Automated Segmentation Exercises with PATKIT

Pertti Palo

3rd June 2025

Land acknowledgement

The University of Alberta, its buildings, labs and research stations are primarily located on the territory of the Néhiyaw (Cree), Niitsitapi (Blackfoot), Métis, Nakoda (Stoney), Dene, Haudenosaunee (Iroquois) and Anishinaabe (Ojibway/Saulteaux), lands that are now known as part of Treaties 6, 7 and 8 and homeland of the Métis. The University of Alberta respects the sovereignty, lands, histories, languages, knowledge systems and cultures of all First Nations, Métis and Inuit nations.

In addition to our university's written land acknowledgement I'd like to speak of my own relation to the lands where I have been working for almost a year now.

Outline

Thank you for your gift of time!

- Land acknowledgement
- This slide
- Introduction: The what and the why
- The Caveats
- Method
- Demo - ish
- Discussion
- Want to have a go yourself?
- MaTiPS ad
- Thanks and references

Introduction: The why

- Even in the age of 'AI' segmentation is a major bottle neck in getting phonetic research done fast.
- Segmentation is part of the essential skill set of a speech researcher in terms of understanding our data.
- Segmentation might even serve as a useful way into learning transcription.
- Segmentation rules vary with context of language, project, etc.

Introduction: The what

- Segmentation can be explained, but actually getting good at it requires hands-on practice.
- While Praat (Boersma and Weenink 2010) has an excellent segmentation interface, it does not provide a segmentation exercise interface.
- The Phonetic Analysis ToolKIT (PATKIT, the program previously known as SATKIT) (Palo et al. 2025) copies (essential parts of) Praat's segmentation interface and adds a resettable exercise feature.

The Caveats

- The exercises are not fully functional yet.
- Everything might change.
- I want your input for making this good.

No plan has ever survived contact with the real world. . .

Method: The workflow

- Segment a bunch of data somewhere - for now probably in Praat.
- Alternatively, get a bunch of segmented data (TextGrids and e.g. wavs) from somewhere.
- Open the data directory in PATKIT as a new assignment.
- PATKIT scrambles the TextGrids.
- Segment away.
- Optional saving of answers to continue later and for things like inter-rater metrics.

Method: Installing PATKIT

- Install uv `https://docs.astral.sh/uv/`
- On the commandline run:
  uv tool install patkit
  patkit
- This should print the commandline help.

Method: Running an example assignment/exercise

- Get the example data from `https://github.com/giuthas-talks/Automated_segmentation_exercises/`
- Put it in a couple of folders of your own choosing.
- Run:
  patkit exercise [folder name]
- This should open the annotator GUI.

Method: Potential command names

- Exercises → Run as assignment
- Exercises → Create assignment...
- Exercises → Save assignment...
- Exercises → New answer
- Exercises → Open answer...
- Exercises → Save answer...
- Exercises → Compare to model
- Exercises → Show model

Demo

Discussion

- So, do you have thoughts?
  - Which terms do you prefer?
  - Exercise vs assignment?
  - Answer, model, solution?
  - Any other ones?
- What would make this more useful to you?
- What other features would be good to have?

Want to have a go yourself?

–

Keep an eye out for version 0.18 and updates (0.18.x) later this week.

MaTiPS ad

Methods and Techniques in Phonetic Sciences: MaTiPS

- `https://matips.org/`
- Title and authors 19 July, anywhere on Earth
- Full one-page abstract 26 July, anywhere on Earth
- Results by beginning of August.
- Conference 17-19 October, University of Alberta, Edmonton, Alberta, Canada.

- Canadian Acoustics Week happens to be in neighbouring Calgary 15-17 October.
  - `https://jcaa.caa-aca.ca/index.php/jcaa/announcement/view/91`
  - There will probably be a bus between the conferences.

Thank you!

References

Boersma, P. and Weenink, D. (2010). Praat: Doing phonetics by computer [Computer program]. Version 5.1.44, retrieved 4 October 2010 from http://www.praat.org/.

Palo, P., Moisik, S. R., and Faytak, M. (2025). PATKIT: Phonetic Analysis ToolKIT [Python software package]. Available in a public software repository, accessed 8 February 2025. https://github.com/giuthas/patkit.