PEPPER GUIDE 2:

PERCEPTUAL DATA REDUCTION AND FORMATTING PROCEDURES

Transcription and Prosody-Voice Coding

Overview

The conversational sample (CS) and all other speech tasks administered as part of the Madison Speech Assessment Protocol (MSAP) are recorded and later transcribed using narrow phonetic transcription conventions and symbols described in *Clinical Phonetics* (for information on the symbol differences between the PEPPER Proportional font used in PEPPER and the CP5e font, see "About PEPPER" on the Phonology Project website). Conversational speech samples are optionally coded in seven prosody and voice domains (Phrasing, Rate, Stress, Loudness, Pitch, Laryngeal Quality, and Resonance) using procedures described in the *Prosody-Voice Screening Profile (PVSP)* manual. Conversational samples and some MSAP tasks may be acoustically analyzed as well. Acoustic analyses procedures are not described in this PEPPER Guide; they are described in PEPPER Guide 5.

This guide provides general transcription and formatting guidelines and some supplemental information on PEPPER formatting conventions and data reduction procedures unique to individual MSAP tasks and/or special participant populations (participants with motor speech disorders, syndromes, dialectal differences, etc.). See the MSAP tab at the link provided above for information specific to the MSAP tasks mentioned here. All transcripts to be entered into PEPPER, including those for MSAP tasks, must conform to the formatting guidelines to be described.

Transcribers with the Phonology Project typically transcribe the speech tasks in the order in which they are administered within the MSAP. Each task is processed using the following steps:

1. *Glossing.* The conversational sample is glossed; strings of connected speech are parsed into utterances using the segmentation rules on pp. 10-11 of the *PVSP* manual. (MSAP speech tasks with stimuli that are administered consistently in a fixed order do not need to be glossed; see p. 71.)
2. *Prosody-Voice coding (conversational sample only)*. After glossing is completed, 24 (or 12) non-excluded utterances from the conversational sample are then coded for prosody and voice using the procedures described in the *PVSP* manual.
3. *Narrow phonetic transcription*. The conversational speech sample and most of the other MSAP tasks are transcribed using narrow phonetic transcription symbols and PEPPER formatting conventions described in *Clinical Phonetics* and in this document, respectively.

After the transcriber has completed all transcription, coding, and scoring (if applicable) of the speech task(s), the transcript(s), PVSP, and other database variables are entered in PEPPER. Because the hard copies of completed transcripts (particularly conversational samples) and associated study materials may contain identifying and/or sensitive information, they are stored in locked filing cabinets.

Glossing

Conversational samples, as mentioned previously, are glossed using the segmentation rules on pp. 10-11 of the PVSP manual. If a PVSP will be completed, the transcriber keeps track of PVSP-excluded utterances, PVSP-codable utterances, and the number of first occurrence words (i.e., word types) while glossing. Figure 1 shows a partially completed gloss of 18 utterances. The transcriber has kept track of the PVSP-codable utterances and exclusion codes with a small note in the upper left hand portion of the beginning of each new utterance. While glossing, she has recorded and tallied the non-questionable words on the First Occurrence Words pepform (see Figure 2). [NOTE: The pepforms are described in greater detail in the following section "How to Transcribe and Format a Speech Sample" beginning on p. 5.]

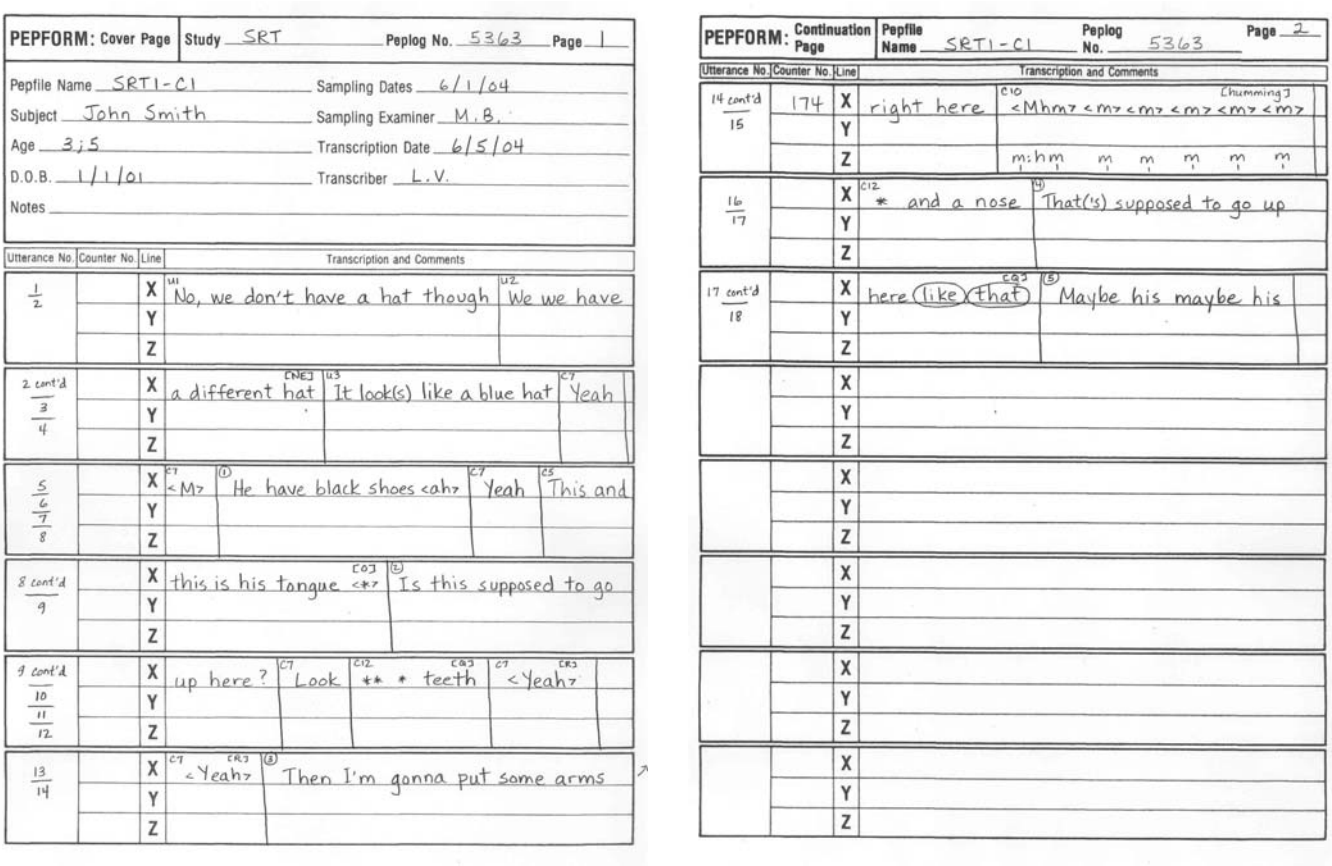
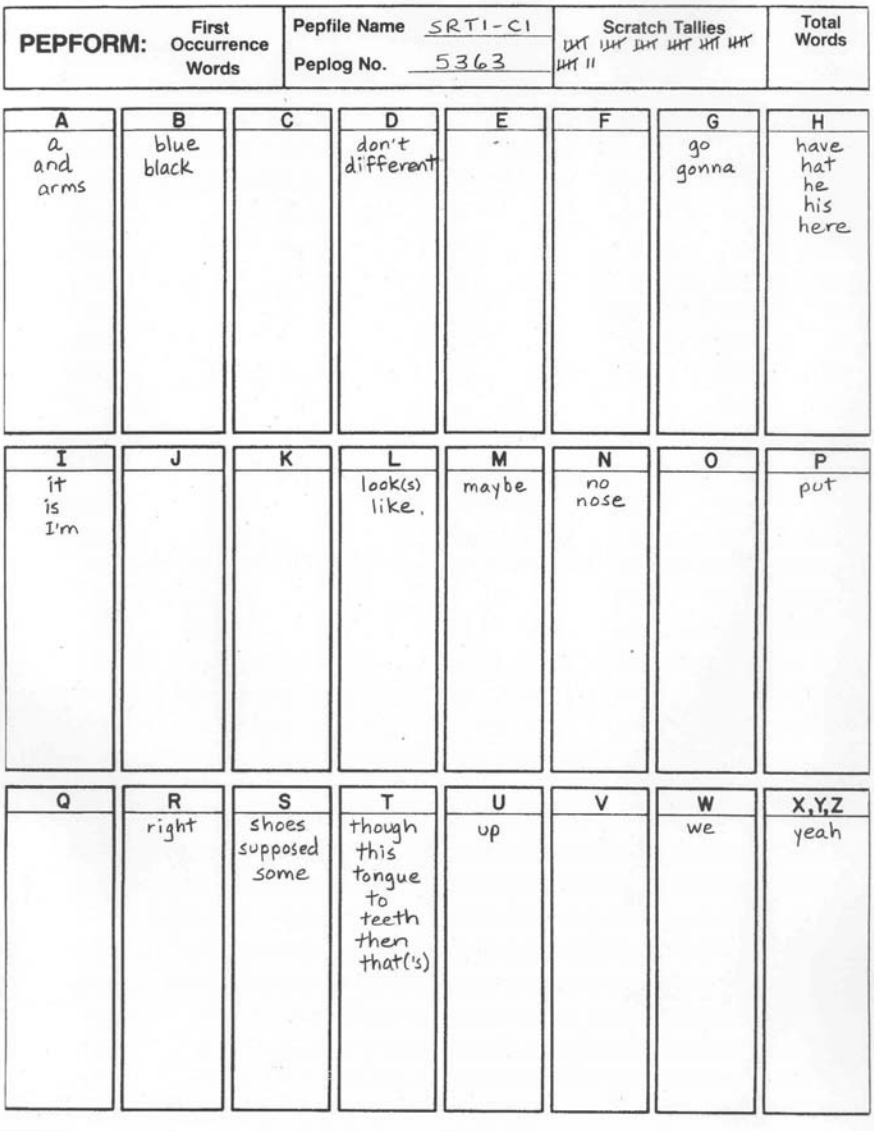


Figure 1. Cover Page and Continuation Page pepforms showing a partial gloss written in the X line.



PEPFORM: First Occurrence Words	Pepfile Name SRT1-C1 Peplog No. 5363	Scratch Tallies	Total Words

A	B	C	D	E	F	G	H
a and arms	blue black		don't different			go gonna	have hat he his here

I	J	K	L	M	N	O	P
it is I'm			look(s) like.	maybe	no nose		put

Q	R	S	T	U	V	W	X,Y,Z
	right	shoes supposed some	though this tongue to teeth then that('s)	up		we	yeah

Figure 2. A partially completed First Occurrence Words pepform that corresponds to the glossed utterances in Figure 1.

Utterances in a continuous speech sample are glossed until a predetermined transcript length is reached. For transcription purposes, we have found that glossing and transcribing to 100 first occurrence words provide samples of sufficient length for most speech analyses. If a PVSP, but no acoustic analyses, will be completed, glossing must continue until the desired number (usually 12 or 24) of PVSP-codable utterances is

reached, and phonetic transcription is completed up to and including the utterance in which the 100th first occurrence word is reached. If a PVSP *and* acoustic analyses will be completed, glossing must continue until the desired number of PVSP-codable utterances is reached, and phonetic transcription continues until 100 first occurrence words *or* 12 PVSP-codable utterances is reached, *whichever comes second*. See item 6 beginning on p. 17 for further information on determining transcript length.

Figure 3 is an example of the last page of a transcript in which 100 first occurrence words were reached before 24 PVSP-codable utterances were reached. One hundred first occurrence words were tallied in utterance 38, but at this point only 21 PVSP-codable utterances had been glossed. The transcriber glossed four more utterances, using the X, Y, and Z lines to save space, in order to reach 24 codable utterances.

PEPFORM: Continuation Page | Pepfile Name UT05-C1 | Peplog No. 6857 | Page 5

Utterance No.	Counter No.	Line	Transcription and Comments
37 cont'd	4:47	X	usually get home and play like Rescue Action
		Y	3 r ʃ
		Z	judəlɪ gɛt hoʊm ɛn pleɪ laɪk ɹɛskjuəksən
38		X	(21) It's this little book that I got for my book
		Y	s ð s ð r
		Z	ɪtsˌ dɪsˌ lɪtəl bʊk ˌət aɪ gɑt fɔɹ maɪ bʊk
38 cont'd		X	order END
		Y	r ð PHONETIC
		Z	ɔɹdɚ
39 / 40 / 41 / 42		X	(22) I can't remember what it is (23) I know the
		Y	rescue book one (CS) I know <the> (COJ) (24) I know
		Z	the end it it has a police one and let's see
42 cont'd		X	a a firetruck one
		Y	
		Z	
		X	
		Y	
		Z	
		X	
		Y	
		Z	
		X	
		Y	
		Z	

Figure 3. An example of the final page of a transcript in which the 100 first occurrence words criterion was satisfied before 24 PVSP-codable utterances were reached.

Prosody-Voice Coding
A conversational speech sample can be used to complete a *PVSP*, a screening instrument that requires perceptual judgments in 7 suprasegmental domains: phrasing, rate, stress, loudness, pitch, laryngeal quality, and resonance. The *PVSP* manual and associated training materials and scoring forms can be found here. If you wish to complete a *PVSP*, we recommend completing prosody-voice coding before proceeding to narrow phonetic transcription, because we have found that suprasegmental decisions can sometimes unintentionally be influenced by segmental judgments.

Narrow Phonetic Transcription
The following section is a revised, updated version of Chapter 3 from the *PEPPER User's Manual (1986)*, a research document used in the Phonology Project and the precursor of the present PEPPER guides. The end of this guide includes additional information about transcribing MSAP tasks and transcribing dialectal variations in speech.

HOW TO TRANSCRIBE AND FORMAT A SPEECH SAMPLE FOR PEPPER

To obtain data for the outputs described in PEPPER Guide 7, PEPPER requires the speech information to be 'formatted' in very particular ways. Accordingly, we will first introduce the formats, terms, and symbols used on the PEPFORM transcription sheets (Section I). These terms and symbols are then used in separate subsections that describe procedures to transcribe and format speech samples (Section II).

I. TERMS, SYMBOLS, AND GENERAL PROCEDURES

A. The Format of a Transcription File

1. PEPFORMS

PEPFORM is the generic name for a series of forms used for efficient handling of various tasks in PEPPER. These pepforms serve several purposes. There are two pepform formats that are used to handwrite a transcription to be entered into the computer: the *Cover Page* or first page (Figure 4; Panel A) and the *Continuation Page* (Figure 4; Panel B) for all subsequent pages in a transcript. The *First Occurrence Words* pepform, as introduced earlier (Figure 2), is used to keep track of the number of word types in a transcript. Let's look at some of the different sections on the *Cover Page* and *Continuation Page* as illustrated in Figure 4; Panels A and B.

[Electronic versions of these 3 pepforms can be accessed and printed from the PEPPER tab on the Pepforms page.

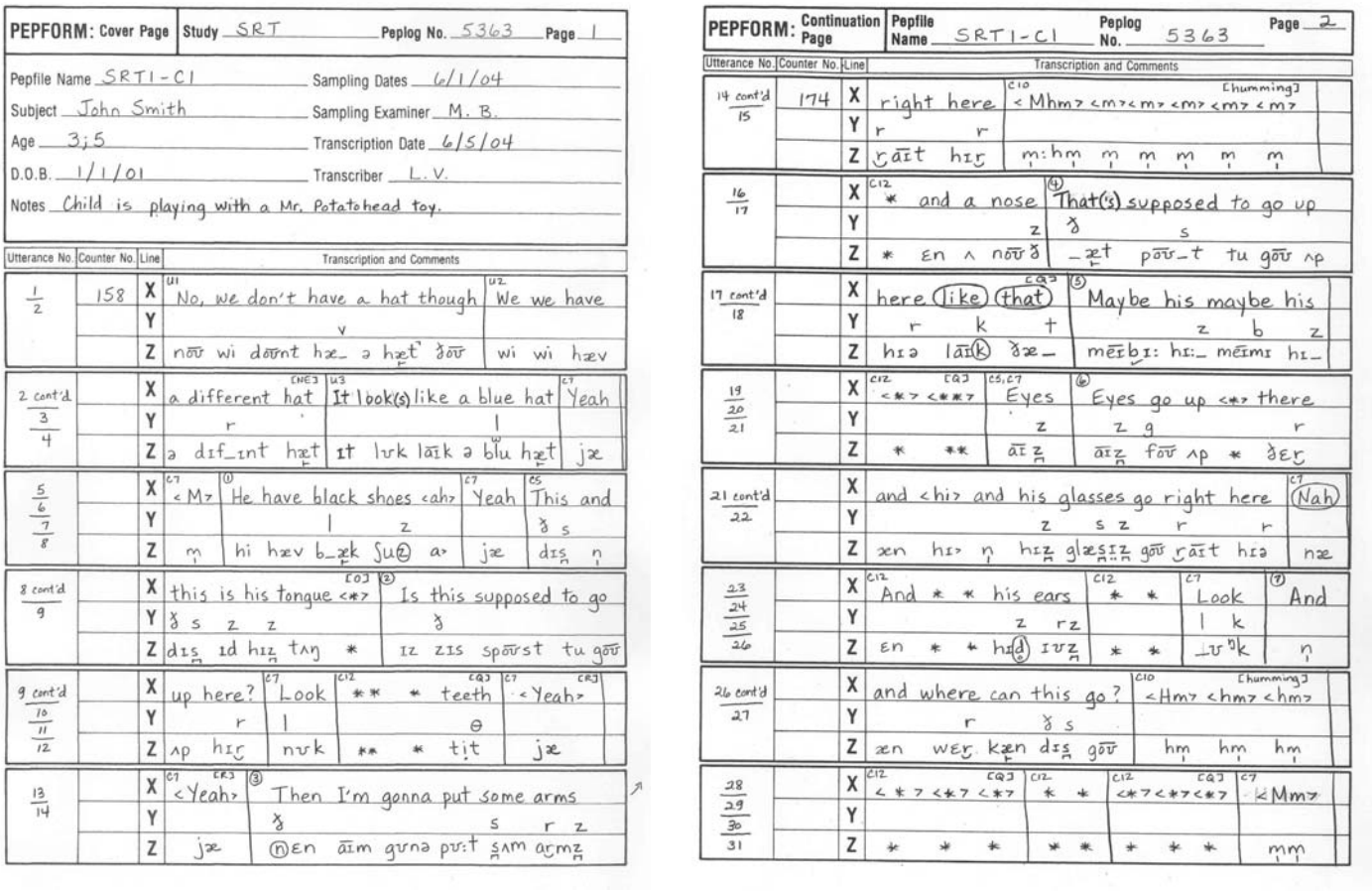


Figure 4. Panel A: PEPFORM: *Cover Page* illustrating the format of transcripts. Panel B: PEPFORM *Continuation Page* illustrating certain transcription and formatting conventions. See text for discussion.

2. Identifying Information

Enter all appropriate identifying information in the top block of the *Cover Page* and also on the top of the *Continuation Pages.* The *Notes* section on the *Cover Page* provides a place for the examiner or transcriber to enter important comments. These include observations on the reliability of the data relative to physical and affective state variables. For example, does the participant have a cold? Was he/she inattentive during all or a portion of the sample? Try to annotate all observations that might affect later interpretation of the speech analyses results.

3. Utterances and Utterance Numbering

An utterance as defined for PEPPER analyses could be an isolated vowel sound, a syllable, a word, a phrase, a string of words, or a sentence. The large blocks on the pepforms (six on the *Cover Page* and eight on *Continuation Pages)* are for entry of speech utterances. Short utterances can be combined in one block (as are all of the utterances in Panels A and B of Figure 4). Utterances longer than one block can be continued on the next block(s). Enter an utterance number in the far left column for each utterance, even if the utterance is only one or more questionable words (see section IIB: "How to Format Questionable Words").

Space is also provided on each pepform to record a playback device's or digital sample's counter number, which indicates where an utterance occurs on the audio (or video) recording (see Utterance 1, Figure 4; Panel A). Continuous speech utterances will be the focus in this guide, but an 'utterance' can be as short as one sound or syllable in, for example, a stimulability test, or a single word, as is the case for an articulation test or many of the imitated tasks in the MSAP.

4. Transcription Lines

Within each utterance block (see Figure 4):

The X line is used for the English orthographic representation of speech,

the Y line is used for the intended (presumed) phonetic output, and

the Z line is used for the actual or observed phonetic output.

The Y line, then, is either the adult ambient form or the child's intended form, including lower-level phonetic feature specifications appropriate to the pragmatic context. The Z line is the participant's realized or surface form. You have the flexibility to represent in many ways how the participant 'intended' to articulate.

Deciding what should be entered in each of these three lines is, in fact, the major task of this guide. As described in detail later, PEPPER compares the entries in each line to one another. Depending on what it finds in each line, speech behaviors can be 'correct' or 'incorrect.'

5. Words, Segments, and Diacritics

Utterances consist of part-words and/or whole words, which in turn consist of phonemes and diacritics. PEPPER views any one or string of segments separated by spaces as a word. As discussed next, the basic units for phonetic transcription are phoneme symbols and diacritic symbols.

B. Phoneme and Diacritic Symbols

PEPPER uses the system of phoneme and diacritic symbols described in *Clinical Phonetics*, editions 1-4 (Shriberg and Kent, 1982, 1995, 2003, 2013). Changes to some symbols in edition 5 of *Clinical Phonetics* (Shriberg, Kent, McAllister, & Preston, 2019) are **not** included in PEPPER. *Clinical Phonetics* is a workbook-audio training series that teaches the perceptual skills needed to transcribe over 100 symbols that comprise the system. Figures 5, 6, and 7 provide, respectively, the organizational system for the consonants, vowels and five diphthongs, and diacritics. If you intend to use the diacritic system extensively, you should be thoroughly familiar with the descriptive materials in *Clinical Phonetics.* Another option, however, is to use only the IPA phoneme symbols and perhaps a few of the diacritic symbols (symbols for dentalized [̪], lateralized [˞], nasalized [̃], derhotacized [̯] and others) you feel confident in using reliably.

C. Additional Symbols Used in PEPPER

In addition to the phoneme and diacritic symbols, PEPPER transcripts use a few special symbols. They will be introduced briefly here and described in detail later.

1. The Three Punctuation Symbols

Only three punctuation marks may be used in the X line:

Comma (,)

Apostrophe (')

Question mark (?)

Do not use periods at the ends of sentences. The three punctuation marks are used solely for descriptive purposes in the orthographic representation (the X line). They do not affect any of the PEPPER analyses. Some examples:

X: Scotty is over here

X: Can't you see it?

X: One, two, three, four, five

Consonant Sounds of English

	MANNER	VOICING	PLACE						
			Bilabial	Labiodental	Lingua-Dental	Lingua-Alveolar	Lingua-Palatal	Velar	Glottal
OBSTRUENTS	Stop	Voiceless	p			t		k	ʔ
		Voiced	b			d		g	
	Fricative	Voiceless		f	θ	s	ʃ		h
		Voiced		v	ð	z	ʒ		
	Affricate	Voiceless					tʃ		
		Voiced					ʤ		
SONORANTS	Nasal	Voiced	m			n	ŋ		
	LIQUID Lateral	Voiced				l			
	LIQUID Rhotic	Voiced					r		
	Glide	Voiced	w				j	(w)	

Figure 5. The 24 English consonants organized by their articulatory features. The liquids /l/ and /r/ are classified as lateral liquids and rhotic liquids, respectively. [Adapted from Shriberg, L. and Kent, R. (2003, 2013). *Clinical Phonetics.* Boston: Allyn and Bacon.]

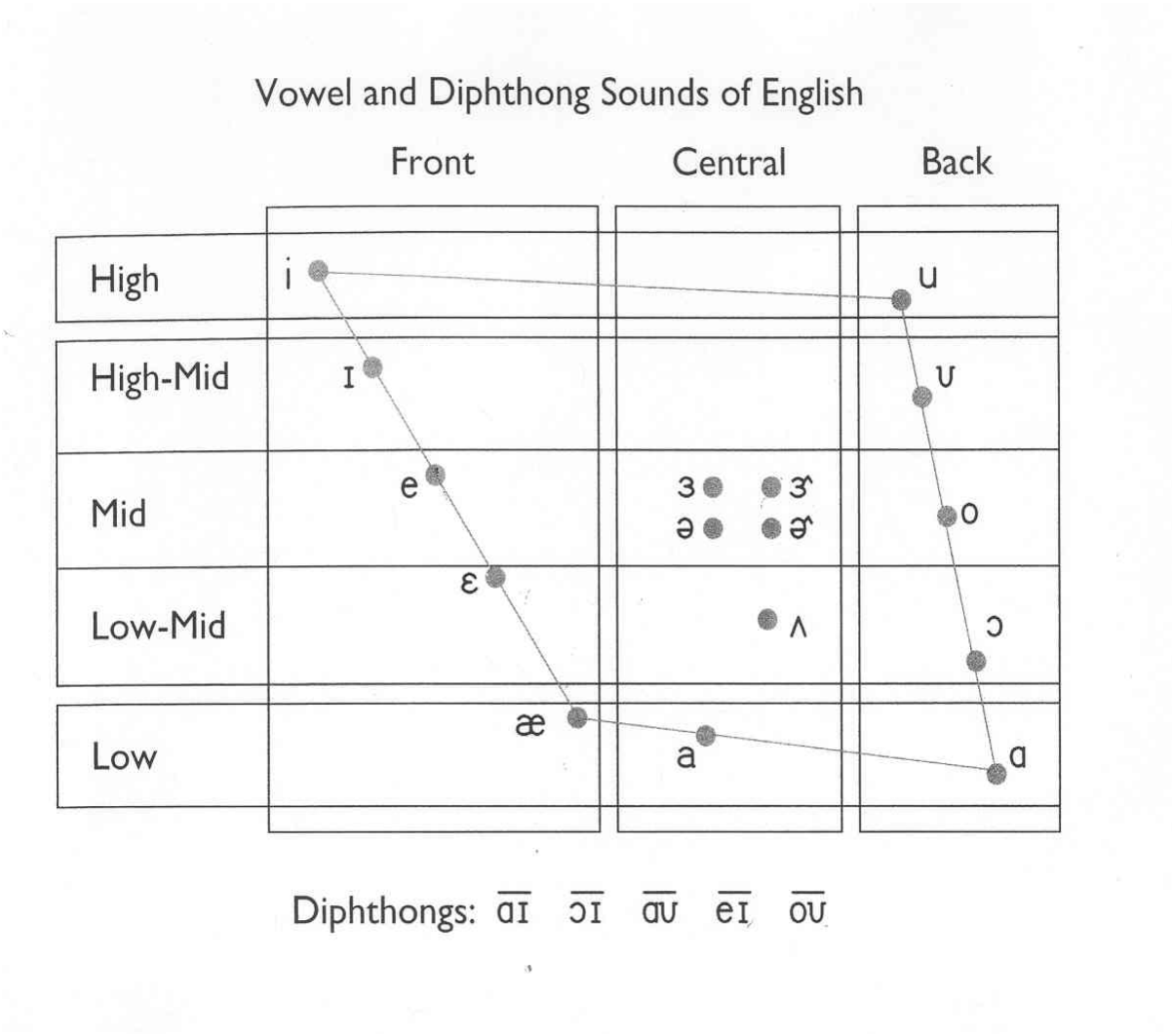


Figure 6. The 16 English vowels organized by their articulatory features (a.k.a. "the vowel quadrilateral"). The alternative IPA symbols for /ɪ/ and /ʊ/ are not included here. The five diphthongs available in PEPPER are /aɪ/, /ɔɪ/, /aʊ/, /eɪ/, and /oʊ/. [Adapted from Shriberg, L. and Kent, R. (2003, 2013). *Clinical Phonetics.* Boston: Allyn and Bacon.]

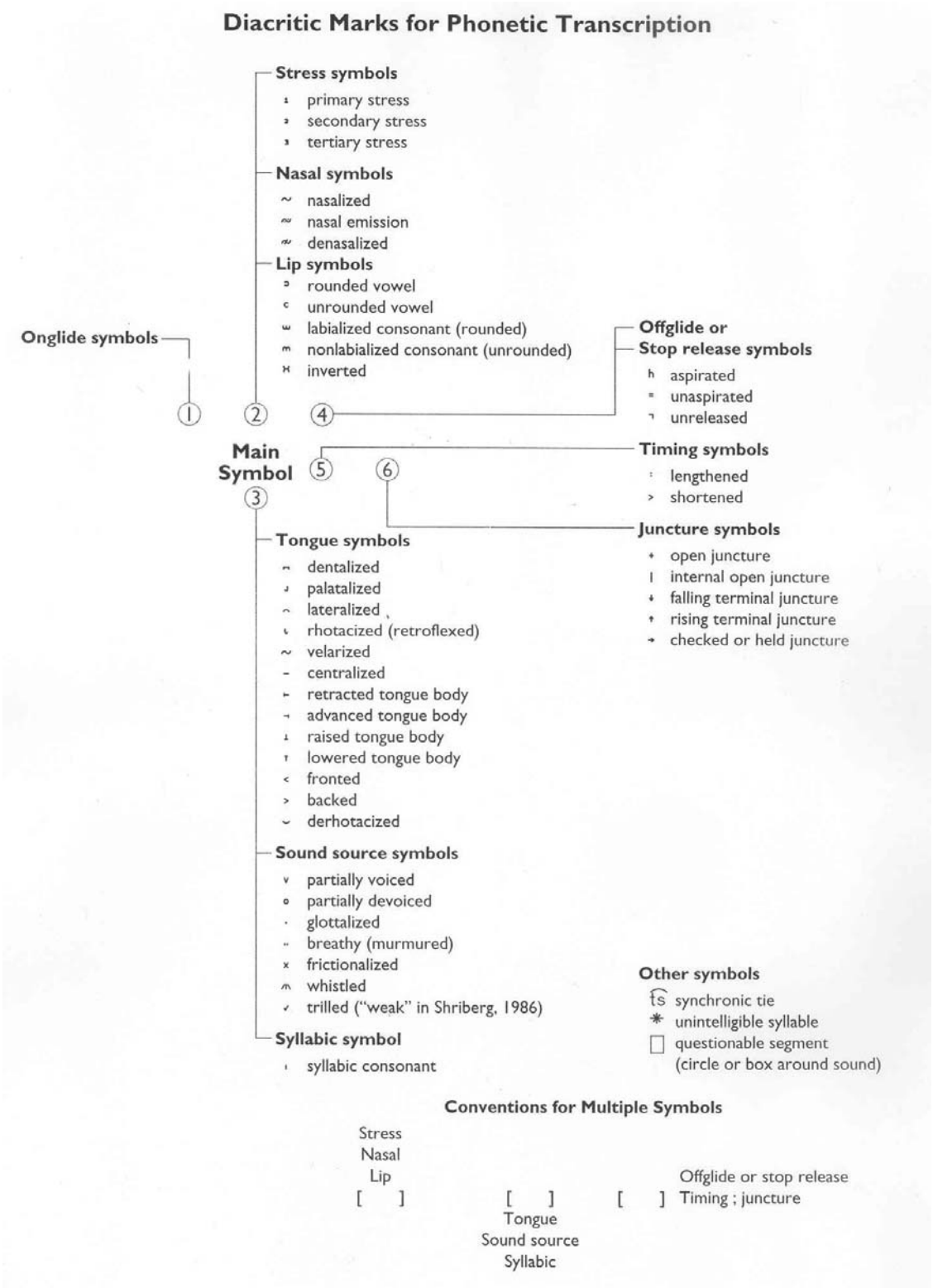


Figure 7. The 47 special characters and diacritics available in PEPPER for narrow phonetic transcription. [Adapted from Shriberg, L. & Kent, R. (2003, 2013). *Clinical Phonetics.* Boston: Allyn and Bacon.]

2. The Comments Symbol [ ] (square brackets)

Use square brackets to enclose your anecdotal comments about the participant or his speech. Square brackets are used only in the X line; they cannot be entered in the Y or Z lines. For example:

X: The car won't move [whiney voice]

Y: ðə k͡ɑ͡r w͡o͡ʊ͡nt m͡u͡v

Z: ðə k͡ɑ͡r w͡o͡ʊ͡nt m͡u͡v

Comments may be used to describe or explain circumstances associated with speech, the audio quality, or the environment; for example, [imitated], [laughing], [whispered], [was too far from mic], [distorted signal], etc. Entered comments will appear on the PepFile transcript and can be viewed as a Comments Summary in PEPPER; however, they are for informational purposes only and do not affect any of the PEPPER analyses. See Table 1 on pp. 22-25 for a list of sample abbreviations for comments. Figure 4 includes some examples of how these abbreviations are used (e.g., [NE] = noise: environmental, [O] = clinician over talk).

3. The Four Symbols for Questionable Words

Four symbols are used to exclude words in the speech transcript from certain quantitative analyses. As a group, words marked by these four symbols are called *questionable* words. Rules and guidelines for using these symbols will be discussed in Section IIB. Words, syllables, or segments coded with these symbols are included in certain PEPPER analyses and excluded from others.

a. *Disregard < > (angle brackets)*

The disregard symbols are called angle brackets (as differentiated from the square brackets used for comments). Angle brackets are used only in the X line for a variety of purposes, such as to indicate words used as fillers, false starts, and sound and syllable (i.e. part-word) repetitions.

b. *Either/Or / (slash)*

Use the Either/Or symbol, a slash, only in the X line when you are unsure about which of two words a participant intended to say.

c. *Unsure O (Circled)*

Draw a circle, the Unsure symbol, around one or more characters in the X line or in the Z line (on the computer screen, the circle appears as a box when entered) when you are unsure about the intended sound(s)/word, and/or whether the sound is present.

d. *Unintelligible * (asterisk)*

Use the unintelligible symbol, an asterisk, in the X, Y, or Z lines when you are unable to gloss or transcribe a syllable or a word.

4. The Bound Morpheme Symbol ( ) *(parentheses)*

Parentheses are used to enclose bound morphemes that are obligatory but were not produced. Their use is described in detail in Section IIC on pp. 33-34.

5. The Missing Symbol __ *(under bar)*

The under bar, a solid line, is used to mark a segment that is missing in the Z line. It acts as a space holder so that PEPPER can inspect the correspondence between what was intended (Y line) and what was said, or omitted (Z line).

D. General Procedures

A discussion of procedures for clinical transcription, including guidelines for efficient use of time, is offered in Chapter 9 of *Clinical Phonetics (5th ed.)* (2019). Topics include recording equipment, listening and scoring options, and strategies to maximize transcription efficiency. Familiarity with materials of this sort is strongly suggested for persons who will be phonetically transcribing speech for PEPPER analyses. The discussion here includes only those procedures that directly affect the formatting of transcription for valid PEPPER entry and analyses.

NOTE: The Phonology Project transcribers use PEPPER to open and play audio files in uncompressed .wav format, and use an X-Keys programmable foot pedal to control audio file playback. For additional information, refer to "Displaying a Sound or Video File" in the Help section of PEPPER; see also PEPPER Guide 5.

1. Determining an Utterance

NOTE: Because we complete prosody-voice coding on all conversational speech

samples, we use the glossing and utterance segmentation rules on pp. 9-11 in the PVSP manual. Therefore, the information below, taken from the 1986 *PEPPER User's Manual*, does not necessarily reflect our current procedure for segmenting and determining utterance boundaries. *If no PVSP will be completed, the following guidelines will suffice.*

a. Basic Definition

Templin's (1957) adaptation of Davis's (1937) definition of an utterance is useful: An utterance is defined as "...the child [comes] to a complete stop, either letting the voice fall, giving interrogatory or exclamatory inflection, or indicating clearly that he [does] not intend to complete the sentence."

b. Complex Utterances

Complex utterances containing conjoined and/or embedded clauses are entered as single utterances. In PEPPER, conjoined utterances are defined as two or more independent clauses joined by a conjunction *(and, but, so).* Sequential utterances that are not joined by a conjunction are entered as separate utterances.

c. Non-complex Utterances

Utterances containing *and* or *or* to join compound subjects or predicates are entered as single utterances. For example:

Susan, Carol *and* Bob, *and* Sam all helped.

I washed *and* dried *and* put away the dishes.

I like cookies *and* cake *and* ice cream *and* donuts.

d. Long Utterance Options

Utterances that do not fit on one of the utterance blocks on a pepform are simply continued to a second block without advancing to a new utterance number (see Figure 4).

Because only one of the PEPPER analyses is concerned with utterance length, it is convenient to divide very long utterances into two or more utterances. Long utterances (longer than 15 words or so) can be difficult to enter on the keyboard and view on the computer screen. Moreover, if the participant makes many speech errors, it becomes more challenging to edit the longer Y and Z lines. Such utterances are given successive numbers. The procedure, then, is to divide the utterance into two utterances using a logical boundary and to assign

the next utterance number to the second utterance. Enter a short note describing what was done in the *Notes* section on the Cover Page. If the speech sample will also be used for language analyses, identify these artificially separated utterances specifically in square-bracketed comments (see section IIA, item 5).

2. Determining Where to Begin Transcribing

Typically, a continuous speech sample is transcribed from the first utterance in the audio sample, but the transcriber can arbitrarily determine where in the sample to begin. For example, if a PVSP will be completed on the sample, the transcriber may wish to begin with the first PVSP-codable utterance. Once a starting point is identified, consecutive utterances should be glossed and transcribed from that point. No utterances should be skipped. This rule controls sample selection so that utterances are not directly selected by the transcriber. It allows the transcriber to select a rich, productive section of the speech sample at which to begin transcribing while still providing a representative sample of the participant's speech.

3. Using Counter Numbers for Digital Samples

Counter numbers are typically recorded on the pepforms (at the beginning of the first utterance and then at the beginning of subsequent pages) and on the cover page of the PVSP Scoring Form (beginning of sample, first coded utterance, last coded utterance). A transcriber may wish to record counter numbers more often, particularly in samples where there are long periods of time when the participant is not talking. If the sample will be acoustically analyzed later, recording a counter number for each utterance that is coded for prosody-voice will make locating these utterances on the recording considerably easier and more efficient.

When playing a digital recording in PEPPER or other playback software, typically the counter number is automatically displayed at the point in the sample that is playing. This makes determining and recording counter numbers an easy, straightforward process.

4. Alternating Among X, Y, and Z Lines

The following sequence seems the most efficient for glossing and transcribing utterances in the X, Y, and Z lines.

a. First determine the gloss, including all questionable segments or words. Write the gloss in the X line, leaving ample space between words. If you are

transcribing a conversational speech sample that will also be coded for prosody-voice, gloss all utterances to be transcribed and/or coded and complete prosody-voice coding before moving on to Step b.

b. Move to the Z line and begin transcribing what you hear. Try to align vertically the symbols within each word so that you can alternate between the Y and Z lines. That is, immediately after you enter your transcription of a segment or word in the Z line, directly above it enter the intended form in the Y line. **It is only necessary to enter the intended form in the Y line if it does not match the realized form in the Z line (i.e., it is a speech error).** As discussed below, this is the critical point at which you decide whether the participant intended to use acceptable allophones--and, therefore, whether the speech in Y is appropriately viewed as containing phonetic or phonologic errors. Continue alternating between Y and Z until all entries in X are accounted for.

c. Finally, and this is a critical step, check your X, Y, and Z lines for accuracy. Did you listen to each segment? Have you observed all the conventions described in this guide? Will your transcription and formatting conventions allow for valid PEPPER input and output?

5. Using PEPFORM: First Occurrence Words

The form called *PEPFORM: First Occurrence Words* is useful to keep track of word types (first occurrence words in a continuous speech sample). For some PEPPER analyses you can elect to look at just these words (*word types*), as opposed to all words (*word tokens*). *PEPFORM: First Occurrence Words* allows you to keep a running tally of first occurrence words as you gloss them. (Remember that conversational samples are typically transcribed to 100 first occurrence words.)

To use this pepform, while glossing, simply enter each new word in the appropriate box, alphabetized by the first letter of the word. You can make tallies at any time by drawing a line after the last entry in each box and summing words across boxes. The scratch area at the top of the form is used for the tally marks. Hash marks are used to keep a running tally of first occurrence words as they are glossed. Obviously, as the sample proceeds you will obtain a successively lower rate of new words per unit of time. In one study that divided a speech sample into fifths, in each successive one fifth of the sample, the proportional occurrence of new words (word types) was as follows: 29%, 21%, 19%, 17%, 14% (Shriberg and Kwiatkowski, 1985).

PEPFORM: First Occurrence Words

Pepfile Name SRT1-C1

Peplog No. 5363

Scratch Tallies

Total Words 100

A	B	C	D	E	F	G	H
a and arms are alot at almost	blue black big but	can chairs chair	don't different didn't	eyes ears every		go gonna glasses green grapes got	have hat he his here had has house home

I	J	K	L	M	N	O	P
it is I'm I in		know kinda keep	look(s) like little lift	maybe muffins many	no nose not	one oat oats of on	put pink play puzzles

Q	R	S	T	U	V	W	X,Y,Z
	right red rice room really remember room	shoes supposed some seeds so sit sometime	though this tongue to teeth then that(s) there they're trucks toys they two them	up used		we where won't was with would	yeah yep

Figure 8. A PEPFORM: *First Occurrence Words* illustrating how a transcriber can keep track of how many first occurrences of words (word types) have been glossed and/or transcribed.

6. Determining the Length of a Transcript

Particularly if multiple continuous speech samples will be compared, it is important to adopt a rule for glossing and transcribing that will result in a reasonably consistent, standardized sample length. The rule that is adopted will depend upon the clinician's or researcher's goals. Potential rules include:

- The 90-70-225 Rule
- The 100 First Occurrence Words Rule
- The 12+ PVSP Codable Utterance Rule

a. The 90-70-225 Rule

One guideline that has been used to standardize the length of a continuous speech sample is referred to as the *90-70-225 rule* (Hoffmann, 1982; Shriberg and Kwiatkowski, 1985; Shriberg, Kwiatkowski, and Hoffmann, 1984). The rule states that you *stop transcribing the speech sample when you have transcribed (whichever occurs first):*

90 non-questionable, first occurrence words (word types)

or

70 utterances, each of which contains at least one non-questionable word

or

225 non-questionable words (word tokens)

The key words in these rules have been defined earlier in this guide. The *90-70-225 rule,* when followed in this sequence, ensures that the sample is neither too short nor too long for PepAnalyses outputs. With normally speaking and speech-delayed children, the *90-70-225 rule* should result in your needing to transcribe approximately 5-8 minutes of continuous speech. Of course, samples that violate this rule can be used, with explanatory comments in the *Notes* section. If you find that you need longer samples to obtain the necessary number of words or utterances, you should examine your speech sampling methods in relation to suggestions discussed in PEPPER Guide 1. Very long samples may contain proportionally more function words, which may jeopardize the validity of the Percentage of Consonants Correct analysis.

1) When using the “90 non-questionable first occurrence words” option in this rule, words that are spelled differently but sound the same *(to, too)* or that mean different things *(times* in multiplication versus the duration of events) are counted as different words. The assumption is that different lexical forms

may be represented or produced differently in child phonology. This convention does not depend on how the child actually said either word, that is, whether or not either or both alternatives were articulated either correctly or the same way.

2) Also, when using the “90 first occurrence words” option, completely transcribe the utterance in which the 90th word occurs.

3) When using the “70 utterances” option in this rule, be sure not to count any utterances in which every word is questionable (every word is in angle brackets or circled) or unintelligible. Also, avoid including too many utterances that consist entirely of words like "yes" or "okay," even if they are not in angle brackets. Doing so would decrease the likelihood that the 70 utterances will contain a sufficient number of different word types and/or total words. Rather, transcribe an additional number of utterances that are not single-word responses.

b. The 100 First Occurrence Words Rule

For many years, we have glossed and transcribed continuous speech samples up to and including the utterance in which the 100th first occurrence word is reached. The final transcript length is dependent upon the speech and language skills of the participant. Generally, the more unintelligible a participant is, the longer the transcript will be, as it takes more utterances to reach 100 intelligible, non-questionable first occurrence words. For children with significant speech and language delays or disorders, 100 first occurrence words might not be reached. In such cases, it is important to keep in mind that the validity of some of the resulting PepAnalyses outputs may be compromised.

c. The 12+ PVSP Codable Utterance Rule

This rule evolved from the need to have a sufficient number of PVSP coded utterances transcribed in order to complete acoustic analyses. Basically, if 100 first occurrence words are reached before 12 PVSP coded utterances are reached, continue to transcribe up to and including the 12th coded utterance.

II. TRANSCRIPTION PROCEDURES AND FORMATTING CONVENTIONS

This section has been organized specifically for later reference. The lists of "do" and "do not" conventions serve several purposes. They are necessary either because of technical constraints in the PEPPER software or reliability constraints in transcription, they are preferable for efficiency, or they are useful for particular questions about the

participant. Hence, the following guidelines are arranged as reference lists to be consulted as needed. You may need to augment these lists with items and new lists of your own. The underlying principles that are described for these conventions may help you develop consistent conventions for any speech transcription and formatting situation. Record each of your own conventions in an appropriate place right in this guide for future reference.

A. How to Use the Alphabetic Characters

1. Upper Case and Lower Case Letters

Only orthographic (alphabet) characters are available in the X line. Attempting to enter any other characters (other than some described below) into the X line on the computer will result in an error message. Capitalize the first letter of the first word in an utterance. Follow standard capitalization rules for all other words (e.g. proper nouns). Capital letters are treated exactly like lower case letters in all quantitative PEPPER analyses.

2. Punctuation

Use standard punctuation rules for comma (,), apostrophe (') and question mark (?). The comma and the question mark are for descriptive purposes only. Because they occur outside of words, they have no bearing on PEPPER analysis programs/outputs. The apostrophe must be used consistently for contractions, possessives, etc. (*it's, we're, John's*), because each unique form will be stored as a different orthographic word type.

3. Spelling

Accurate spelling of all words in the X line is critical for computer recognition. Each word must be spelled the same way each time it occurs. Because number symbols (7, 33) are not available for entry in the X line, they must be spelled out (seven, thirtythree). Below are some guidelines for how to spell frequently occurring forms in the X line. Although unique spellings are used for alternative casual speech forms, conventional spelling is used when word forms are reduced or misarticulated.

Yeah, yes, yep, okay

um, oh, uh, m, mhm, uhuh, uhhuh

donno, gimme, lemme, gonna, gotta, hadta,
hafta, kinda, wanna, whatcha

til (in place of until), ‘*cause* (in place of because

- note: not "cuz")

playing (even though child said "playin")

4. Word Division

Differentiate noun compounds (blackboard) from noun phrases (black board). The former will be treated as a two-syllable word, the latter as two words. Ritualized reduplications (for example, *bye bye)* are entered as one word (*byebye*). Proper names (people's names, character names, games, book titles, movie titles, song titles, etc.), such as *Big Bird* or *Dairy Queen* are treated as one word (and are actually entered in the X line as one word, *BigBird*, *DairyQueen*) only for a young child or a child with limited language. *Disney World,* for example, might be assumed to be represented as only one word /dɪzniwɝld/ for a young child; for an eight-year-old child it would, assumedly, be represented as two morphemes /dɪzni/ /wɝld/. Because a hyphen is not available in the X line, words that are normally hyphenated, for example, t-shirt, forty-five, are spelled out or entered in the X line as one word, for example, *teeshirt*, *fortyfive* (see correspondence constraints in later sections). Also, because the period is not available in the X line, abbreviations are entered without them, as in *LA* or *DC*.

5. Square Bracket Comments

PEPPER's square bracket option allows you to enter comments at any point in the X line. Only orthographic characters can be entered within the square brackets, and square brackets are available only in the X line. All entered comments will be displayed on the PepFile transcript, but they will not affect any of the PepAnalyses outputs. A convention we've used is to place a square-bracketed comment above and immediately after the word, phrase, or utterance it references.

Comments may be of any length and may be used to annotate virtually anything you want. However, they should be kept brief to save space and prevent typing errors. The following alphabetized list includes some abbreviations we have used to keep track of various linguistic behaviors, non-linguistic behaviors, environmental or technical issues, and other details that were important for certain projects. Once again, we recommend developing and keeping a log of abbreviations you may want to use repeatedly.

Table 1. Sample abbreviation codes for comments.

Abbreviation Code	Description/Explanation
[A]	*Addition* indicates that the participant added a word or phrase to an articulation test word or a stimulus phrase in a *protocol*. (See later discussion of articulation tests and protocol files.)
[AE]	*Audible Exhalation*
[AI]	*Audible Inhalation*
[AS]	*Assimilation* across words, e.g., *things that*, is said as [θɪŋz zæt]. PepAnalyses programs compute the occurrence of assimilation (regressive, progressive) within words, but do not provide a way to compute presumed assimilation across words.
[BC]	*Breath Control* problem
[BLO]	*Block*, particularly a dysfluency-type block on an obstruent sound
[BN]	*Background noise* that interferes with reliable transcription
[C]	*Confirmation* of the examiner's gloss; used for discourse analysis when the participant says words like yeah, mhm, etc.
[CLAR]	*Clarification Register*; when not understood in the previous utterance
[D]	*Distance* of the microphone from the participant's mouth was too close or too far, with adverse consequences for transcription and/or prosody-voice coding
[DI]	*Delayed Imitation*; typically used only to contrast a response with those obtained spontaneously, by immediate imitation, etc.
[E]	*Exclamation*, particularly those that involve suprasegmental features of interest or that are associated with changes in

	previous articulation patterns
[ES]	*Exaggerated Stress*
[F]	*Fluency* breakdown
[FP]	Used to indicate that a sound transcribed as [t̼] in the Y or Z lines (see later discussion) sounds like a *flap*. Due to technical constraints in the analyses, the IPA flap symbol, [ɾ] is not available in PEPPER.
[GF]	Glottal Fry, as heard on the preceding vowel(s). Notice that the [̰] (see Figure 7), which indicates *glottalized*, is used to represent a variety of vocal percepts, e.g., *glottal fry*, *hard glottal attack*, *hoarseness*, *aphonic*, etc. This example demonstrates how comments can be used to differentiate among such behaviors.
[H]	*Hoarse* vocal quality (see above)
[HA]	*Hard glottal attack* (see above)
[I]	*Imitation*
[L]	*Laughing* as participant said the word or utterance
[LC]	*Lateral Click*; has been used to describe a stop or fricative that includes a non-linguistic click-like sound
[LREV]	*Language Revisions*
[LSC]	*Long Stop Closure*
[M]	*Mumbled* articulation, often used in association with sounds or utterances that are coded in the Y line
[MIC DISTORT]	*Microphone-related distortion* of the acoustic signal that may interfere with the transcription and/or acoustic analysis of the preceding word(s). [NOTE: Words that are acoustically distorted should be put in angle brackets because the accuracy and validity of the transcription could be compromised.]
[MON]	*Monotone/monopitch*

[MP]	*Misplaced stress*, either lexical or sentential
[N]	*Noise*; vocal, but non-meaningful; e.g., vowel or consonant-like noise that may be transcribed for a child with severe speech involvement
[NA]	*Narrative*; as opposed to conversational register
[NE]	*Noise: environmental*, which is transient, as opposed to consistently present noise from an air conditioner or fluorescent light (see [BN] above)
[NR]	*No Response*; i.e., an item in an articulation test or some other protocol task was administered, but the participant did not respond
[NS]	*Noise: sound effect*, e.g., using voice to simulate sound of a motor (often quite revealing in vocal intensity relative to sound made in 'speech')
[NT]	*Not tested*; i.e., an item was not administered, as in an articulation test or some other task or test wherein items are administered in a fixed sequence
[O]	*Overtalk* by the examiner or someone else which may, but not necessarily, preclude obtaining a gloss or a transcription of the previous word or utterance
[P]	*Playful, sing song* quality in the utterance
[PB]	*Pitch Break*
[PRO]	*Prolongation*, particularly a dysfluency-type prolongation on an obstruent sound
[PRQ]	*Poor Recording Quality*
[Q]	*Quiet* volume as if talking to oneself
[QLC]	*Questionable lateral click*; i.e., a slight or questionable [LC]
[R]	*Repetition*, used to explain certain words that have been angle-bracketed beginning with the third repetition (see later section on use of angle brackets)

[REP]	*Repetition*, particularly a dysfluency-type repetition
[RR]	*Rapid Rate*
[S]	*Searching; groping*
[SR]	*Snap Release* of stop
[TREM]	*Tremulous* voice quality
[VB]	*Voice Break*
[W]	*Whispered*
[Y]	*Yawned*
[YE]	*Yelled*

Additional abbreviation codes have been used to explain many other events, such as why words were placed in angle brackets and the nature of problems in agreement between two transcribers. When comments need to refer to speech sounds such as /ʃ/, you will need to spell out these sounds in orthography, for example, "sh", because phonetic characters cannot be entered in the X line.

This sample list of arbitrary abbreviation codes suggests the range of uses you may develop for comments enclosed in square brackets. By inspecting the number of each type of comment used in a transcript, you gain information about the participant and about your speech sampling procedures. A Comments Summary output available in PepAssess provides a summary of each comment used in a file including percentage data for the proportion of utterances containing each comment type.

B. How to Format Questionable Words

The general rule in transcribing continuous speech samples for PEPPER analyses is to account for everything a participant says. Once you begin transcribing, every successive utterance is transcribed. As introduced earlier in this guide, four categories are used for words that are in some way questionable: disregards, either/ors, unsures, and unintelligibles. In this subsection, rules (things you *must* do) for use of each of these categories are listed, and guidelines (things you may want to do) are presented for situations you may encounter.

1. Rules for the Use of Disregards (words enclosed in angle brackets, < >)

a. Angle brackets are used only in the X line. They may be placed around a unit no larger than a word, and whatever they enclose is treated as a disregarded 'word' in PEPPER computations. Angle brackets may be placed around the unintelligible symbol, <*>, around one sound orthographically spelled in the X line, <ch>, around one or more syllables, <ba> <baba>, or around an entire word, <yeah>.

b. Although a sound, syllable, or word is placed in angle brackets in the X line, the Y and Z lines should be completed in the usual manner. That is, a transcription of the intended sound or word should be entered in the Y line, and a transcription of what the participant said should be entered in the Z line. If the entire word or any part is unintelligible, use asterisk symbols following the rules described later in this section. Observe all capitalization rules, spelling rules, and other rules and conventions for words entered in angle brackets in the X line.

c. Disregarded words, as the term suggests, are words excluded from phonetic and phonologic analyses in PEPPER because of various problems with their validity or intelligibility. However, they are included in some structural computations, such as Average Words Per Utterance.

d. If a word or sound is placed within angle brackets, it does *not* need to contain a vowel or a syllabic consonant in the Y line. (The ERROR CHECK program in PEPPER requires that all nonquestionable words have at least one vowel nucleus.)

2. Guidelines for the Use of Disregards

The following list provides 11 guidelines and sample entries for use of disregard symbols (angle brackets).

a. *Repetitive words.* Routine uses of certain words or phrases are placed in angle brackets beginning with the *third occurrence* in the transcript (examples: *yes, okay, like, uhuh, no, I don't know, nothing*). Words in angle brackets will be excluded from PEPPER's phonetic/phonologic analyses. Such words or phrases must occur as a single unit, not in the context of utterances with additional words, such as "Yes, I can." Put a comment in square brackets to explain any non-obvious use of this guideline. For example, if a child repeatedly said "go go go go go," successive occurrences of "go" would be placed in angle brackets, and the utterance followed by an explanatory comment in square brackets:

X: And I go go <go> <go> <go> [R]

b. *Fillers.* The words *mm, uh* and other place holders are placed in angle brackets from their first occurrence onward.

c. *Imitated utterances*. If you choose to exclude imitated words from analyses, angle-bracket them and explain in a square-bracketed comment:

X: <television> [I]

d. *Exclamations.* Word-like forms such as *ah, oh, wow, whoops* would not be bracketed until their third occurrence in the transcript.

e*. False starts.* A single consonant or syllable that is presumed to be a false start on a word may be placed in angle brackets. For example,

X: <g> monkey [probably started to say gorilla]

Y: g mʌŋkɪ

Z: g mʌŋkɪ

f. *Revisions.* Whenever the participant revises a previous word or phrase, all completed words in both the original and the revised form are entered without angle brackets. However, all incomplete words in the original or revised forms are entered in angle brackets. Typically, but not always, the revised forms are more correct in grammar or articulation. Some examples:

X: And he she went...

X: He <ru> running is running

X: Yeah my bus comes don't come <un> until...

g*. Gloss confirmation.* Words such as *yes* or *mhm* that confirm a clinician's gloss of the previous utterance should be placed in angle brackets, followed by a comment, e.g.

X: <yeah> [C]

h. *Non-meaningful vocalizations.* Angle bracket all sounds made in play to imitate objects, such as cars or musical instruments, and follow with a comment. For example,

X: <rah> <rah> [NS]

X: <vroom> [NS]

i. *Questionable transcription.* If a sound, a whole word, or the whole utterance can be transcribed but is questionable because of participant variables (playfulness, whispered, muffled, peculiar voice quality) or technical/environmental problems (e.g., background noise, microphone distance, clinician over talk), put the X line entries in angle brackets and explain in a comment:

X: <that> <one> [P]

Y: ðæt wʌn

Z: ðæt wʌn

j. *Problem transcription.* If a sound, a whole word, or a whole utterance is difficult or impossible to transcribe because of problems due to participant, technical, and/or environmental variables, use angle brackets for that element and a comment:

X: <stores> [NE]

Y: stɔrz

Z: stɔrz

k. *Indistinct and contradictory forms.* Sometimes a participant's utterance has a meaning in adult language that was not the participant's phonological target. An indistinct utterance ([ɛ], [$\tilde{\text{ʌ}}$]) can be intended to mean *yes, no,* or *what,* and its particular meaning may not be clear from the context and inflectional characteristics of the utterance. In such cases, write the intended message *(yes, no, what)* in angle brackets in the X line and the utterance as produced by the participant in the Y and Z line.

Indistinct utterance: X: <yes>
Y: $\tilde{\text{ʌ}}$
Z: $\tilde{\text{ʌ}}$

Contradictory form: X: <yes>
Y: n$\tilde{\text{ʌ}}$

Z: nʌ̃

Of all the questionable word categories, angle-bracketed words are the most frequently occurring type. These 11 uses of angle brackets pertain primarily to procedures for glossing continuous speech. Other guidelines have been developed for articulation testing and for special circumstances when consensus transcription between two transcribers yields certain results. Once again, although a word is angle bracketed in the X line, be sure to enter a complete transcription for the sound, syllable, or word in the Y and Z lines.

3. Rules for the Use of Either/Ors (alternative glosses separated by a slash, /)

The second category of questionable words, *either/or,* is perhaps the least frequent category of questionable words used in PEPPER transcripts. Whenever the linguistic content and/or your knowledge of the participant suggest two likely glosses for a word in the X line, write your first choice before the slash and your second choice after the slash.

Either/or words, such as this/it, was/why, are excluded from phonetic and phonologic analyses because they are questionable. Because they are included in structural analyses, however, transcriptions in the Y and Z lines should be completed for the word before the slash, for example,

1) X: A/The boy

   Y: ə bɔ̄ɪ̄

   Z: ə bɔ̄ɪ̄

2) X: Her/She is big [pointed to girl]

   Y: hɝ ɪz bɪg

   Z: _ɪ ɪ_ bɪ_

4. Rules and Guidelines for the Use of Unsures (circled segments or whole words, for example, *the* (*cat*))

a. Circle sounds or whole words (circles appear as boxes in PEPPER) for one of two reasons: I) if you're unsure of the partial or complete gloss of a word, circle the appropriate segments in the X line, or 2) if you're unsure of the partial or

complete transcription of a word, circle the appropriate segments in the Z line.

b. On pepforms, an entire word can be circled, but when that word is entered in PEPPER, each sound in the word is circled individually in the X or Z line.

c. Sounds are never circled in the Y line because it would be conceptually inappropriate. An unsure in the X line means you are unsure of the gloss. Unsures in the Z line mean you are unsure about the participant's actual output form. Examples are:

1) Unsure of gloss

X: (Daddy) here?

Y: dædɪ hɪr

Z: dæ__ hɪ_

2) Unsure of transcription

X: At my party

Y: æt maɪ parti

Z: æt maɪ pa͡r ti

As a general rule, whenever you are in doubt about the gloss or your transcription, it is advisable to circle sounds or whole words. Review of a transcript for the occurrence(s) of unsures can tell you about the participant's speech and, perhaps, where you may need additional transcription practice.

5. Rules and Guidelines for the Use of Unintelligibles (non-glossable and/or non-transcribable syllables ✳)

a. Unintelligible symbols (asterisks) are used in essentially the same way as unsures, except that they are the 'symbols of last resort.' An asterisk is placed for each syllable in the X line when it cannot be glossed, and in the Y and Z lines when a syllable cannot be transcribed.

b. An asterisk denotes a syllable and therefore one asterisk is used for each syllable in a word. Syllables in a word are apparent by the train of vowel-like sounds. If you think an unintelligible word has more than one syllable, place an

asterisk next to one another for each vowel heard. Leave a space between asterisks to indicate word boundaries, for example,

X: ** *** *

c. The ratio of monosyllabic words to polysyllabic words is approximately 3:1 in a speech sample (Shriberg & Kwiatkowski, 1980; 1983). Therefore, if the participant presents a continuous string of unintelligible syllables, make the first three syllables monosyllabic words (one asterisk) and make the next two syllables a two-syllable word (two adjoining asterisks). Repeat this convention to the end of the unintelligible string. For example,

X: * * * ** * * * **

d. Asterisks are the only questionable symbols that can be used in the Y line. Each asterisk is tabulated as a syllable and each continuous string of asterisks (i.e. not separated by spaces) is tabulated as a word. Here is an example of a 2-word utterance:

X: <Going> <**> [shouted into mic; distorted]

Y: gou̯ɪŋ ✳✳

Z: ✳✳ ✳✳

e. Any unintelligible syllable within a word can be represented by an asterisk. For example,

X: <mixture> [O]

Y: mɪkstʃəˤ

Z: mɪks✳

f. It is important for the purposes of the Intelligibility Index to represent all speech by either a gloss and transcription or by the unintelligible symbol. Note that if the word cannot be glossed, but can be transcribed, asterisks should be placed only in the X and Y line. For example:

X: *

Y: ✳

Z: ʌpᵊ

g. Remember that asterisks can be used in combination with other questionable categories. For example,

X: <*> [NE]

Y: ✳

Z: ✳

In this case, because the unintelligible word was placed in angle brackets, it will be removed from the Intelligibility Index. The inability to gloss and transcribe the word was due to a technical problem with external noise ([NE]), that is, it was not due to the participant's lack of intelligibility.

To test your understanding of procedures for questionable words, describe the consequences of each entry in the following example--then review the guidelines in the previous pages to check your accuracy:

X: What/where is it? [M]

Y: wɑt ɪz ɪt

Z: ⓦⓐ ⓣ ɪz ɪ□ (circled underbar after /ɪ/

To summarize, the basic principle in the use of questionable words is always to gloss and transcribe speech whenever possible. If some participant or technical factor makes glossing or transcription impossible or unreliable, use the *either/or, unsure,* or *unintelligible* symbols as required. If a sound, syllable, or word should be excluded from the phonetic and phonologic analyses for any of the 11 reasons listed in the discussion of *disregards,* use the angle brackets convention. Finally, the use of square brackets for comments is recommended to provide a useful record in the transcript of the rationale underlying each of your decisions.

C. How to Format Errors Involving Grammatical Morphemes

The interaction between phonology and syntax often poses problems for transcribing and formatting a transcript when a child or an adult has deficits in both areas. Procedures for each of four possibilities are described as follows.

1. Correct Grammar: Correct Speech

If the participant uses a correct grammatical form and also articulates it correctly, transcribe and format in the usual way. Some examples:

X: He's running	X: Two carrots	X: I'll go
Y: hiz rʌnɪŋ	Y: tu kɛrəts	Y: a͡ɪl go͡ʊ
Z: hiz rʌnɪŋ	Z: tu kɛrəts	Z: a͡ɪl go͡ʊ

2. Incorrect Unbound Morphemes: Correct Speech

If the participant uses an incorrect unbound morpheme form, but articulates it correctly, enter the gloss exactly as said in the X line and its intended and obtained transcription in the Y and Z lines. Some examples:

X: Him goed	X: Her putted it here	X: I has [have] two
Y: hɪm go͡ʊd	Y: hɝ pʊtəd ɪt hir	Y: a͡ɪ hæz tu
Z: hɪm go͡ʊd	Z: hɝ pʊtəd ɪt hir	Z: a͡ɪ hæz tu

You can use the square bracket convention to enter the obligatory form after the incorrect or deleted form. This procedure allows for later rapid identification of incorrect grammatical elements, if you are interested in inspecting such data.

3. Correct Grammar: Incorrect Speech

Whenever a participant's error is obviously attributable to speech rather than grammatical deficits, enter the correct grammatical gloss in the X line, the correctly-intended transcription in the Y line, and the incorrectly-articulated speech in the Z line. Some examples:

X: Go running	X: All the bunnies	X: She is fishing
Y: go͡ʊ rʌnɪŋ	Y: ɔl ðə bʌniz	Y: ʃi ɪz fɪʃiŋ
Z: go͡ʊ r͡ʌ_ɪ_	Z: ɔl ðə bʌnid	Z: t̪ʰi ɪ_ fɪ_i_

4. Incorrect Bound Morphemes: Incorrect Speech

If the participant did not appear to intend a bound morpheme in an obligatory context, gloss the root word in the X line, placing the grammatical morpheme within

parentheses. In the Y line, enter only the root word and transcribe what the participant says in the Z line. Some examples:

X: two cat(s)	X: he('s) happy	X: Adam walk(ed) home
Y: t͡u kæt	Y: hi hæpɪ	Y: ædm̩ wɑk hoʊ̃m̃
Z: t͡u kæt	Z: hi hæpɪ	Z: ædm̩ m̥wɑk hoʊ̃m̃

This last situation involves only sounds or syllables deleted from root words. For instance, deleted auxiliaries would be handled by the procedure listed for "Incorrect Unbound Morphemes: Correct Speech." The parenthesis convention preserves in the transcript instances where a bound morpheme is required but not represented in the participant's speech. (Parentheses are treated like any other orthographic character within the PepAnalyses, that is, just like any other alphabet character or punctuation symbol.) The combined use of square brackets and parentheses will allow for subsequent inspection of potential influences of language on speech and speech on language. [NOTE: Parentheses can only be used to indicate deleted sounds or syllables at the end of a word. PEPPER does not allow parentheses to be entered around sounds or syllables at the beginning or in the middle of a word.]

D. How to Transcribe Phonetically and Format Speech Errors

The five following subsections provide explicit procedures and guidelines for phonetic transcription of speech. These materials have been derived from over thirty years of experience with hundreds of participants from preschool age to adult and many clinician-transcribers. These accumulated guidelines provide both the answers to common situations you will encounter as well as the principles for solutions to novel situations you might encounter. Obviously, these materials cannot be memorized in one or even several readings. But our experience indicates that if you keep referring to these guidelines as you prepare and enter new transcripts, you gradually will assimilate the basic information.

1. General Considerations

a. Broad Versus Narrow Phonetic Transcription

You may use either broad phonetic transcription or narrow phonetic transcription for creating PepFile transcripts for subsequent PepAnalyses outputs. Questions about academic background and training, clinical-research needs, available time, and related issues will need to be considered. If you do

elect to use some form of narrow phonetic transcription, you gain descriptive information at the cost of time and possibly reliability. Such issues (Shriberg & Kent, 2003; Shriberg & Kwiatkowski, 1980; 1985) continue to intrigue workers in child phonology and speech pathology. For the remainder of this section, it is assumed that you have elected to use some form of narrow phonetic transcription and that you are transcribing continuous speech. Even if you will be using only a few diacritics to capture common speech-sound distortions (e.g.[s̪], [l̫], [r̮]), you should study this section carefully. Included are conventions that apply when using broad or narrow transcription, such as how to format deletions, how to evaluate casual speech sounds, and other situations requiring standardized conventions. Applications to articulation testing protocols and other evaluation procedures should be readily discernible, though you will need to standardize procedures for your particular requirements. The emphasis here is creating transcripts that will generate valid and reliable PEPPER data to be used in making clinical and research decisions.

b. Reference Sources

All versions of *Clinical Phonetics* include definitions for all phonetic symbols used in PEPPER and guidelines for their use. Two reference sources for broad phonetic transcription of intended words include, but are not limited to, *A Pronouncing Dictionary of American English* (Kenyon & Knott, 1953) and *Longman Pronunciation Dictionary, 3rd edition* (Wells, 2008). Especially for traditionally troublesome sounds like those in the /r/ family, for instance, *ear* [iɾ] vs. [ɪɾ] vs. [iɚ], PEPPER users will need to become familiar with the transcription procedures recommended in *Clinical Phonetics* and to look up underlying forms in a pronunciation dictionary as needed. As described below, however, the formal citation register that a child might use in an articulation test (assumed in the form listed in Kenyon and Knott) is intended only for some words in continuous speech. Participants typically use more casual speech forms in clinical-research speech samples. For these and other reasons, we need to give them the benefit of the doubt rather than formatting their speech so that PEPPER scores it as incorrect. More on this later.

c. Transcription and Formatting Legibility

If you are handwriting glosses and transcriptions prior to entering a PepFile into PEPPER, two general suggestions about handwriting and neatness warrant brief comment. First, both your orthographic and phonetic characters, including all main characters and diacritics, should be written clearly and consistently. In the press of time, printed letters that resemble one another, such as "r" and "n", can become ambiguous. And phonetic symbols, such as [̪] versus [̮], [̬] versus [̌] and [̩] versus [̣], can become so degraded that you or someone

else cannot tell one from another at some later date. Also, the size and placement of symbols need to be regularized to prevent ambiguity, for example, [ɔrɪnd̪ʒ] rather than [ɔrInd̪ʒ]. Occasionally we have experienced problems with incomplete erasures. Sometimes it is not clear later whether a symbol was just written faintly or had been partially erased.

The second condition concerns the alignment of all words and symbols. Within lines, for example, the placement of onglides and offglides is frequently ambiguous if a transcriber does not place the added sound close to a main character. Hence, [te͡ɪᵐ bo] will be analyzed differently than [te͡ɪ ᵐbo]. Guidelines are presented later for determining where an on-/offglide belongs relative to the main symbols on either side of it. Between the Y and Z lines, align words and symbols so that the correspondences are easy to view, as also discussed later.

The point here is that once you have decided on the appropriate transcription, be sure that your handwriting and neatness make the transcription legible and formally correct for the keyboarding process described in PEPPER Guide 3. We have found that a stock of No. 1 pencils, comfortably soft erasers, and an electric pencil sharpener are indispensable tools for transcription. (This sounds a bit antiquated, but it's still true!) (See also the 4th edition of *Clinical Phonetics,* Chapter 7, for specific technical and procedural suggestions.)

2. Basic Formatting and Transcription Constraints

a. The Two Basic Formatting Constraints

PEPPER's analysis programs compare entries in the X, Y, and Z lines. Therefore, you must follow two basic formatting constraints:

1) The number of words must correspond among all three lines (X, Y, and Z).

2) The number of phonetic segments must correspond in the Y and Z lines. Methods to assure correspondence matches and examples are provided later in this section.

b. The Two Basic Transcription Guidelines

Two principles underlie the methodological approach to transcription and formatting of normal and disordered speech for PEPPER analyses.

Principle No. 1: Give the participant the benefit of the doubt.

The benefit-of-the-doubt rule guides much of the material in this section. Avoid formatting speech behavior such that PEPPER will view certain speech behaviors as speech errors when they are not. Rather, attempt to find a rationale to give the participant the benefit of the doubt.

Principle No. 2: Follow the error hierarchy - distortion, substitution, deletion.

This rule is consistent with the benefit-of-the-doubt notion. For example, it is more conservative to infer that a participant has devoiced /z/, [z̥] (a distortion) rather than to infer that the participant has made a substitution of /s/ for /z/, [s/z]. The difference translates to an articulatory or a phonetic error [z̥] versus a phonologic error [s/z]. Given the limitations of perceptual phonetics and the assumed direction of a child's developing system, phonetic differences are the more likely. Always ask the following question as you are transcribing: "Can this speech difference be transcribed using the same main characters in Y and Z?" You should use a different main character in Z to indicate a substitution *only* if your perceptual judgment is more certain that it is an s/z, w/r, w/l, or θ/s, for example. If you are uncertain, a distortion such as z̥/z, r̝/r, l̟/l, s̻/s, etc. should be indicated.

c. General Guidelines for Including Normally Occurring Allophones

Use of narrow phonetic transcription allows a level of descriptive detail useful for certain questions about normal and disordered speech. However, this level of description should not bias PepAnalyses outputs such that the participant's speech is seen to have many phonetic and phonologic "errors." To allow for detail (i.e., diacritics) in the Z line, without causing a mismatch in Y-Z correspondence, the solution is to enter the Z line behavior (diacritic) also in the Y line. By entering the same symbols in both the Y line and the Z line, these added elements of detail will not be viewed as incorrect in PepAnalyses outputs.

To apply the correct principles, the transcriber must first be familiar with the obligatory and the optional allophones of speech sounds. The following eight English allophone rules cover allophone situations we have accounted for using the diacritic characters available in the PEPPER program.

1) In a stressed syllable, initial singleton voiceless stops are aspirated, e.g. *pat* [pʰat], *car* [kʰar].

2) In an unstressed syllable, singleton voiceless stops are optionally unaspirated, e.g. *city* [sɪt̚ɪ] or [sɪtʰɪ]. Note: The alveolar flap [ɾ] is not included as a main character in PEPPER. In its place, use the unaspirated alveolar stop [t̚].

3) Voiceless stops in all clusters preceding a vowel are unaspirated, e.g. *stand* [st̚ænd], *spoon* [sp̚un], or *iceskates* [a͡ɪsk̚e͡ɪts].

4) In medial clusters and word-final position, all stops are optionally released, e.g. *cupcake* [kʰʌpʰkʰe͡ɪkʰ] or [kʌp̚kʰe͡ɪk̚], *bed* [bɛdʰ] or [bɛd̚].

5) Voiced stops and fricatives in word-final position normally are partially devoiced, e.g. *sled* [slɛd̥], *bees* [biz̥].

6) Postvocalic /l/ is normally velarized, e.g. *almost* [a͡l̰ mo͡ʊst], *tail* [te͡ɪl̰].

7) Vowels and diphthongs that occur before nasal consonants are optionally nasalized, e.g. *ring* [rɪŋ] or [r̃ɪŋ].

8) Vowels and diphthongs are longer in duration in open syllables (V, CV) than in syllables closed by *voiceless* obstruents (VC, CVC), but shorter in duration than syllables closed by *voiced* obstruents. Consider: *bee, beet, bead*.

Given the list of allophone rules above, when and how should obligatory and optional diacritics be transcribed in Y and Z? The transcriber really must deal with two issues or questions:

1) Q: How much phonetic detail should be preserved in a PEPPER transcript?

A: You need not use diacritics in Z and Y to indicate the normally occurring allophones described in Rules 1-8 above. The normal allophones will be assumed to occur. However, you should transcribe such behavior when a research protocol requires it or when transcription of these allophones will clarify an associated error. For example, the transcriptions [kɑːˍ] *car* and [pæ̃ˍ] *pan,* show that the participant 'marked' the deleted final consonant. Enter [kɑːr] or [pæ̃n] in Y and [kɑːˍ] or [pæ̃ˍ] in Z. Another example:

[ˍt̻̚ɛrwe͡ɪ] *stairway* in Z and [st̻̚ɛrwe͡ɪ] in Y shows that although the child deleted the /s/ and dentalized the /t/, the obligatory unaspirated /t/ was preserved. (Dentalized alveolar stops are non-errors; see Table 5.)

However, you may wish to enter normally occurring allophones in both Y and Z when the participant has previously used incorrect allophones or exaggerated articulation. For example, enter *car* [kʰɑr], with a square bracketed note to indicate the reason, i.e., [exaggerated aspiration].

2) Q: What response definitions should be used to have PEPPER recognize and keep track of phonetic and phonologic errors?

A: You should use diacritics when they describe a nonpermissible allophone. Nonpermissible allophones differ from those permitted in the language, but they are not always considered speech errors. For example, a lateralized /s/ [s͡ˠ] is generally considered a speech error, whereas a palatalized /s/ [sʲ] is not.

In this section, the suggestion is to transcribe normally occurring allophones only when they may be of some descriptive value. Subsequent sections of this chapter concern the diverse rationales used to determine if a speech behavior is acceptable (entered in both Y and Z) or unacceptable (an 'error' entered only in the Z line).

d. A Time-Saving Convention

It is efficient on a pepform to enter transcription in the Y line only when it differs from the transcription in the Z line (see Figure 3 for example). Depending on the type of error, you may elect to enter in the Y line just the sounds that differ or, for clarity, the whole word. Note these alternatives:

X: the cat	X: the cat	X: the cat
Y:	Y: ð kæt	Y: k
Z: ðə kæt	Z: də kæːˌ_	Z: ðə tæt

3. Procedures for the Use of Special Phonetic Symbols

PEPPER Guide 3 provides detailed instructions for entering symbols in the computer. This subsection provides rationales for several restrictions on symbol use. Reference to Figure 7 (p. 11) will be useful for the following discussion of diacritic symbols.

a. The Syllabic Symbol

1) Every non-questionable word in the Y or Z line must have either a vowel or a syllabic consonant. PEPPER's entry program inspects each word for this criterion. This restriction does not apply to questionable words, such as those that are placed in angle brackets. For example, [m] as a nonquestionable word in the Y or Z line would be an entry error, but [m̩] would be acceptable. Moreover, <ts> transcribed as [ts] in the Y and Z lines would be acceptable because angle-bracketed questionable words do not have to have at least one vowel or syllabic.

2) When a syllabic consonant is used, a vowel segment is assumed to precede the syllabic consonant. Hence, the canonical form of [h̃m̩] is CVC, and of [m̩] is VC.

3) Syllabic consonants are most likely to occur when the preceding sound has the same place of articulation, e.g. [bʌtn̩], [wɪsl̩] although the reduced speed of young participants' articulation typically includes a [ə] even in these situations, e.g., [bʌtən], [wɪsəl].

b. The Five Juncture and Three Stress Symbols

The 5 juncture symbols [+], [|], [↓], [↑], [→] and the 3 stress symbols [ˈ], [ˌ], [ˈˈ] are used for descriptive purposes only. Although the 5 juncture and 3 stress symbols are useful to describe prosodic aspects of speech, the reliability of perceptual judgments of prosody is not high. Instrumentally-aided approaches that enable these 8 symbols to be used reliably would justify the development of associated quantitative analysis programs.

You may freely put the 8 juncture and stress symbols in the Z line only or in both the Y and Z lines, although placing them in the Z line is usually sufficient for descriptive purposes. Juncture symbols are frequently used to indicate pauses occurring within words. (Note how [sp+un] differs from [s:pun]. Owing to their mutually exclusive referents, only one diacritic from each group should be used to transcribe a prosodic pattern, e.g., one from the 5 juncture symbols and/or one from the 3 stress symbols.

c .Number of Diacritics Allowed in Each Position

The number of diacritics that can be entered within each position is restricted by both logical and technical considerations. The following table describes the limitations (see Figure 7 on p. 11 for symbols).

You may select up to the listed number of diacritics for each position group. All groups are independent of one another except for the position shared by

offglides and stop-release symbols, as mentioned previously. Thus, you must limit the number of symbols you use to describe what you hear. For example, although you may want to transcribe a distorted /z/ as [z̥̹̺̻] (pretend here that all 3 of the diacritics are centered under /z/), you will have to choose the two diacritics that are least redundant and most different from the expected features (probably [z̹̺̻]) (again, pretend that both diacritics are centered under /z/) for PEPPER entry and analyses. The three groups restricted to only one diacritic include diacritics that are mutually exclusive (e.g. either [tʰ], [t⁼], or [t˺]).

Finally, as alluded to in item 7 in Table 2 above, even if you could demonstrate that such fine transcription can be made reliably, onglides and offglides cannot be modified by a diacritic.

Table 2. Maximum number of diacritics allowed in each position.

The 7 Symbol/Position Groups	Maximum Number of Diacritics Allowed Within Each Position Group
1) The 3 Nasality & 5 Lip Symbols [̃], [͌], [͌], [̑], [ͨ], [͡], [ͫ], [ͯ]	2
2) The 3 Stress Symbols [1], [2], [3]	1
3) The 3 Stop Release Symbols [ʰ], [⁼], [˺]	1
4) The 2 Timing Symbols [ː], [>]	1
5) The 5 Juncture Symbols [+], [|], [↓], [↑], [→]	1
6) The 21 Tongue and Sound Source Symbols [⌐], [˩], [⌢], [˪], [~], [-], [⊢], [⊣], [⊥], [⊤], [>], [<], [⌣], [v], [∘], [·], [··], [×], [/|\\ ], [✓], [ı]	2
7) Onglides, offglides (any vowel or consonant character)	1 onglide; 1 offglide; Diacritics may not be used to modify on- or offglides. Also, the offglide position is phonetically incompatible with stop release symbols. Therefore, only one offglide *or* one stop release symbol can be used.

d. Use of Onglides and Offglides

Phonetic additions may occur for a variety of reasons, such as during a period when a child is adding a new sound to her inventory, or perhaps as phonetic revisions when a participant with a speech disorder may 'grope' for the correct motor sequence. Formatting on- and offglides for valid PepAnalyses outputs requires adherence to some guidelines that may seem arbitrary. Rationales are provided below. The basic rule is to assign the on- or offglide to the sound that best matches its phonetic class. Hence, wherever possible, vowel on- and offglides are assigned to adjacent vowel main characters, and consonant on- and offglides are assigned to adjacent consonant segments. Some examples:

	1	2	3	4
X:	brake	chair	fish	sand
Y:	[bre͡ɪk]	[tʃe͡ɪr]	[fɪʃ]	[sænd]
Z:	[b ᵊre͡ɪk]	[ᵗse͡ɪr]	[fɪ ᵗs]	[sᵗ ænd]

Example 1 provides an illustration of the importance of correct formatting for PEPPER analyses. The schwa [ə] is entered as an onglide to the /r/ [ᵊr] rather than as an offglide to the /b/, [bᵊ]. The important rule here is always to assign the on- or offglide to the most likely character it modifies, typically, the more 'difficult' sound to articulate. Examples 2 and 3 illustrate the importance of the positioning of the main character and the diacritic. To anticipate phonologic coding in the series of PepAnalyes outputs dealing with Natural Phonological Process Analyses (Shriberg & Kwiatkowski, 1980), the child's [ᵗse͡ɪr] for [tʃe͡ɪr], *chair,* meets the definition for Palatal Fronting. Notice that if the /t/ were made the main character and the /s/ the offglide, i.e. [tˢe͡ɪr], this addition would be coded as Stopping. Another example: [fɪtˢ] for *fish* would be Stopping, but the correct (more advanced) entry would be [fɪ ᵗs] which would be coded as Palatal Fronting. The principle here is to always give the child the more advanced form as the main character, that is, the benefit of the doubt. Again in Example 4, making /s/ the main character will result in this addition being phonetically incorrect, but phonologically correct, that is, the child is in an advanced stage of fricative development wherein the fricative appears together with the old form, a stop, in the surface form.

e. The Synchronic Tie Symbol [‿]

The main function of the synchronic tie symbol [͡] is to allow one sound to be 'added' to another sound in the Z line. The tie symbol cannot be used in the Y line; for technical and theoretical reasons, it is used only in the Z line. Guidelines and constraints on the use of the synchronic tie are as follows.

1) Tied sounds are viewed by PepAnalyses programs as either substitutions or as uncoded, depending on the output. Hence, you should use the tie symbol only when you want to indicate that another phoneme has been added to the word rather than an on- or offglide. As discussed presently, the error hierarchy principle suggests that sound additions should be transcribed as *phonetic additions* (on- or offglides) wherever they are plausible, rather than phonemic additions (tied sounds). In those transcription situations where you think the participant does intend an additional phoneme, use the tie symbol. PEPPER will process the tied sound as one sound, hence avoiding a correspondence error (i.e., a mismatch) between the number of sound segments in the Y line and the Z line.

2) The principles above suggest two specific guidelines. When a word begins or ends with a vowel and a consonant is perceived as a sound addition, tie the consonant to the vowel. Alternatively, if the added sound is a vowel, make it an onglide or an offglide to a main character vowel or consonant. This arbitrary convention assumes that whereas vowels may occur as by-products as the tongue proceeds from one posture to another, consonant-like sounds would more likely be intended and hence can be given more phonemic status by representing them as main characters.

3) Any two sounds that are connected by a synchronic tie can be modified by only one diacritic from the Nasality-Lip Group. The reason for this constraint is that the tie is considered the second member of this group, and only two diacritics are allowed in this position. You may choose additional diacritics from all other groups but you cannot use on- or offglides between the two tied sounds. All diacritics modifying tied sounds are tabulated as modifying both of the sounds. (When you transcribe a diacritic above or below tied sounds, center it above or below the two tied characters, for example, [s͡t̲] not [s̲͡t]. This is the way it will appear on the computer screen and PepFile print-out.)

4) Transcription should distinguish between a lengthened or delayed release sound, [dɪvːa͞ɪd], and a sound that has two releases, [dɪv͡va͞ɪd]. The synchronic tie is useful to capture the second of these two situations. In some participants, the perception of a double stop [d͡gɑg], can be captured with the tie diacritic, followed by a comment.

5) The following examples provide additional illustrations of how the tie symbol is used to format Z-line entries.

	1	2	3	4
X:	air	up	train	outside
Y:	[e͞ɪr]	[ʌp]	[tre͞ɪn]	[a͞ʊtsa͞ɪd]
Z:	[h͡e͞ɪr]	[ʔ͡ʌʔ]	[tre͞ɪn͡tˢ]	[a͞ʊt͡ss͡a͞ɪd]

Example 1 is a typical application of the use of the tie diacritic. By tying the additional segment to the adjacent segment, correspondence between the Y and Z lines is preserved. Example 2 illustrates this principle, as well as a glottal substitution (ʔ/p). Example 3 illustrates another situation when two segments must be tied to the last segment in Y. Because only one main character can be tied, the final /s/ is made a diacritic (offglide) of the /t/. Example 4 was an unusual occurrence wherein the child actually said two distinct /s/'s. This was retained for PEPPER analyses and 'finessed' past the correspondence editor by using the tie diacritic to join the two segments. Recall that the tie symbol may not be entered in the Y line. As described later, neither can the glottal stop--which is why a tie was used for the first glottal stop in Example 2 above.

f. Use of Unaspirated /t/ for Alveolar Flap

The symbol for an alveolar flap [ɾ] is not available in PEPPER. Alveolar flaps tend to occur between stressed and unstressed syllables, for example, [lɛɾɚ] *letter,* and have a shorter closure time than the voiceless, unaspirated alveolar stop [t̄]. If you want to differentiate [t̄] from a flap, you can follow a perceived flap with a comment, as in [lɛt̄ɚ] [flap] (or the abbreviated comment [FP]).

g. The Glottal Stop Symbol

Glottal stops are not available in the Y line due to technical and conceptual issues that concern the status of glottal stops as acceptable allophones in American English dialects. The following guidelines for their use in the Z line should be observed.

1) Glottal stops are difficult to perceive reliably in word-initial position, and may be confused with a hard glottal attack on a word-initial vowel, for example,

okay [o͡ʊ̯ke͡ɪ] [glottal attack] versus [ʔo͡ʊ̯ke͡ɪ]. They are most reliably perceived in word-medial position, particularly when the first syllable is stressed and the second is unstressed, [ʌ̃¹ ʔʌ̃²] *uhuh.* In word-final position, glottal stops are most reliably perceived when they are followed by a schwa offglide, [pɪʔᵊ] *pig*. Notice how these latter two examples can be handled in the Y and Z lines:

X: Uhuh	X: pig
Y: ʌ ʌ	Y: pɪg
Z: ʔʌ͡ʔʌ	Z: pɪʔᵊ

2) Here are some examples of words that have required glottal stop transcriptions:

X: fish	X: toothbrush	X: zipper	X: matches
Y: fɪʃ	Y: tuθbrʌ̹ʃ	Y: zɪpɚ	Y: mætʃəz
Z: fɪ ʔʃ	Z: tutbwʌ̹ʔs	Z: dɪʔtɚ̯	Z: mæʔəz

In the first three examples, where the glottal stop occurs in addition to another consonant, the glottal stop is treated as an onglide (a phonetic distortion). In the last example, the glottal stop is treated as a substitution for /tʃ/.

4. Reduced Word Forms

Continuous speech sampling differs from articulation testing in many respects. A major difference in the present context is that participants make many speech sound changes that they might not make in the formal speech register used in naming or imitating articulation test words or other stimuli. For validity reasons, therefore, if certain word forms are reduced in continuous speech samples (casual speech forms), they must be formatted as acceptable, that is, as non-error data.

If the X line may be taken to represent what the participant intended to say, the Y line may be considered to represent how the participant intended to say it.

Consider the following questions in deciding how a participant intends to articulate a word.

1) How might a normal participant intend to say the words if in a casual speech context?

2) What dialect may be natural to the participant in question, and what are the context-sensitive allophones the participant may intend?

3) What effects might linguistic stress, pragmatics and speech register (playful speech) have on the intended word form?

4) What effects might a cold, fatigue, or other physical or emotional state variables have on intended forms?

Each of these four considerations may suggest that the intended form in the Y line is not the one used in "formal" General American English, as listed in pronunciation dictionaries. This section lists, by category, some reduced word forms that we believe may qualify as non-errors or acceptable forms in continuous speech samples. You will need to add, to the lists presented, those word forms that may be unique to both the group and individual speech characteristics of your clients or study participants.

a. Possible Disregard Words and Catenatives

Some of the words in this category may be placed within angle brackets after they appear twice (see section on formatting questionable words). Regardless, the following transcription guidelines should be followed.

1) yeah [jɛ] [jæ] [jʌ]
   yep [jɛp] [jʌp]
   yes [jɛs] [jʌs]

Enter in the Y line whatever vowel the participant used in the Z line, provided it is reasonable. This exemplifies the principle mentioned previously: *Give the participant the benefit of the doubt.* However, a form such as [jus] (i.e., a high back stressed vowel) would not be considered a reasonable entry as the intended vowel for *yes* in the Y line.

2) aw [ɔ̃]
   hm [hm̩]
   ihi [ɪ̃ hɪ̃]

m	[m̩]	
mhm	[m̩hm̩]	positive
mm	[m̩m̩]	negative
nn	[n̩n̩]	negative
oh	[o͡ʊ]	
uhhuh	[ʌhʌ]	positive
uhuh	[ʌ̃ʌ̃]	negative

These are the X line spellings and Y and Z line phonetic transcriptions (the former might have been placed within angle brackets in the X line).

3) donno	[do͡ʊno]	[dəno͡ʊ]	[əno͡ʊ]
gimme	[gɪmɪ]	[gɪmi̍]	
gonna	[gɑnɑ]	[gʌnə]	[gʊnə]
gotta	[gɑtə]	[gɔtə]	
hadta	[hætə]	[hætʰə]	[hætˉə]
hafta	[hæftə]	[hæftˉə]	
kinda	[kɑ͡ɪnə]		
needta	[nidə]		
wanna	[wɔnə]	[wɑnə]	[wʌnə]
whata	[wʌtˉə]		

Again, under the benefit-of-the-doubt rule, the exact vowel heard and transcribed in the Z line is presumed to be the one intended and entered in the Y line. That is, it is an acceptable variant, one that a participant might say in casual speech, fast speech, or under special circumstances such as being tired or having a cold.

b. Dialectal Differences

All dialectal variants, as transcribed in Z, should also be entered as the intended forms in the Y line. Examples include [æks] for *ask*, and [pɪtʃɚ] for *picture*. The transcriber will need to be thoroughly familiar with the sound change rules of a participant's ethnic or regional dialect, particularly the vowel

and diphthong rules. Chapter 11 in the 5th edition of *Clinical Phonetics* (Shriberg et al., 2019) includes discussion and examples of regional and social dialects, and dialectal variations among those who speak English as a second language.

c. Casual and Fast Speech Forms

When people speak in a casual or fast register, speech is simplified in systematic and predictable ways: (1) Vowels and diphthongs are simplified such that they become more centralized, unstressed vowels, and (2) consonants, especially within clusters in unstressed words and syllables, may be deleted or assimilated to other consonants. Interestingly, casual speech does not include certain consonant distortions, such as lateralized /s/, [s̪]. Rather, it is more associated with substitutions and deletions. Try to observe these principles as you transcribe casual speech forms. The following are some casual speech forms that occur quite frequently in speech samples of both participants with normal speech and delayed speech.

1) Vowel substitutions: [ə] for any unstressed vowel, e.g. *has* [həz]; *I* [ə]; *I'll* [əl]

2) Vowel replacement: *for* [fɚ] or [fɚ]; *they're, there, their* [ðɚ] or [ðɚ]; *you're, your* [jɚ] or [jɚ]

3) n/ŋ substitution in an unstressed syllable, e.g. *singing* [sɪŋɪn]

4) Deletion of initial and final consonants in unstressed words or syllables. For any segment that is deleted in word initial or word final position, determine if the word or syllable is unstressed. If so, a casual/fast speech form might be entered in the Y line. This guideline is particularly applicable to final nasal clusters. For example, if *friend* [frɛn] is heard as a stressed word, it is not an acceptable casual speech form. In an unstressed environment, however, because it is considered a casual speech form, [frɛn] can be entered in both the Y and Z lines. Sentence-final word position is likely to be unstressed.

Table 3 below includes alphabetized lists of casual speech forms that we have judged to be acceptable (listed first) and unacceptable (listed second). For reference convenience, all words listed in this section (including the previous examples) are included in the two lists. The bases of these decisions include

our observations of normal participants of all ages and what has seemed phonetically natural in our own social speech. Notice that most reductions involve clusters (phoneme substitutions and deletions) and syllable deletions.

The forms considered in Table 3 are frequent in casual speech, or are dialectal variants or common child forms. The principle underlying entry of these forms in the Y line is that PEPPER looks only at a participant's output phonology. To give the participant (even a young child) the benefit of the doubt, you may assume that certain word forms have a reduced 'underlying form' for speech output. Caregivers are likely to use some of these forms when speaking to young children. Among the word forms judged to be unacceptable, the most common principle is that a singleton, cluster, or a syllable should not be reduced in a stressed context. You may need to add, delete, or rearrange words on these lists according to local dialect rules and other observations made in the clinical or research setting.

Table 3. Acceptable and unacceptable word forms.

WORD	ACCEPTABLE FORMS Transcription in Both Y and Z lines
about	[bɑ͡ʊt]
acts	[æks]
actually	[æk ʃəlɪ]
ask	[æks]
asked	[ækst]
ain't	[e͡ɪn]
airplane	[e͡ɪrople͡ɪn]
afraid	[fre͡ɪd]
almost	[ɔmo͡ʊs]
already	[ɔrɛdɪ]
alright	[ɔrɑ͡ɪt]
always	[ɔwe͡ɪz]
and	[ən] [n̩]
animal	[æmɪnəl] [ænən̩l]
another	[nʌðɚ]
anything	[ɛnɪθɪn]
around	[ɚɑ͡ʊn]
balloon(s)	[blun] [blunz]
bananas	[nænəz]
basketball	[bæskɪbɔl]
breakfast	[brɛkfəs] (regardless of stress)
computer	[pjutɚ]
couldn't	[kʊdn̩] [kʊnt]
crayons	[ke͡ɪənz] [kræns]
didn't	[dɪnt] [dɪn̩t] [ɪnt]
doesn't	[dʌzən]

don't	[dō͡ʊn] [ō͡ʊnt] (I don't know)
don't know ("donno")	[dō͡ʊno] [ənō͡ʊ] (I don't know)
except	[sɛpt]
find	[fɑ̄ɪn]
first	[fɝs]
for	[fɝ] [fɚ]
friend	[frɛn] (unstressed only; see text)
gimme	[gɪmɪ] [gɪmi]
gonna	[ɔnə] [gɑnɑ] [gʌnɑ] [gənə] [gʊnə]
goodbye	[gʊbɑ̄ɪ]
gotta	[gɑtə] [gɔtə]
grandma	[græmə] [græmɑ]
grandpa	[græmpə] [græmpɑ]
hadta	[hætə] [hæt̄ə] [hætʰə]
hafta	[hæftə] [ævtə] ([jə ævtə])
hand	[hæn]
hanger	[hē͡ɪŋgɚ] [hæŋgɚ]
has	[həz]
hasn't	[hæzən]
her	[ɝ] ("I'll feed her")
him	[ɪm]
his	[ɪz] ("on his arm," i.e. only when unstressed)
I	[ə]
I'll	[əl] [ɑl]
isn't	[ɪzən]
just	[ʤʌs] ("Just now")
kept	[kɛp]
kind	[kɑ̄ɪn]
kind of	[kɑ̄ɪndə] [kɑ̄ɪnə] ("kinda")
kindergarten	[kɪnɚgɑrtən]
last	[læs] ("last night")
library	[lɑ̄ɪbɛrɪ]
macaroni	[rō͡ʊnɪ]

McDonald's	[məkdɑnəlz]
milk	[mɛlk]
mirror	[mir] (dialectal)
months	[mʌnts] [mʌns]
must	[mʌs] ("Must be")
needta	[nidə]
new	[nʲu] (pedantic form)
night night	[nɑ͞ɪ nɑ͞ɪ]
Nintendo	[tɛndo͞ʊ]
of	[ə] ("piece of pie")
okay	[n̩ ke͞ɪ] [ñ̩ ke͞ɪ] [ke͞ɪ]
only	[o͞ʊni]
pajamas	[ʤɑməz] [ʤæməz]
picture	[pɪtʃɚ]
potatoes	[te͞ɪtoz]
pretty	[pɹtɪ]
probably	[prɑblɪ]
pumpkin	[pʌŋkən]
refrigerator	[frɪʤɚe͞ɪtɚ]
remember	[mɛmbɚ]
round and round	[ra͞ʊn n̩ ra͞ʊn]
sandbars	[sænbɑrz]
sandbox	[sænbɑks]
sandwich	[sæmwɪtʃ] [sænwɪtʃ]
second grade(r)	[sɛkən gre͞ɪd] [sɛkən gre͞ɪdɚ]
shouldn't	[ʃʊdən] [ʃʊn̩t]
softball	[sɔfbɔl]
something	[sʌm̩] (not [sʌʔm̩] because the glottal stop symbol is not allowed in the Y line)
supposed	[spo͞ʊst] [po͞ʊst]
that	[ðɛt] [ðət]
that	[ən næt] ("and that" said quickly without pausing between words)

the	[ən nə] ("and the" said quickly without pausing between words)
them	[əm] ("he got 'em")
then	[ən nɛn] ("and then" said quickly without pausing between words)
there	[ən nɛr] ("and there" said quickly without pausing between words)
they	[ən nēɪ] ("and they" said quickly without pausing between words)
this	[ən nɪs] ("and this" said quickly without pausing between words)
there, their, they're	[ðɝ] [ðɚ]
told	[tōʊl]
tomato	[mēɪdo] [mēɪt̄o]
towards	[tɔrdz] [tordz]
twenty	[twɛnɪ] [twʌnɪ]
usually	[juʒlɪ]
wanna	[wɑnə] [wʌnə]
wasn't	[wʌzən]
went	[wɛn] ("Went on it")
yes	[nɛs]
you	[jə] ("You have to")
your, you're	[jɝ] [jɚ]

WORD	UNACCEPTABLE FORMS Transcription in Z Line only
afraid	[frēɪd]
also	[ɔsōʊ]
always	[ɔwēɪ]
another	[nʌvɚ]

apartment	[əpɑrmɛnt]
around	[rɑ͡ʊnd]
August	[ɔgʌs]
birthday	[bɝde͡ɪ]
breakfast	[brɛfəst]
don't put	[do͡ʊn pʊt]
eleven	[lɛvən]
expensive	[pɛnsɪv]
except	[sɛp]
friend	[frɛn] (if stressed; see text)
fifth	[fɪf] [fɪθ]
football	[fʊbɔl]
hundred	[hʌnɚd]
it	[ɪ] (needs a stop)
next	[nɛks] (if not followed by a word that begins with /t/ or /d/)
piano	[pæno]
project	[prɑʤɛk]
sandwich	[sæmɪtʃ] (needs some lip rounding for /w/)
sixth	[sɪks]
soft	[sɔf]
that's	[ðæs] ("that's all")
the	[ə] ("all the time")
want	[wɑn]

5. Transcription and Formatting of Deletions

a. Alignment Rules

1) When a participant deletes speech sounds, the transcriber must follow two guidelines to align the missing segments: a) syllables within a word are preserved, and b) segments that remain in the Z line are matched by common place-manner features to their most likely correspondent in the Y line. Consider the following examples:

	1	2	3	4
X:	little	sleeping	standing	airplane
Y:	[lɪt̩]	[slipɪn]	[st‾ændɪŋ]	[e‾ɪrple‾ɪn]
Z:	[lɪ_o]	[t̠ _ipɪ_]	[_k‾ændɪ_]	[p͡e‾ɪ_p^{h}_ʌ_]

Example 1 above illustrates how syllables are maintained. Notice that a syllabic is placed under /l/ to give the subject the benefit of the doubt. That is, if the intended form [lɪtəl] was placed in Y, alignment requirements would force the participant's [lɪ_o] to be missing two segments [lɪ_o_] rather than one. Examples 2 and 3 present more difficult correspondence problems, which can be resolved using the two guidelines. Example 4 is the most complex example. Use of the tie symbol to add on the /p/ was explained in a previous section. Notice how alignment of vowel-diphthongs yields this solution. If the matching task makes the phonetic transcription too questionable, use the disregard brackets around the word in X and enter an appropriate square-bracketed comment. For example,

X: <telephone> [Y Z correspondence?]

Y: tɛləfon

Z: pʌ_____

2) Following are some additional alignment examples and rationale for each:

a) Child says [m̩] for the word *spoon:*

X: spoon The spoken nasal is aligned with the intended nasal.

Y: spun	(The word form change is from CCVC to VC)
Z: __əm	Using the [ə] rather than the syllabic diacritic gives the participant the benefit of the doubt, i.e. two omission errors instead of three.

b) Child says [tsēɪʃən] for *station:*

X: station	This transcription and formatting guideline follows from the benefit of the doubt rule. PEPPER
Y: stēɪʃən	will credit the child with two substitutions rather than a distortion and deletion, i.e. [ˢs_ēɪʃən].
Z: tsēɪʃən	Related issues are discussed in a following section about transcribing distortions.

c) Child says [pˉunz] for *spoon:*

X: spoon	This is the most conservative alignment. Whether the /z/ is intended as a plural morpheme or has
Y: spˉun	simply migrated to the end of the word (metathesis) cannot be determined from this one example of the
Z: _pˉun͡z	child's speech.

d) Child says [tʃɛs] for *chair:*

X: chair	An alternative in Z would have been to presume a plural intention with an /r/ deletion, i.e. [tʃɛ_s], but
Y: tʃɛr	that would be less 'conservative' on the error hierarchy.
Z: tʃɛs	

e) Child says [ənt͡əm] for *thank you:*

X: thank you	The three-consonant medial cluster is aligned and the added final consonant is tied to the final vowel.
Y: θēɪŋkju	
Z: _ənt_ə͡m	

f) Child says [m̩hm̩z] for *balloons:*

X: balloons
Y: bəlunz
Z: m̩_həmz

The four consonants are aligned, and the second syllabic nasal is written in VC form so that one vowel deletion is counted instead of two (benefit-of-the-doubt rule).

g) Child says [ɛlfnənt] for *elephant:*

X: elephant
Y: ɛləfənt
Z: ɛl͡f_nənt

Use of the tie here will capture the syllable deletion and allow the best sound correspondence. The n/f substitution that results will probably represent a unique event in the transcript, and be discernible as a lexically specific 'misarticulation.'

h) Child says [a͡ɪunmə] for *vacuum:*

X: vacuum
Y: vækjum
Z: _a͡ɪ__un͡mᵊ

In order to align the vowels in the two syllables, a tie is used to add the nasal, and the schwa is treated as an offglide.

i) Child says [ka͡ʊ] for *car:*

X: car
Y: kɑr
Z: kɑʊ

Although this sounds like a consonant plus diphthong, [ka͡ʊ], we assume the second vowel replaces the /r/ (/r/ and /ʊ/ are both sonorants), so it is transcribed as a substitution, not a deletion of /r/.

j) Child says [bɛə] for *bed:*

X: bed	As in the example above, rather than assume the less conservative /d/ deletion, [bɛᵊ_], the alignment assumes a schwa replacement for a consonant.
Y: bɛd	
Z: bɛə	

k) Child says [kɝfɪs] for *cursive:*

X: cursive [metathesis]	This nice example of metathesis should be noted in a comment.
Y: kɝsɪv	
Z: kɝfɪs	

In these examples, arbitrary conventions are used to determine alignment, rather than using criteria such as acoustic similarities between spoken forms and original targets or looking for patterns in the participant's speech. Such coding decisions, though based on a logic that yields reliable decisions, may be viewed as inappropriate by linguists who favor more analytic, data-driven approaches. This issue is central to a number of questions about linguistic analysis.

b. Deletions versus Other Error Types

Deletion errors, especially in word-final position and those involving nasals, are among the least reliable error classes in phonetic transcription. Specific transcription guidelines and practice are presented in *Clinical Phonetics* (2003, 2013, 2019). Essentially, you should question whether the participant is deleting a sound, replacing it with a glottal stop, or saying the sound weakly, for example, [s_ænd] rather than [st̚ænd] for *stand.* In word-final position, listening to whether the vowel seems to end abruptly or trails off can help in your judgment of whether a weak sound follows (abrupt) or not (trails off). It is always more appropriate to circle your best decision when unsure than to arbitrarily assign the most conservative transcription. Instrumentally-aided transcription allows decisions about perceptually ambiguous apparent speech-sound deletions to be made with more confidence.

c. Post Vocalic /r/ Deletions

A deletion error will appear to occur when three-segment words ending in /r/,

such as *here* [hɪr], are reduced to a two segment word, [hɝ]. Or sometimes a two segment word will be reduced to [ɝ], as in *or* [ɝ] and *are* [ɝ]. To avoid a segment correspondence error, add a consonant /r/ in the Z line:

X: here

Y: hɪr

Z: hɝr

Casual forms, such as *your* [jɝ] and *there* [ðɝ], can simply be entered in the Y line as two-segment forms (see previous discussion). Dialectal variants, such as [kɑː] for *car* (Southern, Bostonian), should be entered as such in both the Y and Z lines, that is, the intended form is [kɑː].

6. Transcription and Formatting of Substitutions

a. General Principles

Transcriptions reflecting speech sound deletions (discussed above) and speech sound substitutions make claims about a person's phonology. Therefore, to invoke the benefit-of-the-doubt rule, it is prudent whenever practicable to transcribe sound changes as distortions rather than as phoneme substitutions. Do not transcribe w/r, for example, unless the assumed replacement really sounds like a /w/, rather than some form of a distorted (i.e. derhotacized) /r/. See if your auditory percept can be represented by using some diacritical modification(s) of the intended sound, rather than by using another sound as the main character.

b. Guidelines/Examples

1) If you think you perceive a sound that is 'somewhere between' the target sound and some other phoneme, always modify the target sound to match your percept, rather than modify another phoneme. Examples:

X: shoe	X: easy	X: had
Y: ʃu	Y: izi	Y: hæd
Z: ʃ̟u (not [s̠u])	Z: iz̬i (not [is̬i])	Z: hæ̞d (not [hɛ̞d])

[NOTE: In the examples above, the diacritics are not necessarily errors, and the transcriber may want to enter them in the Y line too.]

2) Several distortion errors can be symbolized using alternative main characters: [l̴] and [r̴], [s̠] and [ʃ̠], [s̻] and [ʃ̻]. These pairs capture essentially the same percept. Avoid transcribing such sounds as substitutions; use the main character of the intended sound. For example, a lateralized fricative in *soup* would be transcribed as [s̠up] not [ʃ̠up].

3) When a sound substitution does best fit your perception of the error be sure to include any modifying diacritics in the Z line. For example,

X: dog

Y: dɑg

Z: t̪ɑg

4) Children often say a "ts" like sound for the /st/ cluster as in *stove* and for the /tʃ/ phoneme as in *chair.* Should this replacement be treated as two phonemes (/t/ + /s/, [t͡s]), as a single phoneme with an onglide ([ᵗs]), or as an offglide ([tˢ])? The basic principles allow for an arbitrary resolution: transcribe the "ts" as substitutions when they replace a cluster ([tsoʊ̄v] *stove*) and as a distortion when they replace a singleton ([ᵗseɪ̄r] *chair).* Each transcription reflects the most conservative claim: a deletion is avoided ([ts] rather than [ᵗs_]) in *stove*, and an addition is avoided in *chair* ([ᵗseɪ̄r] rather than [t͡seɪ̄r]).

One exception to this occurs when children produce a "tsh" like sound for an /s/ + stop cluster. While in the previous example the sound was "ts"-like, here it sounds like the affricate, [tʃ]. In this situation, it seems more appropriate to assume that the affricate replaces /s/. Hence, the following transcriptions are entered in the Z line:

X: school X: steps X: skip

Y: skul	Y: stɛps	Y: skɪp
Z: tʃ_ul	Z: tʃ_ɛps	Z: tʃ_ɪp

7. Transcription and Formatting of Distortions and Additions

(For PEPPER users who make extensive use of narrow phonetic transcription)

a. Definitions

This category of speech sound errors consists of two related but different error types. A speech sound *distortion* occurs when all or a portion of the sound is misarticulated, or not said precisely. An *addition* occurs when a correct or distorted sound begins (onglide) or ends (offglide) with another sound. In both cases, the main character of the intended phoneme is retained in the Z line and modified by characters, including onglides and/or offglides if necessary.

Principled formatting conventions for the infinite variety of distortions/additions that occur in normal and disordered speech are difficult to derive. The following guidelines represent our best efforts in transcribing errors in this category.

b. Additions to Shriberg and Kent (2013)

The following suggestions capture speech sound distortions that have been observed clinically in speech delayed children and were not specifically discussed in Shriberg and Kent (2013). As presented at the end of this section (see Table 5), only some of these phonetic behaviors on the list of distortions are considered common clinical distortions (Shriberg, 1993).

Table 4. Additional transcription symbolization.

Sound Class		Suggested Transcription Description and Approximate Frequency
Glides	[w̞]	A rhotacized /w/. Infrequent.
	[m̥w]	An unrounded /w/. Quite frequent.
	[ᵛw], [wᵛ]	A /v/-like addition. Somewhat frequent.
	[w̪]	A /v/-like sound as a single segment. Infrequent.
Stops	[˯]	Weakly articulated stop (any stop). This symbol was originally used for trill (Shriberg & Kent, 2003), but has been redefined. Very frequent.

	[̪]	A /ⱱ/-like bilabial. Infrequent.
	[ˑ]	A delayed-release stop, long closure time. Infrequent.
	[t̝]	An alveolar flap (see discussion on pp. 44-45). Very frequent.
Fricatives & Affricates	[s̪˭]	An inspirated /s/-like sound. Very infrequent.
	[f̪˭]	Inverted lower lip is pulled back under front teeth. Infrequent.
	[k̟ˣ]	A velarized fricative. Infrequent.
	[k̠ˤ]	A pharyngealized fricative. Infrequent.
	[h̝]	A pharyngealized fricative, more backed than [k̠ˤ]. Infrequent.
Liquids	[l̴]	An 'overarticulated' /l/. Infrequent.
	[l̪]	A palatalized /l/. Infrequent.
	[l̪ʷ]	A labialized /l/. Very frequent in prevocalic and intervocalic position.
	[r̪ʷ]	A derhotacized /r/ with /w/ coloring. Frequent.
	[r̝], [ɜ˕]	A trilled /r/ sound. Prefer to use a square-bracketed comment, but the fronted symbol is another option. Infrequent.
Vowels	[ɛ̽]	A centralized vowel, more stressed than [ə] and higher than [ʌ] or [ɑ̽]. Somewhat frequent.
	[ɛ˞]	An /r/-like vowel, slightly rhotacized. Infrequent.

The entries in the above table should suggest ways in which the diacritic symbols in Figure 7 can be used to capture other distortion/addition changes in special populations. Again, these suggestions only supplement the materials and discussion in Shriberg and Kent (2013).

c. Tables of Non-Error and Error Distortions/Additions

Table 5 below sorts distortions and additions into two categories, non-errors

and errors. The non-error entries meet one or more of the following three criteria:

1) They symbolize optional allophones in casual speech forms, such as an unreleased stop [p̚];

2) They are difficult to transcribe reliably (Hoffmann, 1982; Shriberg & Kent, 2003; Shriberg & Kwiatkowski, 1980); or

3) They presumably would not be perceived by the lay person as articulation errors, as *bat* [b̃æt].

In contrast, entries in the error column

1) symbolize non-optional allophones,

2) are more reliably transcribed using only perceptual phonetics, and/or

3) are likely to be included in the set of socially relevant articulation errors for which speech management services are provided.

Entries in these two lists have emerged from studies of normal speech sound acquisition (Hoffmann, 1982), speech delay (Shriberg & Kwiatkowski, 1982b; 1986), and associated literature in normal and delayed speech acquisition. The table is organized first by rules affecting all sounds, then rules affecting only each manner class, and finally, rules affecting only specific diacritics.

Table 5. Guidelines for determining speech errors vs. non-errors for PepAnalyses outputs.

Target Sound Class	Non-Error Enter in both Y & Z lines	Error Enter only in Z line
Rules Affecting All Sounds		
	If a sound begins incorrectly, but ends correctly, enter in both Y and Z and note, in a square-bracketed comment, the distortion or substitution error. For example, X: red [derhot to correct] Y: ˠrɛd Z: ˠrɛd (Recall that diacritics cannot be attached to on-glides.)	If a sound begins correctly, but ends incorrectly, enter the incorrect form in Z only and note in a square-bracketed comment. For example, X: red [correct to derhot] Y: rɛd Z: ˠr͡ɛd
	If a participant revises a segment or word from incorrect to correct, enter the incorrect form in angle brackets and the revised correct sound as the target word and comment: X: <r> red <tar> car [two corr rev] Y: r rɛd tar kar Z: r͡ rɛd tar kar	If the participant revises a segment or word from correct to incorrect, enter the correct form in angle brackets and the revised form as the target word and comment: X: <r> red <car> tar [two incorr rev] Y: r rɛd kar tar Z: r r͡ɛd kar tar
Rules By Manner Class		
Nasals	[̃] denasalized [nᵈ] epenthetic stop (off-glide); also [nᵗ], [mᵇ], [mᵖ], [ŋᵍ], [ŋᵏ]	

[̬] weak

Glides

[̬] weak

[w̢] r-like, retroflexed /w/

[w̫] labialized /w/

[w͡] nonlabialized /w/

[ʰw] wh-like sound for /w/

[ᵛw] or [wᵛ] /v/-like sound as onglide or offglide

[w̪] /v/-like sound as a single segment

Stops

[ʰ] aspirated

[̪] dentalized bilabials

[̪] dentalized alveolars

[͊] nasal emission

[̽] frictionalized

[̃] nasalized, e.g. [d̃]

[̥] partially devoiced, in word-final position

[̥] partially devoiced, in word-medial position

[⁼] unaspirated when aspiration is *not* obligatory

[̥] partially devoiced, in word-initial position

[˺] unreleased

[⁼] unaspirated when aspiration is obligatory

[̬] weak

[͈] lateralized

[˂] fronted

[˃] backed

Fricatives/Affricates

* palatalized alveolar: standard, e.g., [sʲ], [zʲ]

[̪] dentalized sibilants

[˂] fronted, e.g., [s˂], [z˂], [ʃ˂], [ʒ˂], [tʃ˂], [ʤ˂]

[͈] lateralized sibilants

[˃] backed, e.g., [s˃], [z˃], [ʃ˃], [ʒ˃], [tʃ˃], [ʤ˃]

[͊] nasal emission

*[̥] partially devoiced:

* palatalized alveolar: notable,

	standard	e.g., [s̻], [z̻]
	[̬] weak	* [̥] partially devoiced: notable
	[̼] whistled	[̺] retroflexed sibilant, fricatives and affricates
		[ħ˂] pharyngealized fricative
		[f̽] inverted lower lip under front teeth
Liquids	[l˂] overarticulated /l/, as if tongue tip is too far forward	[r͡] derhotacized /r/
	[̬] weak	[l̽] labialized /l/
	[l̰] velarized postvocalic /l/ medially and finally	[l̻] palatalized /l/
		[l̰] velarized prevocalic /l/
		[l̃] nasalized /l/
Vowels	[̤] breathy or whispered	[͡] derhotacized /ɝ/, /ɚ/
	[ˑ] glottalized	[˞] r-like vowel
	[ː] lengthened	*[̃] nasalized: notable
	*[̃] nasalized anywhere	*[˔] raised
	[̹] rounded	*[˕] lowered
	[˃] shortened	*[˖] advanced
	[̜] unrounded	*[˗] retracted
		*[̽] centralized

Rules for Diacritics

[ʰ] aspiration after final consonant, e.g., [tʌbʰ]	*[ə] notable on-/offglide anywhere including within a cluster, e.g., *[bəlu]
[ː] describes increased duration of a stop closure or of a continuant sound	[ⁿ] [ᵐ] nasal consonant on-glides, any position
*[ə] slight on-/offglide within a cluster	[ʰ] [ˉ] to note whether aspiration feature is correct for the stop in consonant clusters (the non-error voicing diacritics [ᵛ] and [ₒ] are used to describe differences in all other situations)
*[ə] slight on-/offglide as if articulating carefully, as a residual from speech therapy, e.g., [kʌpə], or due to stuffy nose	*[ᵖ] [ᵇ], [ᵗ] [ᵈ], [ᵏ] [ᵍ] stops; typically follow word-final nasals. Do not sound like true clusters, but more pronounced than mechanical lip or tongue clicks.
[ᵛ] partial voicing, any word position	
[ʰ] when optional, aspirated release, e.g., [kʌpʰ+kēīk] as opposed to [kʌpkēīk] (juncture used to differentiate these two transcriptions)	
*[ᵖ] [ᵇ], [ᵗ] [ᵈ], [ᵏ] [ᵍ] stop-like clicks, weak, brief--sound like lips or tongue opening/removal gesture typically follow word-final nasals (i.e., homorganic nasals). When participant has a cold all stop on-/off-glides are non-error diacritics.	

* Entries marked with an asterisk involve decisions based on relative severity, i.e., the

degree to which the distortion/addition is socially notable. Because the reliabilities for these judgments are not high, the same symbols are used for *standard* and *notable*, but notables are not entered in the Y line (i.e., are considered errors).

For certain clinical and/or research questions, an alternative to the system presented in the table above is to count only common clinical distortions as errors and all other distortions (i.e., uncommon clinical distortions) as errors. See the Appendix in Shriberg (1993) for a list of nonclinical and clinical distortions (common and uncommon).

Table 6. Additional response definitions used in the reference data.

Target	Perceived	Description	Non-Error: Enter	Substitution
Segment	Sound		in Both Y/Z Lines	Enter Only in Z Line
[n]	[n̟]	sounds primarily n-like	X	
[n]	[n̠]	sounds primarily ŋ-like		X
[ŋ]	[ŋ̠]	sounds primarily ŋ-like	X	
[ŋ]	[n̟]	sounds primarily n-like		X
[t]	[t̟]	sounds primarily t-like	X	
[t]	[k̠]	sounds primarily k-like		X
[k]	[k̠]	sounds primarily k-like	X	
[k]	[t̟]	sounds primarily t-like		X
[ʃ]	[ʃ̠]	sounds almost like /s/, but retains [ʃ] quality	X	
[ʃ]	[ʃ̠]	sounds lateralized in addition to above		X
[ʃ]	[ʃ̟]	sounds clearly ʃ-like, but broad and diffuse	X	

* Application of Backing [̟] and Fronting [̠] diacritics for phoneme substitutions and non-error distortions (note the decision logic underlying these 11 conventions).

Special Considerations in Transcription

Madison Speech Assessment Protocol Tasks

The imitation tasks that comprise the computerized portion of the Madison Speech Assessment Protocol (MSAP) are transcribed using the transcription guidelines and formatting conventions described in this guide, with a few modifications:

- The transcriber uses a preprinted pepform, with the X and Y lines already completed, to record transcriptions in the Z lines. The preprinted pepform is generated from a protocol PepFile that has been entered and saved in PEPPER. Protocol PepFiles for the following 12 MSAP tasks are available through PEPPER (see the "MSAP" section in the PEPPER Help Index for links to the PepFiles): LST, CWT, VT1, VT2, VT3, SRT, NRT, EST, RST, MWT1, MWT2, and SPT. Transcriptions may sometimes require finessing to conform to the preprinted protocol. If acoustic analyses will be done on the MSAP task(s), the transcriber must adhere to the glosses and intended forms as much as possible, even when the participant's intended response is clearly different from the stimulus. A square-bracketed comment can be used to indicate that the participant may have misunderstood the stimulus and/or imitated it incorrectly.
- Carrier phrase words on the LST and RST are not transcribed. They are not included in the X lines of the MSAP protocol PepFiles.
- If a participant provides more than one response for a given stimulus, the first response is transcribed whenever possible *unless acoustic analyses will be completed.* If acoustic analyses will be done, do not transcribe the first response if it is unlikely that it can be analyzed acoustically (because of over talk, environmental noise, signal distortion, etc.). If the first response is not usable, the next usable subsequent response is transcribed. The transcriber indicates in a square-bracketed comment which response she transcribed, e.g., [second attempt].
- If a stimulus item was not administered or is not on the recording, a square-bracketed comment is entered in the X line ([NT], [not on the recording], etc.), and a single asterisk for each word is entered in the Y and Z lines, regardless of how many syllables are in each word. For multi-word stimuli (like those in the VT3, EST, and SPT), assign one asterisk per word.

To print pepforms from protocol files:

1) Download the protocol PepFiles by clicking on the links in the MSAP section of the PEPPER Help Index and save on your computer.
2) In PEPPER, go to *File>Open* and choose the protocol PepFile for the task of interest (CWT, NRT, etc.).
3) With the PepFile open, go to the PepFunctions drop-down menu, choose PepTools, and select "Generate PepForm from X Lines."
4) Print the pepform that is generated.

5) Close the generated pepform and the PepFile when finished.

Transcription and Prosody-Voice Coding Guidelines for Unique Populations

When speech samples are collected from participants with dialects that differ from General American English, it is important to establish guidelines for transcribing dialectal speech differences (accents). Some speech, prosody, and voice characteristics that would typically be transcribed or coded for prosody-voice (Shriberg, Kwiatkowski, & Rasmussen, 1990) as unacceptable or inappropriate when compared to General American English dialect may be acceptable/appropriate for a particular participant's dialect. See Chapter 11 in *Clinical Phonetics* (Shriberg et al., 2019) for discussion and examples of dialects that may be encountered by clinicians and researchers in the United States. In general, it is important for a transcriber to be familiar with the dialectal rules and characteristics of the participant whose speech is being transcribed, and transcription and prosody-voice coding decisions are made based on knowledge and awareness of these differences. Figure 9 is an example of a form that was created for use in one study that included participants from several states in the southwestern U.S. A form like this may be useful in documenting the transcriber's decision-making process and can later be useful in interpreting PepAnalyses output(s).

For participants with motor speech disorders (MSD) who have relatively more severe speech and prosody-voice involvement (with or without cognitive, affective, and/or behavioral deficits), transcription and prosody-voice coding can be very challenging. In our research, our transcribers are purposely not provided with diagnoses so as not to be biased during data reduction. After reviewing the literature and our current system for narrow phonetic transcription and prosody-voice coding of conversational speech samples, we are confident that most if not all speech and prosody-voice behaviors/characteristics associated with MSD can be captured using narrow phonetic transcription and the prosody-voice codes (see *PVSP*) currently available.

References

Davis, E. (1937). The development of linguistic skill in twins, singletons with siblings, and only children from age five to ten years. *University of Minnesota institute of child welfare monograph,* (No. 14), University of Minnesota Press, Minneapolis.

Hoffmann, K. (1982). *Speech sound acquisition and natural process occurrence in the continuous speech of three- to six-year-old children*. Unpublished Master's thesis, University of Wisconsin-Madison.

Kenyon, J.S., & Knott, T.A. (1953). *A Pronouncing Dictionary of American English.* (2nd ed.). Springfield, Mass.: Merriam-Webster.

Shriberg, L.D. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Madison, WI: Software Development and Distribution Center, University of Wisconsin-Madison.

Shriberg, L.D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36*, 105–140.

Shriberg, L.D., & Kent, R.D. (1982). *Clinical Phonetics*. New York: Macmillan.

Shriberg, L.D., & Kent, R.D. (1995). *Clinical Phonetics* (2nd ed.). Boston, MA: Allyn & Bacon.

Shriberg, L.D., & Kent, R.D. (2003). *Clinical Phonetics* (3rd ed.). Boston, MA: Allyn & Bacon.

Shriberg, L.D., & Kent, R.D. (2013). *Clinical Phonetics* (4th ed.). Boston, MA: Pearson Education.

Shriberg, L.D., Kent, R.D., McAllister, T., & Preston, J.L. (2019). *Clinical Phonetics.* (5th ed.). Boston, MA: Pearson Education.

Shriberg, L.D., & Kwiatkowski, J. (1980). *Natural Process Analysis: A procedure for phonological analysis of continuous speech samples*. New York: Macmillan.

Shriberg, L.D., & Kwiatkowski, J. (1982). Phonological disorders II: A conceptual framework for management. *Journal of Speech and Hearing Disorders, 47*, 242–256.

Shriberg, L.D., & Kwiatkowski, J. (1983). Computer-assisted Natural Process Analysis (NPA): Recent issues and data. In J. Locke (Ed.), *Assessing and*

treating phonological disorders: Current approaches. Seminars in Speech and Language, 4, 389–406.

Shriberg, L.D., & Kwiatkowski, J. (1985). Continuous speech sampling for phonologic analyses of speech-delayed children. *Journal of Speech and Hearing Disorders, 50*, 323–334.

Shriberg, L.D., & Kwiatkowski, J. (1986, November). *A follow-up study of 36 children with developmental phonologic disorders.* Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Detroit, MI.

Shriberg, L.D., Kwiatkowski, J., & Hoffmann, K.A. (1984). A procedure for phonetic transcription by consensus. *Journal of Speech and Hearing Research, 27*, 456–465.

Shriberg, L.D., Kwiatkowski, J., & Rasmussen, C. (1990). *The Prosody-Voice Screening Profile*. Originally published by Communication Skill Builders, Tucson, AZ; rights returned to The Phonology Project, University of Wisconsin-Madison, Madison, WI.

Templin, M. (1957). *Certain language skills in children*. Minneapolis: University of Minnesota Press.

Wells, J.C. (2008). *Longman Pronunciation Dictionary* (3rd ed.). Harlow: Pearson Education Limited.

PHONOLOGY PROJECT DIALECT FORM

PepFile Name: _____ Peplog #: _____

Study Name: _____ Transcriber Name: _____

Date: _____

Does this study participant appear to have a dialect other than General American English?

Yes ☐ No ☐ *If yes, complete the rest of this form.*

Have you allowed (i.e., not corrected in the Y line) any or all dialectal differences on vowels?

Yes ☐ No ☐

List examples of vowel differences you allowed (examples: i/ɪ, ɔ/ɑ, raised /ɪ/):

_____

_____

_____

List examples of vowel differences you did not allow:

_____

_____

_____

Have you allowed (i.e., not corrected in the Y line) any or all dialectal differences on consonants?

Yes ☐ No ☐

List examples of consonant differences you allowed (examples: f/θ, ks/sk, d/ð):_____

_____

List examples of consonant differences you did not allow:

_____

_____

_____

Figure 9. The Phonology Project Dialect Form.