events and states, to be verbalized as nouns and pronouns.

From this point on, in order to avoid the awkward phrases *event idea* and *state idea* I will often say simply *event* and *state. It* is important to keep in mind, however, that in this usage events, states, and referents are all ideas that exist in the minds of speakers and listeners. Whether or not they have correlates in the "real world" is irrelevant. I can (and do) think of the feats of Scarlett O'Hara as naturally as those of Marilyn Monroe. That only one of these referents ever existed in "reality" makes no difference to my thought or speech, at least with respect to the phenomena being discussed.

It is not unusual for an intonation unit to verbalize little or nothing more than a referent, as in intonation units a, c, f, i, and j of (13), originally presented as (3) above:

(13) a(A) ... (0.4) Have the .. ánimals,
b(A) ... (0.1) ever attacked anyone ín a car?
c(B) ... (1.2) Well I
d(B) well Í heard of an élephant,
e(B) .. that sat dawn on a VW one time.
f(B) ... (0.9) There's a gìr
g(B) .. Did you éver hear thát?
h(C) „. (0.1) No,

i(B) (0.3) Some élephants and these
j(B) (0.1) they
k(B) ... (0.7) there
l(B) these gáls were in a Vólkswagen,
m(B) ... (0.4) and uh,
n(B) ... (0.3) they uh kept hónkin' the hórn,
o(B) ... (0.2) hóotin' the hóoter,
p(B) .. (0.6) and uh,
q(B) ... (0.4) and the … élephant was in fróntof em,
r(B) so = he jùst procèeded to sit down on the VW.
s(B) ... (0.3) But thèy … had ... mànaged to get óut first.

Typically such *isolated referents* (expressed as so-called free NPs) are subsequently included as participants in events and states. But intonation units like these show that it is quite possible for speakers to focus on a referent alone.

Whereas events and states are activated transiently, *many referents* remain active for longer periods than any of the events or states in which they participate. For example, the idea of the speaker himself must have been active well before the sequence in (12) began, and it must have remained active well after that sequence ended. The idea of the *thick patch of barley* was activated in (12)a and remained active at least through (12)j. This is not to say that referents cannot be as transient as the events or states in which they participate. *The kitchen and living* room in (12)c provides one example of such a transient referent. Conversely, some events and states may remain active beyond a single intonation unit or may subsequently be reactivated. There are at least two ways in which the activation of an event or state may be made less transient than would normally be expected. Both ways are illustrated in the sequence in (13). Most obviously, the event verbalized in (13)n was *reverbalized* in (l3)o. The speaker dwelt on the same event over the space of two intonation units, expressing it with different language. It is more frequently the case that an event or state persists by being converted into a referent—by being reified or *nominalized*. The event verbalized in the sequence (13)d-e was nominalized with the word *that* in (13)g. People are able to conceptualize events and states as if they had temporal persistence. Once an event or state has been given this derived status as a referent, it may then, like other referents, participate in and persist through a series of other events or states.

Later we will see the importance of recognizing that activated ideas do not immediately recede into the inactive state but remain for a time semiactive. For the moment, however, our major interest is in the fact that events and states are highly transient in fully active consciousness,

each remaining for no. more than a brief interval, whereas some but not all referents persist longer in *the* fully active state.

Summary and Prospects

Information in the mind may be in any one of at least three activation states: active, semiactive, or inactive. There may well be more activation states than these, and the boundaries between them may be less categorical than this division suggests.

Spoken language lends itself to segmentation into intonation units. Such units are identifiable on the basis of a variety of criteria, among which are pauses or breaks in timing, acceleration and deceleration, changes in overall pitch level, terminal pitch contours, and changes in voice quality. Intonation units are hypothesized to be the linguistic expression of information that is, at first, active in the consciousness of the speaker and then, by the utterance of the intonation unit, in the consciousness of the listener, or at least that is the speaker's intent. Intonation units may be substantive, regulatory, or fragmentary. Regulatory units tend to be one word long, while substantive units are fairly strongly constrained to a modal length of four words in English, a fact that suggests a cognitive constraint on how much information can be fully active in the mind at one time. Regulatory units tend to be simple particles, fragmentary units have no determinate structure, but the majority of substantive intonation units have the form of single clauses, though many others are parts of clauses.

Each clause verbalizes the idea of an event or state, and usually each intonation unit verbalizes a different event or state from the preceding, which is to say that events and states tend to be highly transient in consciousness. Most events and states include within them one or more referents—ideas of people, objects, or abstractions that participate in them. Many referents persist, remaining active through a series of intonation units, although some are transient, remaining active only during the activation of a single event or state. Conversely, events and states are sometimes converted into referents, or nominalized, a process that allows them to persist and appear as participants in other events or states.

Looking toward the future 1 would note that the properties of intonation units, both acoustic and perceptual, need to be more definitively established as part of a larger effort to relate physical sound to the perception of prosody. This is an ideal area in which to combine observations of natural speech with relevant experimental manipulations. The measurement of intonation units in terms of time, number of words, and grammatical composition will obviously benefit from access to more extensive and

varied samples of speaking. Their classification into substantive, regulatory, and fragmentary intonation units can be elaborated and the bases of such classifications made more precise. Varying transitions between intonation units will be touched on in later chapters in terms of sentence-internal, sentence-external, and topic boundaries at various levels, but much more can be done in the way of relating strengths of intonation unit boundaries to the flow of consciousness. Finally, the relation of intonation units to clauses needs further study, both within and across languages.

6

Activation Cost

There is an intuitive plausibility in the notion that some of the information expressed in an intonation unit or clause is "new" while other information is "old." Because of the misleading connotations of the word *old,* the term *given* has often been used instead, but a different term does not in itself answer the questions that arise as soon as one tries to give more precise content to these words.

There is, for example, the question of the *domain* of newness or givenness. If someone said in a relevant context

(1) -... I tàlked to a láwyer last nìght,

it might be supposed that the entire intonation unit expressed the idea of an event that was in some sense new within this discourse. In that sense the whole of (1) could be said to have conveyed new information. On the other hand, the distinction between new and given information can be applied independently to the referents that participate in events and states. Thus, the referent expressed by *I* might be thought to be information that was already given, because of the obvious presence of the speaker in the conversation, whereas that expressed by *a lawyer* might be thought to be new. This more local view of givenness and newness is appealing because it helps to explain why the idea of the lawyer was expressed in a full noun phrase with a primary accent, whereas the idea of the speaker was expressed in a pronoun with a weak accent. Examples like these suggest that language gives more prominence to new ideas than to given ones, prominence being recognizable in terms of full nouns (more prominent) versus pronouns (less prominent), and strong accent (more prominent) versus weak accent (less prominent).

This way of viewing things, however, still does not answer the question of just what is meant by new and given. In what sense was the idea of the lawyer new and that of the speaker given? An initial hypothesis might be that a new idea is an idea the speaker thought was previously unknown to the listener. Its newness could then be identified with its status as a new entry into the listener's mind, or at least the speaker's judgment that it had such a status. A given idea would then be one that the speaker thought was already known to the listener. The example in (1) fits such an interpretation, since the listener in this case could be assumed not to