



e-Learning Samoyedic

Part I: Topics in Samoyedic phonetics

Tasks and explanations

By Alexandre Arkhipov and Gerrit Jawinsky

Co-funded by the
Erasmus+ Programme
of the European Union



The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Online course

Topics in Samoyedic phonetics

Foreword and suggested reading

This part of the course presents several empirical tasks on selected topics in phonetics of Samoyedic languages. Some of them only need reading and listening, while others will require using Praat, a widely acknowledged software for phonetic analysis.

Basic preparation in general phonetics is a prerequisite (while knowledge of Samoyedic or other Uralic languages is not); selected textbook chapters will be suggested for those who will have to revise it first.

Previous experience with Praat is welcome but minimally sufficient help (in addition to built-in Help and tutorials in Praat itself) will be provided.

General phonetics

1. Ladefoged, Peter; Johnson, Keith. 2011. *A Course in Phonetics*. 6th edition. [Or any previous edition, by Peter Ladefoged alone.] Wadsworth/Cengage Learning. (*the whole book is highly recommended but esp. ch. 1 and 6-9.*)
2. Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Technique*. Oxford: Blackwell.
3. *Handbook of the International Phonetic Association: A guide to the use of the international phonetic alphabet*. 2005. Cambridge University Press. (*Parts 1 and 3.*)

Further reading:

4. Brian Gick, Ian Wilson, Donald Derrick. 2013. *Articulatory Phonetics*. Wiley-Blackwell. (*Ch. 1, 5, 8, 9.*)
5. Harrington, John. 2010. Acoustic Phonetics. In: William J. Hardcastle, John Laver, Fiona E. Gibbon (eds.). *The Handbook of Phonetic Sciences*. 2nd ed. Blackwell.
6. Fullop, Sean. 2011. *Speech Spectrum Analysis*. Springer. (*Ch. 1-4, but also ch. 6 as an advanced topic.*)
7. Henning Reetz, Allard Jongman. Phonetics: Transcription, Production, Acoustics and Perception. Online Ressource. (*Ch. 10*)
<http://isites.harvard.edu/fs/docs/icb.topic482062.files/Reetz-Jongman%20%20Chapter%2010%20%20Acoustic%20Characteristics%20of%20Speech%20Sounds.pdf>

Praat manuals

Taken from <http://www.fon.hum.uva.nl/praat/manualsByOthers.html> (visit for overview and further languages)

1. PRAAT – German
- Jörg Mayer, *Phonetische Analysen mit PRAAT*
http://praatpfanne.lingphon.net/downloads/praat_manual.pdf
2. PRAAT – English
- Will Styler, *Using PRAAT for Linguistic Research*

<http://savethevowels.org/praat/UsingPraatforLinguisticResearchLatest.pdf> (last update: 02.01.2017)

3. PRAAT – Hungarian
 - Szekrényes István, Bevezetés a Praat program használatába
<http://lingua.arts.unideb.hu/compling/praat/praat-tutorial.html>
4. PRAAT-internal (taken from homepage, 19.03.2017)
 - Introductory tutorial: choose **Intro** from Praat's **Help** menus.
 - Extensive manuals and tutorials: in Praat's **Help** menus.
 - FAQ:
http://www.fon.hum.uva.nl/praat/manual/FAQ__Frequently_Asked_Questions_.html

Fora / websites

1. [Rob Hagiwara's How to read a spectrogram](#) (engl.)
2. <https://www.phonetik.uni-muenchen.de/studium/skripten/SGL/SGLKap2.html> (ger.)
3. https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf (eng.; very detailed manual, not all is needed)
4. <https://www.phonetik.uni-muenchen.de/~reichelu/kurse/akustik/PraatIntro.pdf> (ger.)
http://www.phonetik.uni-muenchen.de/~hoole/kurse/sockel/akustik/sockel_akustik_ur_ph.pdf (ger.)

Videos

1. [FloridaLinguistics on YouTube](#) (engl.)
2. [Andy McMillin on YouTube](#) (engl.)
3. Playlist by Virtual Linguistics Campus:
<https://www.youtube.com/playlist?list=PLRIMXVU7SGRIeAGqIQqzbsErDMgZWkuqh>
recommended: series on "Reading Spectrograms" (four parts: vowels, consonants, words I and II)

Online Course

Topics in Samoyedic Phonetics

TASKS

TASK 1. FILL IN THE GAPS

This first task intends to give you an impression of what the Nganasan language actually sounds like. Try to identify the single sounds in the realizations.

STEP 1. Download the .zip-file “Sounds 1” and unpack the content (three single folders, one for each part).

STEP 2. Listen to the utterances attentively and fill in the gaps!

Created by: Wagner-Nagy and Szeverényi (2013)

1. Where is the glottal stop (only one per word) – if there is at all?

ani_ka_a ‘big, large’

bi_i_ ‘ten’

bi_ə_ ‘wind ’

d'e_hi_ə ‘old (reindeer)’

sa_la_sa ‘to soil with mud’

2. What do you hear: n or ŋ(k) or maybe ń?

basaða_ ‘brass; something made of brass’

_ü_kə ‘nose’

_u_kəgə ‘many, much’

s’e_ ‘cover’

tə_tud’a ‘to catch up’

a_iʔə ‘big’

ku_i_i ‘where’

_ə_d’isi ‘to stand’

_aagəə ‘good, fine’

_e_amaʔ ‘neighbour’

ńe_sə ‘down, fluff, cotton wool’

3. What do you hear: s or s'?

a_a ‘Dolgan’

d’en_i ‘price’

_akiüə ‘fierce, savage (dog)’

_umiü ‘quiet, silence’

TASK 2. TRANSCRIBE ON YOUR OWN

This exercise intends to reactivate your knowledge on articulatory phonetics. Try to phonetically transcribe what you hear using the IPA transcription system in broad transcription. Please don't use PRAAT at this point, even if you're familiar with the program, try to focus on the perceived impression. The solution will show, how the Hamburg project on Nganasan uses to transcribe the language phonologically.

Transcription

STEP 1. Download the .zip-file “Sounds 2” and unpack the content (= one folder).

STEP 2. Listen to the single recordings attentively and transcribe them, using the IPA system (broad transcription). Do NOT open the .pdf yet!

STEP 3. When you're done transcribing all the single utterances, open the .pdf-file “Solutions Phonological Task 2” and compare your results with the general, phonological transcription provided there. Are there striking differences or is there anything striking in what you heard?

TASK 3. DISTRIBUTION OF THE GLOTTAL STOP

Please describe the distribution of the glottal stop in Nganasan (/ʔ/) and in Tundra Nenets (/q/, /h/) with respect to positions within the word and to the occurrence in clusters with other consonants.

Use the following examples:

Nganasan

/kuʔbasaʔarəku/ "as (big) glass beads"	/Kuəʔmaʔku/ „Kuomaku“ (Name)
/ŋənʔuʔtətu/ „(s/he) is wondering“	/təðuʔsutə/ „s/he will give“
/taatʔuʔsʔi/ „reindeer shepherd“	/hoiʔhinə/ "when/if I chop"
/kɪruʔmusʔa/ „(to) break up“	/ŋənuʔnamuʔə/ „s/he / it stopped in surprise“
/tʔuʔŋaləʔ/ „(several) fats“	/tʔeðiaʔliʔə/ „very insolent“
/sʔuaruʔlʔiʔəm/ "having become my friend“	/bəjmuraʔa/ "abfällig"

Tundra Nenets

Two glottal stop phonemes present: /q/ (“non-nasalizable”) and /h/ (“nasalizable”). Transcription follows Nikolaeva (2014): glottal stops are denoted with <q> and <h>; <ʔ> denotes palatalization; <°> to be regarded as a vowel; long vowels are marked with acute accent (e.g. /i/ = [i:]). Note: Transcription may deviate from the general, cyrillic-based orthography with latin characters.

/tʔuqe-/ “up” (adv. stem)	/padtaqma/ “drawing, pattern”	/yaq/ “strand of hair”
/ńumq/ “names” (elic.)	/ŋumh/ “grass”	/ʂiqn°/ “my”
/tox°q/ “cloth”	/pəqmʔa/ “sharp-sighted”	/melq/ “(skilled) craftsman”
/ilh/ “life” (i.t.o ~time)	/ŋobh/ “at once, suddenly”	/səqla/ “moron”
/sıqw°/ “seven”	/səqńuy/ “wet, dump, humid”	/padarq/ “paper”
/yemp°qŋa/ “s/he dressed”	/ŋarkaqa/ “very big”	/toh/ “blanket”
/wanşerh/ “incorrectly, wrong”	/ŋobq/ “one, single, unique”	/yah/ “soot”
/wəqlʔoy/ “sly, intricate, complicated, complex”		

TASK 4. EXPLORE PHONETIC REALIZATIONS OF A PHONEME [in Praat]

This task will (finally) allow you to use Praat (perhaps for the very first time). If you're new to it, please read the [[>>General notes on using Praat](#)] first.

Realizations of /q/ in Northern Selkup

STEP 1. Launch Praat. Open the **Collection** file which will become a number of Sound objects and corresponding TextGrids. Each Sound contains (usually three) pronunciations of a single word, all featuring a uvular obstruent /q/.

STEP 2. Explore each pronunciation and use the appropriate symbol(s) as the phonetic transcription of /q/ in the **Phones tier** of the TextGrid.

NB: You may choose between two difficulty levels:

Basic task

1. Use the **q-aligned.Collection** file. The TextGrids will be already aligned at the Phones tier.
2. While exploring the pronunciations, you will only have to **change the q symbol in the Phones tier** to match the sound you observe in each case (do not alter the Words tier).

Advanced task

1. Use the **q-nonaligned.Collection** file. The TextGrids will only be aligned at the Words tier.
2. Select all the TextGrids and run the **scr-SegmentIPA.praat** script. [[>>See notes on scr-SegmentIPA script](#)]. The **Phones tier** will now have word transcriptions automatically divided into a number of intervals of equal length.
3. You will first need to **adjust the boundaries in the Phones tier** (by dragging them around) to get the correct alignment for each phone. [[>>See notes on segmentation!](#)]
4. Explore the pronunciations, **changing the q symbol in the Phones tier** to match the sound you observe in each case (do not alter the Words tier).

STEP 3. Briefly describe the realizations you have observed. Which features vary between instances of /q/? Is it related to the phonetic environment (how)? **Save this as q-comments file** in either plain text or MS Word or odt format (q-comments.txt, q-comments.doc, ...). When you're finished, select all the TextGrids (without the Sounds!) and **save them as q-transcribed.Collection**.

TASK 5. MEASURE AND PLOT VOWEL FORMANTS [in Praat]

Providing a vowel plot in the F1-F2 formant space is commonplace in phonetic descriptions. However, reliably compiling such a plot is often a real challenge. In this task, we will only sketch an approach to plotting vowels, unveiling some of the underlying complexities.

Drawing a vowel chart for Northern Selkup

STEP 1. Launch Praat. Open the **Collection** file which will become a number of Sound objects and corresponding TextGrids. Each Sound contains (usually two or three) pronunciations of a single word, featuring different vowels in different positions in the word.

STEP 2. Explore each pronunciation. You will need to only measure formants for vowels that are pronounced clearly and are not shorter than 40-50 ms. Take note of the best setting for the formant ceiling you've found for the current speaker [[>>See notes on scr-GetFormants script](#)]. Some vowels will present difficulties for formant measurements whatever settings you use (especially unstressed vowels pronounced indistinctly and shortly, or vowels in the context of nasalization). The plotting script (**scr-PlotFormants.praat**) will group vowels by transcription symbol; so you can alter the transcription in order to exclude some less reliable vowel instances from the plot, or to group some of them apart (e.g. create a separate group for unstressed /i/ vowels).

- To leave out a vowel instance, add a dash “-” just before the vowel symbol in the **Phones tier** of the TextGrid.
- To group some vowels apart, either **change the vowel symbol** in the **Phones tier**, (e.g. ə > E) or **add something** to the vowel symbol (e.g. i > i0).

*Be aware that the script calculating the formant values only looks for strings beginning with vowel symbols listed in the regular expression "**^[iɪɪiyeɛəäöõuü]**", so in the first case you'll need to add characters to that regular expression as well.*

NB: You may choose between two difficulty levels:

Basic task

1. Use the **v-aligned.Collection** file. The TextGrids will be already aligned at the **Phones tier**.
2. While exploring the pronunciations, you will only have to **put a dash in the Phones tier** in front of vowels you wish to omit from calculations, or **change the transcription symbol** if you wish to group some vowel instances apart (**do not alter the Words tier**).

Advanced task

1. Use the **v-nonaligned.Collection** file. The TextGrids will only be aligned at the Words tier.
2. Select all the TextGrids and run the **scr-SegmentIPA.praat** script. [[>>See notes on scr-SegmentIPA script](#)]. The **Phones tier** will now have word transcriptions automatically divided into a number of intervals of equal length.
3. You will first need to **adjust the boundaries in the Phones tier** (by dragging them around) to get the correct alignment for each phone. [[>>See notes on segmentation!](#)]
4. Explore the pronunciations, **putting a dash in the Phones tier** in front of vowels you wish to omit from calculations or **changing the transcription symbol** if you wish to group some vowel instances apart (**do not alter the Words tier**).

When you've finished, select all the TextGrids (without the Sounds!) and **save them as v-sorted.Collection**.

STEP 3. Select each pair of Sound + TextGrid and run the **scr-GetFormants.praat** script, making sure you use the setting noted earlier for formant ceiling. A Table object will appear with the calculated formant values for each vowel. [[>>See notes on scr-GetFormants script](#)]. When you've finished, select all the Tables and run command Synthesize > Append. This will create a single Table with all the values. You can explore it by clicking View & Edit. Save this table as tab-separated file (**v-formants.Table**).

STEP 4. With this single Table selected, run **scr-PlotFormants.praat** script. You should get two plots one under another, the first one in traditional F1xF2 space, the second one showing the third formant (F3xF2). [[>>See notes on scr-GetFormants script](#)]. Adjust the F1, F2 and F3 ranges if the vowels do not fit the picture frame well. Save the plot as 300dpi PNG file (**v-plot.png**).

Briefly describe the plot you have just generated. Does it match your expectations? Compare to other vowel charts you've dealt with.

What kind of problems do you observe? How could one alleviate these problems to arrive at more reliable and informative plots?

Save this as v-comments file in either plain text or MS Word or odt format (v-comments.txt, v-comments.doc, ...).

TASK 6: TRANSCRIPTION OF SENTENCES [in Praat]

Now that you've become familiar with several functions of PRAAT try combining your acquired skills: Create an Object from the sound files, create a TextGrid and transcribe the sentences setting your own borders and eventually save your progress. If you have so far only chosen the basic tasks or encounter problems concerning the single steps, consult the [[>>General notes on using Praat](#)] or the manuals for detailed information.

Transcription of two sentences in Nganasan using PRAAT

STEP 1. Download the .zip-file and unpack the content (two audio files).

STEP 2. Open these sound files in PRAAT.

STEP 3. For each of them, create a TextGrid with two interval tiers ("Words" and "Phones")

STEP 4. Annotate the utterances.

STEP 5. When you're done, save your progress by the name "TASK 6 final YOUR_NAME".
Upload the file until the deadline (to be found on the moodle).

NOTES ON USING PRAAT

Explanation – What are we (i.e. you) going to do?

The tasks you are going to deal with in this course will also allow you to get to know analysis and annotation of speech using the PRAAT phonetic analyzer. If you are already familiar with PRAAT, you may skip the following preliminaries and start with the exercises. Before getting to use PRAAT, we'll try to get to know the language a bit, as the phonetic realization of the words we are going to deal with often differs quite drastically from the adopted transcription.

Preliminaries

PRAAT is a phonetics tool, developed by the Dutch researchers Paul Boersma and David Weenink. It is probably the most widely used and the most easily accessible program for doing phonetics. While its interface is somewhat simplistic, it has several major advantages which make it an unbeatable tool:

- Free and open source
- Regularly updated (up to several times a month)
- Extremely lightweight -- just one executable file of under 20 Mb
- Does not require installation — just download from <http://www.fon.hum.uva.nl/praat/> (or simply <http://praat.org/>), unpack if zipped and run
- Has an impressive array of functions for linguists, including speech synthesis, perception experiments, statistics, neural networks and much more!
- Is scriptable, so you can easily automate running similar task sequences over numerous files/fragments (through sound annotation with TextGrids)
- Can draw plots to save as images and save data as tables to use in statistical packages like R or SPSS (or Excel/Calc)

So hurry up to <http://praat.org/>, download, unpack and run immediately to start your analysis!

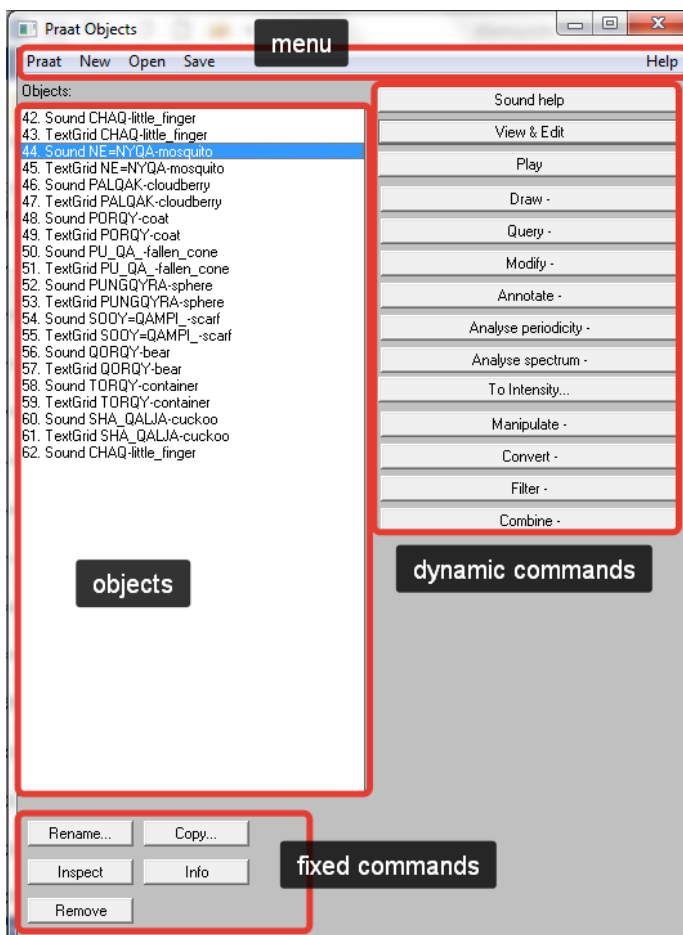
Besides extensive built-in help including several tutorials, you can also find Praat **manuals** in several languages (right [HERE](#)). They also contain explanations on certain analysis methods, giving you step-by-step instructions, which you are of course welcome to use, although most relevant information will be given below.

Basic elements of using Praat

Windows, objects and files

When launched, Praat will open two windows from the start: the Objects window and the Picture window. You can close the Picture window rightaway, it will reappear as soon as you draw something.

The Objects window is the “command center” for all further actions in Praat. It has a menu on top, a list of Objects (blank when you start) on the left, a few command buttons common for all objects in the bottom, and dynamically changing command buttons specific for selected objects on the right.



NB: Every file you open (except scripts) appears in the Objects list. Other things you extract from the files or create anew will also be added to this list. Everything that Praat does happens with these objects which are stored in Praat’s memory, not with files on disk. **Thus you must remember to save things — unless you explicitly save your work on disk, all the objects will disappear when Praat is closed, and cannot be restored.**

How to open/save a file

To open files of any type, use [Open > Read from file](#) from the menu, or [Ctrl+O](#). If you select e.g. a WAV sound file it will appear in the Objects list as a Sound object. If you open a .Collection file, it will add a number of objects (e.g. Sounds and corresponding TextGrids) stored in it to the Objects list at once. Thus a .Collection file is a pretty decent way to share files with other people or to continue the analysis at another point.

For saving work progress Praat, as basically every other program, needs a clear instruction on which files you want to save and how — so it's not just clicking SAVE. Please stick to this description, when saving your progress!

If you want to save an individual file, select it in the Objects list and use [Save](#) menu and the most appropriate option there (for sound files there will be a number of sound formats offered with WAV among them; for annotation in TextGrids use text formats).

If you want to save your progress with a number of files you're working on at the moment, you can save them all as a Collection. To do this, **mark all the files** in the Objects window and use [Save > Save as binary file](#) (takes least space) or [Save > Save as short text file](#) (good for text files only, e.g. TextGrids).

Eventually, resume saving as with any other file (choose a location and a name).

How to view/work with files

To open and analyse a file, mark BOTH belonging to it (Sound and TextGrid) and then click [VIEW & EDIT](#) command button.

Another (third) window will pop up — this is the editor (analysis) window, showing you the **waveform** of the signal plus any annotation in the TextGrid. You may turn on or off the different analysis parameters in the designated menus above, such as:

- **SPECTROGRAM:** The spectrogram is calculated from the waveform and demonstrates the various frequency components present in the signal at each point. Darker areas show concentration of energy at certain frequencies, lighter areas show low energy.
- **FORMANTS:** are such concentrations of acoustic energy resulting from resonance due to the specific vocal tract shapes. They are especially important for vowel analysis. When Show formants option is on, the estimated formant frequencies are shown in red (overlaid on spectrogram). We will deal specifically with formants in one of the later tasks.
- **PITCH:** estimated fundamental frequency measured in Hz. Changes in pitch give an impression of "high" or "low" sounds.

It is highly recommended to keep at least the spectrogram activated since it offers you the possibility of analyzing the single sounds in detail.

Eventually, use either the pre-segmented TextGrid for the annotation on basic level, the non-segmented for a higher difficulty or create one on your own (see below). Now, have fun!

How to create a TextGrid

Choose a file from the given list in the objects window by marking it blue. Click on [ANNOTATE>TO TEXT GRID](#). You will be given the possibility of choosing both *point* as well as *intervall tiers*. We won't need point tiers, so write LEXEME, PHONE into the interval tier column, each term will create a single tier.

How to segment the audio using TextGrid

After having detected a single sound (i.e. its quality), you can add borders. Therefore, click onto the corresponding point in the spectrogram or wave form (e.g. where the sound itself begins / ends). An axis will appear, follow the y-axis (top to bottom) and

click onto the UPPER tiny dot you can find on each annotation tier to add a boundary in that tier. First, segment the whole word (its beginning and end), then move on to segment the word into its single phones, using exactly the same mechanism as described above.

Scripts in Praat

Praat offers a possibility to create and run scripts — sequences of commands performing various operations as if done by clicking Praat buttons and selecting commands from menus (and more than that). Scripts are particularly great if you need to repeat some actions with many files or many portions of a file – e.g. do the same measurements for each vowel in a sound file.

Many Praat users share their scripts on the web, so many you can find many useful scripts that can be used directly or adapted to your needs.

How to open and run scripts

1. Most of the commands in scripts are exactly the same commands that can be run by clicking a button or choosing a command from a menu. As you already know, the list of available commands depends on the active window (the main, i.e. the Objects window, or an editor window for a particular object) and the objects currently selected. This is why it is important to know where you call a script from and what objects are selected before you call a script. In this course we will only use scripts called from the Objects window.
2. While in the Objects window, you can open a script either from the Open menu as any other file ([Open > Read from file](#), or Ctrl+O) or from the Praat menu ([Praat > Open Praat script](#)). If the script is stored with a .praat extension, double clicking it will probably make it open in Praat as well.
3. It's a good idea to have a dedicated folder for scripts and avoid opening scripts directly from an email attachment or from a download folder. First save a script to some location you can easily find. This will avoid you the risks of running the wrong version of a script you're editing, etc. Also, if your script is supposed to save any files on your computer as a result, it will probably try to do that in the folder in which it is itself located.
4. When a script is open, it will appear in a separate editor. Make sure you have the right objects selected in your Objects window and switch back to the script window. Run the script with [Run > Run](#) (Ctrl+R). For testing / debugging purposes, you can also run a fragment of a script with [Run > Run selection](#) (Ctrl+T).

How to create and edit scripts

Although not a topic of this course, you might be interested in creating your own scripts or adapting/enhancing some of the scripts you find on the course pages or elsewhere. Feel free to refer to the scripting tutorials offered in the Help menu ([Help > Scripting tutorial](#)) or found on the web.

1. To create a new script, use [Praat > New Praat script](#) and start typing. For instance, type:

```
Get total duration
```

Save the script with Ctrl+S as duration.praat. Check that a Sound object is selected in the Objects window, switch back to the script editor and run the script with Ctrl+R.

An Info window will appear with something like the following:

```
48.13673469387754 seconds
```

Congratulations! You've successfully run your first Praat script. Carry on!

2. A very convenient feature for writing scripts is the **command history**. Whatever useful you did in Praat at some moment, you can switch to the script editor and paste the command history into the script ([Edit > Paste history](#), or Ctrl+H). You will then have all the actions you have done since you opened Praat listed in your script, ready to be run again. In case you've done too much already, use [Edit > Clear history](#), then do the thing you want to see in your script and paste the history.
3. Use **comments** (lines started with #) to help yourself and others understand what your script is supposed to do. It is also a good idea to specify what kind of objects should be selected before running the script and what other requirements should be met (e.g. annotation tiers, sound duration, whatever).

Notes on segmentation

Various tasks dealing with acoustics require splitting the speech signal (the “sound chain”) into segments, or phones, or else “speech sounds,” each of which can be viewed as representing a unit of the language’s sound system, or phoneme.

Phone-level segmentation is needed not only to know the duration (e.g. of a vowel or sibilant), but also for measuring e.g. formant frequencies, since formants are often calculated on a central portion of a vowel (for instance, the central 1/3 of the total duration).

In some cases, borderlines between segments are pretty clear; in some others, no evident border can be drawn that would not be arbitrary. Very much depends on the classes of adjacent sounds. We will address some of the relevant issues below.

How to segment speech — General directions

1. There are no absolute rules for segmentation. A complete segmentation requires taking many decisions which will always be to some extent arbitrary. It is thus important to implement them consciously and consistently.
2. Segmentation should rely on more than one aspect of the signal. A common approach is to use the following in combination: (i) the waveform, (ii) the spectrogram, and (iii) the auditive impression (listening). For this, you’ll need to check the [Spectrum > Show spectrogram](#) option in TextGrid editor.
3. In an ideal case, the border between two segments will be drawn at a point such that (i) there are significant changes in both the waveform and the spectrum at this point, and (ii) when listening to the sound before and after this point, one hears two neatly separate sounds. However, it is often impossible to approach this idealization, partly because human perception is not at all trained to listen to individual sounds, and a segment cut out of a word may sound to a human ear as something completely different from the same segment inside a word where it belongs. Then try to find a point matching at least two of three criteria.
4. If there is a long and fuzzy transition between two segments instead of an abrupt change, a practical solution would be to mark the midpoint of the transition as the border.

Case by case notes

1. As a rule, the clearest cases for segmentation purposes are sequences of phones of maximally different classes, like vowels vs. voiceless obstruents. The worst cases are sequences of vowels, or vowels + sonorants.

See examples of segmentation (from Russian), referred to below as (1) and (2)) for an illustration.

2. The main acoustically relevant classes of sounds include:
 - stops (voiceless vs. voiced)
(silent vs. voiced) closure + burst

- affricates (voiceless vs. voiced)
(silent vs. voiced) closure + burst + turbulent noise
- fricatives (voiceless vs. voiced)
(pure vs. voiced) turbulent noise
- vowels and approximants
higher intensity periodic structure (regular oscillation cycles)
- nasals
lower intensity periodic structure, more steady intensity throughout the whole duration; sharp formant changes on edges
- trills
very similar to (adjacent) vowels but with one or more “gaps” in the middle (closure or voiced fricative)

3. Beginning and end of a word

(Here we're speaking of isolated words as in the examples given; otherwise it is only valid for phrases/utterances delimited by pauses.)

Word-initially, no real start point can generally be found in voiceless stops since their first component, a closure, acoustically only produces silence. One can then set an arbitrary limit, e.g. of 50ms before the burst.

- cf. initial [p] in **(1)**

Word-finally, most sounds can have quite a long “tail” of waning intensity. It is then a technical decision to draw a border somewhere across this tail. For vowels, it can be decided e.g. by visibility of higher formants.

4. Vowels and Co.

Vowels are not the only stretches of sound produced with the voice source and no (or very little) noise; such are the approximants, nasals and the voiced closures of the voiced stops. They can be distinguished to some extent based on their relative intensity:

open vowels > close vowels > approximants > nasals > voiced closures

While full vowels will normally have higher intensity (larger amplitude in the waveform and more dark matter in the spectrogram), usually with more than 2-3 formants, the latter (voiced closures) will only have a very subtle periodic signal in the waveform with just a “voice bar” in the bottom of the spectrogram (no formants).

- cf. [d] in **(2)**

Nasals will have the second-lowest intensity, a relatively even shape in the waveform (since the nasal resonance does not change as much as the oral one), and their formants can begin and end abruptly when transitioning from/to adjacent vowels (as the nasal resonance is switched on/off).

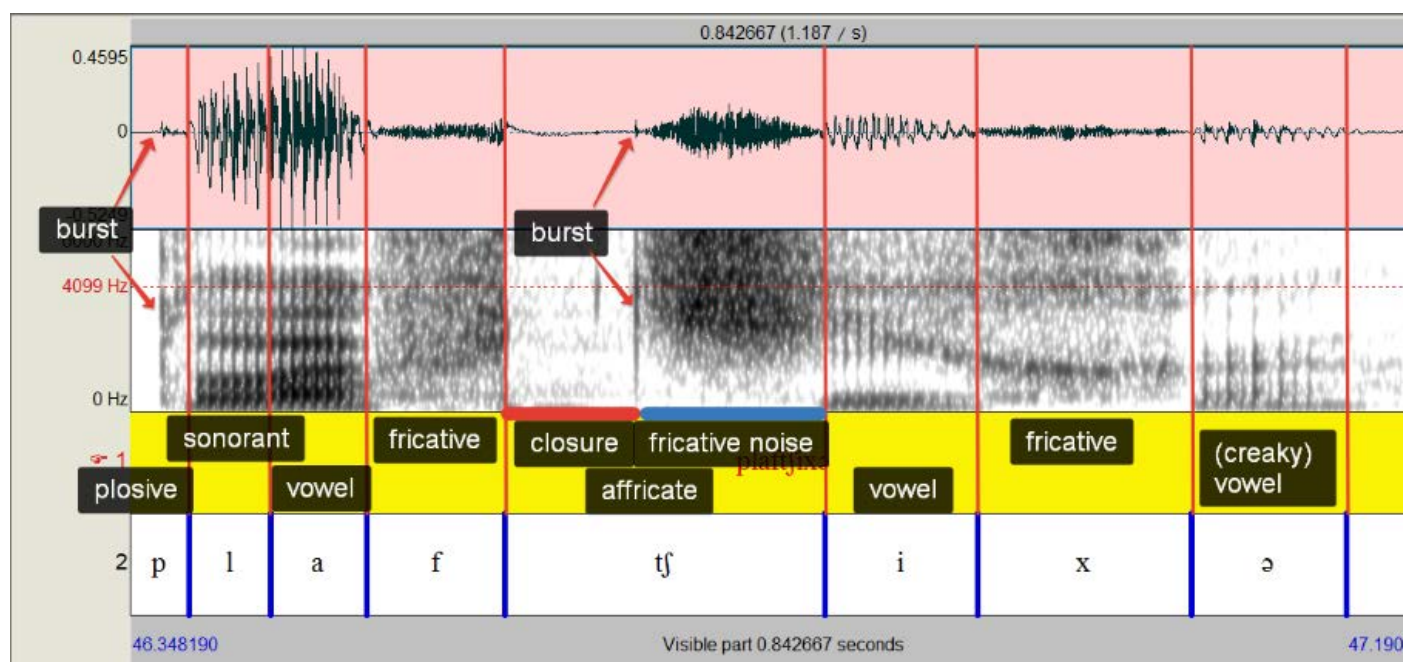
- cf. [m] in **(2)**

So a sequence of vowel + sonorant can be segmented based on drop of intensity (lower amplitude in the waveform, formants become pale in the spectrogram) combined with changes in formant frequencies. However,

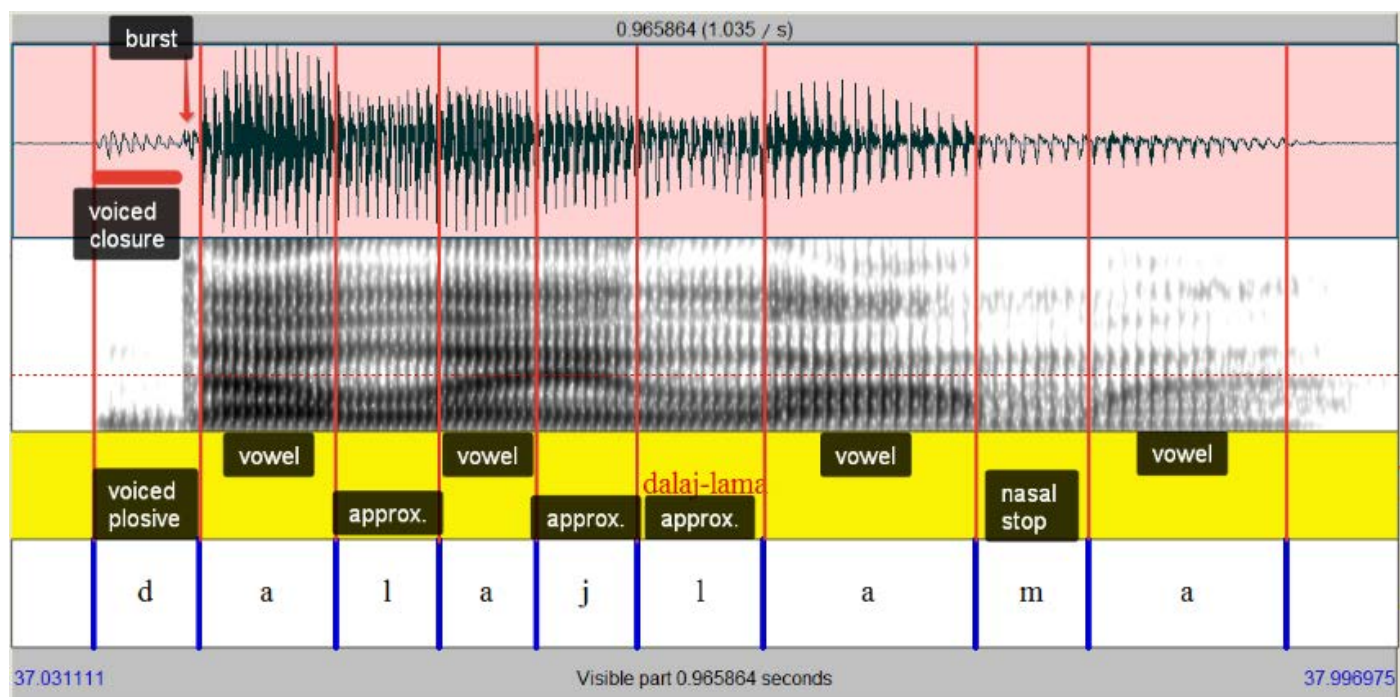
central approximants ([j], [w]) can hardly be distinguished by intensity from close vowels, the only remaining cue being the formant frequencies.

- cf. [l], [m], [j] in **(2)**; notice the slow transition between [a] and [j]

Segmentation Example 1: пловчиха [plaf'tʃixə] '(female) swimmer'



Segmentation Example 2: далай-лама [dalaj'lama] 'dalai-lama'



Notes on Selkup vowels

Northern Selkup boasts a fairly rich vowel inventory compared to other Samoyedic languages. Authors of the Essays on Selkup [Kuznecova et al. 1980] posit 13 basic vowels with additional length distinction, totalling 25 vowel phonemes. Compare this to the Nganasan system as seen in the Tasks 1-2, which only counts 8 vowels.

NB: The sample in Task 5 does not distinguish between long and short vowels, and also lacks some of them, namely e, ε and ö.

Table 1. Vowel phonemes in Selkup after [Kuznecova et al. 1980]

	Front		Central	Back		
	Unrounded		Rounded	Unrounded	Rounded	
	Tense	Lax			Tense	Lax
Close	i ī	ɪ ȷ	ü ũ	y ŷ	u ū	
Mid	e ē	ɛ ĕ		ə ǣ	o ō	ɔ
Open	ä ā			a ā		

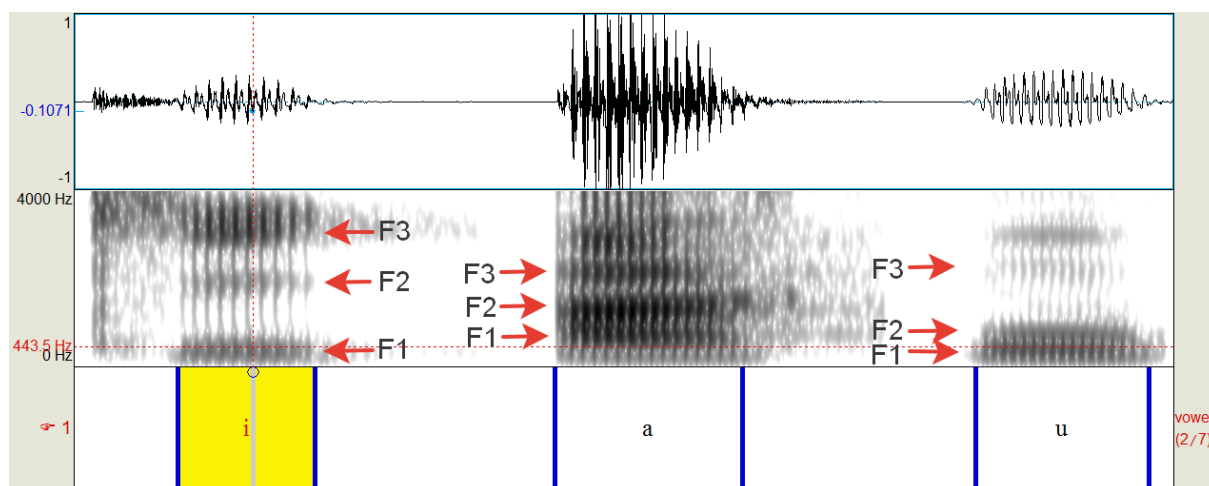
Note 1. In Task 5, dotless i (i) is used for lax close front vowel, and barred i (ᵢ) for close central vowel.

Vowel formants

Apart from the length contrast disregarded here, the main distinctions are expected to lie in the vowel qualities, manifested acoustically most notably in formant frequencies. This is what we are going to investigate in Task 5.

Formants are frequency ranges in which the excitation generated by the source (i.e. the vocal folds) is boosted due to resonances in the vocal tract. Changes in the shape of the vocal tract affect the resonant frequencies, resulting in different formant values depending on the position of the tongue and lips, among other factors.

Figure 1. First three formants in the spectrum of Russian vowels [i], [a], [u]



Formants are characteristics of the (resonances in) the **vocal tract**, as opposed to the **harmonics** which together compose the signal generated at the **source** (the vocal folds). They can thus change independently: raising and lowering the pitch affects harmonics but not the formants, while changing the position of the tongue and lips affects the formants but not the harmonics.

Although a formant is properly described by several parameters, such as its (central) frequency, bandwidth and amplitude, in many discussions only one parameter, namely the **formant frequency** (i.e. the centre of the frequency range), is mentioned.

Any vowel can be demonstrated to have many formants but only a limited number can be perceived by humans, and only three or four are effectively used to differentiate vowels. Most descriptions will only cite and plot values for the first two formants, F1 and F2, since for many languages their combined information is enough to tell apart the basic vowel qualities. In Task 5, we will use a script which also calculates values for F3, and we'll see whether F3 contributes to the overall picture.

Formants are characteristic of vocalic segments, i.e. vowels and sonorants, but sometimes the similar resonance-driven high-energy frequency ranges in noise are also called formants.

Formant values and articulatory features

It has been observed that the first two formants, F1 and F2, correlate with such articulatory features of vowels as height and backness. Namely, **F1** is inversely related with **height** (high F1 ~ open (low) vowels), while **F2** is inversely related with **backness** (higher F2 ~ front vowels). This is why F1xF2 plots with reversed axes are generally taken to represent vowels in language descriptions.

In a triangular vowel system, the extreme points of the vowel space are usually occupied by vowels **i**, **u** and **a**, with highest F2 and lowest F1 for [i], lowest F2 and also F1 for [u], and highest F1 for [a]. To have a rough idea of the possible values in Hertz (Hz), compare these approximate values for Russian:

Table 2. Approximate formant values for Russian (after L.M. Zakharov)

		F2 (Hz)				
		2000+	1700	1300	900	600
F1 (Hz)	300	i		i		u
	500		e		o	
	700			a		

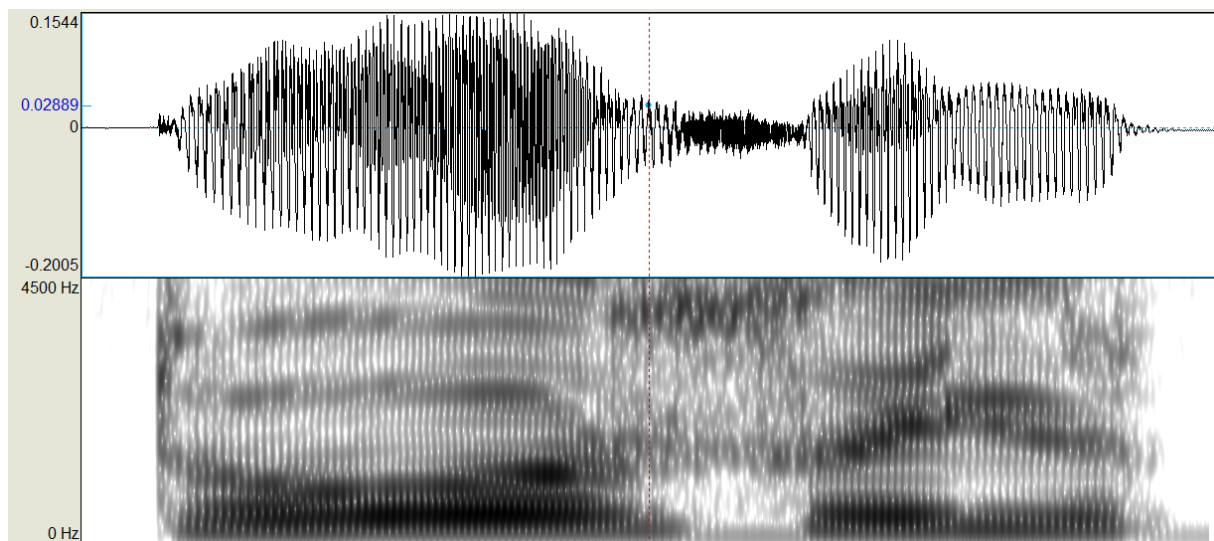
However, the whole story is much more complicated, since all the formant values are related to the length and shape of the whole vocal tract. Thus, female speakers generally have higher formant values than males because of smaller size (including length) of the vocal tract. An articulatory feature such as lip rounding has a lowering effect on all the formants. On the other hand, some features such as r-colouring (like on vowels preceding /r/ in American English) are mostly connected with F3 (which is lowered in this particular case). This is why, in addition to the traditional F1xF2 plot, the script for Task 5 will also produce an F3xF2 plot.

How to see formants in Praat

In order to see formants in Praat, open a [Sound](#) object with [View & Edit](#) (possibly together with the corresponding [TextGrid](#)) and make sure the [Spectrogram](#) is displayed: [Spectrum > Show Spectrogram](#) must be checked, and the window should not span more than 10 seconds (this limit can be changed in [View > Show analyses > Longest analysis](#)).

To see formants in the spectrogram well, set the upper frequency range limit to something between 3500 and 6000 Hz ([Spectrum > Spectrogram settings > View range \(Hz\)](#)). Formants will appear in vowels as darker horizontal bands, bending upwards and downwards as the articulators move (**Fig. 2**).

Figure 2. Spectrogram with up to four visible formants (view range 0 to 4500 Hz).



You may also need to tune the darkness of the spectrogram. If it is too dark *between* formants, try setting the dynamic range lower to see distinct formants clearly. If, on the other hand, you can hardly see anything in place where the waveform shows good signal, try setting the dynamic range higher ([Spectrum > Spectrogram settings > Dynamic range \(Db\)](#)).

Now, Praat can also estimate formant frequencies for you. To do that, check [Formant > Show formants](#). The estimated formant tracks will appear as red dots overlaid on the spectrogram. If the settings are suitable for the given speaker and vowels, the red dots will follow the dark bands of formants you see (**Fig. 3**).

However, it can happen that the current settings are not optimal for the data in question. Most often that means that the algorithm finds *more* or *less* formants than there actually are. The main setting you'd need to adjust in such cases is the frequency limit within which the algorithm looks for a certain number of formants ([Formant > Formant settings > Maximum formant \(Hz\)](#)). It is known that this setting can vary not also between speakers, but also between different vowels in the same speaker. Most notably, front vowels may require higher threshold than back vowels.

By default Praat uses the value of 5500 Hz for 5 formants, which is adapted for an average female voice. For a male voice, 5000 Hz is normally a better guess. For the data in Task 5, it appears that 6000 Hz as the value for [Maximum formant](#) suits most of the vowel instances, though not all of them. This is what you'll need to check for the vowels which you're going to submit to the script calculating

formants. (Care especially for back vowels, where F1 and F2 tend to be very close to each other).

If you see two tracks of red dots where you only expect one formant, as in **Fig. 4**, your **Maximum formant** setting is probably too low, and Praat tries to squeeze in too many formants. Try setting the threshold higher. If, on the contrary, you only see one track somewhere in the middle between two actual formants, like in **Fig. 5**, then the threshold is probably too high and Praat shows the mean between the two. Try changing the settings to get clear, steady tracks following the formant bands. Take note of the best setting to use it later with the script.

Figure 3. Formant tracks estimated by Praat.

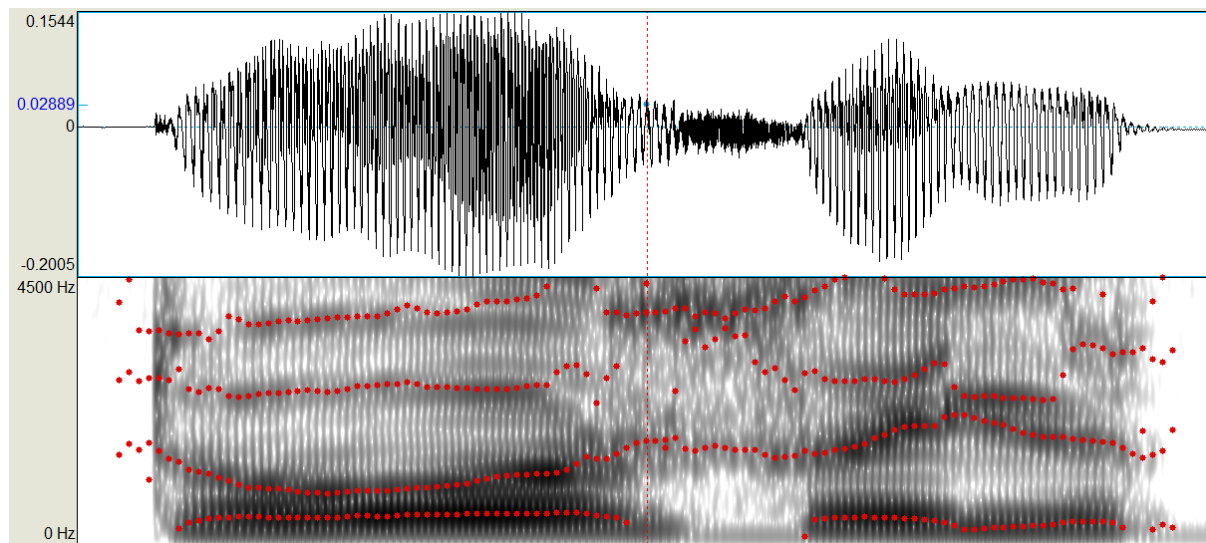


Figure 4. Max. formant too low (too many formants)

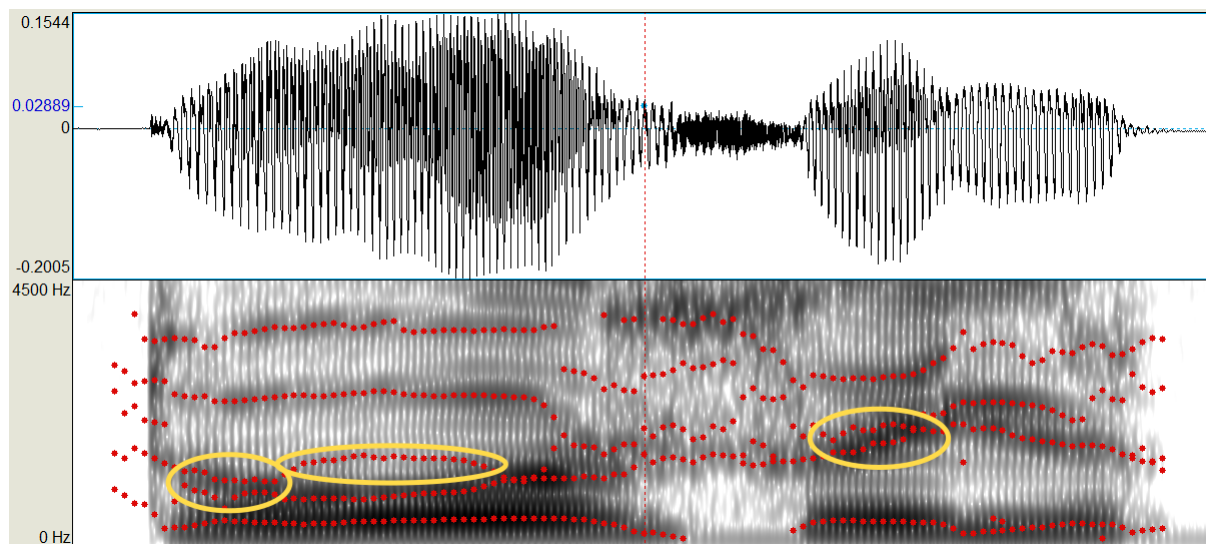
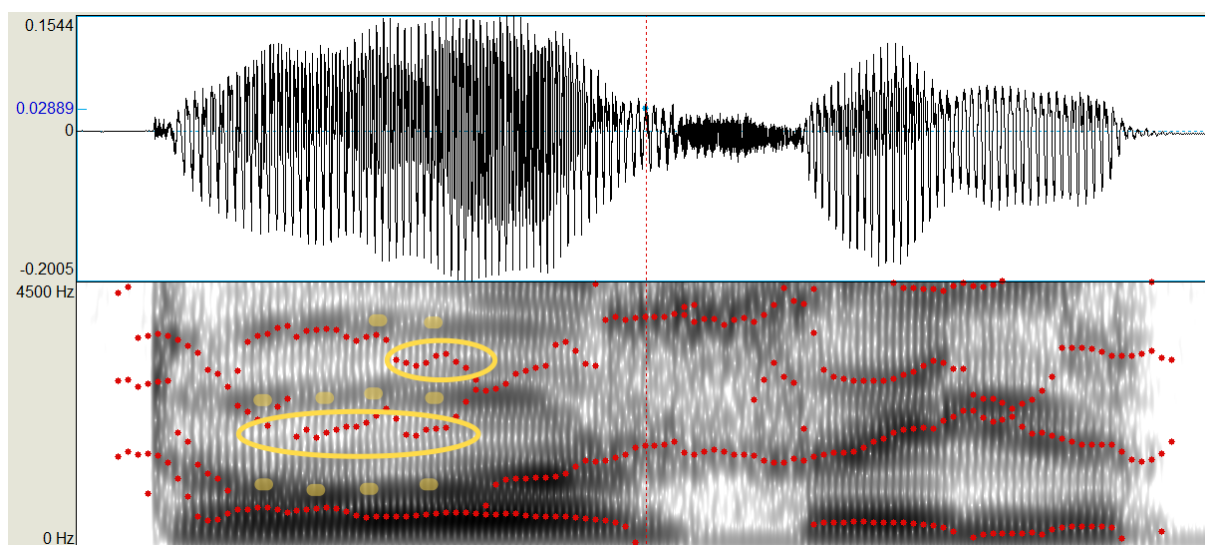


Figure 5. Max. formant too high (too few formants). Yellow dots indicate the actual formants missed by the algorithm. Yellow circles show misplaced formant tracks.



See also [Notes on scr-GetFormants](#) script.

Questions for Task 5

(bear in mind while writing *v-comments*)

1. Can all the vowels in the sample be distinguished on the basis of F1 and F2 formant frequencies?
2. Does F3 contribute to distinguish any vowel pairs? Are there any vowel pairs that cannot be well distinguished with all of F1, F2 and F3?
3. Do the values of the F1 and F2 suggest the same layout of vowel phonemes as presented in the inventory in Table 1?

Reference

Kuznecova A.I., Helimskiy E.A., Grushkina E.V. Ocherki po sel'kupskomu jazyku. Tazovskij dialekt. [Essays on Selkup. Dialect of Taz] Vol. 1. Moscow, 1980.

Task 5: Summary

Comments on variation

Variation is inherent in any living language. This also concerns phonetic realizations of phonemes. It should be kept in mind that phonemes and phonemic inventories are products of analysis performed by linguists, while in the natural data only varying realizations can be found.

One of the goals of this task was actually to show the variable and overlapping vowel qualities that we as linguists have to deal with. The next question is, how one can deal with it. In the present task, we had an aprioristic knowledge of the phonemes to look for. However in fieldwork it is normally the other way round: we start with the observed variation to arrive at idealized phonemes and abstract features at a much later stage.

As one of the students put it, *“The plot does not match with the other vowel charts I have seen. The phoneme boundaries should not get into the areas of each other, at least not in this extent.”* This is true, but this is only true after some tedious work on isolating variation-inducing impact of different factors — only then can we establish the phoneme inventory and point at some averaged locations of phonemes in the acoustic space.

One of important steps to isolate more or less “pure” vowel qualities is to restrain the chart to only stressed vowel instances — of course limited to languages which have stress. (This could be done in this task by putting dashes before unstressed vowels in the TextGrids.) As stressed vowels tend to be articulated in a more distinct manner, the corresponding formant values tend to lie closer together for any particular phoneme. Thus when selecting only stressed vowels, we can get a much clearer picture.

For instance, in the present sample the unstressed /ɪ/ and /ə/ overlap and scatter over a large area, while the stressed ones do not. However, this does not help much with separating /o/ and /ɔ/ (more on that below). *See Fig. A and B below for comparison.*

Another important consideration is taking the (immediate) context into account, most importantly the preceding consonant (but also the following one). This could not be done with the very limited sample at hand, but it is another necessary step to pinpoint the regularities of the phonetic variation. In other words, assuming we have lots of sound samples at our disposal, we would make the calculations separately e.g. for contexts such as “word-initial”, “following a labial”, “following a palatal”, “following a uvular”, “preceding a nasal”, etc.

Remarks on particular vowels

/i/ (tense) vs. /ɪ/ (lax)

The two front high unrounded vowels lie close together, but are still distinct in F2 (lower in lax /ɪ/, which points to a more back articulation) and to a lesser extent in F1 (also lower in lax /ɪ/, which would unexpectedly suggest a closer articulation).

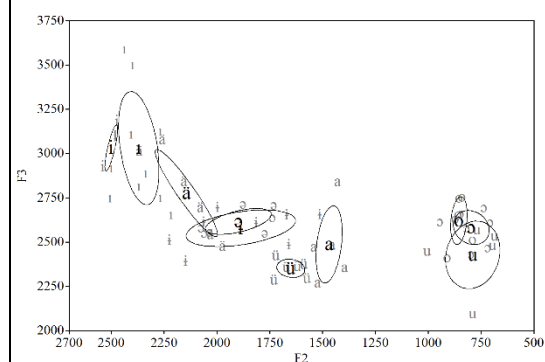
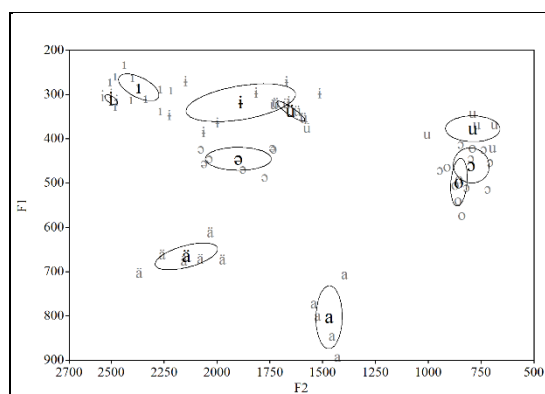


Fig. A. Vowel chart with only stressed vowels

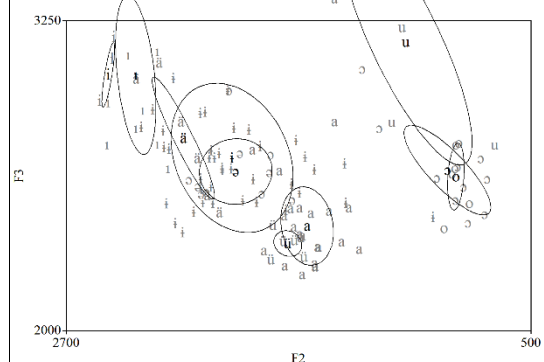
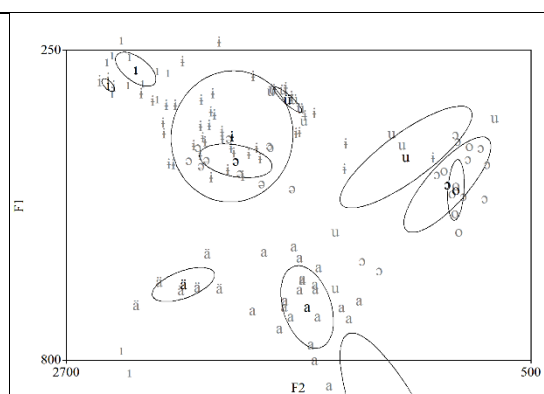


Fig. B. Vowel chart with all vowel instances

/ɪ/ vs. /ə/

The two central vowels tend to neutralize when unstressed (see Fig. B), but are neatly distinguished by F1 (lower in /ɪ/ as expected for a high vowel).

/ü/

As pointed out by two students, the high front rounded /ü/ shows an extremely low F2, even lower than the central /ɪ/ and /ə/ discussed above. While it is expected for a rounded vowel to have lower formant values, as we see here for both F2 and F3, this degree of F2 lowering seems unexpected. One possible reason for this is the context of preceding uvular /q/ in which this particular vowel had been reported to have a more retracted articulation [Prokofyev 1935: 20].

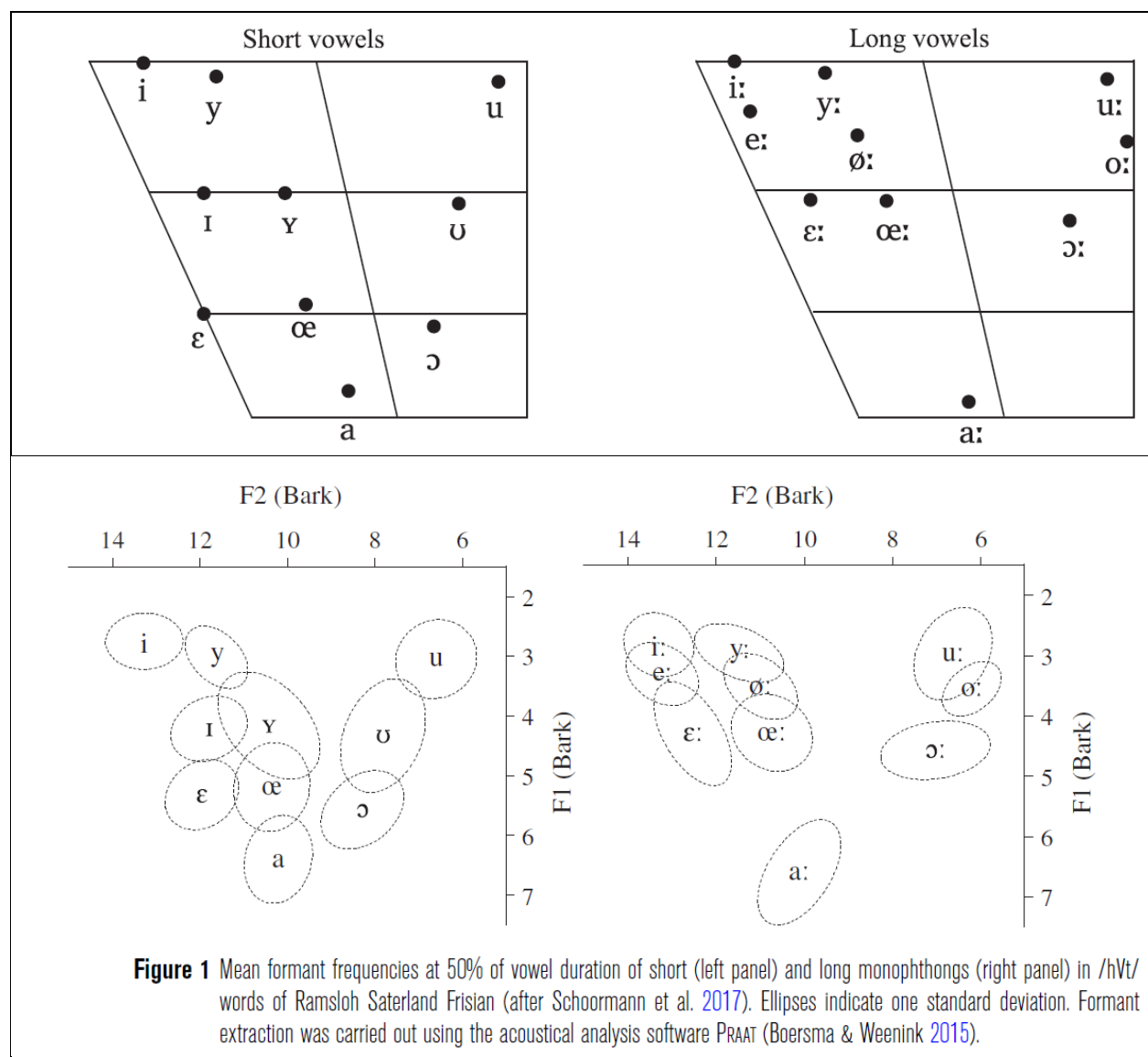
/o/ vs. /ɔ/

These two vowels do overlap significantly both in F1xF2 and F3xF2 dimensions, even when only taken in stressed position. Moreover, the speaker who produced the words for the recording did not have any intuition to distinguish these two sounds. Most likely these two phonemes merged into a single mid-back rounded vowel in her individual variety.

/u/ is distinguished from /o/s by both lower F1 and lower F3, although remaining relatively close to them.

Just for comparison

See these charts for Saterland Frisian vowels [Peters 2017] as averaged points and as ellipses showing variation.



References

[Prokofyev 1935] — Прокофьев Г. Н. *Селькупская грамматика*. Л., 1935.

[Peters 2017] — Peters, J. *Saterland Frisian*. Journal of the International Phonetic Association, 2017, 1-8. doi:10.1017/S0025100317000226