PEPPER GUIDE 8:

PEPPER UTILITIES

Overview

This guide provides information for the PEPPER tabs labeled PepAgree, PepStat, and PepGraph. PepAgree software computes agreement percentages for phonetic transcription and prosody-voice coding. Point-to-point agreement percentages can be obtained for 2 transcribers/coders (interjudge agreement) or for 1 transcriber/coder who has reduced the same speech and prosody-voice data at 2 points in time (intrajudge agreement). PepStat allows calculation of various statistics of data from a table of values. PepGraph allows plotting of data from a table of values. Such tables are usually generated by retrieving values from PepStore (see PepStore under the PepFunctions tab), though they can also come from other spreadsheet, statistical analysis, or database programs.

PepAgree

Overview

The PepAgree utility calculates agreement percentages and provides detailed comparisons for broad and narrow phonetic transcription and, for conversational samples (CS), prosody-voice coding decisions obtained from the *Prosody-Voice Screening Profile* (PVSP; Shriberg, Kwiatkowski, & Rasmussen, 1990). It has potential applications in academic settings (e.g., to assess transcription competency and progress in students taking a phonetics course); clinical settings (e.g., to support the intrajudge and interjudge stability of clinical assessment or treatment findings); and research settings (e.g., to obtain agreement percentages for perceptual data reduction methods used in a study; see Table 3 in Baylis & Shriberg, 2019).

The PepAgree menu consists of 4 options:

- *Agreement Analysis...* produces broad and narrow phonetic transcription agreement calculations and data for consonants, vowels, and diacritics. (Each pair of transcripts that are compared must have the same number of utterances and words to be included in the *Agreement Analysis*.)
- *Gloss Agreement* compares X lines for agreement on text and type (disregard, either/or, unsure, unintelligible, and non-questionable).
- *PVSP Agreement* compares Prosody-Voice Screening Profile (PVSP) coding data for agreement on exclusion codes and, for coded utterances, agreement on judgments of appropriate vs. inappropriate and agreement on inappropriate codes.
- *X & Y Lines Agreement* compares the X lines and Y lines between each PepFile transcript pair and identifies any disagreements.

PepAgree allows PEPPER users to obtain agreement percentages on pairs of PepFile transcripts with or without Prosody-Voice Screening Profile (PVSP) data. PepAgree procedures include the following steps:

1. Transcribe one or more speech samples using the formatting conventions described in PEPPER Guide 2. Optionally, complete PVSP coding on one or more conversational samples (CS).
2. Create PepFile(s) by entering and saving the transcript(s) using procedures described in PEPPER Guide 3 (PG3).
3. Enter PepStore variables and PVSP data, if available, using procedures described in PEPPER Guide 4 (PG4).
4. Complete Steps 1 to 3 a second time for the same speech sample(s). If interjudge agreement calculations are of interest, the transcription (and optionally PVSP coding) is completed by a second transcriber. If intrajudge agreement calculations are of interest, the transcription (and optionally PVSP coding) is completed by the same transcriber after several weeks to several months (or longer) have passed following the first transcription.
5. Create and save a PepFile List following the procedures described in this guide.
6. Run PepAgree analyses.
7. Interpret results.

Prepare Transcripts/PepFiles and PVSP Materials for PepAgree

Following is a description of the procedures used in the Phonology Project for preparing CS transcripts/PepFiles for PepAgree analyses (for both phonetic transcription and PVSP coding). These procedures may require some modification, depending upon the goals of the PEPPER user. Instructions in this section are specific to Step 4 in the previous section; in other words, the assumption is that Steps 1-3 for the Time 1 transcript have already been completed and that the user is familiar with the content in PEPPER Guides 1 through 4 and with the PVSP.

Generate a Pepform From an Existing PepFile

To save time and insure that the X line glosses for the Time 1 and Time 2 transcript pairs match as closely as possible, generate a pepform using the Time 1 PepFile transcript that has already been entered in PEPPER and saved. This pepform, which contains the X line glosses, is printed and used by the second transcriber (or the first transcriber a second time) to record their transcriptions in the Z line. To generate and print a pepform, do the following:

1) In PEPPER, go to *File>Open* and choose the PepFile of interest.
2) With the PepFile open, go to the PepFunctions drop-down menu, choose PepTools, and select "Generate PepForm from X Lines."
3) Print the pepform that is generated.

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

If the pepforms are generated from CS PepFiles with stored PVSP data, the utterance number for each PVSP-codable utterance appears in small type at the beginning of each utterance in the X line, beginning with U1. This insures that the same utterances are identified and coded for prosody and voice in both the Time 1 and Time 2 transcripts for a given sample. (If an utterance does not have a printed codable utterance number, it means that it was assigned one or more exclusion codes and was not judged for prosody and voice.)

If agreement analyses will be completed on PVSP data, and the highest codable utterance number labeled on the pepform is lower than the total number of utterances coded for prosody-voice (typically 24 utterances), it is necessary to check the hard copy of the Time 1 transcript for glossed utterances beyond the final utterance entered in the Time 1 transcript. For example, if the first transcriber transcribed the sample up to the utterance containing the 100th first-occurrence word, but hadn't yet reached 24 PVSP-codable utterances, she would have continued to gloss additional utterances in order to reach the required 24 PVSP-codable utterances. These additional glossed utterances, because they weren't transcribed, weren't entered in PEPPER, and so will not appear on the pepform generated from the PepFile. If this is the case, the transcriber completing the PVSP coding for Time 2 must be given the additional glossed utterances, with notation indicating which of the utterances were coded for prosody and voice.

The printed pepform (plus additional glossed utterances, if applicable) is given to the transcriber who is completing the second transcription. If prosody-voice coding must also be completed, the transcriber is given a blank PVSP Scoring Form.

Complete Second Transcription and Prosody-Voice Coding for PepAgree

If PepAgree analyses will be run on PVSP data, the transcriber first completes the PVSP Scoring Form, including the Sequential Utterance Log with exclusion codes and the Prosody-Voice Coding Log, including all appropriate and inappropriate codes/judgments in the 7 prosody and voice domains. As mentioned previously, the PVSP-codable utterance number appears at the beginning of each codable utterance in the X lines of the preprinted pepforms; each utterance without a codable utterance number indicates that the utterance was assigned one or more exclusion codes during the first prosody-voice coding session. These numbers, or lack thereof, should be used as a guide to insure that the same utterances are coded and excluded in each transcript pair.

After prosody-voice coding is completed, the transcriber then records phonetic transcription in the Z lines on the preprinted pepform(s), using the X-line glosses as a guide. (Entering Y-line "intended" target sounds is optional.) There may be instances when the second transcriber doesn't agree with the gloss from the Time 1 transcript. The transcriber can cross out and replace the disagreed-upon word(s) in the X line and complete the Y and Z lines based on the modified gloss. (If the number of segments in

the Y line changes as a result of the modified gloss, the word will be skipped during the PepAgree analysis.)

Enter and Save Second Transcripts/PepFiles for PepAgree

Enter and Save the Time 2 Transcripts/Pepfiles

Entering the Time 2 transcripts for PepAgree can be accomplished in one of several ways:

1. Create a protocol file from the Time 1 PepFile transcript and use the protocol to enter/edit the Time 2 transcript. Information on creating and editing protocol PepFiles can be found in PG3.
2. Open the Time 1 PepFile transcript and choose *Save As...* from the file menu to save a copy of the PepFile that can be edited based on the transcription and PVSP coding decisions from the Time 2 transcript.
3. Enter the Time 2 transcription (and PVSP data if applicable) as a new PepFile; including relevant PepStore values.

Clearly, procedures 1 and 2 are more efficient for keyboarding than procedure 3 and are likely to maximize the number of words that are eligible for PepAgree analyses. Procedure 3, while more time-consuming, can reveal differences in glossing and/or transcription procedures that may need to be addressed to improve overall agreement between/among transcribers.

While entering/editing the Time 2 transcript, it's helpful to keep in mind that when the PepAgree utility is run, PEPPER compares X lines in the Time 1 and Time 2 PepFile pairs to make sure that the number of words in each utterance agrees. Even though the number of words in each utterance must match, the words themselves don't have to match (i.e., the glossed words can be different). PEPPER also compares the number of segments in each Y-line word in the Time 1 and Time 2 PepFile pairs. If the number of segments for a given word in the Time 1 and Time 2 transcript pair doesn't match, the word is skipped in the agreement analysis. After constructing a PepFile list that contains the 2 matched groups of transcripts, the *X & Y Lines Agreement* analysis can be run to get a list of X and Y line disagreements/mismatches; this allows inspection of words that might be skipped in the agreement analysis.

Suggestions for PepFile Names for PepAgree Transcripts

Although there are no specific PepFile naming requirements for PepAgree analyses, this section describes the PepFile naming procedures used in the Phonology Project. PepFile names can be constructed to suit the specific needs of the PepAgree user.

PG3 includes a section on general PepFile naming tips. It is helpful if the names for PepFiles used in PepAgree analyses include information that help identify them as Time 1 or Time 2 transcripts, and whether they are interjudge or intrajudge agreement transcripts. For example, 1234-c1_JM1 and 1234-c1_JM2 are examples of PepFile names that can be used to identify the transcripts as conversational samples that will be used for intrajudge agreement analyses for transcriber/coder "JM." The "1" and "2" after the transcriber's initials represent the Time 1 and Time 2 transcriptions, respectively. For interjudge agreement analyses, the PepFile name for the Time 1 transcript could again be 1234-c1_JM1, and the PepFile name for the Time 2 transcript could be 1234-c1_SH1. Here, the Time 2 transcript has been completed by another transcriber "SH," and her transcription (and optionally PVSP coding) will be compared to JM's. Because this is SH's first transcription/coding of this sample, a "1" appears after her initials instead of a "2."

How to Edit Functionally Equivalent Transcriptions

Some phonetic characters are considered functionally equivalent; even though the symbols themselves are different, they represent the same speech sound(s):

- /ʌ/ and /ə/
- /ɝ/ and /ɚ/
- /ən/ and /n̩/ (i.e., /ə/ + consonant vs. a syllabic consonant)

When entering/editing transcripts using procedure 2 described in the previous section, you will be able to see the Y and Z lines for Transcriber 1/Time 1. When editing the PepFile for Transcriber 2/Time 2, enter or edit the functionally equivalent phonemes so that they're the same. For example, if Transcriber 1 transcribed the word *her* as /hɚ/, and Transcriber 2 transcribed the same token of *her* as /hɝ/, do not change the /ɚ/ to /ɝ/ in the Time 2 transcript.

Enter/Edit and Save PepStore Variables and PVSP Data

Information on entering and editing PepStore variables and PVSP data is covered in PG4. If the Time 2 PepFile is an edited version of the Time 1 PepFile (i.e., it was entered using procedure 1 or 2 described on p. 4), the entered variables and PVSP data from the Time 1 PepFile are preserved and can be edited as needed. If the transcriber/coder for Time 1 and Time 2 are not the same person, variables 40 (transcriber's last name) and 41 (PVSP coder's last name) will need to be changed in the Time 2 PepFile (*PepFunctions* > *PepStore* > *Enter/Edit/View PepStore*). The PVSP data in the Time 2 PepFile will need to be edited to reflect any differences between the Time 1 and Time 2 PVSP exclusion codes and prosody-voice codes (*PepFunctions* > *PepStore* > *Enter/Edit/View PVSP Log*). Save all edited data before exiting the Time 2 PepFile or when prompted by PEPPER to do so.

Create a PepFile List for PepAgree

PepFile Lists, which can be either manually entered or generated by PEPPER, include one or more groups of PepFiles. The line for each PepFile in the list includes the drive/directory, folder(s)/subfolder(s), and the PepFile name (i.e., the path for the unique location where the PepFile is stored). Detailed procedures for constructing PepFile Lists (*.pfl*) are available in PEPPER Guide 6 and in PEPPER Help. The instructions in this section describe how to manually type a PepFile list for the PepAgree utility.

A PepFile List constructed for PepAgree contains 2 matched groups: the first group contains a list of the Time 1 PepFile transcripts for Transcriber 1, who is considered to be the standard or expert; and the second group contains a list of the PepFile transcripts for Transcriber 2 (for interjudge agreement) or a list of the PepFile transcripts for Transcriber 1's second transcription/coding of the same samples (for intrajudge agreement). For example:

```
=====T1
C:\pepper\agreement\amr10-f3_jm1.pep
C:\pepper\agreement\amr18-m3_jm1.pep
C:\pepper\agreement\amr19-m3_jm1.pep
C:\pepper\agreement\amr20-m4_jm1.pep
C:\pepper\agreement\amr29-f4_jm1.pep
=====Inter
C:\pepper\agreement\intr10f3_sh1.pep
C:\pepper\agreement\intr18m3_sh1.pep
C:\pepper\agreement\intr19m3_sh1.pep
C:\pepper\agreement\intr20m4_sh1.pep
C:\pepper\agreement\intr29f4_sh1.pep
```

The 4 equal signs separating the 2 groups above signifies in PEPPER that the PepFile transcripts in group 1 are matched with, or paired with, the PepFile transcripts in group 2 and that the data in each pair of PepFiles will be compared. When PepAgree is run, the first PepFile in list 1 will be compared to the first PepFile in list 2, the second PepFile in list 1 will be compared to the second PepFile in list 2, etc. The list for each group must have the same number of PepFiles.

The equal signs above group 1 are optional; they are included in this example so that the "T1" label can be used. The "T1" and "Inter" labels to the right of the equal signs are optional; they function as identifiers for each group of PepFiles. These labels will appear on PepAgree outputs.

To construct a PepFile list for PepAgree, open PEPPER and go to *File>New* and select PEPPER File List (*.pfl) from the dropdown list of file types. Type the lines in the following order, using the example above as a guide:

1. Type the first series of 4 equal signs and the group label (this line is optional).
2. Type the paths for all PepFiles for group 1. Type "Enter" at the end of each path to move to the next line.

3. After all group 1 PepFiles have been typed, type the line of 4 equal signs with the optional label.
4. Type the paths for all PepFiles for group 2, following the same order as the corresponding transcripts in group 1.
5. Save the PepFile list when finished.

Run PepAgree Analyses

With the PepFile list open, go to the PepAgree menu and select the agreement output of interest from the drop-down menu. If running the *Agreement Analysis...* (for transcription agreement calculations on vowels, consonants, and diacritics), you will be prompted to select whether to run the analysis on all words in the transcripts or first occurrence words only.

As mentioned previously, PEPPER compares the X and Y lines in the transcript pairs when PepAgree is run; before the output is generated, there may be messages that pop up on the screen indicating mismatches that PEPPER has identified. Click "OK" to move through each message. You may wish to attempt to resolve, if possible, any mismatches that are detailed in the message(s) by editing the PepFiles that PEPPER has flagged.

Interpret PepAgree Reports

For detailed information on the content of each of the PepAgree reports, how the transcription and PVSP agreement percentages are calculated, and how to interpret the agreement data for each PepAgree output, see the "PepAgree menu commands" section under the "Contents" tab in PEPPER Help.

PepStat and PepGraph

PepStat includes some basic statistical analyses options, and PepGraph includes some basic graphing options for tabular data generated in PEPPER or imported from elsewhere. For detailed information on PepStat and PepGraph, see the "PepStat menu commands" and "PepGraph menu commands" sections, respectively, under the "Contents" tab in PEPPER Help.

References

Baylis, A. L., & Shriberg, L. D. (2019). Estimates of the prevalence of speech and motor speech disorders in youth with 22q11.2 Deletion syndrome. *American Journal of Speech-Language Pathology*, *28*, 53-82. doi: 10.1044/2018_AJSLP-18-0037.

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.