

ANALYSIS OF LANGUAGE-SPEECH SAMPLES WITH SALT AND PEPPER

AUDREY D. WESTON LAWRENCE D. SHRIBERG JON F. MILLER
*The Waisman Center on Mental Retardation and Human Development
University of Wisconsin-Madison*

Microcomputer-aided analysis of spontaneous language-speech samples offers researchers an efficient means of analyzing large amounts of data. It may be necessary, however, to format samples for more than one software program in order to obtain comprehensive morpho-syntactic and phonetic/phonologic analyses. This paper suggests a procedure for the combined use of SALT (Systematic Analysis of Language Transcripts, Miller & Chapman, 1985) and PEPPER (Programs to Examine Phonetic and Phonologic Evaluation Records, Shriberg, 1986) that is designed to minimize the duplication of effort involved in following two different formatting procedures. Results of a study undertaken to explore methodological issues in the combined use of SALT and PEPPER generally support the validity, reliability, and efficiency of the procedure. Results also raise some issues concerning the use of narrow phonetic transcription as opposed to standard orthographic transcription of continuous language-speech samples.

KEY WORDS: language, speech, phonology, articulation, computers

Considerable recent attention has been focused on methodological alternatives for spontaneous language-speech sampling analysis (e.g., Hubbell, 1988; Klee, 1985; Lahey, 1988; Pye & Ingram, 1988). Although microcomputer-aided software appears to offer efficient means for phonologic, syntactic, and semantic analyses (cf. Schwartz, 1985), an important limitation of currently available options concerns their separate focus. At present, if a microcomputer user wishes both comprehensive morpho-syntactic and phonologic analyses of the same language-speech sample, it may be necessary to use more than one software package and hence, more than one transcription and formatting procedure. Thus, although detailed analysis procedures can be carried out efficiently by microcomputers, the duplication of effort involved in meeting certain analysis needs may limit the clinical and research applications of language and speech analysis software. Although recent research frequently addresses interactions among phonetic, phonologic, morphemic, semantic, and pragmatic levels of language (e.g., Camarata & Leonard, 1986; Campbell & Shriberg, 1982; Donahue, 1986; Healy & Madison, 1987; Kamhi, Catts, & Davis, 1984; Nelson & Kamhi, 1984; Paul & Shriberg, 1982; Schwartz, Leonard, Loeb, & Swanson, 1987), at present no systematic approaches to the combined use of computer-aided software for research purposes have been reported.

Two independently developed software programs that might be used to analyze the same continuous speech sample are Systematic Analysis of Language Transcripts (SALT) (Miller & Chapman, 1985) and Programs to Examine Phonetic and Phonologic Evaluations Records (PEPPER) (Shriberg, 1986). Each program has been used singly to analyze research data (e.g., Holland et al, 1985; MacLachlin & Chapman, 1988; Shriberg & Kwiatkowski, 1985, 1988; Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986), but they have not, to date, been used in combination. The programs are compatible in

encouraging users to employ a variety of sample elicitation techniques consistent with obtaining representative samples of the type desired, e.g., conversation or narration. If it were determined that the transcription and formatting guidelines for these programs are also compatible and that the same language-speech sample can be processed efficiently for both analysis procedures, the benefits in data generated may warrant the time required for an additional formatting and keyboarding procedure. This double-duty benefit may be especially apparent in the time-consuming transcription of poorly intelligible speech.

The purposes of this study are to present a procedure designed to minimize required transcription time in the combined use of the microcomputer versions of SALT and PEPPER and to describe a study undertaken to explore the validity, reliability, and efficiency of the procedure.

DESCRIPTION OF SALT AND PEPPER

SALT, which processes data collected from one or more speakers during a sampling session, is designed to analyze morphemic, pragmatic, and semantic aspects of language. Speech samples submitted for SALT analysis are entered into the computer using standard English orthography. Analysis options include a number of pre-set analyses that address three areas important to a child's language status: structural forms, pragmatic behavior, and semantic development. Users may specify additional analyses to meet their unique needs. SALT procedures are based on guidelines for sample elicitation, transcription, and analysis established by Miller (1981). When transcription is done at the keyboard the time required for transcript entry is approximately seven minutes for every one minute of conversation (Miller & Smith, 1983).

PEPPER, which also processes data collected from one or more speakers during a sampling session, provides 10 pre-set quantitative analyses and facilitates user-determined analyses of a speaker's phonetic inventory and phonologic system. PEPPER programs can be used to analyze samples of any type or length, including continuous speech samples, sentences, phrases, words, syllables, or isolated segments. The program accommodates the narrow phonetic transcription system described in Shriberg and Kent (1982), although analyses can also be done on samples transcribed using the broad phonetic symbols of the International Phonetic Alphabet (Principles of the International Phonetic Association, 1978). The PEPPER User Manual provides guidelines for sample elicitation, transcription and formatting procedures, and reference data for interpretation. Given a previously transcribed sample of 90 first-

occurrence words, keyboard entry is approximately 1 hour, depending on the speaker's severity of involvement.

Table 1 is a description of the major operational features of SALT and PEPPER, including hardware requirements and how each addresses different stages of the sampling process. Although a number of similar transcription issues listed in Table 1 are addressed by both SALT and PEPPER, the two programs are quite different from one another in the level of data submitted for analysis and the scope of their analysis options.

Table 2 is a summary of some of the analysis options offered by the two programs. Formatting the same language-speech sample for keyboard entry for both programs would make it possible to obtain analyses that address these levels as well as levels not available within either program alone.

TABLE 1. Major features of SALT and PEPPER.

<i>Feature</i>	<i>SALT</i>	<i>PEPPER</i>
Hardware requirements	Requires an IBM PC-XT/AT (or true compatible) with 256K RAM, two floppy disk drives, and a DOS operating system, 2.0 or later.	Requires an IBM PC-XT/AT (or true compatible) with 640K RAM, a math coprocessor, either two 5-1/4" floppy disk drives or a combination hard disk and floppy disk drive, a display device, and a DOS operating system.
Collecting the sample	Procedures are based upon guidelines presented by Miller (1981) for sampling spontaneous language in a variety of situations.	User manual includes technical procedures for obtaining quality recordings and for structuring the sampling session. Samples may consist of sounds, syllables, words, phrases, or continuous speech.
Transcribing the sample	Transcription manual includes guidelines for incorporating sample duration and pause times. Uses standard English orthography. Requires approximately 7 min of transcription and formatting for every 1 min of recorded conversation. Nine transcription issues are addressed by both programs: a. number of speakers transcribed per sampling session b. where to begin transcribing (from recorded utterances) c. length of samples d. segmenting utterances e. incomplete utterances or interruptions f. words or utterances set apart from body of utterance g. unintelligible words/utterances h. excluded utterances i. grammatical morpheme errors	Requires broad transcription and encourages narrow phonetic transcription. Includes a system of diacritic symbols. Transcribing and formatting a continuous speech sample of 90 first-occurrence words requires approximately 90 min.
Formatting for computer entry	Both programs offer the option of doing transcription and computer entry in parallel or in series. Conventions for coding, punctuating, and spelling are specific to each program. An error-checking routine that scans for formatting errors is included in each program.	
Analyzing the sample	A number of analysis options are available in each program. Both programs include analysis procedures that quantify user-specified codes. Analyses can be run on individual or grouped files. SALT offers user-designed options in a "SEARCH" subprogram. Both programs facilitate user-designed quantitative analyses as well as qualitative analyses. Some of the options available in each program are summarized in Table 2.	

TABLE 2. Summary of standard analyses available in SALT and PEPPER.

SALT	PEPPER
Summary of utterance types <ul style="list-style-type: none"> breakdown of utterances based on completeness, intelligibility, and type (question vs. statement) 	Structural statistics <ul style="list-style-type: none"> analyses of intended and realized canonical forms, average words per utterance, type-token ratio
Table of mazes and overlaps <ul style="list-style-type: none"> number of utterances containing at least one maze and/or overlap 	Phoneme analysis (consonants, vowels and diphthongs) <ul style="list-style-type: none"> inventory of 24 consonants, 12 vowels, 5 diphthongs percentage each phoneme correct, deleted, substituted or distorted by initial, medial, final position, and overall
Distribution by speaker turns <ul style="list-style-type: none"> number of turns that are 1,2,... 6+ utterances in length 	Feature analysis (consonants, vowels and diphthongs) <ul style="list-style-type: none"> classifies intended consonants by sonorant vs. obstruent, voicing, place and manner errors classified and percentaged as above substitution summary by features vowels classified by place-height
Pause summary <ul style="list-style-type: none"> number and length of pauses within and between utterances 	Item analysis <ul style="list-style-type: none"> error analysis of each word produced in each utterance
Timing and rate summary <ul style="list-style-type: none"> number of utterance attempts per minute and total words per utterance attempt 	Percentage consonants correct <ul style="list-style-type: none"> total percent intended and percent correct by manner class (nasals, glides, stops, fricatives, affricates, liquids) percent correct by word position, number of syllables, singleton vs. cluster severity rating and adjective label intelligibility index
Utterance and line code table <ul style="list-style-type: none"> frequency information on any verbal or nonverbal data of interest (i.e., irregular past or topic change) 	Natural process analysis <ul style="list-style-type: none"> phonetic inventory of sounds targeted, correctly realized, used as substitutions, never used, or never targeted percent occurrence of 8 Natural processes and other Uncoded processes word lists searched by consonants, vowels, diacritics, canonical forms, or coded processes
Utterance and maze distribution tables <ul style="list-style-type: none"> number of utterances by length in word roots or morphemes number of mazes by utterance length 	
Word and morpheme summary <ul style="list-style-type: none"> includes Type-Token Ratio, MLU, Brown's Stages, Expected Age Range 	
Bound morpheme table <ul style="list-style-type: none"> percent occurrence of bound morphemes in obligatory contexts 	
Word tables <ul style="list-style-type: none"> lists of all words, lists by category, or user-designed lists 	
User configured analyses <ul style="list-style-type: none"> provides options for coding morphemes or utterances employing any coding taxonomy 	

Many potential research questions require integration across the domains of language reflected in a combination of SALT and PEPPER analysis options (for a relevant discussion of possible interactions, see Crystal, 1987). Table 3 is a sample list of eight such speech-language questions. Suggested analyses reflect only some potential approaches to the questions posed, with each program suggesting complementary analysis options.

Procedures for Combined Use

The use of phonetic transcription symbols for PEPPER suggests a logical approach to combined use of the two programs. Transcribing for PEPPER *first* will also represent utterances at the level required for morpho-syntactic analysis, so that once a sample has been transcribed and formatted for PEPPER much of the work necessary for SALT has already been done. Accordingly, the source for

keyboard entry of a SALT transcript can be a hard copy PEPPER transcript, rather than an audiotaped sample.

Table 4 is a summary of the transcription guidelines for SALT and PEPPER and the specific guidelines, developed for this project, for converting PEPPER transcripts to SALT transcripts—termed the PEPPER-SALT procedure. Items are ordered first to address some general concerns relevant to the continuous speech samples (Items 1–4), followed by items that address concerns about specific utterance types within samples (Items 5–11). For the first four general concerns in Table 4, PEPPER procedures are adequate to meet SALT guidelines with only minor adaptations. For example, when the transcription of both speakers in a sampling session is desired for SALT analyses, the data can be incorporated readily into the original PEPPER transcription. Moreover, timing and rate information, data not formally addressed in PEPPER conventions, also can be incorporated within PEPPER guidelines.

For five of the seven concerns in Table 4 related to the

TABLE 3. Hypothetical clinical and research questions and SALT/PEPPER analyses that might be used to address them.

<i>Question</i>	<i>SALT</i>	<i>PEPPER</i>
Do children with low intelligibility exhibit formulation problems?	Examine number of utterances containing mazes in Summary of Utterance Types and Mazes by Utterance Length. Code types of nonfluencies.	Use analyses such as Percentage Consonants Correct, which includes an Intelligibility Index.
Do unintelligible words constrain language?	Examine partly intelligible utterances to determine MLU and compare with MLU in complete and intelligible utterances. Use Bound Morpheme Summary. Use SEARCH to identify utterances for qualitative analysis.	Examine number of questionable, or unsure but transcribable, words.
What contexts facilitate intelligibility?	Use SEARCH option to identify utterances containing target words.	Use Item Analysis to identify words with no or few errors. Compare to unintelligible contexts.
What changes in syntactic complexity accompany improved articulation?	Use information in Word and Morpheme Summary such as Brown's Stage Assignment based upon use of 14 grammatical structures. Code use of particular morphemes such as verb tense.	Use Feature Analysis to describe changes in phonologic organization in context of syntactically simple versus syntactically complex structures.
Do children who avoid certain word forms and sounds constrain language in other ways?	Use Table of Utterance Distribution by Length and Summary of Utterance Types. Check Type-Token Ratio and vocabulary lists.	Use Structural Statistics and phonetic inventory from Natural Process Analysis.
Do children who exhibit many sound changes have lowered productivity?	Use Timing and Rate Summary to determine number of utterance attempts per minute and number of words per utterance.	Use Natural Process Analysis to determine use of eight natural processes as well as any Uncoded Processes.
What changes in social language accompany improved sound production?	Examine Distribution of Utterances by Speaker Turns; code and examine topic shifts; examine pause times.	Use Artic Test protocol to describe improvement in articulation.

formatting of specific utterances, PEPPER procedures facilitate SALT transcript entry. For example, note PEPPER's coding conventions for identifying incomplete utterances (Item 5) and for representing unintelligible words (Item 7). In the coding of bound morpheme omissions PEPPER differentiates omissions apparently due to speech from those more likely due to grammatical constraints (Item 9). If a child appears not to have intended production of a final bound morpheme (e.g., as in third person singular "walks" produced as [wɔk]), the omission of the speech sound involved is not coded as a speech error such as final cluster reduction. The gloss of the child's speech, however, indicates all such omissions, so that what is appropriately *not* counted as an error for PEPPER readily and appropriately can be identified as an omission for subsequent SALT analysis. PEPPER transcription attempts to represent every vocalization a child makes, even though some may be coded for exclusion from certain PEPPER analyses. Consequently, all the productive language information necessary for a SALT representation is present in a PEPPER transcript.

In addition to determining that PEPPER transcripts incorporate required SALT information, three interrelated issues of validity, reliability, and efficiency must also be considered to establish the procedural efficacy of the PEPPER-SALT approach. First, the orthographic transcription approach of SALT and the narrow phonetic transcription approach of PEPPER are alternative descriptions to represent different levels of productive language. To the extent that continuous language-speech sampling is viewed as a means of "measuring" or profiling a speaker's productive language ability, the form of the description may affect content. Because phonetic transcription and orthographic transcription may reflect different underlying metrics, the concurrent validity of the PEPPER-SALT approach as a shortcut to conventional SALT transcription needs to be established.

Second, the PEPPER-SALT reformatting approach to obtaining SALT analyses has two major components. A clerical component primarily involves the copying of utterances into SALT format, whereas a transcription-like component involves decisions that include, as noted in

TABLE 4. Summary of transcription guideline for SALT and PEPPER and description of the PEPPER-SALT procedure for combined use.

Item	Transcription Guidelines		
	SALT	PEPPER	PEPPER-SALT Procedure
1. Number of speakers transcribed per sampling session	Utterances of 1 or 2 speakers per sample can be analyzed. Transcription of both child and other speaker is recommended so that important discourse and contextual information is available.	Only the child's speech is transcribed. Utterances from only 1 speaker per sample are entered into the computer for analysis.	Initial transcription for PEPPER can incorporate transcription of examiner utterances or use systematic annotations while transcribing. If only the child's speech is transcribed reference to audiotapes may be used as necessary to interpret errors or intended forms.
2. Where to begin transcribing	Transcribing begins with the first intelligible utterance longer than 1 word and includes consecutive utterances from that point.	Typically, the first utterance recorded is the first utterance transcribed, but an arbitrary starting point can be identified. From this point, although some utterances may be coded for exclusion from analyses (see Excluded utterances), consecutive utterances are transcribed until the required sample length is obtained.	The first utterance transcribed for PEPPER may or may not meet SALT guidelines. Transcription for SALT should begin with the first PEPPER utterance longer than one word and include consecutive utterances from this point.
3. Length of sample	No specific guidelines are included. Either a particular number of utterances or a target duration can be used. Except for TTR, all analyses are based on the total number of utterances or the number of complete and intelligible utterances in the transcript.	Criteria for a sample of adequate length for analysis are 90 first-occurrence words, 70 utterances, or 225 words, whichever occurs first.	The criterion of 90 first-occurrence words may not yield 50-utterance samples. It may be necessary, using SALT procedures for parallel transcription and computer entry, to transcribe more of the taped samples for SALT than for PEPPER.
4. Segmenting utterances	Utterances are segmented according to final intonation contour, pauses in the flow of speech, or grammatical criteria.	Segmentation is based on falling intonation contour. Very long utterances (15 or more words) may have to be artificially divided because of space limitations on the terminal screen.	Procedures are compatible. Any artificially divided utterances for PEPPER (unlikely in child speech) can be entered as one utterance for SALT.
5. Incomplete utterances or interruptions	These are coded for SALT by punctuation at the end of the utterance. They may be excluded from some analyses as the user desires.	Comment codes may be used to quantify interruptions such as clinician overtalk, but procedures do not routinely address interruptions or abandoned utterances.	Relevant information from the speech sample should be included in the PEPPER transcription by using codes developed for this purpose. Suggestions for using various codes are given in PEPPER guidelines for the Comment Summary Analysis.
6. Words or utterances set apart from body of utterance	These are identified as "mazes" and are placed in parentheses. They include filler words, hesitations (uh, um), part and whole word repetitions, revisions, and false starts.	Words to be disregarded are placed in angle brackets. These include words such as <i>yes</i> or <i>mhm</i> used to confirm a clinician's gloss of a previous utterance, nonmeaningful vocalizations, hesitations, false starts, and the incomplete words or revisions. Whole word revisions are not disregarded. Other disregarded words represent interference with intelligibility, e.g., overtalk by the clinician.	Disregarded words for PEPPER are generally compatible with maze elements for SALT. Some utterances, such as gloss confirmations, may be omitted from SALT entry as the user decides. Some words not disregarded by PEPPER (e.g., whole word revisions and repetitions) may be identified as maze elements for SALT and some words disregarded by PEPPER (e.g., produced with interfering background noise) may be included in the body of an utterance for SALT. PEPPER's angle bracketed code convention can be used to flag words appropriately for SALT transcription.

TABLE 4. (Continued)

Item	Transcription Guidelines		
	SALT	PEPPER	PEPPER-SALT Procedure
7. Handling unintelligible words and utterances or words with questionable gloss	Unintelligible words are coded with Xs. Estimation of the number of morphemes used corresponds to the number of Xs used. Both completely and partially unintelligible utterances are entered. The user can decide whether analyses will be based on only complete and intelligible utterances or will include incomplete and/or partially to completely unintelligible utterances.	Asterisks are used to code unintelligible words or utterances, with one syllable represented by each asterisk. Careful representation of unintelligible words is important for accurate estimation of intelligibility. Words for which the gloss provided is questionable are indicated by specified conventions and excluded from some analyses.	The PEPPER asterisk system readily translates to SALT entry, keeping in mind what each program is counting. Codes must be used accurately so as not to over- or under-inflate obtained estimates. Decisions about questionable words need to be made, for example, in the case of bound morphemes and relevant to Brown's Stage assignment, will such words be included in coding target behaviors?
8. Excluded utterances	Utterances are excluded from analyses based on guidelines in Miller (1981). These are primarily repetitions and routines. Such utterances may or may not be included in the transcript.	Repeated words beginning with the third occurrence in the transcript are excluded from some analyses. One may or may not include imitations. Routines such as counting and nursery rhymes are excluded. All such utterances are entered at the keyboard for their potential information value and coded for exclusion from all or some analyses.	Guidelines for exclusions are generally compatible. However, each word coded for exclusion from PEPPER needs to be considered individually to determine if it should or should not be excluded from SALT analyses.
9. Grammatical morpheme omissions	Bound morphemes, present and omitted in obligatory contexts, are coded at keyboard entry.	Transcription requires determining if bound morpheme errors are attributable to language or speech. If a bound morpheme does not appear to be intended, its omission will not be coded as a speech error.	The gloss of the child's utterance in the PEPPER transcript indicates omission/substitution of forms, or correct use of bound morphemes, so that use of grammatical forms can be correctly coded for SALT.
10. Coding pause time	Both within- and between-utterance pauses can be entered for quantitative analyses.	Not formally addressed, but can be done using the Comments feature.	If one desires information from relevant SALT analyses, initial transcription should incorporate pause time.
11. Sample duration	Duration of sample in minutes can be entered along with other relevant data in information lines at the beginning of a transcript.	Same procedure as used in SALT.	Measures such as number of utterance attempts per minute depend on time information. This can readily be added to PEPPER procedures.

Table 4, utterance segmentation, determination of appropriate end-of-utterance punctuation, identification of mazes, and recoding of PEPPER-excluded items. It needs to be established how reliably both the clerical and the transcription components of the PEPPER-SALT procedure can be implemented.

Third, the fact that the PEPPER-SALT procedure involves more than simply reproducing utterances at the keyboard is relevant to the question of the procedure's efficiency. The time required to make SALT transcription decisions such as those mentioned above from printed PEPPER transcripts, and the need to evaluate PEPPER entry conventions in light of SALT entry conventions, may make the PEPPER-SALT procedure no more efficient than conventional SALT transcription, especially when SALT transcribers are very skilled.

A study was undertaken to explore the above issues in the use of the PEPPER-SALT procedure for generating SALT transcripts. Three questions addressing the validity, reliability, and efficiency of the procedure were posed:

1. Does the PEPPER-SALT procedure yield language analysis data equivalent to that of the conventional SALT procedure?
2. How do the percentages of agreement between (a) PEPPER-SALT transcripts prepared by different transcribers and (b) SALT and PEPPER-SALT transcripts prepared by the same transcriber compare to the current interjudge SALT percentage of agreement?
3. Does use of the PEPPER-SALT procedure take less time for experienced SALT transcribers

than the time required for conventional SALT transcription?

METHOD

Speech Samples

The speech samples used in this study were randomly selected from a database of 64 samples previously collected from children with speech delays ranging in severity from mild-moderate to moderate-severe. The samples had been transcribed by consensus and formatted for PEPPER analyses by two transcribers skilled in the use of a narrow phonetic symbol system (Shriberg, Hincke, & Trost-Steffen, 1987; Shriberg & Kent, 1982; Shriberg, Kwiatkowski, & Hoffmann, 1984). The samples were subsequently entered into a Harris 800 minicomputer by keyboarders trained in the use of PEPPER data-entry conventions.

The Intelligibility Index from the Percentage of Consonants Correct (PCC) analysis option in PEPPER was used to divide the speech samples into three subgroups. A high intelligibility group included samples with an Intelligibility Index score greater than one standard deviation above the mean score for the total group. A low intelligibility group included samples with an Intelligibility Index score less than one standard deviation below the mean. Samples falling between these two extremes were considered to have medium intelligibility. Transcripts of two samples were randomly selected from each intelligibility level to ensure that the three levels would be represented in the data; however, this small number of transcripts at each level precluded statistical comparison across levels. Table 5 is a description of low, medium, and high intelligibility Transcript 1 and Transcript 2. The Average Words per Utterance (AWU) from PEPPER's Structural Statistics analysis option are included along with the Percentage of Consonants Correct data for each transcript at each intelligibility level.

Validity and Reliability

Two master's students in communicative disorders who were experienced SALT transcribers participated in the project. Each processed the six samples for SALT analysis using both the established SALT procedure and the PEPPER-SALT conversion procedure. Completion of each set of transcripts was separated by 1 week, with order counterbalanced by transcriber.

The transcribers were provided the PEPPER-SALT guidelines described in Table 4 as well as a more concise statement of these guidelines, included in an appendix, that also included some relevant SALT guidelines. An original audiotaped sample was the source for each SALT procedure and a PEPPER transcript was the source for each PEPPER-SALT procedure. The transcribers did not refer to the audiotaped samples while using the PEPPER-SALT approach.

The dependent measures for validity included the number of words, utterances, bound morphemes, mazes, words in mazes, and unintelligible words occurring in the samples. "Mazes" (Loban, 1976) included hesitations, false starts, repetitions, and revisions, and "words" included hesitations and fragments, as well as whole words. As noted in Table 4, these behaviors are identified and coded for SALT analysis purposes.

The reliability of the PEPPER-SALT procedure, considering both its clerical and transcription components, was assessed by comparing the obtained percentages of agreement for both interjudge PEPPER-SALT and mean intrajudge SALT with PEPPER-SALT to obtained percentages of agreement for interjudge SALT. The term *intrajudge* for SALT with PEPPER-SALT agreement appropriately indicates that the same transcriber completed both the SALT transcription and the reformatting of PEPPER transcripts. However, obtained percentage of agreement figures also reflect interjudge agreement between the SALT transcribers and the original PEPPER transcribers for each of the samples. A point by point comparison of transcripts, similar to the approach described by Miller and Smith (1983), was completed to determine all interjudge and intrajudge percentages of

TABLE 5. Description of the 6 continuous language-speech samples selected from the research database.

<i>Intelligibility group</i>	<i>Intelligibility index</i>	<i>Percentage of consonants correct (PCC)</i>	<i>Average words per utterance</i>
Low			
Transcript 1	66.67	62.82	2.20
Transcript 2	75.90	61.38	3.48
Medium			
Transcript 1	86.50	65.09	3.58
Transcript 2	92.35	68.09	4.06
High			
Transcript 1	98.65	78.14	7.67
Transcript 2	97.73	72.75	5.48
Total	<i>M</i> 86.30	68.04	4.41
	<i>SD</i> 12.75	6.40	1.92

agreement for each sample. The formula to calculate agreement was

Percentage of Agreement =

$$\frac{\text{Total Number of Agreements}}{\text{Total Number of Agreements} + \text{Disagreements}} \times 100$$

Because transcription of language samples does not involve the use of a closed set of behavioral descriptors, obtained percentage of agreement figures could not be chance-corrected using Cohen's Kappa (e.g., Suen & Lee, 1985) or a similar statistic.

Efficiency

The efficiency of the PEPPER-SALT procedure was evaluated by comparing the PEPPER-SALT entry times with the SALT entry times for each sample. The time required to complete PEPPER transcription and data entry was assumed to represent a constant for each sample and was not considered in this evaluation. Each transcriber processed the samples without interruption and recorded elapsed time for each sample completed.

RESULTS

Validity and Reliability

Table 6 is summary of the obtained validity data. Means and standard deviations for the number of words, utterances, bound morphemes, mazes, words in mazes, and unintelligible words for the SALT and PEPPER-SALT transcripts are reported. Using Wilcoxon signed-ranks matched-pairs tests, significant differences were found only for number of words ($T = 4$; $n = 12$; $p < .01$) and number of words in mazes ($T = 8$; $n = 11$; $p < .05$). As shown in the means columns, PEPPER-SALT transcripts in this sample yielded more words in both categories than tallied in the comparison SALT transcripts.

Reliability data are displayed in Figure 1. The dotted line at 81% in both panels represents the overall mean interjudge percentage of agreement for SALT transcripts ($n = 6$) and is an estimate of the current average interjudge SALT agreement for transcription of speech-de-

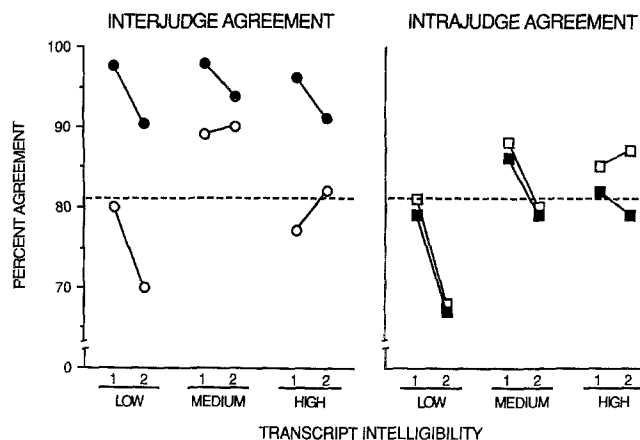


FIGURE 1. Comparative reliability data for the PEPPER-SALT transcription procedure. The left panel includes SALT (open circles) and PEPPER-SALT (filled circles) interjudge agreement data for the two transcripts (1,2) at each level of intelligibility. The right panel includes intrajudge SALT with PEPPER-SALT agreement data for each of the two transcribers (open and filled squares). The dotted line in both panels is the mean overall interjudge percentage of agreement for SALT transcription.

layed children. The left panel is a display of the obtained interjudge SALT percentages of agreement (open circles) and the comparative interjudge PEPPER-SALT percentages of agreement (filled circles). For each of the two transcripts (1,2) at each intelligibility level, obtained percentages for PEPPER-SALT were consistently higher, with Wilcoxon signed-ranks matched-pairs tests resulting in significant T values ($T = 0$; $n = 6$; $p < .05$) for differences in obtained percentages.

The right panel is a display of intrajudge SALT with PEPPER-SALT percentages of agreement for each of the two transcribers (open and filled squares) for the six transcripts. Wilcoxon signed-ranks matched-pairs tests resulted in nonsignificant T values ($T = 5$; $n = 6$ and $T = 10$; $n = 6$) for obtained differences in percentages of agreement between interjudge SALT and intrajudge SALT with PEPPER-SALT for both transcribers. The greatest difference in obtained percentages of agreement occurred on medium intelligibility Transcript 2, with 90% agreement for interjudge SALT (left panel) compared with 80% and 79% agreement for the two transcribers for intrajudge SALT with PEPPER-SALT (right panel). For all other transcripts, interjudge and intrajudge agreement figures are within 8 percentage points of each other.

TABLE 6. Validity data for the PEPPER-SALT procedure.

Measure	SALT		PEPPER-SALT		Wilcoxon Test T	ρ
	M	SD	M	SD		
Words	183.2	26.3	197.3	22.6	4.0	.01
Utterances	59.8	14.8	57.4	15.6	13.5	ns
Bound morphemes	11.9	6.7	11.9	5.3	19.5	ns
Mazes	8.4	6.0	9.0	6.1	16.5	ns
Words in mazes	17.8	12.9	23.8	16.7	8.0	.05
Unintelligible words	12.4	12.0	12.7	12.8	21.0	ns

Efficiency

Figure 2 is a display of mean transcription time in minutes for the SALT and PEPPER-SALT procedures for each transcript at each intelligibility level. As shown, mean transcript preparation time using the PEPPER-SALT procedure was consistently lower than the mean preparation time using the SALT procedure. The obtained Wilcoxon T value ($T = 0$; $n = 11$) for differences in transcription time was significant at the .01 alpha level. The increase in efficiency for PEPPER-SALT transcription is most obvious for samples of low and medium intelligibility, with low intelligibility Transcript 1 differing by 14.5 min in mean transcription time between the two methods.

DISCUSSION

Results of the comparison of obtained language analysis data provide support for the validity of the PEPPER-SALT procedure. The data for the number of utterances, bound morphemes, mazes, and unintelligible words were not statistically different for PEPPER-SALT in comparison to SALT, indicating that one can obtain SALT-equivalent output from converted PEPPER samples.

The obtained levels of transcript agreement for interjudge PEPPER-SALT and intrajudge SALT with PEPPER-SALT provide support for the reliability of the procedure. The higher levels of interjudge agreement obtained for PEPPER-SALT transcription reflect the fact that initial transcription decisions had already been made, and potential sources of disagreement primarily involved reformatting or clerical issues. For example, in the comparison of transcripts, differences in the representation of unintelligible words accounted for the greatest number of disagreements tallied. As noted in Table 4, the PEPPER asterisk system readily translates to the SALT "X" system; however, if the segmentation of unintelligible words in sequences of unintelligible words is not

maintained, the result is a discrepancy in total words entered. That discrepancy in addition to end-of-utterance punctuation discrepancies may have accounted for the lowest obtained interjudge PEPPER-SALT percentage of agreement, approximately 91% for both low intelligibility and high intelligibility Transcript 2.

The obtained percentages of transcript agreement for intrajudge SALT with PEPPER-SALT were not statistically different from interjudge SALT agreement figures. In both cases, potential disagreements between comparison transcripts involve actual transcription decisions more than reformatting procedures. Consequently, obtained percentage figures are consistently lower than those obtained for interjudge PEPPER-SALT. The discrepancy in obtained agreement figures for medium intelligibility Transcript 2 may represent a particular problem area in the transcription of language samples. For this transcript the interjudge SALT percentage of agreement was 90, whereas intrajudge PEPPER-SALT percentages of agreement were 80 and 79 respectively for the two transcribers. Disagreements on mazes within this transcript were noted to be especially frequent for both transcribers using the PEPPER-SALT procedure in comparison to the SALT procedure. However, for all comparisons, regardless of transcript entry procedures, the percentages of agreement on mazes were low when compared with overall percentages of agreement. Loban (1976) has observed that mazes may represent behaviors inherently difficult to categorize. In contrast, the comparatively low percentages of agreement for both procedures for low intelligibility Transcript 2 may reflect problems specific to low intelligibility samples.

Finally, the transcription time data indicate that the PEPPER-SALT procedure can be used efficiently to obtain SALT analysis data. PEPPER-SALT transcription times were consistently lower than SALT times for all samples; moreover, the differences in time shown by the PEPPER-SALT procedure were especially apparent for the low intelligibility samples. It should be noted, however, that the mean PEPPER-SALT transcription time of approximately 15 min for all transcripts is an estimate of the time required to complete the procedure for 90 first-occurrence PEPPER words. For all but one of the six transcripts in this sample, 90 first-occurrence words yielded at least 50 SALT utterances. The exception, high intelligibility Transcript 1, as described in Table 5, had an obtained AWU score of 7.7. This transcript, which yielded 36 utterances, may not be typical of the continuous speech-language samples collected from most speech delayed children. It may be that for samples containing longer utterances, more of the audiotaped sample needs to be transcribed to obtain the required number of utterances for some SALT analysis options such as MLU and Brown's Stages.

The findings of statistically significant differences for words and words in mazes provide data more generally relevant to transcription issues in speech and language sampling. PEPPER-SALT transcripts had both more words and more words in mazes than comparison SALT transcripts; therefore, the source of the disagreements in

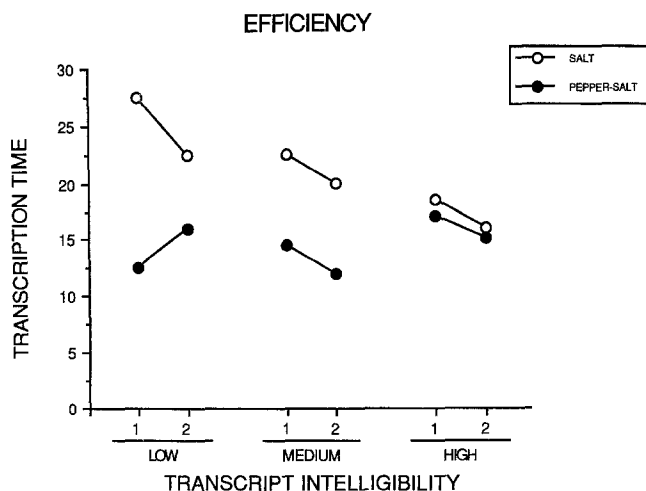


FIGURE 2. Mean transcription times in minutes for SALT and PEPPER-SALT transcription for each transcript at each intelligibility level.

word counts was not simply a consequence of how words were divided by the transcribers between the two categories. Rather, the differences in words transcribed may reflect subtle differences in the perceptual sets associated with transcription by standard orthography and transcription involving the use of narrow phonetic symbols. In related discussions, Buckingham and Yule (1987) present many relevant examples of the influence of listener expectations on segmental decisions, and Ochs (1979) addresses several biases that may be introduced into a data set as a consequence of the transcription approach used. The use of standard orthography, for example, may force one to make a literal interpretation of utterances that may represent only phonological manipulation, as in the case of soundplay. Other researchers have suggested that the use of standard orthography in the transcription of very early language increases the likelihood that important data may be overlooked. David Ingram (personal communication) has speculated that the use of a narrow phonetic symbol system versus standard orthography in transcribing the speech of very young children may result in as many as 20% more words being identified.

The greater number of words and words in mazes in PEPPER-SALT transcripts provides some support for an assumption that phonetic transcription may be a more powerful tool than orthographic transcription when intelligibility is a problem. That is, the use of narrow phonetic symbols may aid in the identification of words in conversational samples of older speech-delayed children as well as in samples of early speech. One might also assume that the more unintelligible the speech samples, the more obvious would be the benefit of narrow phonetic transcription in word recognition. Figures for the low intelligibility transcripts support this assumption. Overall, PEPPER-SALT transcripts yielded 9.3% higher means for number of words; however, the between-method difference in percentages for low intelligibility transcripts was 14.7% compared to 6.1% and 3.9% for medium and high intelligibility transcripts, respectively. The figures for the mean number of words in mazes do not show the same pattern of difference in percentages between intelligibility levels. The overall difference was 33.7%, with differences of 36.7%, 37.0%, and 31.2% for low, medium, and high intelligibility groups respectively. These differences too, may be a consequence of the more fine-grained transcription system used by PEPPER, but not be specific to intelligibility problems. Words in mazes, as previously noted, included hesitations (e.g., uh, um) and word fragments (e.g., he-he-heman) as well as whole words. Disagreements noted in the comparison of individual transcripts suggest that differences in the number of words in mazes may have been due to the transcription of more hesitations and part-word repetitions across all intelligibility levels by PEPPER transcribers.

One other reported result relating to the issue of transcription approaches warrants further discussion. Because of the narrow phonetic transcription of the original PEPPER samples, it had been anticipated that PEPPER-SALT transcripts would contain fewer unintelligible words than comparison SALT transcripts. Further, a de-

cision was made to transcribe "questionable" words in PEPPER as though they were intelligible words for SALT. As noted earlier, however, differences in the number of unintelligible words in PEPPER-SALT transcripts compared to SALT transcripts were not statistically significant. In fact, in the comparison of individual transcripts, it was noted that some unintelligible portions of PEPPER-SALT transcripts often had more unintelligible words, owing to segmentation differences, than corresponding segments of SALT transcripts. In PEPPER-SALT, for example, a string of single-syllable unintelligible words (i.e., XXX) may have been represented as one multisyllable unintelligible word in SALT (i.e., XXX). The representation of unintelligible portions of a sample by fewer multisyllable words in SALT transcripts, may have resulted in a cancelling effect when the number of unintelligible words was compared, even though less of the sample had actually been transcribed.

Clearly, adequate assessment of the power of narrow phonetic transcription versus standard orthography in transcribing samples of delayed speech and language will require controlled study. There presently are no empirical data addressing subtle differences due to transcription approaches in the processing of low intelligibility language-speech samples. In the present study, it is not known how results might have been influenced by differences in the skill levels of the original PEPPER transcribers and the SALT transcribers who participated in the project. However, the finding that two independent judges made similar decisions 81% of the time is both encouraging and a reminder that the transcription task involves many difficult decisions. Although statistical tests were not applied, percentage of agreement figures suggest that some decisions are more difficult to make for low intelligibility samples compared to high intelligibility samples. For example, differences in the segmentation of words, as well as utterances, especially when identification of lexical items involved some transcriber guesswork, were frequently observed for these samples.

Both the general difficulties in the transcription process and the implications of using orthographic versus narrow phonetic transcription suggest that a decision to use software packages in combination must involve more than a superficial consideration of transcription and formatting compatibility. An unintelligible word, for example, may represent either an unintelligible lexical item or an untranscribable noise. It is important to specify how such behaviors will be reliably identified and handled according to the guidelines of each program.

ACKNOWLEDGMENTS

We wish to express our appreciation to all who shared in the work of this project. Joan Kwiatkowski provided valuable assistance in the development of the PEPPER-SALT procedure, Claudia Dorr and Emily Coon contributed their expert SALT transcription and keyboarding skills, and Maria Cavicchio spent many hours in the comparison of transcripts. We also thank Patty Engebosc and Annette Ortiz for their very able assistance with tables and figures. Portions of this work were supported by the

United States Department of Education, Research in Education of the Handicapped Program, grant number G008400633 and the National Institutes of Neurological and Communicative Disorders and Stroke, grant number NS26246-01.

REFERENCES

- BUCKINGHAM, H., & YULE, G. (1987). Phonemic false evaluation: Theoretical and clinical aspects. *Clinical Linguistics & Phonetics*, 1, 113-125.
- CAMARATA, S., & LEONARD, L. (1986). Young children pronounce object words more accurately than action words. *Journal of Child Language*, 13, 51-66.
- CAMPBELL, T., & SHRIBERG, L. (1982). Association among pragmatic function, linguistic stress and phonological processes in speech-delayed children. *Journal of Speech and Hearing Research*, 25, 547-553.
- CRYSTAL, D. (1987). Towards a 'bucket' theory of language disability: Taking account of interaction between linguistic levels. *Clinical Linguistics & Phonetics*, 1, 7-22.
- DONAHUE, M. (1986). Phonological constraints on emergence of two-word utterances. *Journal of Child Language*, 13, 209-218.
- HEALY, T., & MADISON, C. (1987). Articulation error migration: A comparison of single word and connected speech samples. *Journal of Communication Disorders*, 20, 129-136.
- HOLLAND, A., MILLER, J., REINMUTH, O., BARTLETT, C., FROMM, D., PASHEK, G., STEIN, D., & SWINDELL, C. (1985). Rapid recovery from aphasia: A detailed language analysis. *Brain and Language*, 24, 156-173.
- HUBBELL, R. (1988). *A handbook of English grammar and language sampling*. Englewood Cliffs, NJ: Prentice-Hall.
- KAMHI, A., CATTS, H., & DAVIS, M. (1984). Management of sentence production demands. *Journal of Speech and Hearing Research*, 27, 329-338.
- KLEE, T. (1985). Clinical language sampling: Analyzing the analyses. *Child Language Teaching and Therapy*, 1, 182-198.
- LAHEY, M. (1988). *Language disorders and language development*. New York: Macmillan.
- LOBAN, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.
- MACLACHLAN, B., & CHAPMAN, R. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders*, 53, 2-7.
- MILLER, J. (1981). *Assessing language production in children*. Baltimore: University Park Press.
- MILLER J., & CHAPMAN, R. (1985). *Systematic analysis of language transcripts: User's manual*. Madison: University of Wisconsin.
- MILLER, J., & SMITH, A. (1983). *SALT transcription manual: Guidelines for transcribing free speech samples*. Madison: University of Wisconsin.
- NELSON, L., & KAMHI, A. (1984, November). *Syntactic, semantic, and phonological trade-offs in preschool children's utterances*. Paper presented at the American Speech-Language-Hearing Association annual convention, San Francisco, CA.
- OCHS, E. (1979). Transcription as theory. In E. Ochs & B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43-72). New York: Academic Press.
- PRINCIPLES OF THE INTERNATIONAL PHONETIC ASSOCIATION. (1978, Rev. ed.). London: University College Department of Phonetics.
- PAUL, R., & SHRIBERG, L. (1982). Associations between phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research*, 25, 536-547.
- PYE, C., & INGRAM, D. (1988). Automating the analysis of child phonology. *Clinical Linguistics & Phonetics*, 2, 115-137.
- SCHWARTZ, A. (1985). Microcomputer-assisted assessment of linguistic and phonological processes. *Topics in Language Disorders*, 6, 26-40.
- SCHWARTZ, R., LEONARD, L., LOEB, D., & SWANSON, L. (1987). Attempted sounds are sometimes not: An expanded view of phonological selection and avoidance. *Journal of Child Language*, 14, 411-418.
- SHRIBERG, L. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Madison: Software Development and Distribution Center, University of Wisconsin.
- SHRIBERG, L., HINCKE, R., & TROST-STEFFEN, C. (1987). A procedure to select and train persons for narrow phonetic transcription by consensus. *Clinical Linguistics & Phonetics*, 1, 171-189.
- SHRIBERG, L., & KENT, R. (1982). *Clinical phonetics*. New York: Macmillan.
- SCRIBERG, L., & KWIATKOWSKI, J. (1985). Continuous speech sampling for phonologic analyses of speech-delayed children. *Journal of Speech and Hearing Disorders*, 50, 323-334.
- SHRIBERG, L., & KWIATKOWSKI, J. (1988). A follow-up study of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders*, 53, 144-155.
- SHRIBERG, L., KWIATKOWSKI, J., BEST, S., HENGST, J., & TERSELIC-WEBER, B. (1986). Characteristics of children with speech delays of unknown origin. *Journal of Speech and Hearing Disorders*, 51, 140-161.
- SHRIBERG, L., KWIATKOWSKI, J., & HOFFMAN, K. (1984). A procedure for phonetic transcription by consensus. *Journal of Speech and Hearing Research*, 27, 456-465.
- SUEN, H., & LEE, P. (1985). Effects of the use of percentage of agreement on behavioral observation reliabilities: A reassessment. *Journal of Psychopathology and Behavioral Assessment*, 7, 221-234.

Received March 3, 1988

Accepted March 8, 1989

Requests for reprints should be addressed to Audrey Weston, Phonology Project, Waisman Center on Mental Retardation and Human Development, 1500 Highland Avenue, Madison, WI 53705.

APPENDIX

DIRECTIONS FOR GENERATING SALT TRANSCRIPTS FROM PEPPER TRANSCRIPTS

Sources

Orthographic X-line entries from PEPPER transcripts,
audiotaped samples,
guidelines for combined use of SALT and PEPPER,
Assessing language production in children (Miller, 1981) pp. 24-25,
SALT user's manual (Miller & Chapman, 1985).

Procedures

1. Identify the first child utterance longer than one word from the PEPPER transcript or the audiotaped sample and enter consecutive utterances from this point. Enter all transcribed PEPPER utterances and all utterances from the audiotape up to and including the identified (from PEPPER transcript) final utterance.
2. When entering each word into SALT follow the guidelines for counting morphemes (p. 24, Miller, and SALT guidelines).
3. PEPPER transcription did not include consistent coding of abandoned and otherwise incomplete utterances, but conversion to SALT format should indicate these utterances whenever possible. Follow SALT conventions for coding utterances in these categories with appropriate punctuation.
4. Place in parentheses:
 - fillers, part and whole word
 - repetitions, part and whole word or phrases
 - reformulations
 - false starts

Watch for words that are placed in < >. These are words disregarded for PEPPER but not necessarily disregarded for SALT. These words may or may not be maze items for SALT. Do not include those "words" coded [N] for noise or [NS] for noise/sound effects. Omitted bound morphemes are indicated in the PEPPER X-line entry with parentheses. Also check Z-line entry to see if evidence of morpheme is present in child's production.

5. Identify all bound morpheme usage from X-line entry and code as appropriate for SALT.
6. Work without interruption at a normal keyboarding rate and record transcription minutes for each sample.
7. When transcribing from audiotape, attempt to transcribe all part-word repetitions of the child, for example, (M M M) Mommy bake cookies, as well as all hesitations.

Analysis of Language-Speech Samples with Salt and Pepper

Audrey D. Weston, Lawrence D. Shriberg, and Jon F. Miller
J Speech Hear Res 1989;32;755-766

This information is current as of June 19, 2012

This article, along with updated information and services, is
located on the World Wide Web at:

<http://jslhr.asha.org/cgi/content/abstract/32/4/755>



AMERICAN
SPEECH-LANGUAGE-
HEARING
ASSOCIATION