

Annotating speech in Praat

Why is speech annotated (i.e., segmented and labeled)?

- In principle, a written transcript of a speech sample is a secondary interpretation of the original events. When speech is used as data for research purposes, a transcript often needs to be supported by an audio or video recording of the original event. If the transcript is at least coarsely aligned with the timeline of the recording, you can quickly locate the part of the audio that corresponds to any given fragment of the written transcript. This makes your work much easier.
- Thus, an annotated speech corpus is more accessible, since you can utilise the annotations when searching for specific locations in the sound (or video) files.
- When using a speech annotation tool, you can quickly and easily locate, visualise, delineate and play back sections of the speech sample. This can make your transcription work significantly faster, easier and more accurate. For instance, if you are coding intonation patterns, it can be useful to compare your perceptual impression with the pitch curve calculated from the corresponding section of the sound signal.
- Once a sound file has been segmented and annotated, analyses can be performed consistently on the annotated portions of the sound. Data and measurements can even be collected automatically and, if required, the analysis can be repeated over and over again. For instance, it is possible to measure the precise durations of all speech sound segments, syllables, pauses, sentences, hesitation noises, etc. in a large speech corpus – given that the relevant units have already been delineated with sufficient accuracy and labeled in a suitable way.
- It is not necessary to splice the original audio into shorter files, but it can be stored in whole as required. For instance, a conversation recording does not need to be split up into individual utterances just for analytical purposes, and no information about pause durations is lost.
- The digital sound sample can be preserved for a long time and identical copies can be made at any time. The annotation files can be modified, enriched and shared between researchers, as long as a reference is maintained to the original recording. The annotated speech recording can be included in a well-described speech database that can be used and recycled by many researchers.

Typical annotation features in Praat

- possibility to create multiple annotation tiers: you can describe several different types of events and properties that may occur simultaneously
- possibility to label time points (in tiers of type **PointTier**) as well as time intervals (in **IntervalTiers**)
- “stand-off” annotation files: the TextGrid files are separate from the audio files; i.e., you do not change the sound file or add anything into it when annotating, you only change the annotation file.
- possibility to annotate very long sound files, even if your computer is short on memory (by using the LongSound object type, you can open and annotate files that include several hours of audio)
- portability and open format of the annotation files (TextGrid objects can be saved in regular text files, and the default UTF-8 encoding helps to ensure that special characters are correctly interpreted by other applications)

- possibility to use scripts for automatically processing and analysing both the annotations and the annotated sound samples.

How to annotate speech in Praat?

Before you begin annotating a sound file, make sure that the file has been edited to its final duration. For the annotations, we need to create a TextGrid object that must have the **same duration as the sound file to which it corresponds**.

If the total duration of the sound is changed later (for instance, if you decide to remove a useless silent portion from the beginning of the file) after you have started annotating, you will be in trouble, since the time stamps of the boundaries will no longer match the same time points in the sound file. It may turn out to be tedious to change the locations of all the boundaries in the TextGrid...

Creating an annotation object in Praat

1. Open the desired sound file in Praat, either as a Sound object (**Open: Read from file...**) or as a LongSound object (**Open: Open long sound file...**). If using LongSound, only the required portion of the sound file is loaded into memory for viewing it on the screen or for playback. LongSound is a good choice in case you are working on a lengthy file and your computer becomes slow. Make sure that the sound object is selected in the Object window.
2. On the dynamic menu, displayed on the right side of the Object window, click on the button **Annotate: To TextGrid...** A form appears. In the top row (*Tier names*), type the names of the annotation tiers you wish to include in the TextGrid, separated by spaces. For instance, if you typed “phone word utterance”, you would get three annotation tiers. You will be able to modify, add and remove tier names later.
3. By default, all the tiers mentioned in the top row will be created as IntervalTiers. In IntervalTiers, it is only possible to label segments or time spans that have a start and end time. However, if you wish to create PointTiers, include the names of those tiers both in the top row and in the row below it (*Point tiers*). In PointTiers, it is possible to label individual time points. This option might be useful for some purposes.
4. Click OK. A new TextGrid object will appear in the Object list. Now you are ready to begin annotating.

Annotating

1. In the Object window, select both the original Sound (or LongSound) object and the corresponding TextGrid object (which you just created). You can select or activate two or more objects by holding down Shift or Control while clicking on an object.
2. Click on the **Edit** button, shown on the right. A new TextGrid editor window will open. The top part of the window contains various displays of the sound. The topmost section shows the sound *waveform*, also known as the *oscillogram*. In addition, other visualizations and analyses of the sound (e.g., the *spectrogram*, shown in grayscale) may be displayed under the waveform. Above the sound displays, there is a white box for typing annotation text. Below the sound displays, there are white horizontal bars: the *annotation tiers*. The currently active annotation tier is displayed in yellow. In the right side of the editor window, you might also be seeing the Selection Viewer for inputting phonetic characters, in case the Selection Viewer option has been set in the

preferences. If you like, you may switch the Selection Viewer off (**File: Preferences...**).

3. In the TextGrid editor, you can play parts of the sound in the same way as in the Sound and LongSound editors. On the bottom of the editor window, under the annotation tiers, you can see two or three horizontal bars. They are for playback. By clicking on the bar on the bottom, you can listen to the entire sound object in one go. (If you clicked on it by mistake, the playback can be stopped by hitting *Esc.*) Clicking right above it, on the second bar, the visible portion of the sound is played. This is useful when you zoom in and out in the sound and scroll right or left in it. The topmost playback bar only appears when you click somewhere within the sound waveform or analysis displays, or when you select a portion of the sound by dragging over it with the mouse in the sound displays. Click on the topmost bar to listen to any portion, short or long, just change the location of the cursor or the selected area. Go ahead and try!
4. As soon as you find a good location for a new segment boundary (after listening and inspecting the displays), for example at the beginning of an utterance, click on the corresponding time within the sound displays. You can see a thin vertical line with a circle on top of each annotation tier. Click carefully on the circle at the tier where you wish to insert the boundary. The new boundary will appear as a thick vertical line.
 - The selected segment boundary is shown in red, the rest of the boundaries will be blue. You can select or activate a boundary by clicking on it.
 - You can remove the selected boundary by selecting **Point/Boundary: Remove.**
 - A segment boundary can be moved right or left by dragging it with the mouse (but not beyond any adjacent boundaries). In order to simultaneously move two or more boundaries that are located at exactly the same time but in different tiers, hold down the *Shift* key and start dragging one of the boundaries.
 - If you want two boundaries in different tiers to be located in the exact same time position, drag one of the boundaries and drop it on top of the desired boundary in another tier. This ensures that the boundaries share the same time stamp. You should now see a thin yellow line in the middle of both boundaries. (This method is useful since it is difficult to drag two boundaries to identical locations just by manual trial and error.)
 - The name of the selected annotation tier is displayed in red text in the right side of the TextGrid editor. On the left, the number of the tier will be red as well.
5. After delineating the utterance, speech sound or some other unit, you need to label it. Click between the boundaries in the desired annotation tier. Then type a label or transcription for the segment. The text will simultaneously appear within the selected segment in the annotation tier and in the white box on top of the window. If you need to edit text you inserted previously, click on the segment in the annotation tier first, and then click in the text within the text editing box to continue typing.

Tips

- **Note that there are usually no pauses between consecutive speech sounds or words!** In speech, sounds and words are actually produced continuously and you may even discover that sounds tend to overlap to some extent. Therefore, within each annotated utterance, the end boundary of each segment would generally be the

starting point for the following segment as well. (However, note that annotation practices can vary according to the purpose of the annotations.)

- There are speech sounds that naturally include a brief moment of silence, including, e.g., the voiceless stop consonants [kpt]. The silent phase reflects the articulatory closure that takes place during a stop, and the closure should not be confused with a pause. If you zoom in the waveform and look very closely, you can often find a small “burst” of noise where the closure of the stop consonant was released. (Voiceless stops can sometimes occur at the beginnings of utterances. In this special case, the quiet closure phase is preceded by a pause, and there is no way to determine where the closure phase of the stop articulation began. In such a situation, it is often safest to only include the release phase in the stop segment.)
- During speech, the sound producing organs of the speaker are constantly moving. It is no wonder that finding segment boundaries requires patience and some compromises!
- If you need to annotate a lot of material, it might be useful to learn a few keyboard shortcuts for the menu commands in the TextGrid editor, since that will save you a lot of mouse clicks. You can find the shortcuts in the menus, after the corresponding command. For instance, pressing Enter will insert a boundary at cursor in the selected tier (**Boundary: Add on selected tier**), and there is no need to try and hit the mini circle with the mouse. The tab key will play the selected section of the sound. *Ctrl-Right arrow* will select the next segment in the selected tier (**View: Select next interval**), and so on.
- You can also search for specific text in the selected annotation tier (**Edit: Find...**). This will come in handy when you annotate lengthy files.

Inserting and removing tiers

Within the TextGrid editor, you can insert more annotation tiers and copy or remove existing ones.

- To add a new empty IntervalTier, select **Tier: Add interval tier...** (for a PointTier, use **Add point tier...**).
- In the *Position* field, type the number of the new tier, counting tiers from top to bottom (default is to add the new tier below the previous ones). Type a *Name* for the new tier and click OK.
- To completely remove an annotation tier from the TextGrid object, including all the annotations in the tier, select the tier to remove by clicking on it. Make sure you selected the correct tier! Then select **Tier: Remove entire tier**. Be careful: you will not be asked for a confirmation! (In some cases, **Edit: Undo** can save you if accidents happen.)
- You can create a copy of an existing tier by selecting **Tier: Duplicate tier**.
- If you just want to remove the text written in an annotation tier, select the tier and **Tier: Remove all text from tier**. The boundaries will remain.

IPA characters

In Praat, you can use all Unicode characters for annotations, including IPA characters (IPA = *International Phonetic Alphabet*). In newer Praat versions (since version 6.1.04) you

can also use the Selection Viewer for picking phonetic symbols. You can switch the Selection Viewer on and off from **File: Preferences...** > *Show selection viewer*.

Praat also has a different way of including special characters. This works by typing the backslash \ character followed by other characters. The symbols will only be displayed in Praat TextGrid editors and the Picture window. You can check the internal key combinations for the special characters by selecting **Help:About special symbols**. You can also try searching the internal manual.

For instance, an *arrow-up* can be typed as a combination by typing \^ (a space is required after the ^ character). An *arrow-down* can be typed as _

Underlining, bolding and italics are not possible in Praat TextGrid editor, since the annotations are not intended for typesetting. If you must, you can of course invent your own mark-up. For instance,

```
here is some text, \_this\_ is underlined
```

will be shown in the annotated interval within the TextGrid editor as:

```
here is some text, _this_ is underlined
```

Note that the plain underscore character _ in Praat will make the next character to be displayed in subscript, e.g., typing **H_2O** you will see the chemical formula for water.

Saving the TextGrid

- If you are in the TextGrid editor window, select **File: Write TextGrid to text file...** to save the TextGrid object.
- If you are in the Object window, select the TextGrid object in the Object list first, and then select **Save: Write to text file...**

Alternative formats

- **Save: Write to short text file...** will make the annotation text file smaller in size and it will be easier to edit it in a regular text editor if necessary. The file will still open again in Praat in the usual way.
- In some cases you may want to try **Save: Write to binary file...**, especially if you are creating very rich annotations and big annotation files. The binary files are small in size, but they cannot be opened in any other program but Praat.