Tools rPraat and mPraat

Interfacing phonetic analyses with signal processing

Tomáš Bořil, Radek Skarnitzl





FACULTY OF ARTS CHARLES UNIVERSITY IN PRAGUE

Introduction

Someone would think that everything in phonetics could be solved in software Praat



Which is obviously not true...

🔳 Pra	at Obje	cts			-		×	
Praat	New	Open	Save				Help	
Objects:					Formant help			
1. TextG 2. Sound 3. TextG	rid R09_ R09_1 rid R09	1_DAM/ _DAMA 1_EISA	4		Draw -			
4. Sound	R09_1	FISA			Tabulate -			
6. LPC R	R09_1_Fl ant R09_1	ISA 1_FISA			Query -			
8. Forma					Query time domain			>
				t -	Query time sampling			>
					Get number of forman	ts		
					Get minimum number	of form	nants	
					Get maximum number of formants			
					Get value at time			
				1	Get bandwidth at time			
					Get minimum			
					Get time of minimum.			
					Get maximum			
					Get time of maximum.			
					Get quantile			
_					Get quantile of bandwi	idth		
Ren	ame		Сору		Get mean			
Ins	pect		Info		Get standard deviation			
Rer	move	1						

What makes phonetic analysis unique?

- Precise processing manual segmentation
 - Time-consuming, exacting, elaborate
 - Consistent, well-defined rules
 - Valuable



◆ Thorough understanding of perceived sounds
 ↔ physiological and linguistic processes



 Languages – dialects – sound change – social variation – individual differences

Phonetics

Praat is a primary tool for manual segmentation

♦ And for that, we have qualified students ☺



Praat

- Praat has many built-in functions for analyses
- Sut how to apply something special?
- New methods, modifications, detailed settings?
- Then, one has to switch to a more elaborate tool and create it on its own
- Matlab, R, etc.





Signal processing, statistics, graphics...

F₀ contours

- Another important fundamental in speech science
- Praat is reputable in
 F₀ contours detection



Anyhow, subsequent interventions are necessary

- Manual correction of errors
 - Octave jumps, creaky phonation, etc.
- Stylization of contours
 - Perception modelling

The valuable work

Manual segmentation and text annotations

TextGrids

Manually corrected F₀ contours

PitchTiers

Missing tools

Native support of these files in Matlab, R

What did we create?

Package / toolbox rPraat & mPraat

- Opensource github: everyone can contribute
- Implement new functions? Modifications? pythonPraat?



rPraat & mPraat

Import / export of TextGrids and PitchTiers

Complete set of functions for their processing

Identical to Praat

And also extras: what we missed in Praat

- New algorithms may be created and the results can be passed back into Praat
- Matlab and R languages are more flexible and wellarranged, more comfortable even for tasks soluble in Praat

rPraat & mPraat

Table 1. Summary of rPraat functions. mPraat features similar names, but without fullstops and with the following letter capitalized, e.g., tg.getLabel() vs. tgGetLabel().

tg.read / tg.write / tg.plot / tg.createNewTextGrid / tg.repairContinuity pt.read / pt.write (*PitchTiers containing F0 tracks*)

tg.getStartTime / tg.getEndTime / tg.getTotalDuration

tg.getNumberOfTiers

tg.getTierName / tg.setTierName / tg.isIntervalTier / tg.isPointTier

 $tg.insertNewIntervalTier\ /\ tg.insertNewPointTier\ /\ tg.duplicateTier\ /\ tg.removeTier$

tg.getNumberOfIntervals / tg.getIntervalStartTime / tg.getIntervalEndTime tg.getIntervalDuration / tg.getIntervalIndexAtTime

 $tg.insertBoundary\ /\ tg.insertInterval\ /\ tg.removeIntervalLeftBoundary$

tg.removeIntervalRightBoundary / tg.removeIntervalBothBoundaries

tg.getNumberOfPoints / tg.getPointTime / tg.insertPoint / tg.removePoint tg.getPointIndexLowerThanTime / tg.getPointIndexHigherThanTime / tg.getPointIndexNearestTime

tg.getLabel / tg.setLabel / tg.countLabels

Help and samples are included

Example: mean energy [e, e:]

```
tgID = Read from file... R01_3_PILA.TextGrid
    sndID = Read from file... R01 3 PILA.wav
 2
 3
    numInt = Get number of intervals: 2
 4
 5
 6
    for i from 1 to numInt
7
         select tqID
8
         lab$ = Get label of interval: 2, i
         if lab$ == "e" or lab$ == "e:"
9
10
             t1 = Get start point: 2, i
                                                                                               phoneme
                                                         kne Šiimtsos im
11
             t2 = Get end point: 2, i
                                                                                               (19)
12
             select sndID
                                                                        s imi
                                                                                               phone
                                                    P\
                                                         kne S
                                                                                S
                                                                                 1 i:
                                               o<del>∽</del> 2
                                                                  m t
                                                                                         S
                                                                                               (3/21)
13
             en = Get energy: t1, t2
                                                                                               syllable
                                                            neS
                                                                 iim ts
                                                       р
                                                3
                                                                        si
                                                                                    sli:S
14
                                                                            mi
                                                                                               (9)
15
              sum = sum + en
                                                                                               word
                                                        Řekneš
                                                                 jim co
                                                                                  myslíš
                                                                         si
                                                                                               (7)
16
             count = count + 1
17
         endif
18
    endfor
                            tg <- tg.read("R01_3_PILA.TextGrid")</pre>
19
                        2
20
   select tgID
                         3 snd <- readWave("R01_3_PILA.wav")</pre>
21
    Remove
                           fs <- snd@samp.rate; signal <- snd@left / (2^(snd@bit - 1))</pre>
                         4
22
    select sndID
                         5
23
    Remove
                        6
                          condition <- tg$phone$label %in% c("e", "e:")</pre>
                        7
                           t2 <- tg$phone$t2[condition]</pre>
                            t1 <- tg$phone$t1[condition]</pre>
                         8
                         9
                       10 - for (S in seqM(1, length(t1))) {
                                 segment <- signal[(trunc(t1[S]*fs+1)) : (trunc(t2[S]*fs+1))]
                       11
                       12
                                 SUM <- SUM + sum(segment^2)
                        13
                                 COUNT < - COUNT + 1
                                                                                                  12
                        14
```

Overview

"Low-level" Traditional "Praat-style"
TG[[2]]\$t1[7] tg.getIntervalStartTime(TG, 2, 7) # 2nd tier, 7th interval
length(TG[[2]]\$t1) tg.getNumberOfIntervals(TG, 2)



access by name
TG\$phone\$t1[7] tg.getIntervalStartTime(TG, "phone", 7)

```
# pipeline operator
tg.read("H.TextGrid") %>% tg.removeTier("word") %>% tg.write("out.TextGrid")
```

```
# vectorized operations
unique(TG$phone$label)  # unique labels
hist(TG$phone$t2 - TG$phone$t1) # interval durat
```

```
hist(TG$phone$t2 - TG$phone$t1) # interval duration
# cut PitchTiers according to TextGrids
- no similar interface with Praat is available
```



Smart code completion

5\$		
4	phone	٦
4	word	
2	phrase	

TG\$word\$

name
type
t 1
t2
label

tg.

-9-1			
tg.getrocarburation tg.checkTierInd	{rPraat}	^	tg.insertInterval(tg, tierInd, tStart, tEnd, label = "") Inserts new interval into an empty space in interval tier: a) Into an already existing interval with empty label (most common situation because, e.g., a new interval tier has one empty interval from beginning to the end. b) Outside og existing intervals (left or right), this may create another empty
↓ tg.insertBoundary	{rPraat}		
⊘ tg.insertInterval	{rPraat}		
↓ tg.insertNewIntervalTier	{rPraat}		
🧼 tg.insertNewPointTier	{rPraat}		interval between.
🧄 tg.insertPoint	{rPraat}		
◆ tg.isIntervalTier	{rPraat}		Press F1 for additional help
tg.isPointTier	{rPraat}	$\mathbf{\vee}$	

Performance analysis

Data: 1000 files, mean of 10 repetitions

 TextGrids: Full, short, binary format: slightly different structure, identical content

Experiments

- Ex1: Mean average duration of all [e/e:] vowels (TG)
- Ex2: Mean energy of all [e/e:] vowels (TG + Wav)
- Ex3: List all labels of the phone tier from all TextGrids into one single file, each label on a separate line (TG)

Performance analysis



 Ex1: Mean average duration of all [e/e:] vowels (TG)

Ex2: Mean energy of all [e/e:]
 vowels (TG + Wav)

 Ex3: List all labels of the phone tier from all TextGrids into one single file, each label on a separate line (TG)

Real vs synthetic speech

 Is there any kind of declination (effect of position in sentence) in postalveolar voiceless fricative [ʃ]?

Duration – Intensity – Centre of gravity (COG)





Real speech vs synthetic (dynamic unit selection)
 ARTIC system, University of West Bohemia

Real vs synthetic speech

• Type 1 sentence (9×)

◆ 1st [ʃ] within sentence, 2nd [ʃ] (near-)final

E.g., "V naší vile občas straší."

Our villa is sometimes haunted.

Type 2 sentence (22×)

• 1^{st} and 2^{nd} [f] both non-final

E.g., "Budeš jist, že dojdeš jistě k cíli."

You'll know for sure that you'll reach the goal.

Real speech: 8 females, synthetic: 3 female voices



Conclusions

- Declination of acoustic properties of the fricative []] is measurable
- There is a distinct difference between real and synthetic speech
 - Helpful parameter for detection of spoofing (speaker verification), manipulations to the speech signal
- rPraat and mPraat tools are open-source, available at http://fu.ff.cuni.cz/praat/

Thank you for your attention.

http://fu.ff.cuni.cz/praat/