Phonetic fieldwork and experiments with phonfieldwork package for R

George Moroz

Linguistic Convergence Laboratory, NRU HSE, Moscow, Russia

15 November 2019

Science seminar, NRU HSE, School of linguistics, Moscow



Presentation is available here: https://tinyurl.com/y5xoks3l

Most phonetic research consists of the following steps:

- Formulate a research question. Think of what kind of data is necessary to answer this question, what is the appropriate amount of data, what kind of annotation you will do, what kind of statistical models and visualizations you will use, etc.
- 2. Create a list of stimuli.
- Elicite list of stimuli from speakers who signed an Informed Consent statement, agreeing to participate in the experiment to be recorded on audio and/or video.
- 4. Annotate the collected data.
- 5. Extract the information from annotated data.
- 6. Create visualizations and evaluate your statistical models.
- Report your results.
- 8. Publish your data.



Most phonetic research consists of the following steps:

- Formulate a research question. Think of what kind of data is necessary to answer this question, what is the appropriate amount of data, what kind of annotation you will do, what kind of statistical models and visualizations you will use, etc.
- 2. Create a list of stimuli.
- Elicite list of stimuli from speakers who signed an Informed Consent statement, agreeing to participate in the experiment to be recorded on audio and/or video.
- 4. Annotate the collected data.
- 5. Extract the information from annotated data.
- 6. Create visualizations and evaluate your statistical models.
- 7. Report your results.
- 8. Publish your data.

The phonfieldwork package is created for helping with items 3, partially with 4, and 5 and 8.



Why/when do you need the phonfieldwork package?

These ideal plan hides some technical subtasks:

- creating a presentation for elicitation task
- renaming and concatenating multiple sound files recorded during a session
- automatic annotation in Praat TextGrids [Boersma and Weenink 2019]
- creating a searchable .html table with annotations, spectrograms and ability to hear sound
- converting multiple formats (Praat, EXMARaLDA and ELAN)



Why/when do you need the phonfieldwork package?

These ideal plan hides some technical subtasks:

- creating a presentation for elicitation task
- renaming and concatenating multiple sound files recorded during a session
- automatic annotation in Praat TextGrids [Boersma and Weenink 2019]
- creating a searchable .html table with annotations, spectrograms and ability to hear sound
- converting multiple formats (Praat, EXMARaLDA and ELAN)

All of these tasks can be solved by a mixture of different tools:

- any programming language can handle automatic file renaming
- Praat contains scripts for concatenating and renaming files



Why/when do you need the phonfieldwork package?

These ideal plan hides some technical subtasks:

- creating a presentation for elicitation task
- renaming and concatenating multiple sound files recorded during a session
- automatic annotation in Praat TextGrids [Boersma and Weenink 2019]
- creating a searchable .html table with annotations, spectrograms and ability to hear sound
- converting multiple formats (Praat, EXMARaLDA and ELAN)

All of these tasks can be solved by a mixture of different tools:

- any programming language can handle automatic file renaming
- Praat contains scripts for concatenating and renaming files

phonfieldwork provides a functionality that will make it easier to solve those tasks independently of any additional tools. You can also compare the functionality with other R packages: rPraat, textgRid [Reidy 2016].



Send me a letter! agricolamz@gmail.com

Presentation is available here: tinyurl.com/y3wtkcbq



References

Boersma, P. and D. Weenink (2019). Praat: doing phonetics by computer (version 5:3:51)[computer program] version 6:0.25, retrieved 15 november 2019 from http://www.praat.org/.

Reidy, P. (2016). textgRid: Praat TextGrid Objects in R. R package version 1.0.1.