SOME CURRENT TRANSCRIPTION SYSTEMS FOR SPOKEN DISCOURSE: A CRITICAL ANALYSIS[1]

Daniel C. O'Connell and Sabine Kowal

1. Introduction

In recent decades, a rather disparate array of transcription systems, all alike intended to make the transient reality of spoken discourse accessible to the eyes of researchers, have found their way into the literature of several related fields of research, including linguistics, sociolinguistics, psycholinguistics, and ethnomethodology.

Until recently, the usefulness and adequacy of these various transcription systems have been largely taken for granted, but in the past several years they have begun to attract attention to themselves as sources of research problems in their own right rather than as practical intermediate steps toward making data accessible.

In the following, we first present the various criteria required by the authors of these systems if they are to be used effectively and adequately. We then review in detail the use of one specific sign for the notation of oral communicative behavior. For this review we have chosen the sign "h" (or, in some cases, "H") because in the various transcription systems, it is the one sign that happens to be used to represent in one way or another all four aspects of oral communicative behavior: verbal, prosodic, paralinguistic, and extralinguistic (see Posner 1985: 238 ff.). The primary use, however, is paralinguistic. In a third step we analyze various aspects of "h" in view of its usefulness and adequacy as a sign for the transcription of spoken discourse. And finally we derive conclusions from our review and analyses for the further development of transcription systems.

In general, the sign under consideration ("h", alone or in specified collocations) is used to transcribe varieties of: phonemes; eye dialect; aspiration; inhalation; exhalation; audible breathing; breathlessness; "breathiness" (French & Local 1983: 36); laughter; "sighing, hearable as unvoiced laughter" (Hopper, Koch, & Mandelbaum 1986: 184); "High key" (Coulthard & Montgomery 1981: ix); guttural quality; and the temporal duration of some of the preceding phenomena. An uppercase "H" is used if the first word of a speaker's utterance begins with that letter, or to indicate loudness, even if the "H" may not be pronounced, e.g., BRIGHT. It is also used to indicate a male speaker (e.g., paradoxically, as "H" [in uppercase lettering] to represent *homme* [Cosnier & Kerbrat-Orecchioni 1987: 371] and "h" [in lowercase lettering] to represent *Herr* [Henne & Rehbock 1982: 126]).

Finally, an uppercase "H" is used as notation for "vertical head movement" (Maynard 1989: 18).

The following analyses are based on a review of more than fifty published versions of transcription systems intended for research in spoken discourse. All have been published in the course of the past two decades and many of them are derived from the system devised by Gail Jefferson.

2. Basic criteria for the design of the four transcription systems

We have extracted criteria for the usefulness and adequacy of transcription systems from a review of the four current transcription systems which have been most explicitly spelled out in the following publications: (1) a recent article by John Du Bois (1991; see also Du Bois, Schuetze-Coburn, Cumming & Paolino 1993) entitled "Transcription design principles for spoken discourse," in which he describes his own system of Discourse Transcription (DT); (2) a series of publications by the German linguist Konrad Ehlich (Ehlich & Rehbein 1976, 1979, 1981; Ehlich 1993) in which he has developed his *Halbinterpretative Arbeitstranskriptionen* (HIAT); (3) Gail Jefferson's (e.g., 1984a, 1989) transcription notation, which has been published in a wide variety of appendices and introductions (e.g., Atkinson & Heritage 1984: ix-xvi) and adapted in a number of ways over the past two decades; and (4) the final report by Brian MacWhinney (1991) on *The CHILDES project: Tools for analyzing talk*, in which the definitive version of CHAT (Codes for the Human Analysis of Transcripts), a transcription system tested over a number of years and adopted by many developmental psycholinguists worldwide, is presented.

(1) Of the four systems, Du Bois's criteria are by far the most extensive, require accordingly the most detailed commentary, and raise a number of serious problems. In order to understand his "Transcription design principles for spoken discourse research," one must first engage his "underlying philosophy of the ideal of discourse transcription". His explicit goal is "to present the multifaceted flux of discourse in a way that is as accessible to the analyst as it is to the participant" (97). The phrase "multifaceted flux of discourse" suggests that the accessibility of the discourse to the participant is somehow quite comprehensive. However, even were we to grant this point, it would still remain difficult to imagine how a transcription system could be capable of making this flux accessible to the analyst just *as it is to the participant*. For one thing, there is no unitary sense in which this accessibility appertains to *the* participant as such. Any discourse is accessible to each participant quite *differently*, precisely because the cognitive, affective, and social roles of each participant are idiosyncratic. Thus, the analyst's access to any participant's "multifaceted flux of discourse" is severely limited. Nor is there any unitary sense in which this accessibility appertains to *the* analyst as such, since there are in principle any number of possible transcribers, researchers, and readers of any transcript.

According to Du Bois, "a good transcription system should be flexible enough to accommodate the needs" (74) of a variety of users. Such a flexible transcription system must then somehow envision the needs of any and all researchers. In this case too, it is not at all clear why one system must bear this burden, especially in view of Du Bois's basic tenet that "there is not, nor ever can be a single way of putting spoken word to paper" (73). It is quite conceivable that a very *inflexible*

system would prove quite adequate and useful for the needs of specific research projects.

Du Bois's design principles for transcription are summarized in a set of five maxims:

1. *Category definition*: Define good categories.
2. *Accessibility*: Make the system accessible.
3. *Robustness*: Make representations robust.
4. *Economy*: Make representations economical.
5. *Adaptability*: Make the system adaptable.

In introducing the first of these maxims, Du Bois states:

"The most basic issue in designing a discourse transcription system is not choosing symbols, but defining the analytical categories for which the symbols will stand." (78)

In doing so, Du Bois deliberately deemphasizes the choosing of symbols relative to the importance of category definition. We would argue that category definition must be carried out as a function of the specific purposes of a research project or design. In other words, there is no general a priori way to determine relevant categories independent of the finality of a given research project. In this sense, category definition is both independent of and antecedent to transcription design, and therefore clearly *not* "the most basic issue in designing a discourse transcription system".

In any event, the problem of choosing symbols reemerges almost immediately under his second maxim (accessibility), which is essentially a vote for familiar, easily learnable notations. However, such notations as Du Bois suggests most often carry with them their own baggage. They have been used for other purposes in other settings. Readers (and researchers themselves) cannot and do not easily prescind from the history of a given notation. "A sequence of three dots to mark a pause" (82) is offered by Du Bois as an example of appropriate notation adopted from literary tradition to transcribe speech pauses of all lengths. Apart from the fact that the convention transcribes pauses of any length indiscriminately, it is not at all clear that its traditional usage in literature is reliably diagnosable or recognizable as a pause, nor that its usage is reliably discriminable from the use of the em dash or sequence of two hyphens for truncation or from the use of the triple ellipse to indicate an omission (e.g., Button & Lee 1987: 17).

The problem of choosing symbols again makes an entry under the third maxim (robustness). The advice to "use widely available characters" (Du Bois 1991: 88) is indeed very practical; it is actually a vote for the currently available, state of the art ASCII (American Standard Codes For Information Interchange) character set. But more relevant to robustness may be the question: Is the current pool of characters the best set, or should it be supplemented or replaced by other characters? This empirical question remains unanswered. The canonization of ASCII simply begs the question of choice.

The fourth (economy) and fifth (adaptability) maxims are both in conflict with the underlying philosophy of Du Bois's discourse transcription. Economy and adaptability can only be implemented in terms of the *specific* goals of a given

research project.

There is, in fact, a certain conflict built into Du Bois's five maxims. Adhering to all of his advice at once is impossible; adhering to some of it throws the researcher right back on his or her own methodological prudence. Nonetheless, as we will observe in the course of our discussion, a number of Du Bois's detailed transcription design principles do provide excellent guidelines - guidelines that have been consistently violated in other transcription systems.

(2) Konrad Ehlich's (1993) most recent presentation of his HIAT system of transcription is fortunately in English. But we must return to the original presentation of the system for an explanation of the notion "*Halbinterpretative Arbeitstranskription*" (semi-interpretative working transcription; or, to maintain the acronym, "Heuristic Interpretative Auditory Transcription" [Ehlich 1993: 125]). This notion is further described as follows:

> "This type of transcription is partially interpretative insofar as the transcriber on the one hand already struc-tures the spoken corpus in terms of both segmentation and commentary on the basis of his own reflectively applied everyday knowledge of language, while on the other hand he imposes no further structure on the corpus by way of interpretation." (Ehlich & Rehbein 1976: 23; our translation)

The quotation incorporates very aptly the basic problematic of transcription systems - meaningful interpretation of the acoustic raw data on the one hand and fidelity to the spoken record on the other. Ehlich's own criteria for an adequate and useful transcription system are the following: "Simplicity, easy application of the transcripts thus produced, and quick learnability" (Ehlich & Rehbein 1976: 22). These have been altered in the most recent presentation of the system to:

> "(a) simplicity and validity, (b) good readability and correctability, and (c) minimum of transcriber and user training." (Ehlich 1993: 125)

These criteria remain far too general to be of much use in the actual design of a transcription system. One can hardly object to them, except perhaps to suggest the need for more than a minimum of training. Ehlich intends these criteria to be practical; his research goal was from the beginning intended to have "a pragmatic orientation" (Ehlich & Rehbein 1976: 22).

(3) To our knowledge, Jefferson herself has nowhere provided explanatory criteria for the design of her transcription system. As we have already indicated, the Jeffersonian transcription system has shifted in a variety of ways over the past two decades and has been adapted for many different research purposes. The practical nature of these alterations is well expressed in the following passage from Zimmerman and West (1975):

> "The transcript techniques and symbols were devised by Gail Jefferson in the course of research undertaken with Harvey Sacks. Techniques are revised, symbols added or dropped as they seem useful to the work. There is no guarantee or suggestion that the symbols or transcripts alone would permit the doing of any unspecified research tasks; they are properly used as an adjunct to the tape recorded materials." (128; see also West & Zimmerman 1982: 535)

Atkinson and Heritage (1984) reiterate the notion that the Jeffersonian

system is one "that continues to evolve in response to current research interests" (ix). This is quite in keeping with our own conviction, stated in regard to Du Bois's emphasis on category definition, that categories should indeed be established by the finality of individual research projects, not in some a priori fashion that in turn leads to notational selection. Schenkein (1978) emphasizes the importance of another aspect of Jefferson's system, readability:

> "She has developed a system of notation and transcript design intending to produce a reader's transcript - one that will look to the eye how it sounds to the ear." (xi)

It might be well at this point to indicate some diversity of opinion regarding what "the data of conversation analysis as such" (Button & Lee 1987: 9) really are. Or to put it another way, how does the eye (the transcription) get back to the ear (the recording)? Zimmermann and West as well as Schenkein emphasize "the tape recorded materials" as data. Button & Lee add:

> "Nor should it be thought that transcripts are the data of conversation analysis as such. The data is naturally occurring conversation as a feature of social life, and the use of tape-recordings and transcripts is a practical strategy for apprehending it, and making it available for extended analysis, and in the case of transcripts, they are a convenient way of presenting the material that was analyzed in research reports." (9)

How the "naturally occurring conversation as a feature of social life" is captured in transcriptions is, of course, precisely the question. If "the material that was analyzed" is not the transcriptions, but must instead be the source of the transcriptions, then only someone who has heard that material - a listener - can make an analysis. Working from transcripts alone would not suffice.

Although the terminology has varied, representatives of the Jeffersonian tradition clearly consider their transcript notations to be systematic. Some of the more recent summaries of transcription systems in this tradition (e.g., Potter & Wetherell 1987: 188; Streeck 1989: 96; Watson & Seiler 1992: iv) have adopted the term transcription or transcript "notation" to designate their system of transcriptional signs; others refer to the system as transcription or transcribing "conventions" (e.g., Hafez 1991: 78; Nofsinger 1991: 167). As we shall see, the systematic nature of their notations is controvertible.

(4) There are several characteristics peculiar to MacWhinney's (1991) CHILDES (Child Language Data Exchange System) project and its CHAT (Codes for Human Analysis of Transcripts) transcription system. Both the project and the transcription system have been undertaken specifically for use with child language development. Furthermore, the transcription system, unlike most of the others developed in recent decades, is explicitly "a system for computerized transcription" (5). In other words, the "Human Analysis" involves as an essential step *making the data compatible for computer data entry*. Hence, to the familiar criteria of clarity and readability, MacWhinney adds "ease of data entry" (5). The notation system can indeed be used without computer entry, but this is not at all the rationale for the development of this specific notation system.

Critical evaluations of the CHILDES transcription system have already appeared in the archival literature. Edwards' (1989) criticisms are based on the biases of the system toward three views outdated in the field of child language

research:

> "toward written rather than spoken language; toward a reductionist, compositional perspective on language and discourse; and toward the view of child language as defective adult language rather than as systematic in its own right." (abstract: unnumbered page)

According to O'Connell's (1991) critique, data derived from applications of the system are not mutually compatible:

> "The logic is quite straightforward. Since the two extreme cases of uniformity and comprehensiveness are both excluded, we are left with selectivity - variously 'for different purposes.' The direct result of this is noncomparability of data sets. But comparability of data sets is the very raison d'être of data archiving." (279)

3. The graphemic use of "h" in alphabetic writing systems

In order to clarify the complexity of transcription systems, we wish first to go back to *ordinary written usage* to examine there the graphemic use of the sign "h". Alphabetic printing and writing are the usual means of recording and preserving everyday spoken discourse. This system underlies all four of the research transcription systems introduced in the preceding section. An exception to this traditional usage is to be found in MacWhinney's (1991) transcription system: "Words that begin sentences are not capitalized" (39). This is in keeping with his requirement of "ease of data entry" (5). Brinker and Sager's (1989) reason for generally avoiding uppercase lettering is "to minimize biased decisions" (54; our translation); in their own transcripts, however, they do use some uppercase lettering. Generally speaking, the alphabet system is obviously both useful and adequate for most everyday purposes and also for many quite complex purposes, e.g., for scientific monographs and court proceedings.

And yet, a great deal of the richness of spoken discourse is necessarily lost in such usage. A very complex phonememegrapheme set of relationships allows "h" to contribute to the transcription of such disparate spoken words (in standard pronunciation) as: bright, cherish, cholera, happy, heir, physics, rough, show, though, through, verandah, where; (in French) *cherchez, l'hotel*; and (in German) *ich, Schule*, and *Theologie*. The acronym NIH (National Institutes of Health) illustrates still another phoneme/grapheme relationship. The examples entail specific phonemic categories; moreover, their written transcriptions neglect allophonic and prosodic variations, and voice qualities. In particular, the prosodic features of pitch, duration, and loudness (Cruttenden 1986: 2) constitute important continuous variables which are by and large eliminated from ordinary written records, although some languages do represent long vowels graphemically. But even in the case of consonants, for example, the written form "withhold" does not at all indicate a spoken phoneme /h/ of twice the duration of a single letter "h".

All four of the transcription systems under consideration here adopt alphabetic notations which, for the practical purposes of written discourse, neglect many of the features important for one or another kind of research. These features must, therefore, be represented by means of some supplementary signs.

In addition, there are other features in ordinary writing which may or may

not contribute to scientific usefulness. These include punctuation, various typographical conventions (e.g., use of uppercase letters), and various literary conventions (e.g., eye-dialect spellings such as *"thuh* for *the"* [see Chafe 1993: 34] and "Wuhjeh" for "What did you" [Sacks, Schegloff, & Jefferson 1974: 731]).

And finally, there are also other ways of transcribing the phoneme /h/, e.g., in Greek with a superscript rough breathing over a vowel and in Spanish by means of the letter "x," as in *Oaxaca*. One could well argue that the use of the IPA (International Phonetic Alphabet) might solve some of these problems. Nonetheless, its use has been generally rejected in research on spoken discourse in favor of readability and ease of use (see, e.g., Ehlich 1993).

The complex of spoken-to-written relationships reminds one more of the biblical Babel than of an orderly transcription system easily adoptable for scientific purposes. The lack of a one-to-one relationship between sign and what it is intended to denote is matched by a general tendency in researchers, embedded as they are in their own literate culture, to neglect features of the spoken in favor of the written, with what Ong (1982) refers to as "the chirographic and typographic bias" (166) and Linell (1982) refers to in the title of his book as *The written language bias in linguistics*.

4. "h" in current transcription systems for spoken discourse

Our preoccupation in this section is with the *supplementary* usage of "h" in transcription systems for research in spoken discourse. This usage is both prosodic and paralinguistic. In the following discussion, we will make no further mention of the one extralinguistic usage we have found, the use of "H" as notation for "vertical head movement" (Maynard 1989: 18). We should make clear at the very beginning that the prosodic usage is not peculiar to "h," but is common to all the graphemes, as we note in Table 1 below; we have chosen to exemplify the usage with "h" because it emerges saliently in the notation of paralinguistic features. As we shall see, there is much disparity in usage among the four transcription systems we have been considering. Even more surprising, there is quite considerable disparity even within one of the systems (the Jeffersonian tradition) in this regard. This disparity extends to a number of other systems which are either independent of the Jeffersonian tradition or at least do not explicitly acknowledge their dependence.

In what follows, the discussion of three of the transcription systems is based solely on the publications of the originators of the system: Du Bois (1991) and Du Bois et al. (1993); Ehlich (1993) and Ehlich and Rehbein (1976, 1979, 1981); and MacWhinney (1991). Our discussion of the Jeffersonian tradition, however, is based on all the publications available to us of ethnomethodologists who explicitly refer to Gail Jefferson as the source of their transcription system. Finally, for purposes of comparison, we refer to other available publications in which transcription systems are presented, although no dependence on Jefferson is acknowledged therein.

4.1. *"h" as an example of notation for prosodic features*

In the research to be discussed in this section, the prosodic features of pitch, duration, and loudness, or more generically, emphasis or stress, are generally indicated by making use of graphemes in their segmental syllabic positions while changing the typography by adding italics or uppercase lettering or underlining. Thus, prosodic notation becomes intrinsic to the segmental notation. The single discrete change in the segmental notation is in turn unable to reflect the continuous variation of a given prosodic feature.

It is important to note that, of the four systems, the Jeffersonian tradition is the major one that makes use of this method of indicating prosodic features. Ehlich (1993) does so in only one instance, to indicate stress by underlining (129); neither Du Bois nor MacWhinney use these means at all to indicate prosodic features. In pursuit of his *accessibility* maxim, Du Bois (1991) emphasizes the need for "having the prosody (mostly) segregated in a separate non-alphabetic 'channel'" so that just the words uttered may "come through loud and clear," e.g., "going to ^ Littleton" (87) to indicate primary accent on the first syllable of the third word. MacWhinney (1991) allows a variety of nongraphemic signs within words as notation for "prosody within words," e.g., **"rhi//noceros"** (46) to indicate primary stress on the syllable following the double slash.

Since the relationships between typographical notation and prosodic features vary considerably from author to author, we have summarized them in Table 1. As the title indicates, only incremental prosodic features are included. In other words, only unidirectional variation in pitch, duration, and/or loudness (or generically, emphasis/stress) is considered. The authors represented there are from the Jeffersonian tradition with the exception of Bublitz (1988), Coulthard and Montgomery (1981), Coupland, Giles, and Wiemann (1991), Ehlich (1993), Fritz (1982), Jucker (1986), Lenz (1989), and Scannell (1991). These authors are among those who do not explicitly acknowledge dependence on the Jeffersonian notation system.

It should be noted in advance that notation for stress or emphasis is essentially indeterminate and can therefore overlap with notations for any of the three prosodic features: pitch, duration, and loudness. Or, as Cruttenden (1986) puts it, "They are all used to make some syllables more prominent in words and in sentences" (16). This confounding is reflected in Table 1 by a repetition of the heading Emphasis/Stress. Authors whose names appear more than once have at least partially confounded the notation of prosodic features. For example, French and Local (1983), in the first line of Table 1, use italics for both pitch and loudness. Bublitz (1988), however, uses italics to transcribe the indeterminate concept of emphasis/stress; which of the three (pitch, duration, and/or loudness) is intended in the specific case is unclear in such a transcription.

In Table 1, we have reduced the material to simple categories. In the following discussion, the details of the individual notation systems are more nuanced. We use quotations generously because the terminology of the various authors for the prosodic features is sometimes in terms of acoustic and sometimes in terms of auditory analysis. Authors representing the Jeffersonian tradition are presented chronologically and then the others.

We wish here to insert a *caveat lector*: The remainder of this section and the following section are unavoidably dense and reader-unfriendly. Apologies in

advance. We know of no other way to convince the reader of the extent of chaos that exists in these systems except to describe it in detail. The impatient reader would do well to move on immediately to section 5.

Davidson's (1978) use of italics to indicate stress indeterminately and his use of uppercase lettering to indicate louder volume partially confound prosodic features. McHoul (1978) allows italics to serve double duty as notation for "accentuation by volume or by intonation" (213). Schenkein's (1978) notation is more generic: "Emphasis is indicated by varieties of italics, the larger the italics the greater is the relative local stress" (xxx). His examples, however, introduce both boldface and uppercase lettering in addition to italics as notation for increased emphasis, but with no acknowledgement or explanation. Psathas (1979) uses underlining as the notation for "various forms of stress," but then specifies that these "may involve pitch and/or volume" (290).

Goodwin (1981: vii) is generic; italics serve as notation for emphasis. West and Zimmerman (1982: 536), however, explicitly use two alternative notations, uppercase lettering and underlining, to transcribe emphatic words. French & Local (1983) limit their italic notation to emphasis "achieved by loudness and/or dynamic pitch movement" (36).

Levinson (1983) uses uppercase lettering as notation for "relatively high amplitude," but then confounds the notation by introducing italics as notation for "syllables stressed by amplitude, pitch and duration" (370). Similarly, Atkinson and Heritage (1984) employ underlining as notation for emphasis without further specification and uppercase lettering as notation for relative loudness.

Hopper, Koch, and Mandelbaum (1986) use underlining as the notation for "stress/emphasis" (184). Button and Lee (1987) use uppercase lettering as notation for relative loudness and underlining as notation for emphasis (14 f.). Potter and Wetherell (1987) introduce partial confounding insofar as they use underlining as notation for "added emphasis" and uppercase lettering as notation for words "uttered louder than the surrounding talk" (188). Helm, Anderson, Meehan, and Rawls (1989) use underlining as notation for emphasis by "pitch or amplitude" (251).

Streeck (1989) repeats the (by now) familiar partial confounding of emphasis and loudness by notation of the former with underlining and of the latter with uppercase lettering (97 f.). Wowk (1989) uses underlining as generic notation for emphasis (71). Hafez (1991) uses uppercase letters as notation for emphasis without, however, differentiating between "increased volume, or high pitch" (79). The most recent author to partially confound the use of underlining as notation for emphasis with the use of uppercase lettering as notation for "noticeable loudness" is Nofsinger (1991: 168).

TABLE 1

Typographic Alteration of Graphemes as Notation for Incremental Prosodic Features in Transcription Systems for Spoken Discourse

Typographic Alteration	Emphasis/ Stress	Pitch	Emphasis/Stress Duration	Loudness
Italics	Bublitz Davidson Fritz Goodwin Levinson Scannell Schenkein	French & Local McHoul		French & Local McHoul
Uppercase Lettering	Coulthard & Montgomery Hafez Scannell Schenkein West & Zim- mermann	Jucker		Atkinson & Heritage Button & Lee Coupland et al. Davidson Helm et al. Jucker Lenz Levinson Nofsinger Potter & Wetherell Streeck
Underlining	Atkinson & Heritage Button & Lee Coupland et al. Ehlich Hopper et al. Nofsinger Potter & Wetherell Streeck West & Zim- mermann Wowk	Helm et al. Psathas	Lenz	Helm et al. Psathas

It should be noted that the only confoundings we have mentioned in the preceding paragraphs are cases in which various typographies of graphemes (i.e., italics, underlining, or uppercase lettering) are involved. In many of these same transcription systems, *additional* notations are used for prosodic features. For example, Schegloff and Sacks (1973) do *not* use italics for emphasis (but instead for "accent"); however, they use a / for "upward intonation" (327). McHoul's (1978) notation uses italics for "accentuation by volume or by intonation," but also uses / for "upward intonation" (213). Similarly, Schenkein (1978: xxx), Goodwin (1981: vii f.), and Levinson (1983: 370) all use italics for emphasis and also punctuation marks

for varieties of intonation. Atkinson and Heritage (1984, xi f.), Hopper et al. (1986: 184), and Watson and Seiler (1992: x) use underlining and punctuation in the same overlapping fashion. Similarly, Streeck (1989: 97) uses underlining and upward or downward pointing arrows, and Hafez (1991: 79) uppercase lettering and punctuation. Some authors who acknowledge no dependence on the Jeffersonian tradition also use italics, uppercase lettering, and/or underlining for the notation of prosodic features and are included in Table 1. Coulthard and Montgomery (1981) use uppercase lettering as notation for "prominent syllables" (ix). Jucker (1986) uses uppercase lettering as notation for "pitch prominence and loudness" (x). Bublitz (1988) uses italics as notation for three disparate phenomena: "words in 'object language'"; "parts of examples used in the text"; and "emphasis" (xi). Lenz (1989) uses uppercase lettering as notation for increasing loudness (*"lauter werdend"*) and underlining for prolongation of duration (*"gedehnt"*), but sings out of chorus by using italics for *less* loudness (*"leiser"*) (256). In order to avoid confusion in the "loudness" column, we have not entered this inverse dimension there. Lenz provides the only entry in the "duration" column in Table 1.

Ozaki (1989) is the only author we have found who uses underlining to indicate "a marked low voice" (23). Since this is the inverse of loudness (as we have just noted in Lenz), we have not entered Ozaki in Table 1 in order to avoid confusion. Coupland et al. (1991) use uppercase letters as notation for "extreme loudness" (315). Scannell (1991) uses italics for stress and uppercase lettering for "extreme stress" (vi). Ehlich (1993) uses underlining as notation for emphasis (129). As we found to be the case in the Jeffersonian transcription tradition, a number of these authors (Coupland et al. 1991; Coulthard & Montgomery 1981; Ehlich 1993; and Scannell 1991) use additional symbols as well for intonation.

There are four cases in which the sign "h" is used as notation for prosodic features, but is *not* simultaneously a segmental grapheme. These cases are not included in Table 1. In all of them, the rationale for the use of "h" is the same: It represents a letter (or letters) in the word or phrase for the designated phenomenon. Coulthard and Montgomery (1981) use an uppercase "H" as notation for "High key" (ix). Brinker and Sager (1989) use "h" before and after a section of spoken discourse that is high pitched: "h---------h" (48). MacWhinney (1991) uses a vertical stroke followed immediately by an uppercase "H" as notation for "hesitation pause" (80) and explicitly includes it in a section on prosodic categories. Du Bois et al. (1993) use "<HI HI>" for a stretch of text that is "higher pitch level" and "<WH WH>" for "whispered" (69). Similarly, they use "<RH RH>" for "rhythmic: stresses in a beatable rhythm," "<ARH ARH>" for "arrhythmic [sic]: halting speech," "<HSK HSK>" for "husky," and "<SGH SGH>" for "sighing" (69). In other words, the angle brackets designate "a marked quality or prosody of some sort" (68).

4.2. *"h" as notation for paralinguistic features*

Although the usage of "h" as notation for paralinguistic features is not limited to the Jeffersonian tradition, it is most commonly represented in transcription systems within that tradition. The paralinguistic features thus transcribed include breathing, breathiness, breathlessness, aspiration, laughter, crying, and sighing. Table 2 summarizes the use of "h," alone and in conjunction with a variety of additional

signs, as notation for paralinguistic features in the Jeffersonian tradition and in other transcription systems. The authors are presented in chronological order. It will be noted that neither Ehlich nor MacWhinney is listed because neither of them employs "h" for these purposes. Once again we will make frequent use of literal citation of authors because of the many terminological inconsistencies.

TABLE 2
"h" as Notation for Paralinguistic Features in Jeffersonian and Other Transcription Systems; Conditions Include: Location of "h" Within and Between Words Indicated by W and B Respectively; Repetition of "h" (Signifying Duration) Indicated by D

Notation	Condition	Paralinguistic Feature	Author(s)
1. Jeffersonian tradition			
(h)	W	explosive aspiration	Sacks et al.
h		audible out-breath	Sacks et al.
.h		audible in-breath	Sacks et al.
hhh	D	breathing	Davidson
hhh		audible aspiration	Schenkein
.hhh		inhalation	Schenkein
hh		exhalation	West & Zimmermann
.hh		inhalation	West & Zimmermann
heh		laugther syllable	West & Zimmermann
henh		laughter syllable	West & Zimmermann
eh		laughter syllable	West & Zimmermann
hhh	W	audible outbreath or breathiness	French & Local
.hh h = 0.1 sec		audible inbreath	French & Local
hh		audible out-breath	Levinson
.hh		in-breath	Levinson
hhh		audible aspiration	Atkinson & Heritage
.hhh*		inhalation	Atkinson & Heritage
gh	W	gutturalness	Atkinson & Heritage
(h)	W	word spoken laughing	Houtkoop & Mazeland
.hhh		inbreath	Houtkoop & Mazeland
hhh		audible out-breath, sighing, hearable as unvoiced laughter	Hopper et al.
(h)		explosive aspiration, sometimes laughter	Hopper et al.
'h or .h		audible in-breath	Hoper et al.
h	D	discernable aspiration	Button & Lee
(h)	W/D	aspiration	Button & Lee
h or .h	D	discernable inhalation	Button & Lee
(.h)	W/D	inhalation	Button & Lee
h		audible inhalation in a filled pause	Cosnier & Kerbrat-Orecchioni
.hh		audible intake of breath	Potter & Wetherell
hh		outbreath	Helm et al.
.hh**		inbreath	Helm et al.
hhh	W	breathiness	Helm et al.
(h)	W	plosive, with e.g.,	Helm et al.

		laughter, crying, breathlessness	
hhh		outbreath	Jefferson (1989)
.hhh		inbreath	Jefferson (1989)
hh	W	breathiness	Jefferson (1989)
(h)		plosiveness, with laughter, crying, breathlessness, etc.	
gh	W	gutturalness	Jefferson (1989)
hr		softening of "r"	Jefferson (1989)
(h.)	D	audible exhalation	Streeck
(.h)	D	audible intake of breath	Streeck
(h.)	D	audible out-breath	Wowk
(.h)	D	audible in-breath	Wowk
hh		audible exhaling, sometimes with laughter	Nofsinger
.hh	D	audible inbreath	Nofsinger
heh		laugh token	Nofsinger
ha		laugh token	Nofsinger
(h)	W	laugh token	Nofsinger
h		outbreath	Watson & Seiler
.h	W	aspiration, commonly laughter	Watson & Seiler

2. Other authors

hhh		audible outbreath	Coupland et al.
.hhh		in-breath	Coupland et al.
(h)	W	laughter	Coupland et al.
(Hx)	B	exhalation	Du Bois
(H)	B	inhalation	Du Bois
hh		audible exhalation	Scannell
.hh		audible inhalation	Scannell
heh		laugh token	Scannell

* As presented in the authors' examples: (hhh)
** Position of the dot is neither specified nor exemplified.

   The first instance in Table 2 is the article on turn-taking by Sacks, Schegloff, and Jefferson (1974):

   "The *h*, within parentheses, within a word or sound, indicates explosive aspiration, e.g. laughter, breathlessness, etc. . . . The *h* without parentheses indicates audible breathing. A dot placed before it indicates an in-breath; no dot indicates out-breath. . ." (733)

In their examples, the dot occurs at the top and to the left of the "h".

   Davidson's (1978: 133) version of an "h" notation for breathing is much simpler:

   hhhh   Indicates breathing, with the number of h's indicating the length of the breath.

Some differences between these first two usages are to be noted. Sacks et al. refer only to *audible* breathing and do not indicate duration at all; Davidson does not specify audibility but does indicate duration, though the unit of duration is not specified. Davidson's version makes no distinction between in-breath and out-breath.
     Another such entry is to be found in Schenkein (1978):

> "Audible aspirations (hhh) and inhalations (.hhh) are inserted in the speech where they occur. . . " (xxx)

It is not explained at all in Schenkein's text, but is nonetheless clear from his examples that the use of parentheses now becomes much more complex. No longer do parentheses indicate "explosive aspiration," as in Sacks et al., but occurrence *within* a word rather than *outside* a word. In addition, the dot no longer occurs at the top, but as a period at the bottom of the line. In West and Zimmermann (1982: 537), the following version occurs:

> .hh hh eh-heh   These are breathing and laughing indicators. A period followed by *hh*'s marks an inhalation. The *hh*'s alone stand for exhalation. The *heh*s and *henh*s are laughter syllables.

There is no further mention of parentheses, and breathing and laughter notation are presented together. Two changes introduce notation from more familiar sign systems: punctuation and syllabification. The dot is now referred to as a period and is thus acknowledged as a familiar punctuation mark. Two laughter sylla-bles are spelled out in the text (heh and henh) and one more is added in the example (eh).
     The version in French and Local (1983: 36) is yet again a different one:

> hhh   audible outbreaths or breathiness in words. The number of 'h's' corresponds to the duration of the outbreath(s) measured in tenths of a second.
>
> .hh   audible inbreaths. The number of 'h's' following the point corresponds to the duration of the inbreaths measured in tenths of a second.

Laughter has now disappeared. "Outbreaths or breathiness" are limited to positions within words. It is not clear whether these two concepts refer to distinct phenomena. Inbreaths are not limited to within-word positions. The identification of the unit of duration lends a new meaning to the set of letters: They now indicate passage of time. The point reverts to a midline dot.
     Contemporaneously with French and Local, Levinson (1983: 370) offers the following simplification:

> hh   indicates an audible out-breath, .hh an in-breath.

Once again, the dot returns to the baseline. There are no parentheses, there is no laughter, there are no limitations of outbreaths to positions within words, and there is no unit of duration.
     We have not yet exhausted variant versions. Atkinson and Heritage (1984)

offer the following:

"Audible aspirations (hhh) and inhalations (·hhh) are inserted in the speech where they occur... A 'gh' placed within a word indicates gutturalness..." (xii)

In the above explanation, the dot is again midline, but in all the examples given on the same page, it occurs at the top of the line. The new element is, of course, the guttural quality.

Now we are ready to return once again to laughing, with a version from Houtkoop & Mazeland (1985: 614):

ve(h)ry - an 'h' between brackets indicates that the word in which it occurs is spoken
          laughing
.hhh   - inbreath

Non-laughing outbreath now disappears from the scene. Laughing is also alternatively designated by description: "((laughs))" (614), presumably when it occurs between words. Parentheses are referred to as brackets. The dot is once again midline.

Hopper et al. (1986: 184) offer the following conventions along with examples:

hhh   h's indicate audible out-breaths, sighing, hearable as unvoiced laughter.
(h)   h in parentheses indicates explosive aspiration, sometimes laughter.
                     BEE: *They* always are(hh)hhh
'h   (in "chicken dinner") or .h (in "mother/daughter") indicates audible in-
           breath.

All three of these notations are difficult to construe and seem to violate both plausibility and readability. They do not correspond, as indicated by Hopper et al., to either Schenkein's (1978) or Atkinson and Heritage's (1984) notations.

Potter and Wetherell (1987) offer a limited variation on breathing:

"A full stop before a word or sound indicates an audible intake of breath, e.g.:
          A: I think .hh I need more"

Helm et al. (1989) present the following details:

"A dot preceding a row or [sic] H's indicates an inbreath. Without the dot a row of H's indicates an outbreath. . . . A row of H's within a word indicates breathiness. . . . H's in parenthesis indicate within-speech plosives - canbe associated with, e.g., laughter, crying, breathlessness."

In this instance, the position of the dot is neither specified nor exemplified by the authors. This is the only instance we have found in which "h" occurs within words with or without parentheses to indicate different paralinguistic phenomena. Despite the use uppercase lettering in their explanation, all their examples use lowercase lettering.

Streeck (1989) phrases her version as follows:

"Audible intake of breath is indicated by (.h) and extended according to duration (.hhh); the same is the case of audible exhalation, indicated by (h.)." (98; our translation)

No unit is given for duration. Wowk's (1989: 71) use of (.h) and (h.) is identical to Streeck's. Their use of a period in postposition to indicate audible aspiration is quite unusual.

Nofsinger (1991: 168f.) spells out his usage in more detail:

.hh	The h preceded by a period represents an audible inbreath. Longer sounds are transcribed using a longer string: .hhhh
hh	The h without a leading period represents audible
st(h)upid	exhaling, sometimes associated with laughter; and laughter itself is transcribed using "heh" or "hah" or something similar. When laugh tokens are embedded in a word, they are often represented by an h in parentheses.

Finally, Watson and Seiler (1992: x) sum up their version as follows:

.h	The letter *h* preceded by a period indicates aspiration in the course of a word, commonly laughter. Without the period, the *h* indicates outbreath.

Once again, the use of "h" for aspiration is restricted to positions *within* words; no parentheses are used.

Since the authors of all of the above usages of "h" for paralinguistic notation explicitly refer to the Jeffersonian system in their presentations, we have reserved Jefferson's (1989: 195) own recent summary of paralinguistic notational applications of "h" to this final position:

.hhh	A *dot-prefixed row* of h*s* indicates an inbreath. Without the dot the *h*s indicate an outbreath.
wohhrd	A *row of* h*s within a word* indicates breathiness.
(h)	A *parenthesized* h indicates plosiveness. This can be associated with laughter, crying, breathlessness, etc. . . .
wghord	A *'gh' stuck into a word* indicates gutturalness.
hr	An *'h' preceding an 'r'* softens the 'r'. This device is used frequently in my transcripts of British talk. Thus, for example, 'part' is shown is 'pahrt', 'court' as 'cohrt', etc.

Before we proceed to a discussion of these various "h" notations for paralinguistic features, we wish to mention several additional authors who make use of such notation without explicitly referring to the Jeffersonian tradition. Coupland et al. (1991: 316) give the following summary:

Ye(h)s	Within speech laughter is shown by (h).
hhh	The h's indicate audible outbreaths.
.hhh	A superscripted period with h's indicates in-breaths.

Du Bois (1991) includes under the generic heading VOCAL NOISES "Inhalation (H)" and "Exhalation (Hx)" (104). Finally, Scannell (1991) adopts .hh for "Audible inhalation," hh for "Audible exhalation," and heh as a laugh token (vi).

5. An analysis of "h" notations

We must now extract from the overwhelming array of detail given in the previous sections an overview of the implications of these notations. Table 3 summarizes the use of the variations and combinations of the sign "h" in the transcription systems analyzed above.

TABLE 3

Use of the Variations and Combinations of the Sign "h" in the Transcription Systems for the Notation of Spoken Discourse Presentend in Tables 1 and 2

Sign	Use			
	Segmental	Supplementary		
		Prosodic	Paralinguistic	Extra-linguistic
h	+		+	
ital.		+		
H	+	+	+	+
h		+		
gh			+	
'h			+	
.h			+	
.h			+	
'h			+	
hh			+	
.hh			+	
.hh			+	
hr			+	
Hx			+	
hhh			+	
'hhh			+	
hhh			+	
.hhh			+	
(h)			+	
(H)			+	
(.h)			+	
(h.)			+	
eh	+			
ha	+			
heh	+			
henh	+			

The first thing to be noted regarding the 26 entries listed in Table 3 is that the rationale for their selection and use has been left almost entirely implicit in the archival literature. The conventional *segmental* use of graphemes in a complicated set of relationships to phonemes is the basic component of all the transcription systems we have considered. Thus, the verbal component of the spoken discourse

is transcribed into syllabic and word units borrowed from ordinary written discourse.

What we have called the *supplementary* uses of lettering (here exemplified by "h") for prosodic, paralinguistic, and extralinguistic components are *not* the traditional ones of ordinary written language. Although the letters and punctuation signs themselves are conventional, the rationale of their use for these purposes is, in fact, partly arbitrary and partly quite implicit. These uses are spelled out in Table 3. A number of relationships emerge from these data.

5.1. *Segmental use of "h"*

Both lowercase and uppercase "h" are used segmentally. The former represents an adoption of the alphabetic conventions of ordinary writing for the *scientific* transcription notation. The latter is similarly used for sentence-initial positions, for proper names and (e.g., in German) nouns.

The only other signs used segmentally are the so-called "laugh tokens" and "laughter syllables". They are segmental insofar as they serve as notation for sequential syllables of spoken discourse. This fact should be carefully noted, since laughter is typically designated as paralinguistic behavior. In accord with their emphasis on the distinction between linguistic and paralinguistic features, Ehlich and Rehbein (1976) recommend *description* of laughter, along with other sounds such as whistling and applauding, rather than *transcription*. Their rationale is the following:

> "As long as there exists no mandatory phonology for the representation of noises and the other acoustic features mentioned above, this practice of describing seems to us more intelligible for purposes of analysis than imitative practices such as one finds in comic strips." (31; our translation)

MacWhinney (1991) also recommends description of laughter by means of an entry in brackets (47). Du Bois et al. (1993) use "a special nonalphabetic symbol," "@," (89) for laughter and thus also emphasize its *para*linguistic nature.

Even though readability has been emphasized as one of the criteria for a good transcription system, the practice in the Jeffersonian tradition of using laughter syllables sometimes yields thoroughly unreadable transcripts, as the following example from Jefferson (1984b: 349) shows:

```
-> M: eh Not the floo:r one      ehh:: h euh   he h-heh-he h
                                  [          [        [
-> G:                             e h h  h e : h  he:h
```

Neither the syllabic segmentation of laughter nor the identity of "he" (as a laughter syllable or pronoun) is clear. The notation violates Du Bois's (1991) advice to "*use space meaningfully*" (93) as part of his *economy* maxim.

There is a noteworthy irony in Jefferson's (1979, 1984b) two articles in which she analyzed laughter. Both appear in edited volumes (Psathas 1979; Atkinson & Heritage 1984) in which the "Jeffersonian" transcription system is spelled out in detail. In *neither* is any sign for laughter included. In this light, Psathas's introductory

comment to Jefferson's earlier study seems overdrawn:

> "Examination of her essay on laughter will show that basic to the study of laughter is such
> a well-developed and precise transcription and notational system." (79)

Without an explicit system of notation, Jefferson's transcription exemplified above remains unmotivated and indecipherable.

5.2. *Prosodic use of "h"*

Italics, uppercase lettering, and underlining are all used to indicate prosody. This usage is literally superimposed on the segmental graphemes and cannot be construed independently of the segmental function. Table 1 indicates a preference among authors for uppercase lettering as notation for loudness. This practice has a certain iconicity though it is also confoundable with the segmental use of "H," e.g., in "Harry" or in the acronym NIH. The prosodic use of italics appears to be arbitrary, whereas underlining to indicate emphasis (the preference of most authors in Table 1) is clearly imported from the ordinary use of written language.

The prosodic use of *under*lining also pinpoints a systematic use of space in these transcription notations. The space beneath the segmental graphemes is here used to indicate prosody, whereas the use of the spaces before and after the segmental phonemes is reserved uniquely for paralinguistic notation.

The application of typography for prosodic notation entails the following problems:
(1) Duration is excluded from these typographic notations of prosody (with the exception of Lenz [1989]; see Table 1) and is instead typically indicated by a different sign system, namely punctuation, e.g., in Jefferson (1989: 194):

::   *Colons* indicate prolongation of the immediately prior sound. The length
             of the colon row indicates length of the prolongation.

The use of punctuation instead of typographical change seems to be motivated by a very simple iconic rationale: Stretching space stands for stretching time; typography itself cannot easily be stretched without distortion and consequently diminution of readability.

Even when typographical changes are used as notation for indeterminate emphasis, the authors clearly do not intend to include duration under this notation. Instead, they generally use punctuation for this purpose, with the single exception of Lenz (1989), who uses italics as notation for an unidirectional change in duration (prolongation). Lenz was therefore excluded from Table 1.
(2) The remaining two prosodic dimensions, pitch and loudness, are partially confounded with generic notation of emphasis or with one another in many of the authors included in Table 1.
(3) Both punctuation and typographical alteration are inadequate to represent the continuous dimensions of prosody. Punctuation always involves discrete elements, and the alteration of typography can reflect continuity (e.g., by thickness of underlining), but without any easily discernible metric.

(4) The typographical notation of prosody has been used unidirectionally. In the transcription systems in Table 1, the *other* directions (lower pitch, shorter duration, less loudness) are either excluded from consideration altogether or are relegated to another notation (e.g., Jefferson's 1989 use of "*right/left* carats" to indicate "speeding up" and the "*degree sign*" to indicate "relatively quieter" [194 f.] utterances; Psathas's 1979 use of "the degree sign" to indicate "that the talk it precedes is low in volume" [291]; and Streeck's 1989 notation of speed [a diminution of duration] by a super-script arrow directed toward the right [98]).

Obviously, "diminishing" typography and/or punctuation is impractical technologically and would be perceptually difficult for the reader to construe. Nonetheless, unidirectional notation runs the risk of biasing transcriptions and of implying a norm involving *standardized* pitch, duration, and loudness in spoken discourse. The cart indeed seems to be pulling the horse: What is practical and accessible in the written medium is dictating what can be considered important in the spoken discourse.

5.3. Paralinguistic use of "h"

A number of general observations emerge from Table 3 regarding the paralinguistic use of "h". The addition of either lettering or punctuation as notation before or after "h," except for the "laughter syllables," always indicates paralinguistic usage. Repetition of "h" indicates prolongation in time. The use of parentheses is typically an indication of embedding within a word. These last two cases both involve a certain iconicity.

As Table 3 further indicates, notations for paralinguistic features also entail some problems. The disparities, lacunae, and indecipherable notations are evident from our previous discussion. More specifically:
(1) Use of a dot or point at the top or middle of the line is impractical for current keyboards. Some authors therefore use instead a period at the bottom of the line.
(2) Letters used as *nongraphemic* notations, especially when they occur *within* words, can be monumentally confusing (e.g., Is "fight" intended as a guttural "fit"? Is "withhold" a construable word "witold" spoken with breathiness or just the word "withhold"? Or what about the previous entry with breathiness - "withhhhold"?).
(3) Jefferson's (1989) own notation of "hr" as a "softening of 'r'" (195) is a misnomer; only the "h" pertains to the notation. It should have been designated as "h" alone in keeping with her examples. There is at least one usage in Irish in which it occurs as part of ordinary written discourse: *aerpohrt*.
(4) The remaining notations are concerned largely with breathing, laughing, breathiness, breathlessness, sighing, plosiveness, and crying, in a somewhat bewildering kaleidoscope of variability. The requirement of audibility is not always clearly indicated. Sometimes the notation is limited to either between- or within-word positions. Sometimes duration is indicated, or a unit of duration (h = 0.1 sec) is even specified, though without any indication of temporal measurements (see French & Local 1983).

5.4. *Some more general observations regarding "h"*

We were unable to find *any* instances in which some of the more esoteric notations offered by authors are actually used for anything more than demonstrational examples, even in the very publication where the detailed notation was thought to be important enough to be explained as part of the transcription system (e.g., Brinker & Sager's 1989 notation before and after a section of high pitched spoken discourse: "h------h" [48]).

A more extensive example involves the question: Why is audible breathing emphasized in the transcription systems? In fact, throughout the actual empirical research literature on spoken discourse, we have not been able to find any mention of audible breathing. There is, in fact, only an occasional animadversion in the empirical literature even to prosodic and paralinguistic features - except in descriptions of the transcription systems themselves. Most of the interest of researchers is in verbal content. The focus on content certainly cannot be faulted; it does reflect, however, a pervasive absence of genuine research interest in the analysis of *performance* variables of spoken discourse.

If various components of transcription systems are, to some notable extent, without scientific motivation, then we are indeed dealing with what one of our colleagues has referred to as "data graveyards" ("*Datenfriedhöfe*"). O'Connell (1991) has also argued that the careful computer storage of transcribed data foreseen in MacWhinney's (1991) CHILDES project may similarly be much ado about nothing, precisely because the system cannot promise "comparability of data sets" (279).

One final note regarding Table 3 is in order. There is one sign - uppercase, unaccompanied "H" - which is used in all four functions considered in this article: segmental, prosodic, paralinguistic, and extralinguistic. It thereby provides an excellent example of the ambiguity of usage from one transcription system to another.

6. Conclusions and suggestions for adequate and useful notation

Somehow or other, transcription systems have become an end in themselves; they look scientific. However, a reasonable economy would dictate that *transcription be limited to features to be subsequently analyzed*. And these in turn are determined not only by what *can* be analyzed, but also by *the specific purpose of the research project*. If the purpose remains at a descriptive or general level, then the data base itself should be determined accordingly. Transcription in IPA, for example, would be monumentally redundant for an ordinary research project to investigate some properties of turn-taking behavior in normal adults with native language in common. A detailed transcription of as many features as can be retrieved becomes bad science. As Elinor Ochs (1979) puts it:

> One of the important features of a transcript is that it should not have too much information. A transcript that is too detailed is difficult to follow and assess. A more useful transcript is a more selective one. (44)

A good example of unmotivated transcription is breathing. We have not been able

to find a single study in the Jeffersonian tradition where breathing has been systematically used in the analysis of data, despite the variety of notations dedicated to it in many of the versions of transcription systems.

The finality of transcription is obviously twofold: for the researcher, it is facilitative of analysis; for the reader, it characterizes the data base. In the Jeffersonian tradition, analysis of the data base has, in any event, been minimal and subordinate to anecdotally based qualitative interpretation. Jefferson's own studies are typically presented with no inferential statistics and with a minimum of descriptive statistics. The logic is dominantly anecdotal. As for the readers, they are only with difficulty able to appreciate the data base, partly because of the many problems entailed in the transcription systems, which we have tried to set forth in the above discussion, and partly because of the anecdotal nature of data presentation in many of the research publications.

At the very least, one must insist that published transcripts be readable. Or as Du Bois (1991) states: A transcription system must "make explicit, consistent, and unambiguous the meaning attached to every notation" (84).

A number of principles can be inferred from our analysis of transcription systems:

(1) *Unitary function of notation.* As far as possible, use of letters should be limited to their basic segmental, graphemic function; and punctuation marks should be used only in their traditional role as "standardized marks or signs in written matter to clarify the meaning and separate structural units" (Webster 1983: 955). Both of these requirements are in keeping with the principle that "use of transcriptional symbols with more than one meaning" (O'Connell & Kowal 1990: 456) must be avoided. More specifically, the principle can be violated in two different ways:

> "(1) The first of these is that of multiple use of symbols within the same transcriptional system of a given researcher. (2) The second is that of appropriating symbols already otherwise in use within the larger language system and conferring new and different functions upon them." (456)

Both of these violations are standard components of most of the systems in current use for the scientific transcription of spoken discourse.

(2) *Integrity of word units.* Within-word use of supernumerary letters or punctuation marks should be eliminated. Du Bois (1991) has included this under his *economy* maxim: *"Minimize word-internal notations"* (92).

(3) *Description of nonphonological phenomena.* Nonphonological acoustic phenomena such as laughter should be described rather than transcribed by special syllables, tokens, or supernumerary lettering (see, e.g., Ehlich & Rehbein 1976). Transcription is appropriate as iconic representation of segmental, phonological material. In all other cases, description is sufficient. These principles assume that discourse is the primary research interest. In the case where the duration of laughter, for example, becomes the phenomenon of primary research interest, some transcription system may indeed be appropriate.

(4) *Measurement.* Pseudo-exactitude through the use of numbers, but without exact measurement (e.g., in the reporting of pause duration or syllable prolongation in tenths of seconds), is to be eliminated entirely. Ordinal scaling is not sufficient in cases in which the units of notation suggest to the reader a ratio or interval scale.

This is precisely the mistake made by Bergmann (1982):

> "Notation of duration of silence in seconds has been chosen only because it allows a precise representation of a conversational event, a representation easier to control than other notations that use, e.g., special signs for "short, medium, and long pauses". (146; our translation)

(5) *Parsimony.* Only what is to contribute systematically to data analysis should be transcribed, and only what makes the data analysis intelligible should be presented to the reader in transcripts.

In fact, the primary goal of transcription is *not* really *readability*, but *usability of the transcribed data for purposes of scientific analysis* (see O'Connell & Kowal 1990: 460). This usability is in turn defined by the specific research purposes of the project at hand. This principle is similar to what Ehlich (1993) refers to as "semiotic plausibility": "The transcript must preserve the most essential information in a clear manner, free from excessive amounts of information that might overload the reader and hinder the analytic process" (124). The principle should not, however, be understood as precluding extraordinarily complex data sets. The notion that such data sets should be readable for *researchers* in the sense that they can be scanned as easily as the discourse itself is unrealistic.

On the other hand, the goal of examples provided in published research reports for the benefit of *readers* is quite different. Here the goal is indeed to make the results of analyses intelligible to the reader. Hence, readability is primary for the reader. The paradox is easily solved: *It is not necessary that the transcription system for a scientific analysis be the same as the notational system for the reader.* In fact, the latter is probably best accomplished by simply using the conventions of ordinary written discourse. Couper-Kuhlen (1990) does indeed use a different transcription for her readers as follows:

> "Note that to facilitate reading the intervals are listed here from first syllable to last syllable, although the figures actually correspond to the amount of time elapsing between the vowel onset of the first syllable in one line up to but not including the vowel onset of the first syllable in the next line." (31 f.)

Atkinson (1984) disagrees with much of what we have said here about parsimony in transcribing. He does indeed use transcriptions "simplified from more detailed originals" (190) for the sake of the reader; however, at the same time he advises:

> "It always pays to transcribe in as much detail as possible, even though the resulting hieroglyphics may eventually have to be simplified in the interests of readability." (191)

His argument is couched in terms of the necessity for the researcher to be open to the unending complexity and subtlety of spoken discourse. We would respond that, as in war so in posing research questions, *divide et impera* has a certain wisdom - especially if "it is not unusual to spend an hour in producing a satisfactorily accurate transcript of one minute's worth of talk" (Atkinson 1984: 190).

The abstract ideal of discourse transcription - "To present the multifaceted flux of discourse in a way that is as accessible to the analyst as it is to the

participant" (Du Bois 1991: 97; see also Psathas & Anderson 1990: 87) - must be forever laid to rest. The ideal transcription is the one that best serves the purposes - however modest - of the researcher. Whether these be speculation, instruction, analysis, or some preliminary exploratory curiosity, they are *not* better served by comprehensive transcription. In this regard, we vote precisely for *inflexible* transcription systems, exactly tailored to research purposes. A good transcription system should *not* be "flexible enough to accommodate the needs" (Du Bois 1991: 74) of all kinds of users. Any realistic "field-wide transcription standard" (Edwards 1989: 5) must be limited to the type of very general principles which we have enuntiated above, but such a standard cannot dictate specific transcription systems.

To return to our four sources from which we sought criteria for the adequacy and usefulness of transcription systems in the first place, we may now ask why we have not found their criteria more useful. As it turns out, all these authors were considering criteria from their own point of view. All of them alike had already invested a tremendous amount of work in the development of *a specific transcription system*. They did not have the distance from their own systems to compare the features of various systems critically. John Du Bois (1991) was right: "There is not, nor ever can be, a single standard way of putting spoken word to paper" (73).

Accordingly, there cannot be a single ideal transcription system for research on spoken discourse.

References

Atkinson, J.M. (1984) *Our masters' voices: The language and body language of politics.* London: Methuen.

Atkinson, J.M., and J. Heritage (eds.) (1984) *Structures of social action: Studies in conversation analysis.* Cambridge: Cambridge University Press.

Bergmann, J.R. (1982) Schweigephasen im Gespräch - Aspekte ihrer interaktiven Organisation. In H.-G. Soeffner (ed.), *Beiträge zu einer empirischen Sprachsoziologie.* Tübingen: Gunter Narr, 143-184.

Brinker, K., and S.F. Sager (1989) *Linguistische Gesprächsanalyse.* Berlin: Erich Schmidt.

Bublitz, W. (1988) *Supportive fellow-speakers and cooperative conversations.* Amsterdam: John Benjamins Publishers.

Button, G., and J.R.E. Lee (eds.) (1987) *Talk and social organisation.* Clevedon, United Kingdom: Multilingual Matters.

Chafe, W. (1993) Prosodic and functional units of language. In J.A. Edwards & M.D. Lampert (eds.), *Talking data: Transcription and coding in discourse research.* Hillsdale, NJ: Lawrence Erlbaum, 33-43.

Cosnier, J., and C. Kerbrat-Orecchioni (1987) *Décrire la conversation.* Lyon: Presses Universitaires de Lyon.

Coulthard, M., and M. Montgomery (eds.) (1981) *Studies in discourse analysis.* London: Routledge & Kegan Paul.

Couper-Kuhlen, E. (1990) *Discovering rhythm in conversational English: Perceptual and acoustic approaches to the analysis of isochrony* (Arbeitspapier Nr. 13). Konstanz, Germany: Fachgruppe

Sprachwissenschaft, Universität Konstanz.

Coupland, N., H. Giles, and J.M. Wiemann (eds.) (1981) *"Miscommunication" and problematic talk*. London: Sage.

Cruttenden, A. (1986) *Intonation*. Cambridge: Cambridge University Press.

Davidson, J. (1978) An instance of negotiating in a call closing. *Sociology* 12: 123-133.

Du Bois, J.W. (1991) Transcription design principles for spoken discourse research. *Pragmatics* 1: 71-106.

Du Bois, J.W., S. Schuetze-Coburn, S. Cumming, and D. Paolino (1993) Outline of discourse transcription. In J.A. Edwards & M.D. Lampert (eds.), *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum, 45-89.

Edwards, J.A. (1989) *Transcription and the new functionalism: A counterproposal to CHILDES' CHAT conventions* (Berkeley Cognitive Science Report No. 58). Berkeley, CA: Institute of Cognitive Studies, University of California, Berkeley.

Ehlich, K. (1993) HIAT - A transcription system for discourse data. In J.A. Edwards & M.D. Lampert (eds.), *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum, 123-148.

Ehlich, K., and J. Rehbein (1976) Halbinterpretative Arbeitstranskriptionen (HIAT). *Linguistische Berichte* 45: 21-41.

Ehlich, K., and J. Rehbein (1979) Erweiterte halbinterpretative Arbeitstranskriptionen (HIAT 2): Intonation. *Linguistische Berichte* 59: 51-75.

Ehlich, K., and J. Rehbein (1981) Zur Notierung nonverbaler Kommunikation für diskursanalytische Zwecke (Erweiterte halbinterpretative Arbeitstranskriptionen [HIAT 2]). In P. Winkler (ed.), *Methoden der Analyse von Face-to-Face Situationen*. Stuttgart: J.B. Metzler, 302-329.

French, P., and J. Local (1983) Turn-competitive incomings. *Journal of Pragmatics* 7: 17-38.

Fritz, G. (1982) *Kohärenz, Grundfragen der linguistischen Kommunikationsanalyse*. Tübingen: Narr.

Goodwin, C. (1981) *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Hafez, O.M. (1991) Turn-taking in Egyptian Arabic: Spontaneous vs. drama dialogue. *Journal of Pragmatics* 15: 59-81.

Helm, D.T., W.T. Anderson, A.J. Meehan, and A.W. Rawls (eds.) (1989) *The interactional order: New directions in the study of social order*. New York: Irvington Publishers.

Henne, H., and H. Rehbock (1982) *Einführung in die Gesprächsanalyse*. Berlin: Walter de Gruyter.

Hopper, R., S. Koch, and J. Mandelbaum (1986) Conversational analysis methods. In D.G. Ellis & W.A. Donohue (eds.), *Contemporary issues in language and discourse processes*. Hillsdale, NJ: Lawrence Erlbaum, 169-186.

Houtkoop, H., and H. Mazeland (1985) Turns and discourse units in everyday conversation. *Journal of Pragmatics* 9: 613-615.

Jefferson, G. (9184a) Notes on a systematic deployment of the acknowledgement tokens "yeah" and "mm hm". *Papers in Linguistics* 17: 197-216.

Jefferson, G. (1984b) On the organization of laughter in talk about troubles. In J.M. Atkinson & J. Heritage (eds.), *Structures of social action. Studies in conversation analysis*. Cambridge: Cambridge University Press, 346-369.

Jefferson, G. (1989) Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. Roger & P. Bull (eds.), *Conversation: An interdisciplinary perspective*. Clevedon, United Kingdom: Multilingual Matters, 166-196.

Jefferson, G., H. Sacks, and E.A. Schegloff (1984) *On laughter in pursuit of intimacy*. (Working Papers and pre-publications No. 135, Series C). Urbino, Italy: Centro Internazionale di Semiotica e di Linguistica, Università di Urbino.

Jucker, A.H. (1986) *News interviews: A pragmalinguistic analysis*. Amsterdam: John Benjamins Publishers.

Lenz, F. (1989) *Organisationsprinzipien in mündlicher Fachkommunikation. Zur Gesprächsorganisation von "Technical Meetings"*. Frankfurt/M: Peter Lang.

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

Linell, P. (1982) *The written language bias in linguistics*. Linköping, Sweden: University of Linköping.

MacWhinney, B. (1991) *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum.

Maynard, S.K. (1989) *Japanese conversation. Self-contextualization through structure and interactional management*. Norwood, NJ: Ablex.

McHoul, A. (1978) The organization of turns at formal talk in the classroom. *Language in Society* 7: 183-213.

Nofsinger, R.E. (1991) *Everyday conversation*. Newbury Park, CA: Sage.

Ochs, E. (1979) Transcription as theory. In E. Ochs & B.B. Schieffelin (eds.), *Developmental pragmatics*. New York: Academic Press, 43-72.

O'Connell, D.C. (1991) The spoken flies away; the written stays put. *Georgetown Journal of Languages & Linguistics* 2: 274-283.

O'Connell, D.C., and S. Kowal (1990) Some sources of error in the transcription of real time in spoken discourse. *Georgetown Journal of Languages & Linguisticss* 1: 453-466.

Ong, W.J. (1982) *Orality and literacy*. London: Methuen.

Ozaki, A. (1989) *Requests for clarification in conversation between Japanese and Non-Japanese* (Pacific Linguistics, Series B-102). Canberra, Australia: Australian National University.

Posner, R. (1985) Nonverbale Zeichen in öffentlicher Kommunikation. *Zeitschrift für Semiotik* 7: 235-271.

Potter, J., and M. Wetherell (1987) *Discourse and social psychology: Beyond attitudes and behaviour*. London: Sage.

Psathas, G. (1979) *Everyday language: Studies in ethnomethodology*. New York: Irvington Publishers.

Psathas, G., and T. Anderson (1990) The 'practices' of transcription in conversation analysis. *Semiotica* 78: 75-99.

Sacks, H., E.A. Schegloff, and G. Jefferson (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696-735.

Scannell, P. (ed.) (1991) *Broadcast talk*. London: Sage.

Schegloff, E.A. and H. Sacks (1973) Opening up closings. *Semiotica* 7: 289-327.

Schenkein, J. (ed.) (1978) *Studies in the organization of conversational interaction*. New York: Academic Press.

Streeck, S. (1989) *Die Fokussierung in Kurzzeittherapie*. Opladen, Germany: Westdeutscher Verlag.

Watson, G., and R.M. Seiler (eds.) (1992) *Text in context: Contributions to ethnomethodology*. Newbury Park, CA: Sage.

Webster's ninth new collegiate dictionary. (1983) Springfield, MA: G & C Merriam.

West, C., and D.H. Zimmermann (1982) Conversation analysis. In K.R. Scherer & P. Ekman (eds.), *Handbook of methods in nonverbal research*. Cambridge, United Kingdom: Cambridge University Press, 506-541.

Wowk, M.T. (1989) Emotion talk. In B. Torode (ed.), *Text and talk as social practice. Discourse difference and division in speech and writing*. Dordrecht, Holland: Foris Publications, 51-71.

Zimmermann, D.H., and C. West (1975) Sex roles, interruptions and silences in conversation. In B. Thorne & N. Henley (eds.), *Language and sex: Difference and dominance*. Rowley, MA: Newbury House, 105-129.