

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/230875969>

Outline of discourse transcription

Chapter · January 1983

CITATIONS

168

READS

411

4 authors, including:



[John Du Bois](#)

University of California, Santa Barbara

23 PUBLICATIONS 1,392 CITATIONS

SEE PROFILE

TALKING DATA: TRANSCRIPTION AND CODING IN DISCOURSE RESEARCH

Edited by

Jane A. Edwards
University of California at Berkeley

and

Martin D. Lampert
*Holy Names College &
University of California at Berkeley*



1993

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Hillsdale, New Jersey Hove and London

3

Outline of Discourse Transcription

John W. Du Bois
University of California at Santa Barbara

Stephan Schuetze-Coburn
University of California at Los Angeles

Susanna Cumming
University of Colorado at Boulder

Danae Paolino
University of California at Santa Barbara

This article presents a set of basic categories, symbols, and conventions for discourse transcription. *Discourse transcription* can be defined as the process of creating a written representation of a speech event so as to make it accessible to discourse research. In the following pages, we present in outline a framework for carrying out such a discourse transcription. (For a discussion of the principles that underlie the design of transcription systems, see Du Bois, 1991, and Edwards, this volume; for an in-depth treatment of the whole process of discourse transcription, see Du Bois, Schuetze-Coburn, Cumming, & Paolino, 1991.)

For each symbol in this discourse transcription system, a brief explanation of usage is given, illustrated with an example drawn from conversational transcriptions. Where appropriate, we comment on why the discourse feature in question should be attended to.

A word about the examples is in order. All examples cited are drawn from actual conversations, which have been transcribed and checked by the authors.¹ The examples are given in a fairly broad transcription, which

¹Most of the examples are from tape recordings made by the authors, but a few have been contributed by other researchers. For each example given, the source is cited immediately preceding the example.

includes the most basic transcription information: the words and who they were spoken by, the division of the stream of speech into turns and intonation units, the truncation of intonation units and words, intonation contours, medium and long pauses, laughter, and uncertain hearings or indecipherable words. In a more detailed, "narrow" transcription, the transcriber would also include notation of, among other things, accent, tone, prosodic lengthening, and breathing and other vocal noises. Such narrow transcription features have been omitted from most of the transcriptions in this chapter, so as not to overload the reader with too many new symbols at once; but for those sections that deal directly with the transcription of these features, the examples will include them. (Narrowly transcribed versions of most of the examples cited here are presented in Du Bois et al., 1991.) For brevity's sake, the examples represent for the most part short stretches of discourse without a great deal of textual context (cotext). Thus the portions cited are not always whole sentences or whole interactions, but each line that is cited is whole—that is, each intonation unit is presented in its entirety—and no omissions have been made within the stretch of transcription that is cited. Whenever a new notational convention is being introduced, the symbol is written in **boldface** letters in the illustrative examples for that section, in order to highlight the feature in question.

1. UNITS

One of the most striking, if elusive, features of conversation is its division into recognizable units at various levels. Any discourse transcription should indicate at least the most fundamental of these. This section presents symbols for boundaries between units of various kinds, including the intonation unit and the word unit, as well as truncated (uncompleted) variants of these units. (The turn, which is a fundamental unit of conversational discourse, is treated later in conjunction with the speaker identification label.)

1.1. Intonation Unit

A carriage return is used to indicate the end of an intonation unit (in effect, the boundary between two intonation units). Thus, each intonation unit appears on a separate line.²

²In general, a speaker's intonation unit should not be broken up into two lines; but for the occasional intonation unit that is so long that it will not fit on one line, see the section on presentation. Also, it may sometimes be necessary to break an intonation unit using the "intonation unit continued" symbol (described in the section on specialized notations).

Roughly speaking, an intonation unit is a stretch of speech uttered under a single coherent intonation contour. It tends to be marked by cues such as a pause and a shift upward in overall pitch level at its beginning, and a lengthening of its final syllable. For a fuller discussion of intonation units, the cues that mark them, and the methods for identifying them, see Du Bois et al. (1991), Chafe (this volume), and Cruttenden (1986, pp. 35–45).

(1) ((Aesthetics))

S: That's interesting,

I mean,

th- that you should pair the word aesthetics,

... with advertising.

J: Yeah.

1.2. Truncated Intonation Unit

A double hyphen (--) indicates that the speaker breaks off the intonation unit before completing its projected contour.

This truncation occurs primarily in cases where a speaker utters the initial portion of a projected intonation unit but abandons it before finishing—that is, a false start. The double hyphen is *not* intended to represent the case of a unit that appears incomplete when measured against the canons of normative grammar. Intonation units that do not constitute complete clauses are commonplace and usually quite normal—and "complete" as intonation units. For example, conjunctions (*and*) and particles (*well*) frequently appear as complete intonation units marked with a comma at the end, which signals "continuing" intonation (see below)—a kind of incompleteness, if you will, but one that is distinct in principle from the truncation signaled by double hyphen. The unit marked with a comma typically constitutes (apparently) all that the speaker projected to say *within the current unit*, whereas in the unit marked with a double hyphen the speaker projected to say more within the current unit but abandoned some portion of the projected utterance. Truncation is thus measured not against normative notions of clause completeness but against the speaker's presumed projection for the current intonation unit.

(2) ((Ranch))

R: He doesn't have any --

... He doesn't know what's going on in this world.

Note that for every intonation unit that *is* complete, the line should end with *some* representation of its intonation contour (see below). If an intonation unit does not so end, it will in general have the double hyphen that marks truncation.

1.3. Word

The space character is used to separate words, as in normal orthographic convention. A space also separates certain other notations, such as those for laughter, pause, inhalation, and so forth.

(3) ((Aesthetics))

S: Hm.

Hm.

... Okay.

1.4. Truncated Word

A single hyphen (-) indicates where the speaker has truncated a word,³ leaving the end of the (projected) word unuttered.³

Truncation is often cued overtly via word-final glottal constriction, but not always—either phenomenon may occur independently of the other. Other truncation cues may include segment shortening, slight rhythmic discontinuities, and so on. (Where it is deemed relevant, the precise pronunciation of the truncated word can be written using phonetic notation.)

(4) ((Friends))

J: ... You know how they do that,
so you can't s- ha- --
you don't have any balance.

This symbol is *not* used to mark words that have been pronounced in an abbreviated fashion as part of an informal speech style. Truncation of a word is measured not against canons of "normal" or "standard" pronunciation but

³ The single hyphen is written at the end of the word or word fragment, with no space intervening. Note that for standard orthographies such as English that use a hyphen to write certain compound words, some care needs to be taken to distinguish this orthographic use of the hyphen from its use for truncated words (and also for truncated intonation units). Because in truncated words the hyphen is followed by a space, whereas in compound words it is generally followed by a letter of the alphabet, this should in general be sufficient to allow automatic discrimination between the two. Similar measures will work for the truncated intonation unit notation.

against the speaker's projected pronunciation for the current word. Only when a speaker projects pronunciation of a word and then fails to complete that projected pronunciation is the phenomenon of word truncation involved.

2. SPEAKERS

2.1. Speaker Identification and Turn Beginning

To identify the speaker of a given turn in the conversation, a code or a proper name (written all in capital letters) is inserted at the beginning of the turn, followed immediately by a colon (:).

(5) ((Door))

A: Now that we have the [side door] fixed,

B: [That's kind of] --

A: he could.

B: Yeah,

C: Yeah.

D: ... Sure.

Although transcribers often assign prosaic codes such as "A" or "B" to their speakers, the reader tends to get a more vivid impression of who the participants are if their utterances are tagged with personal names, which are always more memorable. The name should in general be a pseudonym, since in any transcription destined for public presentation, privacy considerations would ordinarily preclude use of the speakers' actual names. The choice of names becomes especially important if speakers use names to refer to each other during the course of a conversation, in which case, the pseudonym in the speaker identification label should match that in the spoken reference, so that all relevant persons—whether they are speaking or merely spoken about—are clearly distinguished. If possible, pseudonyms should retain some flavor of the actual names.

(6) ((Aesthetics))

JEFF: That's all it does.

It doesn't [even] reach a conclusion.

SARAH: [mhm],

JEFF: The conclusion is up to you.

SARAH: mhm,

JEFF: in going out to --
... to buy the thing.

SARAH: Hm.
Hm.
... Okay.

When it is unclear which of several speakers on a tape is responsible for a particular utterance or noise, the letter *X* is used to label the unidentified speaker.

(7) ((Friends))
X: ((BLOWS WHISTLE))

Note that the stretch of speech between two different speaker labels constitutes, roughly speaking, the discourse unit known as a turn. The picture is somewhat complicated, however, by the listener's interjection of continuative backchannel responses (*mhm*, *yeah*, etc.) into a speaker's extended turn. Although a backchannel response must for clarity's sake bear a speaker label, as must the two sections of the turn it occurs within, one does not want to be misled by this practical consideration into overlooking the essential continuity of the extended turn unit across such fleeting interjections.

2.2. Speech Overlap

Square brackets are used to indicate the beginning (left bracket) and the ending (right bracket) of overlap between the utterances of two speakers. One set of brackets is inserted surrounding the first speaker's overlapping utterance portion, and a second set of brackets surrounds the second speaker's overlapping portion. This notation signals that the two bracketed utterance portions were uttered at the same time.

For the sake of reading clarity, the second speaker's left bracket is aligned vertically under the first speaker's left bracket (by inserting as many spaces as needed). This alignment of space on the page helps to give an iconic sense of the temporal alignment of the two overlapping utterances. Note that only the left bracket need be aligned vertically; for reasons of clarity and practicality it is not advisable to force the right bracket to do so.

(8) ((Depression))
B: Clint is still screaming about that,
...
R: [Because he wanted the stamps],
B: [all those stamps],
... Mom let Ted Kenner have.

Wherever several overlaps occur in rapid succession within a short stretch of speech, distinctive combinations of brackets (e.g., single brackets [] vs. double brackets [[]]) may be needed to make clear what is overlapping with what. This will be necessary whenever two distinct cases of overlap occur without at least one line of nonoverlapped text between them, because if only one kind of bracket were used (e.g., just single brackets) the reader could be misled to think that the first and second bracketed portions (marked with single brackets) were simultaneous with the third and fourth bracketed portions (if also marked with single rather than double brackets). After one full line of speech containing no overlaps—when there is no longer danger of confusion—the use of double brackets can be dropped and single brackets resumed. If distinctive bracketing is needed again later in the same text, the single and double brackets should be used in alternation. Whenever no other overlaps occur nearby, it is best to use just the single brackets alone.

(9) ((Aesthetics))
J: [Yeah].
S: [Which] colors ... all of the communication,
[[after]] that.
J: [[Yeah]].

When there are many overlaps in very close succession, it may occasionally be necessary to use more than two kinds of distinctive brackets. For example, triple brackets ([[[]]]) or brackets indexed with numbers ([3 3]) can be used to create distinctive bracketing.⁴

(10) ((Dinner))
B: Nobody wants [to leave].
A: [They don't] move [[out]].
S: [[Berkeley]] just keeps [3 getting 3] bigger and [4 bigger 4].
B: [3 Yeah 3],
[4 Yeah 4],
... Well it's amazing to me.

Occasionally, it may be useful to employ distinctively marked (e.g., doubled) brackets even when no other overlaps occur nearby, in order to help the reader follow a complicated conversational exchange.

⁴ The numeral 1 is avoided because it is easily mistaken for the lowercase letter *l* or even the capital *I* in many typefaces; numerals 2 through 9 engender no such confusion.

(11) ((Hypochondria))

- G: ... Well,
 the worst [[thing I ever had,
 K: [[He's a medical miracle]].
 G: was brain]] fever,
 when I had proposed to her.
 K: @@@@
 ... From which you haven't recovered.

Often enough, the second speaker in an overlap begins to speak in the middle of a word being uttered by the first speaker. In such cases, it is useful to keep track of precisely where the overlap begins, because this may carry significant information about how the speakers are responding to each other in "real time" (Schegloff, personal communication). To indicate this, the bracket is placed within the word at a point corresponding to the overlap. (When a bracket is written inside a word, no space should be inserted—whether the bracket is indexed or not—since any space would break up the word and cause it to appear as two separate words.)

(12) ((Hypochondria))

- G: ... Then I had,
 uh,
 K: Cytomegalo[virus],
 G: [Don't] forget,
 cytomegalo[[virus]].
 K: [[@]]
 D: [[What is that]].

3. TRANSITIONAL CONTINUITY

In speech, important information is carried in the speaker's intonation, encompassing fluctuations of pitch and other cues. Although a discourse transcription can never capture a complete representation of the infinite variety of possible intonation contours, it can nonetheless provide a useful representation of at least the more critical intonational information by distinguishing broad *classes* of contours. It is useful to distinguish here between *functional* and *phonetic* analyses of intonation, each of which has its place in discourse transcription. The symbols in this section deal with the functional analysis of intonation, whereas a set of symbols that address the

phonetic analysis of intonation is introduced in the section on terminal pitch (see also the section on tone).⁵

The system of categories presented in this section seeks to identify in general terms *one aspect* of intonational function, that of marking *transitional continuity*. When a speaker arrives at the end of an intonation unit, poised to continue on to the next—or not continue—the intonation contour usually gives a fairly clear indication of whether the discourse business at hand will be continued or has finished. This is transitional continuity: the marking of the degree of continuity that occurs at the transition point between one intonation unit and the next. The scope of the continuity—the question of what it is that is being continued or finished—is open-ended: A "final" contour may apply to the end of a sentence, the end of a turn, or the end of some other discourse unit. Whereas it may be possible to make finer discriminations in transitional continuity within the broad class of contours covered by each transitional continuity symbol, the distinctions among "final," "continuing," and "appeal" (see below), at least, seem to be basic.

Although the intonation contour classes in this set are defined in terms of their function, each category is more or less consistently realized by a specific form: a specific phonetic contour, or a set of contours (in effect, intonational allomorphs), where each member of the set is determined by its context. The range of phonetic realizations for a given transitional continuity class differs somewhat from one language to the next, which is one reason for using functionally based categories: They help to ensure that similar intonational functions are written similarly across languages, facilitating comparison even where phonetic realizations differ. Preliminary observations in a limited number of languages suggest that, remarkably, all languages are likely to make intonational distinctions between the transitional continuity classes presented in this section, though their phonetic realizations may vary.

The symbols used to represent transitional continuity here are drawn from those employed in written punctuation. Although using commas and periods in ways that are reminiscent of their function in written language does make it easier to remember them, it also means that the transcriber must guard against slipping into habits of thought associated with written punctuation. In discourse transcription as presented here, the punctuation symbols comma, period, and question mark *always* represent intonation classes and never grammatical or semantic structure per se.

⁵ For researchers who wish to pursue the representation of intonation in discourse further, the work of Couper-Kuhlen (1986), Cruttenden (1986), Crystal (1975), Cutler and Ladd (1983), Gumperz (1982, this volume), McLemore (1991), Svartvik (1990), Svartvik and Quirk (1980), and others should be consulted (see Couper-Kuhlen, 1986, and Cruttenden, 1986, for additional references). For the notion of intonation unit, see Chafe (this volume) and Du Bois et al. (1991).

3.1. Final

The period (.) indicates a class of intonation contours whose transitional continuity is regularly understood as *final* in a given language. For English and many other languages, this means primarily (but not exclusively) a fall to a low pitch at the end of an intonation unit. It is important to recall that, because this symbol represents an intonational category rather than a syntactic one, it can appear in places other than the end of a sentence. Conversely, it need not appear at the end of every (normative) sentence.

(13) ((Depression))

R: For what.

B: ... They make rope of it.

3.2. Continuing

The comma (,) indicates a class of intonation contours whose transitional continuity is regularly understood as *continuing*, in a given language. The contour is often realized in English as a slight rise in pitch at the end of an intonation unit (beginning from a low or mid level), but it may have other realizations as well, each of which presumably has slightly different pragmatic implications. One type of continuing contour is realized by a terminal pitch that remains level; another, by a terminal pitch that falls slightly, but not low enough to be considered final.⁶

(14) ((Ranch))

R: If you think about it,
yeah,
if it rains a lot,
the horse is always wet,
and it's always moist,
it's always on something moist,
... Sure it's going to be softer.

3.3. Appeal

The question mark (?) indicates a class of intonation contours whose transitional continuity is regularly understood as an *appeal*, in a given language. (For English, this is often realized by a marked *high rise* in pitch at

⁶ The significance of the end point of pitch movement is well known (Couper-Kuhlen, 1986, pp. 88–90). As one intonation specialist notes, "The lower the end point, the greater the degree of definiteness and conclusiveness" (Trim, 1970, p. 265, cited in Couper-Kuhlen, 1986, p. 88).

the end of the intonation unit.) "Appeal" here refers to when a speaker, in producing an utterance, seeks a validating response from a listener. The most common type of appeal in this sense is a *yes-no* question, but not all *yes-no* questions are said with the appeal contour, and in such cases the question should not be written with a question mark. Conversely, the appeal contour may be used where there is no *yes-no* question; in such cases, the question mark *is* written. For example, a speaker will often check to see if listeners remember a particular person by uttering that person's name with an appeal contour (high rising pitch), where the response sought from this appeal may be nothing more than a slight nod of recognition. In such cases, the proper name will be written with a question mark following it.

It is important to emphasize that the question mark is *not* used for a grammatical question uttered with intonations other than the appeal contour, such as a final contour. Thus, there will occur grammatical questions (including some *yes-no* questions) that do not carry this type of contour; conversely, the question mark will appear in units that lack the morphosyntactic structure of a (normative) question.

(15) ((Friends))

J: ... Should we waste him?

or should we stop him,

and ... then waste him.

4. TERMINAL PITCH DIRECTION

Whereas analysis according to functional classes (i.e., in terms of transitional continuity and/or other functional classification) captures one kind of information about an intonation contour, there is another kind of intonational information that is worth recording, involving the phonetics of the pitch movement. The symbols in this section iconically represent the movement of pitch at a critical location in the intonation unit: at the end of the unit (i.e., the transition point from one intonation unit to the next). In contrast to the symbols in the last section, which represent a certain aspect of intonational function, these symbols are designed to represent the auditory shape of the pitch movement. Naturally, no finite set of symbols can provide more than a general classification of pitch phonetics, since a complete representation would require an infinitely variable analog display. But when symbols for terminal pitch direction are supplemented with symbols for transitional continuity and tone, the combination is an effective means of capturing key features of intonation at the most reasonable cost in time and effort.

Inevitably, different researchers will wish to take different approaches to representing intonation in discourse. Among the minimalist alternatives available are to use exclusively the categories for transitional continuity, to use some other functional classification, to use just the phonetic categories for terminal pitch direction presented in this section, or to use just the categories for tone. Or the transcriber can use some combination of these, such as the transitional continuity and terminal pitch classifications—a combination that is particularly useful for those primarily interested in how extended discourse is chunked into units, rather than in the subtle and often elusive meanings distinguished by the various intonational tones. (Many other approaches to intonation are represented in the literature; see Cruttenden, 1986 and Couper-Kuhlen, 1986.) The decision about what intonational categories to use will be influenced by one's research goals and theory of intonation, and the degree of delicacy sought for a particular transcription.

4.1. Fall

A backslash (\) indicates that the direction of the terminal pitch movement is falling. This downward-sloping line iconically represents downward movement and is reminiscent of the International Phonetic Association (1989) arrow symbol (↓) for a "global fall" in pitch.

Depending on how low the endpoint of the fall reaches (relative to neighboring pitch levels), such pitch movements may be functionally assigned to the continuing or final contour classes.

(16) ((Aesthetics))

J: ... You're not saying something, \
you're doing something to people. \

4.2. Rise

A slash (/) indicates that the direction of the terminal pitch movement is rising. This upward-sloping line iconically represents upward movement and is reminiscent of the International Phonetic Association (1989) arrow symbol (↑) for a "global rise" in pitch.

Depending on the specific shape and pitch level of the rising movement, the contour may be functionally analyzed as pertaining to the continuing class (often a low or mid rise) or the appeal class (a high rise).

(17) ((Ranch))

R: ... And then, /
they videotape us, /
as we go. \

4.3. Level

Underscore (_) indicates that the direction of terminal pitch movement is level. This pitch movement is most commonly associated with the continuing contour class.

(18) ((Cars))

D: You know, _
call them on the phone, /
and uh, _
... take a lunch, /

5. ACCENT AND LENGTHENING

It is important for a discourse transcription to indicate which words are characterized by accent and length.

5.1. Primary Accent

A caret (^) indicates a word that bears a primary accent. The primary accent is characterized by its prominent pitch movement carrying intonational meaning; it is where the significant intonational "action" is focused, within the intonation unit. Primary accent is broadly comparable to the "nuclear accent" category of Crystal (1975), Cruttenden (1986), and others, which is characterized as "the most prominent syllable in a tone-unit," whose prominence is generally due to the "presence of noticeable pitch movement" (Couper-Kuhlen, 1986, p. 79).

Although there is some tendency for an intonation unit to contain exactly one primary accent, cases of two primary accents within one intonation unit are common enough. It is for this reason that we avoid the term *nuclear accent*, with its apparent presumption that each unit will contain no more than one nucleus. Also, intonation units containing no primary accent are fairly common, especially among minor intonation units (e.g., one-word intonation units) and truncated intonation units.

In English and many other languages, the particular syllable within the word on which prominence is realized is lexically predictable and thus need not be indicated in a *discourse-level* transcription.⁷ Hence the primary accent mark is written immediately before the first letter of the accented word. (For the occasional utterance of a word token in which a prominence is realized on

⁷ However, for those who prefer to write the accent mark immediately before the actual stressed syllable, there is no harm in doing this.

a syllable other than the expected one, this fact can be captured by using the notation provided for phonetic transcription [see below].) But for languages in which a word's stressed syllable is *not* lexically predictable, the primary accent notation should be written immediately before the stressed syllable (which hence may place the symbol within the word).

(19) ((Forces))

B: ^I met `him,
and I `thought he was a `ni=ce ^kid.
S: He ^is a nice `kid,
but he's ^wei=rd.

5.2. Secondary Accent

A raised vertical stroke or grave accent (˘) indicates a word that bears a secondary accent, relative to nearby primary accented and unaccented words.

(20) ((Hypochondria))

G: ... (2.2) `a=nd of course,
a `lot of herb ^tea,
when I'd `rather be drinking ^whiskey.

Because it can be difficult to distinguish reliably among three degrees of accent—primary, secondary, and (implicitly) nonaccent—some researchers may prefer to mark only two degrees of accent, corresponding to “accented” (to be written with the raised stroke, i.e., the grave accent character) versus “unaccented” (unmarked).

5.3. Booster

The exclamation point (!) can be used optionally to mark a high “booster”—very roughly, a higher than expected pitch on a word.⁸ Low booster can optionally be written with a semicolon (;). For a fuller discussion of the concept of booster, see Crystal (1975) and Cruttenden (1986). The booster symbol is written immediately before the word in question and any symbols for accent.

⁸ An extrahigh booster can optionally be written with two exclamation points (!!).

5.4. Lengthening

An equal sign (=) indicates that the preceding segment is lengthened prosodically, to a degree greater than what is expected on the basis of accent and lexical stress patterns. The slight lengthening that is to be expected when a syllable is accented is not marked with the equal sign, being implicit in the accent marking. Similarly, segments that are *phonemically* long (in a language with a contrast between long and short vowels, or long and short consonants) do not on that account receive the equal sign notation: Phonemic length should be written with a different notation (e.g., doubled letters). For sounds that are represented in standard orthography by a digraph (e.g., in English, *ee*, *ea*, *oo*, *ph*, *ch*, *tt*, etc.), the convention is that the equal sign is written after the final letter of the digraph.

Prosodic lengthening is especially important to indicate because of its role as a potential cue for intonation unit boundaries (Cruttenden, 1986, pp. 35–45): It frequently occurs at the end of an intonation unit.

(21) ((Hypochondria))

K: ... ^Greg's never had a a ^co=ld,
or the ^flu=.

6. TONE

Each major intonation unit is in general characterized by some kind of prominent pitch movement, which carries the most significant intonational information about that unit. The locus of this prominent pitch movement is generally centered on the word which bears the primary accent: either the sole primary accent, or if there is more than one in a particular intonation unit, the last one. The various distinctive intonational shapes that are possible in this position are commonly called tones. A tone's pitch contour is often realized across a spread of several words, frequently extending from the last primary accent until the end of the unit. Because the shape of this pitch contour carries the most distinctive intonational meaning in the unit, it is useful to have symbols that can at least partly capture the differences. The classification of tones remains a substantial challenge for intonation specialists, as is attested by the existence of several competing classificatory systems, each with its adherents (see, for example, the various systems described in Couper-Kuhlen, 1986, and Cruttenden, 1986, and in the many references they cite).

This section presents notations for rising, falling, rising-falling, falling-rising, and level tones. The symbol for tone is written immediately before the accented word, with no intervening space. (Alternatively, the marks can be

placed immediately before the *syllable*, rather than the word, that bears the accent.) When tone is written, it may be possible to dispense with as redundant the primary accent mark, at least on one analysis.

6.1. Fall

A backslash (\) before a primary accented word indicates that the contour associated with the accent is falling.

(22) ((Forces))

A: he can't \spell.

6.2. Rise

A slash (/) before a primary accented word indicates that the contour associated with the accent is rising.

(23) ((Hypochondria))

D: Is he going to make her become a /Catholic?

6.3. Fall-Rise

The combination backslash-slash (\/) before a primary accented word indicates that the contour associated with the accent is first falling, then rising. This pitch movement can cooccur with any of the transitional continuity classes, though it is more common with continuing and final than with appeal.

(24) ((Ranch))

R: If you \think about it,
 yeah,
 if it /rains a lot,
 the horse is always \wet,
 and it's always /moist,
 it's always on something \moist,
 ... \Sure it's going to be softer.

6.4. Rise-Fall

The combination slash-backslash (/ \) before a primary accented word indicates that the intonation contour associated with the accent is first rising,

then falling. This pitch movement often cooccurs with a widened pitch range, which may be interpreted as expressing "high involvement". The transitional continuity class it is most often associated with is final.

(25) ((Aesthetics))

S: ... A lot of it's really /\bad.

6.5. Level

An underscore (_) before a primary accented word indicates that the contour associated with the accent is level.

(26) ((Hypochondria))

K: ... (1.2) They just _represent,
 each of the _days,

7. PAUSE

The placement and timing of pauses in spoken discourse conveys significant information about the speaker's discourse production process (Chafe, 1980b) and orientation toward the ongoing conversational interaction (e.g., Goodwin, 1981). Each pause should be indicated explicitly using one of the three notations presented in this section. As the intonational symbols (e.g., comma and single period) do not of themselves denote pause, any pause—even a slight one—that occurs in conjunction with an intonation contour must be specifically indicated using one of the pause notations.

7.1. Long

A sequence of three dots (...) immediately followed by a number in single parentheses is used to represent relatively long pauses (.7 seconds or longer). The approximate duration is indicated within parentheses to the nearest 10th of a second. That is, the duration is indicated as (.7), (.8), (1.6), and so on.

(27) ((Ranch))

R: ... This .. is a type of person,
 ...(.9) that ...(.7) is like ... (1.0) a hermit.

Ordinarily, a pause between two intonation units is written together with the unit that follows it (never with the one that precedes it). However, if a

pause is attributable to more than one speaker (as when, during a long pause, it is unclear who is going to speak next), it is often preferable to place the pause notation on a separate line by itself. In some cases, the questions of who a pause belongs to, how long it lasts, and even whether it has occurred in a specific place, become subtly and inextricably linked to the interpretation of turn-taking and overlapping between speakers.

(28) ((Depression))

B: ... I remember,
 ...(.8) I used to help Billy,
 and I'd get twenty-five cents a week,
 ...(.1.2)
 R: [A week].
 B: [Twenty] --

Whereas some researchers use subjective judgments of pause duration relativized to each speaker's current tempo (a "second" for a fast speaker is objectively shorter than a "second" for a slow speaker), this is in general not advisable, due to the difficulties in making such judgments consistently and reliably and in interpreting the "time" notations that result. Among other things, if a pause occurs at a turn boundary between the utterances of two speakers with different tempos, it is unclear which speaker should be used as the basis for relativizing the duration. Even a pause within the speech of a single speaker can be problematic if it occurs between a rapid stretch of syllables and a slow stretch—a fairly common configuration in everyday speech. Unless such ambiguities can be addressed, the only reliable practice is to indicate the actual pause duration in clock time (preferably as measured instrumentally).

7.2. Medium

A sequence of three dots (...) indicates a pause of medium length—one that is noticeable but not very long, about half a second in duration (specifically, between 0.3 and 0.6 seconds, inclusive).

(29) ((Hypochondria))

G: ...(.1.7) I'd like to have .. my ... lungs,
 ... my entire respiratory tract,
 ... replaced,
 ... with .. asbestos.
 .. or something.

7.3. Short

A sequence of two dots (..) indicates a brief break in speech rhythm, that is, a very short, barely perceptible pause (about 0.2 seconds or less).

(30) ((Ranch))

R: .. a reining pattern is,
 .. a pattern where you .. do sliding stops,
 .. spins,
 ... lead changes,
 .. I know you probably don't know what that is.

It is important to note that not all brief silences are to be classified as pauses. The moment of silence that necessarily occurs during a lexically or phonologically required voiceless stop should not be classified as a pause, even if it is longer than expected (as in an emphatic or "marcato" pronunciation of a word containing a voiceless stop). The reason for this is that for discourse research what matters is the pause as a functional cue to aspects of discourse production and conversational interaction, not as a raw acoustic fact.

7.4. Latching

A zero within single parentheses (0) indicates that the following utterance "latches" onto the preceding utterance, that is, there is no pause (or "zero" pause) between the two speakers' turns. Because it symbolizes a noticeable lack of pause between actual *turns*, mere continuative backchannel responses (*m=hm*, etc.) are not ordinarily marked with this symbol.

(31) ((Cars))

G: .. I was using number seven,
 .. gun number seven,
 D: (0) It broke the[chisel].
 G: [and] it broke my chisel,
 man.
 <X Now X> --
 D: (0) So now you have no chisel.
 G: (0) <X It's X> my only good chisel.
 man,

Since simply not writing in any pause notation—not even a two-dot pause—will already serve to suggest the absence of a pause, the latching notation is to some degree redundant. This plus the fact that determining the presence of latching presupposes a potentially difficult judgment about the turn (or nonturn) status of an utterance leads some researchers to avoid this transcriptional category.

8. VOCAL NOISES

The participants in a conversation do more with their vocal tracts than just utter words: They also cough, yawn, click, inhale, laugh, and produce a variety of other noises. The notations in this section are designed to allow the transcriber to easily notate nonverbal noises produced in the vocal tracts of speech event participants. The reason for distinguishing vocal noises made by speech event participants as a special category is that participants often use this channel to give each other subtle cues about aspects of the ongoing linguistic interaction, as when a speaker takes a sharp in-breath in order to signal the purpose to speak next (Sacks, Schegloff, & Jefferson, 1974). Crickets chirping and microphones rustling do not consistently carry such interpersonal meanings for humans.

8.1. Vocal Noises

Single parentheses surrounding a description written in capital letters (COUGH) are used to indicate nonverbal noises produced in the vocal tract of speech event participants. This kind of notation encompasses coughing, throat clearing, tongue clicking, breathing, and so on, but not dish washing, finger drumming, dogs barking, and so forth (for which double parentheses are available; see below).

The capital letters and parentheses help to make it clear that the words so written were not actually uttered by the speaker; that is, rather than *saying* the word *cough*, the speaker *did* cough. (For some high-frequency vocal noises, a special nonalphabetic symbol is used—e.g., @ for laughter—and in such cases the parentheses are unnecessary.)

The notation (THROAT) indicates the sound made by someone clearing their throat. Similarly, (GULP) can be used to represent a gulping sound, and (SWALLOW), (SNIFF), (SNORT), (BURP), and (YAWN) likewise represent the indicated sounds. Additional notations in this format can be generated as needed for indicating other vocal noises.

(32) ((Aesthetics))

S: (H) (THROAT)

Yeah.

The notation (TSK) indicates the utterance of a click of the tongue—in English this is usually an alveolar click—as an isolated vocal sound, for example, what is commonly written *tsk* in newspaper cartoon style.

(33) ((Ranch))

R: and then,

... (TSK) our job,

is to shape the shoe,

... to the horse's foot.

8.2. Glottal Stop

The percent sign (%) indicates a *paralinguistically* introduced glottal stop or glottal constriction. This notation is *not* used in positions where glottal stop is phonologically predictable, as at the beginning of vowel-initial words (under certain conditions) in English. Nor is this notation used where glottal stop is lexically required, as in certain words in languages with *phonemic* glottal stop. (For “creaky” or glottalized voice quality extending over whole words or longer stretches of speech, see “Quality” below.)

One reason for taking the trouble to transcribe paralinguistic glottal stop is that speakers often seem to produce it when they abandon a word or utterance. To the extent that glottal stop functions as an objective cue for abandoned utterances, it is useful to have it on record. Glottal stop and glottal constriction may act as cues to other aspects of the discourse production process as well.

(34) ((Ranch))

R: it's mandatory,

you have to --

% to graduate,

you know,

% well,

to ... get the degree,

you know,

... you have to take this class.

8.3. Inhalation

A capital *H* in single parentheses (**H**) indicates audible inhalation.⁹

In conversation, breathing is more than just a bodily necessity; it can be used, for example, as a signal that one is about to take a turn at speaking (Jefferson, 1984, p. 353f.; Sacks et al., 1974).

(35) ((Hypochondria))

K: ... (**H**) leukemia,
... (**H**) bronchitis,
... (**H**) uh,
tuberculosis,
@@@@ (**H**)
and he's recovered from all of them.

Where a pause and a quiet inhalation occur in immediate succession, it is often difficult to separate the two (e.g., in order to time the pause). In such cases, it may be preferable to write the pause and inhalation together with no intervening space and to assign any indication of duration to the pause-plus-inhalation complex taken as a whole.

8.4. Exhalation

A capital *H* followed by a small *x* within single parentheses (**Hx**) indicates audible exhalation.¹⁰

(36) ((Depression))

B: ... (4.3) (**Hx**) ... Kids in the city miss so much.

Sometimes a speaker audibly inhales and exhales several times in immediate succession. All of this can be written within a single set of parentheses: (**H Hx H Hx**).

Note that neither the inhalation symbol (**H**) nor the exhalation symbol (**Hx**) is used within a word (e.g., for breathy voiced segments, laughter, etc.).

⁹ For some purposes, it may be useful to make use of multiple *H*s to represent iconically the relative duration of a long inhalation: (HHHHHHHHHH). For example, this may be called for if another speaker overlaps with the inhalation and one wishes to show the exact point where overlap begins and ends.

¹⁰ Again, in some circumstances it may be helpful to make use of a series of *H*s to iconically represent the duration of a long exhalation: (HHHHHx). Note that since the lowercase *x* in this notation acts in effect as a sort of "subscript" attached to the string of *H*s, it is written only once.

Because of the serious potential for confusion that an ambiguous use of *H* would introduce, other notations using discriminable characters (Du Bois, 1991) are preferable (see the next section).

8.5. Laughter

The @ symbol is used to represent laughter. One token of the symbol @ is used for each "syllable," or pulse, of laughter. (The @ symbol bears a certain mnemonic resemblance to the pervasive "smiley face" icon.)

Because the placement of laughter can be of great consequence for a conversational interaction (Jefferson, 1979, 1984), it is important to note it carefully. Although laughter falls in the category of nonverbal vocal tract sounds, and so by present conventions could in principle be written within single parentheses (i.e., as (LAUGH)), it occurs so pervasively that it warrants its own distinctive symbol. The @ symbol has the additional advantages of being easily reiterated in a minimum of space—allowing the duration of the laughter in "syllables" to be represented iconically—and of being readily discriminable when written within a word.

(37) ((Hypochondria))

K: @@@@
... From which you haven't recovered.

For laughter of extended duration, the transcriber may prefer to write just a single laugh symbol followed by an indication of duration—if it is not easy to determine how many syllables of laughter there are (as is often the case when several people are laughing at once) or if the investigator is simply not especially interested in how many laugh syllables have occurred.

(38) ((Miracle))

ALL: @(12.7)

Sometimes it is useful to distinguish between different kinds of laughter. For example, the symbol @N can be used for nasal laughter, a usually voiceless laugh in which the air is emitted through the nose. To the extent that further distinctions among kinds of laughter may be significant (Jefferson, 1979), such distinctions can be indicated by suffixing various characters to the @ symbol as modifiers of it, with the resulting complexes (e.g., @I@I, @A@A) defined by the researcher. (The unmarked symbol for all kinds of laughter, however, is simply @.)

(39) ((Aesthetics))

J: ... You're not supposed to use these powerful[techniques].

S: [@N@N@N]

... Hm.

9. QUALITY

There are many occasions in conversation where, for a stretch of a few words or lines, a speaker's voice takes on some special quality, shifts in pitch, or slows in tempo, and so on. Because this kind of momentary marked quality or prosody can serve important functions in exposing some perhaps un verbalized aspect of the speaker's stance or the speech production process, and because it can have consequences for the ongoing spoken interaction, it is important to record it. But because the special qualities that can occur are so diverse, the notation must be flexible enough to meet any demands that may arise in the discourse material. The notational formula introduced in this section is designed to accommodate this kind of diversity.

9.1. Types of Quality

Angle brackets <Y Y> are used to indicate that the stretch of text which they enclose has a marked quality or prosody of some sort. The particular quality (higher pitch, increased loudness, etc.) is specified by a supplementary symbol, represented here by Y. The text enclosed within these symbols often amounts to several words and may run across several lines. The marked quality is judged relative to the surrounding discourse produced by the same speaker. For example, a sentence would be marked for tempo if it is noticeably quicker or slower than the speaker's current or usual tempo. This set of symbols (partly based on Boase, 1990) is in principle open-ended, and new ones can be developed to suit a particular investigator's needs. For most transcribing purposes, these notations are used sparingly, to indicate just those phenomena that are of special interest and consequence for the spoken interaction.

Listed here are some of the more common types.

Loudness

<F	F>	forte: loud
<P	P>	piano: soft
<CR	CR>	crescendo: gradually louder
<DIM	DIM>	diminuendo: gradually softer

Pitch

<HI	HI>	higher pitch level
<LO	LO>	lowered pitch level
<W	W>	widened pitch range
<N	N>	narrowed pitch range
<PAR	PAR>	parenthetical prosody

Tempo and Rhythm

<A	A>	allegro: rapid speech
<L	L>	lento: slow speech
<RH	RH>	rhythmic: stresses in a beatable rhythm
<MRC	MRC>	marcato: each word distinct and emphasized
<ARH	ARH>	arrhythmic: halting speech

Voice quality

<WH	WH>	whispered
<BR	BR>	breathy
<HSK	HSK>	husky
<%	%>	creaky (or: <CRK CRK>)
<FAL	FAL>	falsestto
<TRM	TRM>	tremulous
<SOB	SOB>	sobbing
<CRY	CRY>	crying
<YWN	YWN>	yawning
<SGH	SGH>	sighing

Following are instances of several of these notations. The angle-bracket pair <F F> is used to enclose a stretch of forte speech (produced with relatively increased loudness).

(40) ((Hypochondria))

A: <F It's not the end of Chanukah F>,
in case you're interested.

The angle-bracket pair <P P> is used to enclose a stretch of piano speech.

(41) ((Ranch))

R: ... But uh,
 ... (3.0) <P What was I going to say P>,
 ... (3.5) X- --
 Oh,
 it's really tiring,
 though.

The angle bracket notation <W W> marks widened pitch range. (This is a marked prosody often interpreted as displaying "high involvement" or "surprise.") The increased pitch range is often accompanied by sudden pitch movement and in English is frequently associated with a pronounced rise-fall tone, which may be accompanied by increased loudness.

(42) ((Cars))

D: ... No basketball.
 G: ... <W Really W>.

The angle-bracket pair <MRC MRC> is used for a stretch of marcato speech, in which each word is uttered distinctly and with emphasis.

(43) ((Friends))

J: ... But the goldfish got stuck,
 ... <MRC halfway into his mouth MRC>.

The angle-bracket pair <WH WH> is used to enclose words uttered in a whisper.

(44) ((Africa))

A: they let us alone.
 ... <WH But we were scared,
 And boy WH>,
 did we ever get in trouble,
 from Mel and Ervin.

The angle-bracket pair <% %> (alternatively <CRK CRK>) indicates creakiness or glottalization of the enclosed words.

(45) ((Aesthetics))

J: <% Tha%- this% --
 I wonder about that though,
 I mean %>,
 when I think of ads,

In cases where it seems useful to specify the precise location of a special quality that begins and/or ends at some point within a word, an underscore can be added to the usual angle bracket notation (e.g., <WH_ _WH> or <%_ _%>) so as to separate the (capital) letters of the quality notation from the letters of the word they enclose. Thus, if just the fourth through sixth syllables of the word *cytomegalovirus* were whispered, this could be written as "cyto<WH_megalo_WH>virus." Where one is not so concerned to avoid ambiguity, the underscore symbols could be dispensed with. This works especially well with nonalphabetic notations like those for creaky quality ("cyto<%megalo%>virus") or laugh quality ("cyto<@megalo@>virus"; see next section). This word-internal quality notation is likely to be used but rarely.

9.2. Laugh Quality

The angle bracket pair <@ @> indicates a laughing quality over a stretch of speaking, that is, laughter during the words enclosed between the two @ symbols. (The angle brackets can be combined with notations for other kinds of laughter as well, e.g., <@N @N>, etc.)

(46) ((Africa))

A: ... and they stepped out in the road,
 and not only did they have uniforms on,
 but they <@ also had guns @>.
 @ @ @

If a laugh occurs during the utterance of just one word, this can also be indicated simply by prefixing the word with one @ sign and dispensing with the angle brackets.

(47) ((Friends))

N: You know,
 this was a rented @snake,
 @

For most transcription purposes, it is sufficient to use the laugh quality brackets to frame whole words or groups of words (the convention followed in this work). But some researchers may wish to indicate on which particular syllables within a word laughter occurs. To do this, each pulse (syllable) of laughter receives one @ token, which is written within the word at the appropriate place, before the laugh-tinged sounds.

(48) ((Lunch))

R: ... (1.0) When they quit going to Littleton,
every week to see his gra@ndmo@the@r @@@,

Sometimes a speaker speaks with a smile rather than a laugh, causing their speech to be tinged with an audible "smile" quality. If desired, this can be written with laugh brackets with the letters *SM* affixed: <@SM @SM>.

9.3. Quotation Quality

The angle-bracket pair <Q Q> indicates a stretch of speech characterized by a "quotation" quality. Its use is warranted where there is some actual shift in the quality of the stretch of quoted speech, as when the quoting speaker imitates some mannerism of the quoted speaker. Where no such shift is audible, this notation should not be used.¹¹

(49) ((Friends))

J: This is a literal quote,
he says to me,
... <Q I'm going to restrain you.
to the fence Q>.

Note that the quotation symbol is not used for metalanguage, such as the name of a letter or a reference to a word (Du Bois et al., 1991), unless this is accompanied by an audible quotation quality.

9.4. Multiple Quality Features

When a stretch of speech is characterized by two or more coextensive special qualities worth noting, these are indicated with multiple angle brackets <Y<Z

¹¹ Some may wish to use plain double quotation marks (" ") for marking quotations that do not carry a special voice quality; but this should be recognized as part of functional coding rather than transcription per se.

Z>Y>. (The several angle-bracket notations are juxtaposed without any space between them.)

(50) ((Friends))

J: So the guy yells at me,
... <Q<F Is that your dog F>Q>?

10. PHONETICS

Although a discourse transcription does not generally seek to represent every variation in pronunciation, there are times when the question of how a word was pronounced takes on immediate significance for the spoken interaction. In such cases, it is useful to have available a way of writing that can unambiguously indicate the actual pronunciation of a particular word or words—without, hopefully, requiring too much in the way of special knowledge or special characters. This section presents a way of citing phonetic (or phonemic) transcriptions for selected words. (A set of symbols that can be used for making precise phonetic transcriptions without requiring special characters is provided in Du Bois et al., 1991.)

10.1. Phonetic/Phonemic Transcription

A symbol complex composed of slashes surrounded by single parentheses (/ /) is used to enclose a representation of the actual pronunciation of a word. The phonetic (or phonemic) transcription is given *in addition* to the traditional orthographic representation of the same word(s), which it follows.

The transcription itself can be written in several different ways, depending on the degree of precision sought and the enthusiasm of the transcriber. The following example illustrates the option of supplementing standard orthography with selected phonetic symbols—in this case, stress marks—in order to represent just enough of the actual pronunciation to allow the interchange to be understood.

(51) ((Cafe))

A: Virago_(/ `Virago/).
C: ... Virago_(/ `Virago/)?
A: ... I don't know how you pronounce it.
B: [I thought it was] Virago_(/Vi`rago/),
A: [<X Does X> this] --

The next example illustrates a more precise, and more ambitious, style of phonemic transcription (Du Bois et al., 1991).

(52) ((Comparative))

G: But this Naiman_(/`naIm6n/) book,
or Naiman_(/`neIm6n/),
I don't know how he says his name,

In general, phonetic transcription is used only where the actual pronunciation of a word is of special significance for the analyst's purposes. Most of the time, standard orthography used alone will be sufficient. A sparing use of phonetic detail notations has the important advantage of making transcriptions easier to read.

11. TRANSCRIBER'S PERSPECTIVE

In addition to symbols for representing speech per se, the transcriber occasionally needs to indicate some aspect of his or her perspective on the transcription—in effect, a metatranscriptional interjection. This section provides several symbols that allow the transcriber to insert useful comments or observations, while keeping such interjections clearly distinct from the actual speech.

11.1. Researcher's Comment

A pair of double parentheses (()) encloses any comment the transcriber or researcher chooses to make. The comment is written all in capital letters, in order to make it quite clear to the reader that the words in question are not actual speech. Comments interjected into the transcription in this way are best kept short, for the sake of a readable transcription.

This notation is also used for indicating any nonlinguistic events that take place within the spoken interaction, such as ambient noises or other noises (excluding vocal noises). But such sounds and other events are usually noted only if they are relevant to the conversational interaction at hand—as when participants comment on or otherwise react to the noise.

(53) ((Friends))

N: and they're,
... you know,
... ((DOG BARKS EXCITEDLY))
...

J: You know --

You know about this piece?

N: She always does that. ((REFERENCE TO DOG))

If a researcher plans to make fairly extended or pervasive commentary—for example, commenting on every turn—it may be preferable to set up a column format, using one side of the page for transcription and the other side for commentary.

11.2. Uncertain Hearing

A pair of angle brackets <X X> marked with the capital letter X—the X suggesting an unknown quantity—is used to mark portions of the text that are not clearly audible to the transcriber, to such an extent that there is some doubt as to what words were spoken. The words so enclosed represent the transcriber's best guess as to what was said, but their accuracy is not assured.¹²

(54) ((Hypochondria))

G: ... Well,

I [don't] normally sound like Lucille Ball.

K: [<X That's X>] --

11.3. Indecipherable Syllable

The capital letter X (again, mnemonically suggesting an unknown quantity) is used to indicate speech that is not audible enough to allow a reasonable guess as to what was said. One X is used for each syllable of indecipherable speech. It is usually possible to make at least a rough estimate of how many syllables were uttered, even when one cannot make out what the words are.

(55) ((Cars))

D: It was basically me,

you know,

X going out.

The problem of going out.

¹² If one is unable to decide between two possible hearings of a stretch of speech, one can indicate both alternatives, as follows:

<X words X><X2 other words X2>

This device should be used most sparingly, however. If the words can be made out at all, it is almost always possible to decide which alternative is the more likely.

The methods and conventions presented in the remaining sections in this article deal either with specialized transcriptional categories or with research practices which, while not strictly speaking part of transcription per se, are closely linked to the production and use of discourse transcriptions.

12. DURATION

12.1. Duration of Simple Events

A number in parentheses (N) may be used to indicate the duration in seconds of any inhalation, hesitation, word, laugh, or other event that is of special interest. For instance, if an inhalation or exhalation is significantly long, its duration can be indicated in the same manner as for pauses, that is, with a number in parentheses immediately following it. In the following example, the notation indicates that the inhalation lasts 0.9 seconds.

(56) ((Forces))

A: ... (1.0) (H) (.9) A=nd,

Similarly, in the following example, the hesitation word *um* (a "filled pause") is held for 0.7 seconds.

(57) ((Aesthetics))

S: u=m(.7),

Aside from its use for notating pause duration, for most transcribing purposes this degree of delicacy will not often be needed.

12.2. Duration of Complex Events

The duration of a complex event (a sequence of pauses and hesitation words, for instance) can be indicated, when it is of special interest, using angle brackets (cf. Chafe, 1980a, p. 301). The duration of the items to be timed (written in parentheses in the usual way) is affixed to both the left and right brackets <(N) (N)>. Because proliferation of this kind of detail can quickly make a transcription difficult to read, for most purposes it will be used but rarely.

(58) ((Ranch))

R: <(1.3) % .. (H) %
... % .. (1.3)> But .. uh,

13. SPECIALIZED NOTATIONS

This section presents a variety of specialized or miscellaneous notations and conventions. Some of the notations are for phenomena that are but rarely encountered, whereas other notations are of specialized interest or application.

13.1. Intonation Unit Continued

An ampersand (&) is used to mark each of the two halves of an intonation unit that for one reason or another the transcriber has split up and written on two lines. (It is not used when a unit is merely too long to fit on one line; for that, see the section on presentation.)

This is a notation that is needed only rarely. But occasionally, the complex realities of conversational interaction bring two fundamental representational principles of the present transcription system into conflict. First, vertical space on the page iconically represents the sequential order of turns (and the passage of time). Second, each intonation unit appears on a single line. But what is to be done when a speaker starts an intonation unit, pauses, and then finishes it, while a second speaker interjects a whole turn during the pause? In order to preserve (as far as possible) the principle that lines written higher on the page represent earlier turns, it is necessary, on rare occasions such as these, to break up an intonation unit into two lines. When this is done, the ampersands are used to mark the continuity of the unit across the intervening material. In such cases, even though the words appear on two separate lines, they should nevertheless be considered part of a single intonation unit.

(59) ((Lunch))

R: When he was real little,
[he] almost died of pneumonia.

L: [Oh].

R: when he was &

L: Oh really?

M: Hey.

R: & three.

13.2. Accent Unit Boundary

The "pipe" symbol (|) is used by some researchers to separate one accent unit from the next, where more than one accent unit occurs within one intonation

unit (Chafe, this volume). The accent unit boundary represents a juncture that marks the scope of a nuclear accent.¹³

13.3. Embedded Intonation Unit

Angle brackets marked with pipe symbols (<| |>) may be used to enclose an embedded intonation unit.¹⁴ This occurs where a larger intonation unit is temporarily interrupted while a parenthetical utterance—at a different pitch level—is inserted, after which the larger intonation unit is resumed. The impression given is that if the interrupting phrase were suppressed, the remaining material would fit together as a single coherent intonation unit. This potentially controversial category sometimes occurs with utterances of hesitation words such as *uh* or phrases such as *you know*.

13.4. Restart

A capital initial letter can be used to roughly indicate a “restart,” that is, the start of a new unit or a restart after a false start. Speakers may signal a restart by shifting to a new baseline intonation level: A higher initial pitch level from which subsequent pitches will gradually tend to drift down over the next stretch of speech (i.e., “declination”; see Schuetze-Coburn, Shapley, & Weber, 1991), until a new restart begins the process all over again. (Capital letters are also used in the standard way for the first letter of a proper noun, the pronoun *I* in English, and so on.)

(60) ((Hypochondria))

K: But he'll recover,

He'll --

D: What is that.

K: He'll be over his leprosy [soon].

G: [Nothing],

it's just dry skin.

... There isn't --

It's no disease,

at all.

¹³ This usage is similar to that of the International Phonetic Association (1989) for marking a “minor (foot) group.”

¹⁴ Regarding embedded tone units, a comparable but distinct category, see Boase (1990) and Svartvik and Quirk (1980).

K: Athletic feet.

... foot.

D: foot.

13.5. False Start

Plain angle brackets <> are used to enclose words that are false starts or “editables”—when such indication is desired.

For a widely known language such as English it is probably best to avoid inserting implicit judgments about correctness and repair at the transcription level (Edwards, 1992). (Such interpretations are commonplace, and fully appropriate, at the more interpretive and theory-bound level of *coding*.) But the picture changes when one considers little-known languages. A linguist who publishes a transcription of a language that is known by only a few individuals in the world would do a decided disservice to simply reproduce all the words as spoken, without any indication of which were considered correct and which were not, in the eyes of the native speaker. This is, after all, the kind of knowledge which native speakers of English make use of implicitly whenever they read and understand an English-language transcription that does not explicitly alert them to the dysfluencies it contains. But in a little-known language, such knowledge may well be inaccessible to any but the linguist who published the text and one or more native speakers in a faraway place.

Thus, while for most purposes one would not specially mark false starts in a transcription of English discourse, one should do so in, for example, a language such as Xinca or Sacapultec Maya. (English examples are presented in the following with this notation just to illustrate how it would be used.)

(61) ((Door))

A: and <they> --

they poked into <the-> the molding,

along the [side].

B: [unhhunh],

13.6. Code Switching

Angle brackets labeled with *L2* (<L2 L2>) may be used to mark stretches where the speaker has shifted into a language different from the one he or she has been speaking, or from the one that dominates the current conversation. If several languages are involved, each can be indicated by its own number: <L3 L3>, <L4 L4>, and so on. Alternatively, more mnemonic (if more cumbersome) codes can be assigned: <L2SP L2SP> for Spanish, <L2XIN

L2XIN> for Xinka, and so on. In either case, a key should be given in the header of the transcription, spelling out the full name of each language so abbreviated. (For example, such a key would indicate that in the transcription from which the following example is taken, L2 equals Spanish.)

(62) ((Dig))

A: So we don't really know if it was the <L2 vice-rector L2>.

Although this notation goes somewhat beyond pure transcription per se, it is useful for making clear to the reader when code switching has taken place and for ensuring that computer searches will not mix up words from two different languages.

13.7. Marginal Words

When listening to ordinary conversation, the transcriber is always confronted with a few words and sounds for which ordinary spelling conventions—designed for written language—offer little or no guidance. The transcriber faced with such a word, rather than simply inventing an ad hoc spelling that may or may not be recognized by other readers, should preferably follow some sort of standard practice.

In this section, we present some suggestions on how to spell various marginal words (or “vocalizations”; see Tottie, 1989)—such as those used in filled pauses, backchannel responses, and so on—so that they can be transcribed consistently, allowing for both ease of reading and automatic identification. For some of these words, spelling can be derived from an already existing informal spelling convention discernable in the practice of playwrights, novelists, and especially cartoonists. The spellings *uh*, *unh*, *um* represent hesitation words (filled pauses); *hm*, *m*, *huh*, and *hunh* express various nuances of awareness, wonder, or other backchannel response; *mhm*, *unhhunh*, and *uhuh*, (all with final syllable stressed) are backchannels or affirmative responses; *unh-unh* (with the first syllable stressed) is a negative response; and *uh-oh* is a mild alarm cry. (The last two words are pronounced with a glottal stop between the two syllables.)

(63) ((Aesthetics))

J: I think of ... aesthetics,
and,

S: *mhm*,

J: *uh*,

...

S: *Hm*.

... @

...

J: ... creation of desire,
for one thing.

S: *mhm*,

14. NONTRANSCRIPTION LINES

It is useful to include a certain amount of background or “bookkeeping” information about the text being transcribed, within the text file itself. When this is done, the lines containing background information should be carefully distinguished from actual transcription lines. Other kinds of nontranscription information, such as interlinear gloss lines, should be distinguished as well.

14.1. Nontranscription Line

The dollar sign (\$) marks any line in a transcription file that is not part of the transcription per se but that encodes other useful information. Examples might include lines indicating the title of a transcribed text, the transcriber's name, the recording date, and so on. In such lines, it is helpful to use the colon to mark the boundary between the information category label and the specific information that falls under that category.

\$ TRANSCRIPTION TITLE:	Door Story
\$ TAPE TITLE:	Door
\$ FILENAME:	door.tn
\$ PRINTOUT DATE:	(etc., etc.)
\$ RECORDING DATE:	
\$ RECORDED BY:	
\$ SPEAKER 1:	
\$ SPEAKER 2:	
(etc.)	

14.2. Interlinear Gloss Line

For many languages (especially relatively little-known ones), it is advisable to include, along with the transcription itself, an interlinear gloss line and/or free

translation, as in the following example from a conversation in Sacapultec Maya:¹⁵

(64) ((SacPear 2))

S: ... k-inijel x-ee-b'eeek?
 ... Erg3pl-all Cp-Abs3pl-go
 ... They all left?

Note that including a pause symbol in the second (gloss) line is not merely redundant: It helps to ensure that each word (and morpheme) in the text line is unambiguously aligned with, and can be automatically matched to, its gloss in the line below.

If for computational purposes it becomes necessary to make clear that the gloss and free translation lines do not represent actual speech, they may be marked with a backslash or other distinctive sign plus capital letter at the beginning of the line: \G for the gloss line, \F for the free translation line, \M for morphosyntactic category, and so on.

(65) ((SacPear 2))

\T I: .. e re-'en x-0-inw-il ta=j.
 \G .. FOC the-Abs1sg Cp-Abs3-Erg1sg-seenot
 \F .. I didn't see it.

If need be, interlinear lines marked with a distinctive sign can also be used to introduce certain types of specialized transcription information. For example, for transcribing videotape, a separate line beginning with \EYE could be placed above each transcription line to record the eye gaze of speech event participants, as iconically synchronized to their simultaneous verbal utterances (see Goodwin, 1981). (This notation must be considered "fragile" [Du Bois, 1991], however, because the indication of temporal synchronization depends on maintaining the vertical alignment of the two lines.)

15. RESERVED SYMBOLS

In any transcription system designed for general discourse research, allowance must be made for recording certain kinds of specialized information, which may differ from language to language and from researcher to researcher. This information may include some kinds that are not strictly speaking part of

¹⁵ For an explanation of the glosses used in this and the next example, see Du Bois (1987).

discourse transcription per se. Language-specific spelling conventions and phonemic orthography, as well as coding of morphosyntactic categories and structure, may each call for the use of some specialized symbols. Some of the symbols that are not used for discourse transcription need to be reserved for this; this section presents suggested notations for each of these domains. In addition, a few symbols are left undefined, free to accommodate the diverse special needs of users of the system. Naturally, different researchers will have different requirements, and even the symbols spoken about here as "reserved" are available to be exploited for other purposes if they are not needed for the purposes described.

15.1. Phonemic/Orthographic Symbols

Apostrophe (') should be reserved for contractions (*she'll*, *don't*) in English and other similar orthographies. In other languages, it may be needed for representing palatalized consonants, ejective consonants, and so on, according to the orthographic conventions of the language in question.

15.2. Morphosyntactic Coding

Researchers who want to study the morphological and/or syntactic structures in their spoken discourse data need to reserve a certain number of symbols for coding purposes. Probably the most important need is for indicating morpheme boundaries (in languages where this is desirable), for which the plus sign (+) can be reserved.¹⁶ For other, more specialized forms of morphosyntactic coding, the following symbols may be reserved: asterisk (*), number sign (#), and curly brackets ({}). (Du Bois & Schuetze-Coburn, this volume). Of course, if these symbols are not needed for morphosyntactic coding they can be freely used for other purposes.

15.3. User-Definable Symbols

Several symbols have deliberately been left without a specific definition in this system to give researchers room to expand the system to meet their special

¹⁶ Alternatively, hyphen (-) can be used for morpheme boundaries. The hyphen is undoubtedly the most widely used symbol for morpheme boundary, but it is also rather widely used for word truncation (as in this system). It is possible to use hyphen for both morpheme boundary and truncation, given their distinct environments—morpheme boundary is generally word-internal and truncation is word-external—as long as the researcher takes care to distinguish the two functions when appropriate (e.g. by using contextual cues). Alternatively, the plus can be used for truncation and the hyphen for morpheme boundary by simply reversing the symbol values proposed here.

needs. The double quotation mark (")¹⁷ or tilde (~) can be combined with numbers, letters, or other symbols to form digraphs (Y", 2", &", ~A, ~B, etc.), and in this way new symbols can be generated as needed. Also, researchers who do not subscribe to a particular transcriptional category (such as the accent unit) can redefine the symbol in question to fit their needs.

Among complex notations, the angle bracket notation discussed earlier allows for constructing an open-ended set of user-defined symbols for features that apply over a stretch of discourse. And the single parenthesis notation allows for creation of an unlimited set of symbols for vocal noises.

16. PRESENTATION

One important use of transcriptions is for illustrating some discourse phenomenon in an article or book, or for presenting some analysis of it. Usually, the attention of the reader is being directed to some particular feature within the discourse extract in question. Thus, in addition to the symbols for transcription per se, certain conventions for the *presentation* of transcriptions are useful, for which we here present some suggestions, based in part on the practice of the Conversation Analysis tradition (Atkinson & Heritage, 1984, p. xvi).

A salient line of text can be indicated by placing a visually prominent symbol, such as an arrow or bullet, in the left margin of the line. To call attention to individual words, they can be boldfaced, underlined, or italicized. To show where some linguistic material (e.g., several lines of nonpertinent text) has been omitted from a discourse example, the number of lines left out can be given within double parentheses ((6 LINES OMITTED)).

Occasionally, an intonation unit is too long to fit on one line. Whenever typographical considerations make it necessary to break a long intonation unit into two successive lines on the page, the remainder (the portion that is shifted down onto the second line) should be set indented five to ten spaces from beneath the first word of the line above.

To allow easy reference to specific places in one's texts, each line can be numbered consecutively in the left or right margin, beginning with the numeral 1 for the first line, 2 for the second, and so on (or, every fifth line can be so marked, etc.).

¹⁷ But see a possible use for the double quotation mark (") under "Quotation Quality."

ACKNOWLEDGMENTS

This chapter was based on research supported by the National Science Foundation under grant No. IST85-19924 ("Information Transfer Constraints and Strategies in Natural Language Communication," John W. Du Bois, Principal Investigator), which we gratefully acknowledge. Additional support was received from the University of California, Santa Barbara Office of Instructional Development and from the Center for the Study of Discourse at the Community and Organization Research Institute of the University of California, Santa Barbara.

The discourse transcription system described in this chapter draws substantially from the work of many people. Discourse researchers from a wide variety of orientations have contributed insights, techniques, and perspective to the present formulation, and many will doubtless recognize in this document their own contributions. Among the most direct influences have been those of Wallace Chafe, Norman McQuown, Elinor Ochs, and Emanuel Schegloff (and indirectly, Gail Jefferson). For their many valuable comments on and contributions to this document and to the system it describes, we thank Karin Aijmer, Bengt Altenberg, Roger Anderson, Ingegerd Bäcklund, Maria Luiza Braga, Wallace Chafe, Patricia Claney, Laurie Crain, Alan Cruttenden, Alessandro Duranti, Jane Edwards, Christine Cox Eriksson, W. Nelson Francis, Christer Geisler, Charles Goodwin, Caroline Henton, John Heritage, Knut Hofland, Marie Iding, Stig Johansson, Geoffrey Leech, Marianne Mithun, Bengt Nordberg, Elinor Ochs, Yoshi Ono, Asa Persson, Janine Scancarelli, Emanuel Schegloff, Emily Sityar, Jan Svartvik, Sandra Thompson, Gunnell Tottie, and Donald Zimmerman. We are also most appreciative of the many comments we have received from the participants in a discourse transcription seminar held at the University of California, Santa Barbara (Summer-Fall 1988), and at presentations given by the first author at the Stockholm Conference on Computers in the Humanities and at the Universities of Lund and Gothenburg (all September 1989). We thank the students in the first author's courses on discourse transcription at the University of California, Santa Barbara (Fall 1988 and Spring 1990) and Uppsala University (Fall 1989). We are especially grateful for the lively representation of diverse viewpoints and the incisive commentary at the conferences on Discourse Transcription (January 1989), Current Issues in Corpus Linguistics (June 1990), and Representing Intonation in Spoken Discourse (July 1990), all held at the University of California, Santa Barbara under the sponsorship of the Linguistics Department and the Center for the Study of Discourse. We are glad to express our thanks to these people and the many others from whom we have gained insights and borrowed ideas—while recognizing that undoubtedly they all would do things at least a little differently. Their contributions to the

formulation of the transcription system have been invaluable and are reflected in virtually every page of this chapter. None of our many benefactors should be held accountable for the choices made in arriving at the final form of the transcription system or its description, for which responsibility rests with us.

REFERENCES

- Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.
- Boase, S. (1990). *London-Lund Corpus: Example text and transcription guides*. Unpublished manuscript, Survey of English Usage, University College, London.
- Chafe, W. L. (Ed.). (1980a). The pear stories: Cognitive, cultural, and linguistic aspects of narrative production. Norwood, NJ: Ablex.
- Chafe, W. L. (1980b). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 169-180). The Hague: Mouton.
- Couper-Kuhlen, E. (1986). *An introduction to English prosody*. London: Edward Arnold.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. (1975). *The English tone of voice: Essays in intonation, prosody, and paralanguage*. London: Edward Arnold.
- Cutler, A., & Ladd, D. R. (Eds.). (1983). *Prosody: Models and measurements*. Berlin: Springer-Verlag.
- Du Bois, J. W. (1987). Absolutive zero: Paradigm adaptivity in Sacapultec Maya. *Lingua*, 71, 203-222. (Reprinted in R. M. W. Dixon, Ed., *Studies in Ergativity*, pp. 203-222. Amsterdam: North-Holland.)
- Du Bois, J. W. (1991). Transcription design principles for spoken discourse research. *Pragmatics*, 1, 71-106.
- Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1991). *Discourse transcription*. Unpublished manuscript, University of California, Santa Barbara.
- Edwards, J. A. (1992). Computer methods in child language research: Four principles for the use of archives data. *Journal of Child Language*, 19, 435-458.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- International Phonetic Association. (1989). Report on the 1989 Kiel Convention. *Journal of the International Phonetic Association*, 19, 67-80.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 79-96). New York: Irvington Publishers.
- Jefferson, G. (1984). On the organization of laughter in talk about troubles. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 346-369). Cambridge: Cambridge University Press.
- McLemore, C. (1991). *The pragmatic interpretation of English intonation: Sorority speech*. Unpublished doctoral dissertation, University of Texas, Austin.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Schuetze-Coburn, S., Shapley, M., & Weber, E. G. (1991). Units of intonation in discourse: A comparison of acoustic and auditory analyses. *Language and Speech*, 34, 207-234.
- Svartvik, J. (Ed.). (1990). *The London-Lund corpus of spoken English: Description and research*. Lund, Sweden: Lund University Press.
- Svartvik, J., & Quirk, R. (Ed.). (1980). *A corpus of English conversation*. Lund, Sweden: C. W. K. Gleerup.
- Tottie, G. (1989). What does *uh-(h)uh* mean: American English vocalizations and the Swedish learner. In B. Odenstedt & G. Persson (Eds.), *Instead of flowers: Papers in honor of Mats Rydén*. Stockholm: Almqvist and Wiksell.
- Trim, J. (1970). Some continuously variable features in British English intonation. *Proceedings of the 10th International Congress of Linguists IV* (pp. 263-268). Bucharest, Romania: Editions de l'Académie de la République Socialiste de Roumanie.

APPENDIX

SYMBOLS FOR DISCOURSE TRANSCRIPTION

Units	
Intonation unit	{carriage return}
Truncated intonation unit	--
Word	{space}
Truncated word	-
Speaker identity/turn start	:
Speech overlap	[]
Transitional Continuity	
Final	.
Continuing	,
Appeal	?
Terminal Pitch Direction	
Fall	\
Rise	/
Level	—
Accent and Lengthening	
Primary accent	^
Secondary accent	`
High booster	!
Low booster	;
Lengthening	=
Tone	
Fall	\
Rise	/
Fall-rise	\ /
Rise-fall	/ \
Level	—
Pause	
Long	...(N)
Medium	...
Short	..
Latching	(0)

(cont.)

Appendix Continued

Vocal Noises	
Vocal noises	()
Inhalation	(H)
Exhalation	(Hx)
Glottal stop	%
Laughter	@
Quality	
Quality	<Y Y>
Laugh quality	<@ @>
Quotation quality	<Q Q>
Phonetics	
Phonetic/phonemic transcription	(/ /)
Transcriber's Perspective	
Researcher's comment	(())
Uncertain hearing	<X X>
Indecipherable syllable	X
Specialized Notations	
Duration	(N)
Intonation unit continued	&
Accent unit boundary	
Embedded intonation unit	< >
Restart	{Capital Initial}
False start	< >
Code switching	<L2 L2>
Nontranscription line	\$
Reserved Symbols	
Phonemic/orthographic symbols	'
Morphosyntactic coding	+ * # { }
User-definable symbols	" ~