ARE TRANSCRIPTS REPRODUCIBLE?[1]

Daniel C. O'Connell and Sabine Kowal

Abstract

The research reported here is part of a larger psycholinguistic project on transcribing and the use of transcripts. It is hypothesized that reproducing transcripts originally prepared on the basis of current transcription systems overloads the capability of those who carry out transcript reproduction and therefore occasions an excessive error rate. Ten reproduced transcripts were taken from (a) three current textbooks (Duranti 1997; Garman 1990; Whitney 1998), and from (b) an earlier textbook (Levinson 1983); and (c) six versions were taken from a German transcript (Keppler 1987). Additions, deletions, substitutions, and relocations of notations were identified according to five categories: Verbal, prosodic, paralinguistic, extralinguistic, and format components. The hypothesis is supported: The overall rate of change is 6.6 syllables per change (2032/308) across all 41 comparisons. Factors underlying this excessive amount of change are discussed. The proposal is made that *only* those notations be made which are to be used for analyses in keeping with the purposes of the research in question.

Keywords: Transcripts, Spoken discourse, Notation systems, Research methodology

1. Introduction

The scientific use of transcription systems is intended to facilitate graphic presentations and scholarly analyses of spoken corpora which would otherwise be less accessible to research. More specifically, excerpts of transcripts have come to be used as a standard methodological device in publications on conversation analysis. Accordingly, Levinson (1983) reproduces 124 such excerpts in one chapter on "Conversational structure" (284-370). In the same chapter, he asserts the importance of such practice: "As anyone who works on conversational data knows, heavy reliance inevitably comes to be placed on transcriptions" (295).

O'Connell and Kowal (1999; Kowal & O'Connell in press b) have recently reviewed the growing literature on current notation systems for transcription in the social sciences. There is considerable evidence that these systems are not as internally coherent, readable,

and interpretable as has been claimed by their authors. We have contended that there is also a need for further empirical research on transcribing and on the use of transcripts.

On the face of it, our contention may seem rather implausible, but historically there have been a number of notable failures in the pursuit of legible representations of oral discourse. For example, output of the Sound Spectrograph was originally referred to as Visible Speech (Denes & Pinson 1963: 120; Potter, Koop, & Green 1947), and it was hoped that it could "transcribe" spoken discourse into legible text for deaf persons. Such was not to be the case; they were unable to learn the system. Similarly, it was hoped that the International Phonetic Alphabet (IPA) would make speech sounds legible to at least professional phoneticians. This project too had very limited success (Crystal 1987: 158; *The New Encyclopedia Britannica* 1985: 353). Nonetheless, Johnson (2000) emphasizes the need for such phonetic transcription. Use of IPA is better suited to elucidate the morphosyntactic development in children's speech because it avoids the danger inherent in orthographic transcription of obscuring "a relatively continuous process of development from formulaic to productive mastery of question forms" (Ratner & Menn 2000: 16).

Much discussion of problems involved in the production and use of transcripts has been focussed on folkloric transcription systems and their theoretical assumptions. Preston (1982) argues for the minimization of deviations from standard orthography such as *eye dialect* insofar as these deviations are already explicable in terms of phonological rules and, in any event, use of *eye dialect* manifests disrespect for the status of the original speaker. Whereas Preston is concerned only with the verbal component of spoken discourse, Fine (1984: xi) engages transcription as "a systematic record of a performance", including nonverbal and contextual features. On the basis of Hirsch's (1976: 32) Piagetian "*corrigible schemata*", she argues for a transcript of dramatic performance detailed enough to allow restoration of the original through readers' performance. However, the sample transcript given by Fine (1984: 184-195) appears to be quite unperformable by reason of readers' overload. Urban (1996) takes an anthropological approach. He compares trancripts produced by two native speakers of a Brazilian indigenous group with the original discourse. Interestingly enough, the changes he observes are quite similar to those identified by O'Connell and Kowal (1994) in untrained German transcribers. In both studies, transcribers intuitively created texts characteristic of formal written prose rather than of the original spoken discourse. In other words, lay transcribers make use of their own implicit criteria to select what they consider "the original instance of discourse" (Urban 1996: 21). At the very least, these findings suggest a warning to researchers who allow untrained personnel to transcribe spoken discourse.

Another way of approaching problems in the production and use of transcripts is to investigate how accurately they are reproduced in scholarly publications. This is clearly an appropriate preoccupation for the social sciences, insofar as the capacities and limitations of scholars, textbook authors, editors, proofreaders, and typesetters determine the adequacy of the reproduction. The citation of an excerpt from a transcript must be trustworthy if it is to be used as a scientific example. Whether such excerpts are indeed trustworthy is therefore a legitimate empirical question. We are not concerned with the generic question regarding occasional typographical errors (which find their way into nearly all printed texts), but only with the more specific question regarding the incidence of changes which occur in transcript reproduction. These are in some instances deliberate on the part of an author whose purposes are somewhat different from those of the original

author. However, the incidence of unmotivated and erroneous changes must also be *investigated.*

In order to familiarize our readers with the sort of phenomenon we will be dealing with, we cite here an excerpt of a transcript from Schegloff, Jefferson & Sacks (1977: 378):

Example 1

Lori:	But y'know single beds'r awfully thin tuh sleep on.
Sam:	What?
Lori:	Single beds. // They're—
Ellen:	→Y'mean narrow?
Lori:	They're awfully narrow // yeah.

Disregarding changes involving the arrow and underlining, we will report below (see Table 3) five changes in this excerpt, as number (8) in our data base reproduced in Levinson (1983: 342).

In questioning the reproducibility of transcripts, we wish to use our own experience as psycholinguists as a pilot study. In fact, our own efforts to copy, use, and interpret such excerpts have demonstrated to us repeatedly that we ourselves misread, miscopy, misperceive, and overlook elements of such transcripts with excessive frequency and must constantly and carefully check one another's drafts and counts. We have noted these errors particularly with respect to prosodic and paralinguistic details, but also with respect to verbal content.

In a recent article, Kitzinger (1998: 137) characterizes these types of transcript errors rather harshly: "Inaccuracy is the norm". In a set of 50 instances taken from Burman and Parker's (1993) edited text *Discourse analytic research*, she finds that "32 percent of the authors' quotations from data already presented in their own chapters are misquoted" (Kitzinger 1998: 138). This she ascribes quite explicitly to authors' carelessness.

Our own purpose in the following is to investigate whether such errors may reveal problems in the transcription systems themselves. Along with Kitzinger, we do not wish to challenge the scholars whose works we cite. But we do wish to derive our data base from their reproduced transcripts, just as she did, in order to test the following hypothesis: Psycholinguistically, reproducing transcripts originally prepared on the basis of current transcription systems overloads the capability of those who carry out transcript reproduction and therefore occasions an excessive error rate. If our hypothesis proves tenable, it follows that the usability of transcripts thus reproduced is severely limited.

As data base, we have chosen three corpora: (a) Ten excerpts from three Current Textbooks (Duranti 1997; Garman 1990; Whitney 1998). We have found these books valuable to our own research and teaching. They are authored by reputable scholars (and published by reputable firms), and they are current examples of deliberate reproduction of transcript excerpts for a scientific readership. (b) Ten excerpts from Levinson's (1983) earlier textbook (chapter 6, "Conversational structure": 284-370). (c) Six versions of a German transcript which originally occurred in Keppler (1987: 291).

2. Method

2.1. *Excerpts from current textbooks*

Altogether, ten excerpts of transcripts from a Current Textbooks corpus were analyzed by comparing each reproduced citation with its source in the archival literature. They were taken from the following three textbooks: (1) eight excerpts from Duranti (1997), (2) one from Garman (1990), and (3) one from Whitney (1998). Types of change included addition, deletion, substitution, and relocation (O'Connell & Kowal 1994: 126; cf. Lindsay & O'Connell 1995: 107; Urban 1996: 27 ff.). The categories involved in these changes consisted of *verbal*, *prosodic*, *paralinguistic*, and *extralinguistic* components (in keeping with the taxonomy of Posner 1986) of the spoken discourse itself, along with a *format* component of the primarily nonverbal graphic representations (e.g., changes in transcript headings, lineation and line designation, speaker designation, upper and lower case, dashes and hyphens, spacing, and alignment of successive lines). The five components constitute the change categories analyzed in all the excerpts of reproduced transcripts.

To return to the Current Textbooks data base, there are additional reasons specific to each book for finding it appropriate for our investigation:

(1)   In the case of Duranti's (1997) *Linguistic anthropology*, it is peculiarly appropriate to use his excerpts because all of them appear in his Chapter 5, which has as its specific topic "Transcription: From writing to digitized images" (122-161). Accordingly, eight excerpts, Nos. 8 (138), 11 (140 f.), 12 (141), 13 (142 f.), 15 (146), 16 (155), 18 (157), and 24 (160), according to Duranti's own enumeration with corresponding pages in parentheses, each from a different source, were compared to their original sources. Selection of the eight was not at all random, but neither was it capricious. We chose the first eight which proved accessible to library search. One reference for an excerpt was missing from Duranti's own bibliography and one was erroneous therein. The remaining excerpts in the chapter were not further pursued.

(2)   Garman's (1990) *Psycholinguistics* offers quite a different opportunity. An excerpt from Crystal and Davy (1975: 19 f.) is given on page 114. The original transcript was made in the now classical tradition of the *London-Lund Corpus* (Svartvik 1990; Svartvik & Quirk 1980). The fact that Garman deliberately presented this excerpt in his textbook to explain "the use of descriptive linguistic techniques" (113) makes its inclusion here all the more appropriate.

(3)   A single excerpt from Whitney's (1998) *Psychology of language* is used. The advantage in this case is that it can be traced back to successive sources: Proximately to Clark (1994) and finally to the *London-Lund Corpus*.

These 10 Current Textbooks excerpts vary in length from a few short clauses to an entire page (in the case of Garman 1990: 114). Range of number of syllables is 18 - 507 (*Mdn* = 62 syl). Since we have the audio recording of the longest excerpt, we know that it represents 147 seconds of dialogic speech.

2.2. *Excerpts from Levinson (1983)*

Levinson's textbook in *Pragmatics* has become since its publication in 1983 a standard

introduction. His chapter 6 on "Conversational structure" provides 124 excerpts from transcripts. As data base, we have selected ten of these according to the following criteria. Accessibility to library search was the basic requisite. In addition, longer rather than shorter excerpts were preferred. The excerpts are considered in the same sequence in which they occur in Levinson. Our excerpts numbered (1) through (10) correspond, respectively, to Levinson's numbers and pages (in parentheses): 31 (312), 40 (322), 49 (329 f.), 70 (340), 74 (341), 76 (342), 78 (342), 79 (342), 84 (344), and 104 (352). They vary in length from 7 to 126 syllables (*Mdn* = 28 syl). With the exception of one excerpt, all of Levinson's transcripts are taken from research in conversation analysis.

2.3. *Keppler (1987) versions*

An excerpt from Keppler's (1987) research on the organization of gossip displays the distinct advantage of occurring in at least six versions in the archival literature. The first chronological occurrence is in German in Keppler (1987: Example 4, 291), then twice in German in Bergmann (1987: Example 6, 124; 150), twice in the English translation of Bergmann (1993: Example 6, 89; 108), and finally, again in German, in Herrmann and Grabowski (1994: 45). The selection of this excerpt was not arbitrary; it came to our attention during the preparation of Kowal and O'Connell (in press a) when we came upon it in Herrmann and Grabowski (1994) and traced it back to its source.

Keppler's (1987) original excerpt is 20 syllables in length. In all the other instances, the original is expanded to more syllables, with the limit set by Bergmann's (1987: 124) use of 40 syllables. The two English excerpts have only 29 syllables.

3. Data analyses

The entire data base consisted of 21 excerpts. However, since each of the six Keppler excerpts was compared with each other Keppler excerpt, including self-comparisons, a total of 41 comparisons were made. These comparisons included a total of 2032 syllables and 308 changes. Hence, the overall syl/change rate is 6.6. For the Current Textbooks corpus, this overall rate was 982/81 = 12.1 syl/change; for the Levinson corpus, 385/56 = 6.9 syl/-change; and for the Keppler corpus, 665/171 = 3.9 syl/change. In other words, the overall syl/change rate in the Keppler corpus is more than three times the rate of the Current Textbooks corpus, with the Levinson corpus intermediate to both.

3.1 *Current Textbooks*

Throughout the following section, the respective ten excerpts are introduced with a paragraph and an initial arabic numeral in parentheses. These designations correspond with our numbering of excerpts in Table 1, where the number of syllables in each of Duranti's original sources, the number of independent changes introduced into the reproduced excerpt of each transcript, and the syllables per change (*Mdn* = 6.6 syl/change) are recorded.

Table 1. Number of Syllables in the Ten Original Sources of Current Textbooks Excerpts, Number of Independent Changes Introduced in the Reproduced Excerpts of the Source Transcripts, and the Syl/Change Ratio

Measure	Reference	Excerpt									
	Reproduced				Duranti (1997)				Garman (1990)	Whitney (1998)	All
	Source	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Syl		40	25	18	57	23	84	77	66	507	85	982
Changes		7	1	8	14	12	1	4	10	4	20	81
Syl/Change		5.7	25.0	2.2	4.1	1.5	84.0	19.2	6.6	126.8	4.2	12.1

(1) Chafe (1980: 304)

(2) Sacks, Schegloff, & Jefferson (1978: 28)

(3) Pomerantz (1984: 96)

(4) Goodwin & Goodwin (1992: 161, 163, 166, 168)

(5) Goodwin (1981: 131-133)

(6) Sherzer (1983: 75)

(7) Briggs (1986: 78)

(8) Kulick (1992: 203)

(9) Crystal & Davy (1975: 19 f.)

(10) Clark (1994: 999)

3.1.1. *Duranti (1997)*

The comparisons of the respective excerpts from Duranti (1997) are presented in the following with their bibliographical sources in the same sequence in which they occur in Duranti's text, but with our enumeration (1) - (8).

(1) The first is from Chafe (1980: 304). The entire excerpt is 40 syllables in length. The following changes (in their original sequence) occur: Deletion of a period, substitution of a pause duration, relocation of a closing square bracket, the substitution of a dash for a double hyphen, and three times the substitution of a triple ellipse for two periods. Altogether there are two changes in prosodic notation and five in format.

The following remarks of Duranti (1997: 138), prefatory to the transcript, seem quite unwarranted in view of the fact that all the notation conventions he mentions are violated in the reproduced transcript:

> It is very readable given that there are only a few extra conventions that a reader must learn, mostly about pauses (between square brackets or with two periods) and lengthened sounds (the symbol "—").

(2) An excerpt from Sacks, Schegloff, and Jefferson (1974: 716) of 25 syllables is essentially intact. An unobtrusive detail of format is the substitution of all capital letters for initial capitals for the speakers' names. Hence, there is only one change in format.

It should be noted that the secondary source (Sacks, Schegloff, & Jefferson 1978: 28) is the only one given by Duranti (1997: 140 f.).

(3) The excerpt from Pomerantz (1984: 96) is the shortest (18 syllables in length). It involves the substitution of an apostrophe for a superscript period, the addition of a comma, the substitution of a single quotation mark for an apostrophe, the deletion of separate underlines for a set of three graphemes and for one separate grapheme, the substitution of a dash for a double hyphen, the substitution of two initial square brackets for a single encompassing one, and the addition of the letter "r" in the final word "*mar*verlous" (Duranti 1997: 141). Altogether there is one change in verbal notation, and there are three changes in prosodic notation and four in format.

(4) An excerpt from Goodwin and Goodwin (1992) is far more complex. The basic component of the excerpt is only 12 syllables long (Goodwin & Goodwin, Example [4]: 161), but in five of the six following versions, an additional three syllables are presented from overlapping talk from another speaker.

One of the problems in this case is that there are unexplained and apparently arbitrary changes in the graphemic presentation of four syllables of these versions in the original source. But the four successive versions in Duranti (1997: 142 f.) do not correspond to the respective versions in the original. All the versions of the four syllables are presented in Table 2.

In Duranti's excerpts, all capitals are substituted for initial capitals of speakers' names, a speaker's name and two italicized commentaries are relocated, "wz" is substituted once for "was" and "waz" once for "was" in the corresponding versions of the original (see Table 2), two successive spaces are deleted from between colons in two cases (see Table 2), and the vertical legs of six (four single and two double) horizontal braces are deleted. Altogether there are five changes in verbal notation and nine in format.

Table 2. Transcripts of the Same Text in the Source (Goodwin & Goodwin 1992) and in the Reproduced Excerpt (Duranti 1997)

	Goodwin & Goodwin (1992)		Duranti (1997)	
1.	it wz s::so: goo:d.	(161)	it wz s::so: goo:d.	(142)
2.	It wz s::so: goo:d.	(162)		
3.	it wz s : :so : goo:d.	(163)	it wz s::so : goo:d.	(143)
4.	it was s : : so : goo:d.	(166)	it wz s::so : goo:d.	(143)
5.	it was s : : so : goo:d.	(168)	it waz s : : so : goo:d.	(143)
6.	it wz s::so : goo:d.	(169)		
7.	it was s : : so : goo:d.	(169)		

(5) One of the excerpts, 23 syllables in length, is referred to two sources (Goodwin 1979; 1981). A difficulty arises in that the transcript does not appear in exactly the same format in the two sources themselves. Another difficulty is that Duranti's (1997: 146) version is a composite of transcript components from various pages in the original sources. We have compared Duranti's excerpt first with the version in Goodwin (1981: 131, 133). In the following sequence, Duranti (1997: 146) introduces three changes associated with gaze movement: Two deletions of a period and a substitution of two periods for two commas; he substitutes a comma for a period and deletes a comma; he deletes the extension of an underline beyond the end of a word and relocates the bracket beneath that line to the right with respect to the following line; finally, three and then two periods are deleted and the remaining periods relocated with respect to the previous line. Two brackets are deleted between the last and the preceding line. Altogether there are two changes in prosodic notation, nine in extralinguistic notation, and one in format. This is the only excerpt in which changes involve extralinguistic notation (specifically for gaze movement).

In Goodwin (1979), the major differences from Goodwin (1981) are the use of underlining instead of italics, the spelling of "t'day" and "acshilly" (Goodwin 1979: 99) instead of "today" and "actually" (Goodwin 1981: 133), and the use of an equal sign (=) after the second line. These differences were introduced by Goodwin himself and are therefore not included in our analysis of Duranti (1997: 146).

(6) The excerpt from Sherzer (1983) is the longest (84 syl) of the eight reproduced in Duranti (1997) and the only one in which a long section of verbal text is deleted:

CC: In truth here.
       In truth I am.
       In good health.
       I utter.
RC: Indeed. (Sherzer 1983: 75)

The Duranti (1997: 155) version simply skipped ahead to text under the *next* following designation "RC: Indeed". This is the longest change (15 syl) introduced into any of the excerpts. However, since the entire verbal deletion was clearly occasioned by a single misreading, it is conservatively counted as only one change in Table 1. Hence there is only one change in verbal notation.

(7) In the case of Briggs' (1986: 78) 77-syllable excerpt, the designation "Disputant" has been moved to the right in Duranti (1997: 157), the word "ese" substituted for the word "eso", the last line relocated as if it were an independent clause, and the word "it" deleted from the last line. The last mentioned change entirely alters the meaning of the English translation. Altogether there are three changes in verbal notation and one in format.

(8) Finally, the 66-syllable excerpt from Kulick (1992: 203), in a New Guinea local vernacular called Taiap, entails 10 changes in Duranti's (1997: 160) excerpt: Relocation of parts of lines in both the Taiap and English versions, deletion of three initial capitals, deletion of two periods, deletion of the letter "a" from a Taiap word, deletion of italics from the first word of the English version, and deletion of the speaker designation ("S:") in the English version. Altogether there is one change in verbal notation and there are three changes in prosodic notation and six in format.

3.1.2. *Garman (1990)*

(9) A comparison of the excerpt from Crystal and Davy (1975: 19 f.) with its reproduction in Garman (1990: 114) revealed two substitutions (a "tone-unit boundary" for a "first prominent syllable of the tone-unit" and "the next syllable is stressed" for "extra strong stress"), one relocation of "level tone" to a preceding grapheme, and one deletion of a "first prominent syllable of the tone-unit" (Crystal & Davy 1975: 17). Hence there are four changes in prosodic notation.

In addition to the excerpt analyzed here, further evidence that prosodic notation is subject to error can be found. In his definition of the brief pause, Garman (1990: 119) substituted a period at the bottom of the line for a centered period. In his examples, he consistently used the lower instead of the centered period (e.g., 119, 120, 123, 124). In referring back to the excerpt used here, he substituted one unit pause for a brief pause (133) and one brief pause for a unit pause (120). It should be noted that Svartvik and Quirk (1980: 25) were themselves not consistent in their use of medial periods for pause notations.

3.1.3. *Whitney (1998)*

(10) The 85-syllable excerpt from Whitney (1998: 285) is characteristic of the occasional usage of transcripts in basic psycholinguistic textbooks. This is the only case in our excerpts from Current Textbooks in which the derivation of the excerpt spans three sources: Whitney reproduced Clark (1994: 999), who in turn had reproduced the excerpt (LLC:o, S.7.1d.1320-1334; see Svartvik, 1990: 31) from the computerized supplement to the *London-Lund Corpus* (Svartvik & Quirk 1980). Consequently, after the analysis of Whitney's excerpt from Clark, an analysis of Clark's excerpt from LLC:o is given below. It should be noted that the latter changes are not incorporated into Table 1.

There are 20 changes in Whitney's (1998: 285) excerpt from Clark (1994: 999): The addition of consecutive numbering of lines, substitution of all capitals for initial capitals for speaker designation, deletion of italics for the entire excerpt, deletion of bold type in two lines at the beginning and two more lines at the end of the excerpt, three substitutions of initial capitals for lower case for the first words of speakers, substitution of "I" for "*i*", six substitutions of hyphens for dashes, deletion of the word "*going*", two word substitutions ("er" for "*are*" and "around" for "*round*"), deletion of a space, and substitution of two stars for two asterisks in "**m* m*". Svartvik and Quirk (1980: 24) used stars in this context, but referred to them as "asterisks", whereas Clark (1994: 987) both refers to "asterisks" and uses them. Altogether there are three changes in verbal notation and 17 in format.

A comparison of Clark's excerpt with its source in LLC:o, S.7.1d.1320-1334; see Svartvik, 1990: 31) revealed a great number of changes. First, it should be noted that the LLC:o transcript (Text S.7.1d.1320-1340) had already been "reduced" so as to omit "all paralinguistic features and certain indications of pitch and stress" (Svartvik 1990: 15). In other words, there is no access to the full original transcript. Clark's (1994: 999) version deleted an additional 12 notation signs (for a total of 72 instances), substituted five (17 instances), and added two (11 instances). He also relocated parts of eight lines, added

italics throughout and bold type for two sets of two lines each, added designations for speakers' names, and deleted the designation for surreptitiousness. After a total of 100 instances of notations had been changed in the text itself, little was left beyond the verbal components themselves. Nonetheless, as indicated above, Whitney's (1998: 285) excerpt entails 20 more changes. The process from the original spoken discourse to the first transcript, through a "reduced" LLC:o transcript and a reproduced transcript in Clark, and finally to Whitney's version is indeed very long - and a marvelous example of the ordeals of transcript reproduction.

3.2. *Levinson*

Table 3 presents the number of syllables in the text of the original sources of the ten Levinson (1983: Chapter 6) excerpts, the number of changes introduced in the reproduced excerpt of each transcript, and syl/change (*Mdn* = 6.8 syl). In keeping with our conservative criteria, Levinson's use of arrows to designate locations of specific interest and his substitution of italics for underlining are not entered in Table 3 as changes. The comparisons of the respective excerpts from Levinson are presented in the following with their bibliographical sources in the same sequence in which they occur in Levinson's text, but with our enumeration (1) - (10).

(1) The first excerpt (29 syl) reproduced in Levinson (1983: 312) is taken from Schegloff (1979: 47). There are two changes: The substitution of a word and a change in speaker designation. Hence, there is one change in verbal notation and one in format.

(2) The second excerpt (16 syl) reproduced in Levinson (1983: 322) is from Schegloff & Sacks (1973: 320). It entails three changes: The substitution of a word and two changes in speaker designation. Hence, there is one change in verbal notation and there are two changes in format.

(3) The third excerpt (126 syl) reproduced in Levinson (1983: 329 f.) is taken originally from an unpublished microfiche (Schegloff 1976), published later in Schegloff (1984: 28). It involves a total of 19 changes. Enumeration of lines is changed into enumeration of turns; four spaces are added and two deleted; a comma and a period are added, and a comma is deleted; two verbal substitutions are made; one speaker designation is deleted and three are added; two slashes are substituted for a bracket; one initial capital is deleted and one is added. Altogether, there are two changes in verbal and three in prosodic notations and 14 in format.

(4) The fourth excerpt (22 syl) reproduced in Levinson (1983: 340) is taken from Schegloff, Jefferson, and Sacks (1977: 364). The three changes include: One addition of an apostrophe, one deletion of a space, and one substitution of word alignment. Altogether, there is one change in prosodic notation and there are two changes in format.

(5) The fifth excerpt (27 syl) reproduced in Levinson (1983: 341) is also from Schegloff, Jefferson, and Sacks (1977: 368). The two changes are: One deletion of prosodic notation and one substitution of format.

Table 3. Number of Syllables in Original Sources of Levinson (1983) Excerpts, Number of Independent Changes Introduced in the Reproduced Excerpts of the Source Transcripts, and Syl/Change Ratio

Measure	Reference	Excerpt										
	Reproduced	Levinson (1983)										
	Source	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	All
Syl		29	16	126	22	27	46	17	28	7	67	385
Changes		2	3	19	3	2	9	5	5	1	7	56
Syl/Change		14.5	5.3	6.6	7.3	13.5	5.1	3.4	5.6	7.0	9.6	6.9

(1) and (9)	Schegloff (1979: 47; 55)
(2)	Schegloff & Sacks (1973: 320)
(3)	Schegloff (1984: 28)
(4), (5), (6), (7), and (8)	Schegloff, Jefferson & Sacks (1977: 364; 368; 364; 377; 378)
(10)	Labov & Fanshel (1977: 363)

(6) The sixth excerpt (46 syl) reproduced in Levinson (1983: 342) is once again taken from Schegloff, Jefferson, and Sacks (1977: 364). The nine changes include: Three verbal substitutions and one verbal addition; one deletion of prosodic notation; and two deletions, one addition, and one substitution of format. Altogether there are four changes in verbal notation, one change in prosodic notation, and four changes in format.

(7) Again, Schegloff, Jefferson, and Sacks (1977: 377) is the source of the seventh excerpt (17 syl) reproduced in Levinson (1983: 342). The five changes include: Two verbal substitutions, and two deletions and one substitution of format. Altogether there are two changes in verbal notation and three in format.

(8) One more excerpt from Schegloff, Jefferson, and Sacks (1977: 378) is the eighth excerpt (28 syl) reproduced in Levinson (1983: 342). The five changes include: One verbal substitution, one substitution of prosodic notation, and one deletion and two substitutions of format. Altogether there is one change in verbal notation and one in prosodic notation, and there are three changes in format.

(9) The ninth excerpt (7 syl) reproduced in Levinson (1983: 344) is from Schegloff (1979: 55). There is only one change, and it is a format substitution.

(10) The last excerpt (67 syl) reproduced in Levinson (1983: 352) is from Labov and Fanshel (1977: 363). There are seven changes: One deletion and one substitution of prosodic notation; and three deletions, one addition, and one substitution of format. Altogether there are two changes in prosodic notation and five in format.

3.3. *Keppler*

Table 4. Syllables Per Change (Syl/Change) for Comparisons of the Same Excerpt in Six Versions: (1) Keppler (1987: 291); (2) Bergmann (1987: 124); (3) Bergmann (1987: 150); (4) Bergmann (1993: 89); (5) Bergmann (1993: 108); and (6) Herrmann & Grabowski (1994: 45)

Excerpt Used as Standard	Number of Syl	Excerpt Used as Comparison					
		(1)	(2)	(3)	(4)	(5)	(6)
(1)	20	6.7	2.8	2.5	2.2	2.5	1.8
(2)	40	-	13.3	40.0	2.7	3.1	5.0
(3)	40	-	-	20.0	2.8	3.1	5.7
(4)	29	-	-	-	9.7	4.1	1.5
(5)	29	-	-	-	-	14.5	1.9
(6)	40	-	-	-	-	-	13.3

In Table 4, the number of syllables and of syllables per change (syl/change) in the six versions of the Keppler text are summarized. Each version is compared with all the others. The number of changes in individual excerpts ranges from 1 to 22 (*Mdn* = 8.0). The comparison of each version with itself will appear to the reader at the very least rather odd. However, in each version, there are elements which are already erroneous or not appropriate to the system of notation according to which the transcript was derived. In other words, there is a de facto form and a canonical form of each excerpt. It is these which are entered as self-comparisons. In the following, each comparison is listed in parentheses at the beginning of a paragraph.

(1 vs. 1)   The very first version (Keppler 1987: 191) contains three changes: A typesetter's error for lines 24 and 25, the notation of impossible simultaneous speech of the speaker with herself, and the notation of an exclamation mark, which is not part of the transcription system. Excerpt (1) is given below as examples 2 and 3 in its de facto form in Keppler and in its canonical form:

Example 2 (Excerpt 1, de facto form in Keppler 1987: 191)

        23 Johanna: Un der Dande Berta hats schon gut
        2425 Martha: gfalln;
        26 Anna: OH: ⌐ob und wie!
                              └Der! Oh Gott=e=Gott

Example 3 (Excerpt 1, canonical form)

        23 Johanna: Un der Dande Berta hats schon gut
        24   gfalln;
        25 Martha: OH: ⌐ob und wie
        26 Anna: └Der Oh Gott=e=Gott

As in our analysis of Levinson's excerpts, Keppler's substitution of italics for underlining is not considered a change according to our conservative criteria. The further comparisons of each excerpt with the other versions are also made conservatively in that they are made with the excerpt in its canonical form.

(2 vs. 2; 3 vs. 3; 4 vs. 4; 5 vs. 5; & 6 vs. 6)   Each of the remaining five self-comparisons manifests two or three changes as well. All but one of these (a deletion) consist in the addition of a prosodic or format notation not in the transcription system.

(2 vs. 3 & 4 vs. 5)   These two comparisons are similar in that they should be expected to manifest the smallest number of changes of any comparisons, since each of them is a successive appearance of the same excerpt in the same publication. In fact, the German comparison (2 vs. 3) manifests only a single prosodic substitution, whereas the English translation (4 vs. 5) manifests seven changes. The English excerpts are given below:

Example 4 (Excerpt 4, Bergmann 1993: 89)

    [6]   *[Sicilian*: AK: EM 14B]
    J: And Aunt Bertha really liked it;
    M: Oh, ⌐and how!
    A: ⌐*Her*! My God=er=my God. Sunday morning I thought,
           "ha - I can't call her yet"...

Example 5 (Excerpt 5, Bergmann 1993: 108)

    [18/6]   *[Sicilian*: AK: EM 14B]
    J: And Aunt Bertha really liked it;
    M: Oh,⌐and how!
    A: ⌐*Her*! Oh my God=er=God. Sunday morning I thought
           "you can't call her yet"...

In example 5, bold print is deleted from the bracket format; in the A line, a second instance of the verbal notation "Oh" is added; the second instance of the verbal notation "my" is deleted; and the prosodic notation "," is deleted; and in the last line, the verbal notation "ha" and the prosodic notation "—" are deleted, and the verbal notation "you" is substituted for "I".

(2 vs. 4; 3 vs. 4; 2 vs. 5; & 3 vs. 5) All these comparisons involve changes between German and English excerpts, all of which, however, are by the same author, i.e., (4) is a translation of (2), and (5) is a translation of (3). Nonetheless, these comparisons introduce 15, 14, 13, and 13 changes, respectively. Since the comparison (2 vs. 3) introduces only one change, a prosodic substitution, (2) and (3) remain very similar to one another. Hence, it is fair to argue that the many changes noted here are due to the translation of the German into English. We have selected comparison (2 vs. 4) to illustrate these differences. Excerpt (4) has been shown already above as example 4 and below is given excerpt (2) as example 6:

Example 6 (Excerpt 2, Bergmann 1987: 124)

    <6>   <*"Sizilianisch"*: AK: EM 14B>

    01   J: Un der Dande Bert*a* hats schon gutgfallen;
    02   M: OH: ⌐Ob und wie!
    03   A: ⌐*Der*! Oh Gott=e=Gott. I hob noch morgens-
    04   am Sonndoch morgens denk i "ha jetz kannsch
    05   noch net ouruafe"...

In example 4, enumeration of lines is deleted; *eye dialect* in German is translated into standard orthography in English; in the name "Bert*a*," a verbal addition and a prosodic deletion occur; a prosodic substitution of "," for ":" occurs; the space before the bracket is deleted; bold format is added to the bracket; "my God" is substituted for "Oh Gott"; "my" is added before the second instance of "God"; "I hob noch morgens am Sonndoch

morgends" is reduced by deletion to "Sunday morning"; a prosodic notation ("-") is deleted; a past tense ("I thought") is substituted for a present tense ("denk i"); a short pause ("—") is added; and two pronouns are changed in the English (the second person "kannsch" becomes "I" and "her" is added).

(1 vs. 4 & 1 vs. 5) Without entering into the same detail as above, we note that these two comparisons involve even higher syl/change rates, apparently because of the language change.

(4 vs. 6 & 5 vs. 6) Both these comparisons involve not only a language change and an author change, but also a deliberate change in notation system. This last type of change is not itself included in our list of changes, but it may well be the source of other changes by reason of confusion. Since (4 vs. 6) manifests the highest syl/change rate (29/19 = 1.5) in the entire set of excerpts, we present it in detail. In example 7, excerpt (6) is reproduced:

Example 7 (Excerpt 6, Herrmann & Grabowski 1994: 45)

J: un der Dande BertA hats schon gutgfallen
M: <o:h> ┌ob und wie:
               └DER oh gott=e=gott\i hob noch morgens ★ am sonndoch
        morgends denk i „ha jetz kannsch noch net ouruafe

The changes introduced from the English to the German are the following: Standard orthography becomes *eye dialect*; the transcript heading is deleted; a prosodic notation (";") is deleted; "Oh," becomes "<o:h>", which involves the addition of syllabic prolongation (":"), prosodic addition (<oh>) and deletion (intonation); deletion of a space after the bracket; prosodic substitution of ":" for the first instance of "!"; prosodic deletion of the second instance of "!"; verbal substitution of "oh" before "gott" for "My"; format substitution of "=" for "="; verbal deletion of "my" before the second instance of "gott"; verbal addition of "i hob noch morgens"; prosodic addition of "*"; prosodic deletion of ","; substitution of present tense in "denk i"; prosodic deletion of "—"; verbal substitution of second person; and verbal deletion of "her".

3.4. *Further analyses of all three corpora*

In Table 5, a summary of the number and percentage of change categories (verbal, prosodic, paralinguistic, extralinguistic, and format) in all three corpora (Current Textbooks, Levinson, and Keppler) of reproduced transcripts is presented. Of the 308 changes, 77 (25%) were verbal, 91 (30%) prosodic, none paralinguistic, 9 (3%) extra-linguistic, and 131 (42%) format. In other words, more than half of the changes involved some aspect of the speakers' behavior. However, a further inspection of the individual corpora indicates that this is true only of the Keppler corpus. In both the Current Textbooks and Levinson corpora more than half the changes involve format.

The types of changes in the reproduced excerpts are also of interest. Of the 308 changes, 113 (37%) were deletions, 72 (23%) additions, 111 (36%) substitutions, and 12 (4%) relocations. Deletions and substitutions together account for almost three quarters of all changes.

Pairings of change categories with types of changes are entered in Table 6. Both the Current Textbooks and Levinson corpora are primarily by format and format substitutions, whereas the Keppler corpus is equally characterized by verbal substitutions, prosodic deletions, and format deletions.

Table 5. Number (No.) and Percentage (%) of Categories (Verbal [Ve], Prosodic [Pr], Paralinguistic [Pa], Extralinguistic [Ex], and Format [Fo]) Involved in Changes and of Types of Changes (Deletions [De], Additions [Ad], Substitutions [Su], and Relocations [Re]) in Three Corpora of Transcript Excerpts (Current Textbooks [Duranti 1997, Garman 1990, & Whitney 1998], Levinson 1983, & Keppler 1987)

Corpus	Statistic	Category Involved					Σ	Type of Change			
		Ve	Pr	Pa	Ex	Fo		De	Ad	Su	Re
Textbooks	No.	14	14	0	9	44	81	35	9	26	11
	%	17	17	0	11	54	100	43	11	32	14
Levinson	No.	11	9	0	0	36	56	18	13	25	0
	%	20	14	0	0	66	100	32	23	45	0
Keppler	No.	52	68	0	0	51	171	60	50	60	1
	%	30	40	0	0	30	100	35	29	35	1
All	No.	77	91	0	9	131	308	113	72	111	12
	%	25	30	0	3	42	100	37	23	36	4

Table 6.   Number (No.) and Percentage (%) of Pairings of
Categories (Verbal [Ve], Prosodic [Pr], Paralinguistic [Pa],
Extralinguistic [Ex], and Format [Fo]) with Types of Changes
(Deletions [De], Additions [Ad], Substitutions [Su], and
Relocations [Re]) in Three Corpora of Transcript Excerpts
(Current Textbooks [Duranti 1997, Garman 1990, & Whitney 1998],
Levinson 1983, & Keppler 1987)

Corpus	Type of Change	Category Involved					
		Ve	Pr	Pa	Ex	Fo	Σ
Textbooks	De	4	8	0	7	16	35
	Ad	2	1	0	0	6	9
	Su	4	4	0	1	17	26
	Re	4	1	0	1	5	11
	Σ	14	14	0	9	44	81
Levinson	De	1	4	0	0	13	18
	Ad	1	2	0	0	10	13
	Su	9	3	0	0	13	25
	Re	0	0	0	0	0	0
	Σ	11	9	0	0	36	56
Keppler	De	9	25	0	0	24	58
	Ad	18	20	0	0	12	50
	Su	25	18	0	0	15	58
	Re	0	3	0	0	0	3
	Σ	52	66	0	0	51	169
All	De	14	37	0	7	55	111
	Ad	21	23	0	0	28	72
	Su	38	25	0	1	45	109
	Re	4	4	0	1	5	14
	ΣΣ	77	89	0	9	133	308
	%	25.0	28.9	0	2.9	43.2	

4. Discussion

The research hypothesis, that transcript reproduction overloads the capability of those who carry out the reproduction and occasions an excessive change rate, is supported by the evidence. Transcript notations seem not to be reliably reproducible and must therefore be used with considerable caution. Were simple typographical reproductions (e.g., quotations from another source) in scholarly textbooks to exhibit anything near such a change rate (*Mdn* = 5.6 syl/change for the entire combined corpus of 41 excerpts; range = 1.5 - 126.8), there would be an outcry in the scientific and publishing communities. The absence of an outcry in the case of reproduced transcripts suggests that the readership has not been paying close attention to transcript notation in any event.

One could well ask why the present authors have not followed Kitzinger's (1998) example in reproducing for their readership the full set of excerpts used in this

investigation, since the most cogent way to demonstrate the untrustworthiness of the excerpts is visual comparison. However, we have found that, in addition to our own tendency to introduce changes into the excerpts, the ability of typesetters and proof-readers to deal with these notations seems also to be severely limited. In other words, reprinting the entire set of excerpts would in all likelihood have reproduced the very same empirical phenomenon we are investigating. Despite our meticulous efforts to record changes accurately, it is not unlikely that errors exist in the present analyses.

The reader will have noted that we have included all types of changes in our listings above. One of the problems which exists for the scholar who reproduces such excerpts and for the typesetter who carries out the reproduction graphically, is the uncertainty as to which notations are crucially important and which are arbitrary, redundant, or trivial. For example, in most transcripts, designation of speakers by all capitals or initial capitals only is a trivial typesetter's option. But even this cannot be taken for granted. In the *London-Lund Corpus* (Svartvik & Quirk 1980: 26), non-surreptitious speech is designated precisely by lower case entries for speakers. And for psycholinguistic analyses, differences between surreptitious and non-surreptitious speech can be very important. What notations *are* important in a given study can only be answered in view of the purposes and methods of the individual research project.

Nonetheless, it is reasonable for one to assume that the adoption of *any* special notation on the part of a scholar - or its deliberate reproduction on the part of another scholar - is invested with importance for a given research project. This assumption may have to yield to contrary evidence, if it is found that much of current notation serves only what must be called a pretentious or cosmetic purpose. Such notation is not scientific. O'Connell and Kowal (1999) have given numerous examples of notation which serves only as embellishment, but remains otherwise unexplained and unused by the researcher who has introduced it into a transcript or has reproduced it in an excerpt.

There are a number of factors which seem to underlie the excessive amount of change observed in the reproduced excerpts investigated in this study.

Density. What characterizes most of the excerpts, in contrast to ordinary prose texts, is a high density of heterogeneous notations. And it is precisely these unusual notation conventions which obviously occasion the vast majority of errors in the reproduction of these excerpts. The density of changes corresponds closely to the density of notation signs. The fact that three of the excerpts from the Current Textbooks corpus, (6), (7), and (8), consist of only minimally dense notations (mostly verbal) makes our analysis of changes all the more conservative.

Unfamiliarity. We are dealing here with a very different phenomenon from simple graphemic notation. As Edwards (1996: 323) has put it, "The initial difficulty of reading CA [Conversation Analysis] transcripts is largely a matter of our overwhelming familiarity with conventional written text", but he remains optimistic regarding "the advantage of reading and producing formal transcripts" and our ability to adjust to them. However, almost our entire experience with reading and writing in the western world is limited to graphemic-phonemic correspondences. In the process of learning our native languages, we have not learned how to process graphic notations of the prosodic and paralinguistic phenomena which accompany spoken verbal material. Punctuation marks themselves are not reliable indicators of prosody. As Altmann (1997) has put it, "How we learn language underlies everything we do with language" (232). We have not yet learned to use current

transcript notation systems in our reading and writing.

Overload. In reading and writing, we literally have better things to do with our time than to process supplementary notations. The original ethnomethodological motivation for detailed notation, particularly of prosodic and paralinguistic components of spoken discourse, has been well expressed by Schenkein (1978): "To produce a reader's transcript - one that will look to the eye how it sounds to the ear" (xi). The changes found in the present research indicate once again that this purpose remains unfulfilled. The evidence indicates that current notation systems are not particularly user friendly.

Remoteness. Not included in our data base as reflected in the successive Tables 1-6 are several supplementary analyses which have been mentioned above. The discrepancies between the original transcript of Goodwin (1979) and that of Goodwin (1981) and the discrepancies within Goodwin and Goodwin (1992) itself both indicate that even transcripts reproduced by the same author are subject to change (see also Kitzinger 1998: 138 f.). In our data base itself, this is the case with several transcripts reproduced by the same author (Bergmann 1987; 1993). Transcripts reproduced by different authors demonstrate another source of remoteness, as in the Whitney (1998) excerpt. This is also the case in comparison (4 vs. 6) in the Keppler corpus. Here multiple authorship combines with a change in language from English to German and with the deliberate shift to another notation system. The result is the highest rate of syl/change in the entire corpus (1.5 syl/change).

5. Conclusions

There were 308 changes overall, across all 2032 syllables of the 41 comparisons from the Current Textbooks, Levinson, and Keppler corpora. This constitutes a rate of one change every 6.6 syllables. This extraordinarily high rate appears *not* to be due to carelessness, as Kitzinger (1998) has suggested, but to notation systems which preclude reproducibility of transcripts by overburdening those involved in their reproduction. This evidence raises questions regarding the practical usability of current notation systems. If the correct reproduction of transcript excerpts is scientifically important, then we must conclude from the present research that these excerpts fail to serve their purpose. Transcripts which are not accurately reproducible do indéed pose serious problems.

In closing, we wish to make a modest proposal, namely that henceforth researchers transcribe spoken discourse with *only* those notations which are to be used for analyses in keeping with the purposes of the research. The resulting transcripts will be less dense and hence easier to reproduce - and an appropriate level of parsimony will be preserved.

References

Altmann, G.T.M. (1997) *The ascent of Babel: An exploration of language, mind, and understanding.* Oxford: Oxford University Press.

Bergmann, J. (1987) *Klatsch: Zur Sozialform der diskreten Indiskretion.* Berlin: de Gruyter.

Bergmann, J. (1993) *Discreet indiscretions: The social organ-zation of gossip.* New York: Aldine de Gruyter.

Briggs, C.L. (1986) *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research.* Cambridge: Cambridge University Press.

Burman, E., & I. Parker (1993) (eds.) *Discourse analytic research: Repertoires and readings of texts in action.* London: Routledge.

Chafe, W. (1980) (ed.) *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Norwood, NJ: Ablex.

Clark, H.H. (1994) Discourse in production. In M.A. Gernsbacher (ed.), *Handbook of psycholinguistics.* San Diego, CA: Academic Press, pp. 985-1021.

Crystal, D. (1987) *The Cambridge encyclopedia of language.* Cambridge: Cambridge University Press.

Crystal, D., & D. Davy (1975) *Advanced conversational English.* London: Longman.

Denes, P.B., & E.N. Pinson (1963) *The speech chain.* Nutley, NJ: Bell Laboratories.

Duranti, A. (1997) *Linguistic anthropology.* Cambridge: Cambridge University Press.

Edwards, D. (1996) *Discourse and cognition.* London: Sage.

Fine, E.C. (1984). *The folklore text: From performance to print.* Bloomington: Indiana University Press.

Garman, M. (1990) *Psycholinguistics.* Cambridge: Cambridge University Press.

Goodwin, C. (1979) The interactive construction of a sentence in natural conversation. In G. Psathas (ed.), *Everyday language: Studies in ethnomethodology.* New York: Irvington Publishers, pp. 97-121.

Goodwin, C. (1981) *Conversational interaction: Interaction be-tween speakers and hearers.* New York: Academic Press.

Goodwin, C. & M.H. Goodwin (1992) Assessments and the construction of context. In A. Duranti & C. Goodwin (eds.), *Rethinking context: Language as interactive phenomenon.* Cambridge: Cambridge Univerity Press, pp. 147-189.

Herrmann, T., & J. Grabowski (1994) *Sprechen: Psychologie der Sprachproduktion.* Heidelberg: Spektrum Akademischer Verlag.

Hirsch, E.D. (1976) *The aims of interpretation.* Chicago: University of Chicago Press.

Johnson, C.E. (2000) What you see is what you get: The importance of transcription for interpreting children's morphosyntactic development. In L. Menn & N. B. Ratner (eds.), *Methods for studying language production.* Mahwah, NJ: Erlbaum, pp. 181-204.

Keppler, A. (1987) Der Verlauf von Klatschgesprächen. *Zeitschrift für Soziologie* 16: 288-302.

Kitzinger, C. (1998) Inaccuracies in quoting from data transcripts: *Or* inaccuracy in quotations from data transcripts. *Discourse & Society* 9: 136-143.

Kowal, S., & D.C. O'Connell (in press a) Datenerhebung und Transkription. In T. Herrmann & J. Grabowski (eds.), *Sprachproduktion.* Göttingen: Hogrefe.

Kowal, S., & D.C. O'Connell (in press b) Psycholinguistische Aspekte der Transkription: Zur Notation von

Pausen in Gesprächstranskripten. *Linguistische Berichte.*

Kulick, D. (1992) *Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean village.* Cambridge: Cambridge University Press.

Labov, W. & D. Fanshel (1977) *Therapeutic discourse: Psychotherapy as conversation.* New York: Academic Press.

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

Lindsay, J., & D.C. O'Connell (1995) How do transcribers deal with audio recordings of spoken discourse? *Journal of Psycholinguistic Research* 24: 101-115.

The New Encyclopedia Britannica. (1985) International Phonetic Alphabet (IPA). Chicago: Encyclopedia Britannica, 5: 353.

O'Connell, D.C., & S. Kowal (1994) The transcriber as language user. In G. Bartelt (ed.), *The dynamics of language processes: Essays in honor of Hans W. Dechert.* Tübingen: Gunter Narr, pp. 119-142.

O'Connell, D. C., & S. Kowal (1999) Transcription and the issue of standardization. *Journal of Psycholinguistic Research* 28: 103-120.

Pomerantz, A. (1984) Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J.M. Atkinson & J. Heritage (eds.), *Structures of social action: Studies in conversation analysis.* Cambridge: Cambridge University Press, pp. 57-101.

Posner, R. (1986) Zur Systematik der Beschreibung verbaler und nonverbaler Kommunikation: Semiotik als Propädeutik der Medienanalyse. In H.-G. Bosshardt (ed.), *Perspektiven auf Sprache. Interdisziplinäre Beiträge zum Gedenken an Hans Hörmann.* Berlin: de Gruyter, pp. 267-313.

Potter, R., G. Koop, & H. Green (1947) *Visible speech.* New York: Van Nostrand Rheinhold.

Preston, D.R. (1982) 'Ritin' fowklower daun 'rong: Folklorists' failures in phonology. *Journal of American Folklore* 95: 304-326.

Ratner, N.B., & L. Menn (2000) In the beginning was the wug: Forty years of language-elicitation studies. In L. Menn & N.B. Ratner (eds.), *Methods for studying language production.* Mahwah, NJ: Erlbaum, pp. 1-26.

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

Sacks, H., E.A. Schegloff, & G. Jefferson (1978) A simplest systematics for the organization of turn-taking for conversation. In J. Schenkein (ed.), *Studies in the organization of conversational interaction.* New York: Academic Press, pp. 7-57.

Schegloff, E.A. (1976) On some questions and ambiguities in conversation. *Pragmatics Microfiche* 2.2, D8-G1.

Schegloff, E.A. (1979) Identification and recognition in telephone conversation openings. In G. Psathas (ed.), *Everyday language: Studies in ethnomethodology.* New York: Irvington, pp. 23-78.

Schegloff, E.A. (1984) On some questions and ambiguities in conversation. In J.M. Atkinson & J. Heritage (eds.), *Structures of social action: Studies in conversation analysis.* Cambridge: Cambridge University Press, pp. 28-52.

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

Schegloff, E.A., G. Jefferson, & H. Sacks (1977) The prefe-rence for self-correction in the organization of repair in conversation. *Language* 53: 361-382.

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

Sherzer, J. (1983) *Kuna ways of speaking: An ethnographic perspective.* Austin: University of Texas Press.

Svartvik, J. (ed.) (1990) *The London-Lund corpus of spoken English: Description and research.* Lund, Sweden: Lund University Press.

Svartvik, J., & R. Quirk (eds.) (1980) *A corpus of English conversation.* Lund, Sweden: CWK Gleerup.

Urban, G. (1996) Entextualization, replication, and power. In M. Silverstein & G. Urban (eds.), *Natural histories of discourse.* Chicago: University of Chicago Press, pp. 21-44.

Whitney, P. (1998) *The psychology of language.* Boston: Houghton Mifflin.