Transcription & Spectrogram Practice

Scott James Perry

1 Instructions

I made this exercise in order for students in undergraduate phonetics classes to practice phonetic transcription and gain more experience analyzing spectrograms. I have prepared an audio file with three different corresponding `.TextGrid` files. The audio, entitled `Prologue_practice_audio.wav` is the prologue of a popular book, 'The Name of the Wind' by Patrick Rothfuss, read by yours truly. The following sections describe different practice exercises that can be completed. Skipping steps is fine, the possible activities are largely independent.

1.1 Transcription practice

In order to practice impressionistic transcription, open just the sound file and transcribe the sounds by ear using the IPA. If you need a little bit of help from the orthography, you can open the audio and the textgrid with the orthographic transcription: `Utterance_aligned.TextGrid`.

1.2 Free segmentation

The next exercise is using this audio to practice segmenting speech into words and phones in Praat. Open the sound file and `Utterance_aligned.TextGrid` in Praat. Add a second and third tier, segmenting all words on tier 2 and all phonemes on tier 3.

1.3 Correcting an automatic segmentation

The next .TextGrid file is an automatic, computer-generated alignment of the words and phones in the text. It has two tiers, one with words segmented out and the other with the phones segmented out. It was created with something called a forced-aligner, specifically the Montreal Forced Aligner (McAuliffe et al., 2017). What I did was give the Montreal Forced Aligner the `.wav` along with `Utterance_aligned.TextGrid`, and it produced `Automatic_alignment_ARPAbet.TextGrid` in about 11 seconds. This is pretty cool, but the alignment is not perfect, you will find mistakes. It also doesn't give the transcription using the International Phonetic Alphabet, but uses another transcription system called ARPAbet. You can practice finding the beginning and ends of sounds by correcting the mistakes made by the Montreal Forced Aligner. If you want to know how to read ARPAbet, the Wikipedia page has a conversion chart that seems to be accurate.

1.4 Checking transcription and segmentation

The third textgrid is called `Corrected_alignment_IPA.TextGrid`. In this file, I have changed the transcription from ARPAbet to the IPA and corrected the alignment of the boundaries on both tiers. You can use this to check your transcriptions, your own segmentation of the audio, or the corrections you have made to the computer-generated boundaries. This file is not perfect, there are mistakes in there that I didn't catch. If you think that I'm wrong, you can discuss it with your instructor, TA, or fellow classmates.

1.5 Important notes for students

As this is connected speech, there is a lot of variability in my productions. Some sounds will not look like you expect them to, and I want to make sure that you are not confused. The most variable segments are the interdental fricatives, which are often pretty stop-like. I have still marked them as interdental fricatives, even though that's not really what they are. For other sounds, I have based the transcription on the specific productions. In some of the textgrids, you will see strings of characters such as ¡unk¿, sil, and sp. These correspond to 'unknown', silence, and silent period. The ¡unk¿ string appears in the automatic alignment when the program did not know a word. If you are doing section 1.3 you will need to create the transcription and the boundaries for the phones from scratch for these words. In segmenting the words and phones is Praat in section 1.2, I asked you to create two new tiers, and put the words on tier 2 and the phones on tier 3. In the corrected version I created, the words are on tier 1 and the phones are on tier 2. There is no utterance tier. If you'd like your textgrid to match mine as closely as possible, you can delete the utterance tier in your segmentation after you no longer need it.

2 Asking questions

If I am your TA, I encourage you to direct your questions to the discussion forum on eclass. If I am not your TA, and your instructor has shared this with you, I recommend directing questions according to whatever your instructor may have told you. If you are a linguistics student who has stumbled upon this exercise, feel free to email me at `sperry1@ualberta.ca`.

References

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., and Sonderegger, M. (2017). Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*.