

Transkrypcja fonetyczna i synteza mowy

Jolanta Bachan

IPA

Międzynarodowy alfabet fonetyczny, MAF (ang. International Phonetic Alphabet, IPA) – alfabet fonetyczny, system transkrypcji fonetycznej przyjęty przez Międzynarodowe Towarzystwo

Fonetyczne jako ujednolicony sposób przedstawiania głosek wszystkich języków.

Składają się na niego zarówno symbole alfabetyczne jak i symbole niealfabetyczne oraz ok. 30 znaków diakrytycznych.

Wikipedia

haʊ kæn aɪ raɪt laɪk ðis

- Zainstaluj klawiaturę IPA
- Google: „ipa keyboard download”
- http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=uniipakeyboard
- [IPA keyboard layout](#)
- Ćwiczenie: Przetranskrybuj
 - "a tiger and a mouse were walking in a field"

Praat



Praat: doing phonetics by computer



Download Praat:

- * [Macintosh](#)
- * [Windows](#)
- * [Linux, FreeBSD](#)
- * [SGI, Solaris, HPUX](#)
- * [the source code](#)



Information on Praat:

- * Introductory tutorial: choose **Intro** from Praat's **Help** menus.
- * Extensive manuals and tutorials: in Praat's **Help** menus.
- * Beginner's manuals by [Sidney Wood](#) and [Pascal van Lieshout](#).
- * Paul Boersma's [publications](#) on algorithms and tutorials.



[Paul](#)

The authors

Paul Boersma and David Weenink
[Institute of Phonetic Sciences](#)
 University of Amsterdam
 Spuistraat 210
 1012VT Amsterdam
 The Netherlands



[David](#)

Downloading Praat for Windows

New: Praat requires the Doulos SIL or Charis SIL phonetic font; see below.

To download the latest **Windows** (2000, XP, Vista...) version of **Praat**, it is easiest to use the self-extracting archive:

[praat5023_winsit.exe](#) (May 9, 2008; 1.7 MB)



(if you have Windows 95/98/NT4/ME, try [praat4601_win98sit.exe](#))

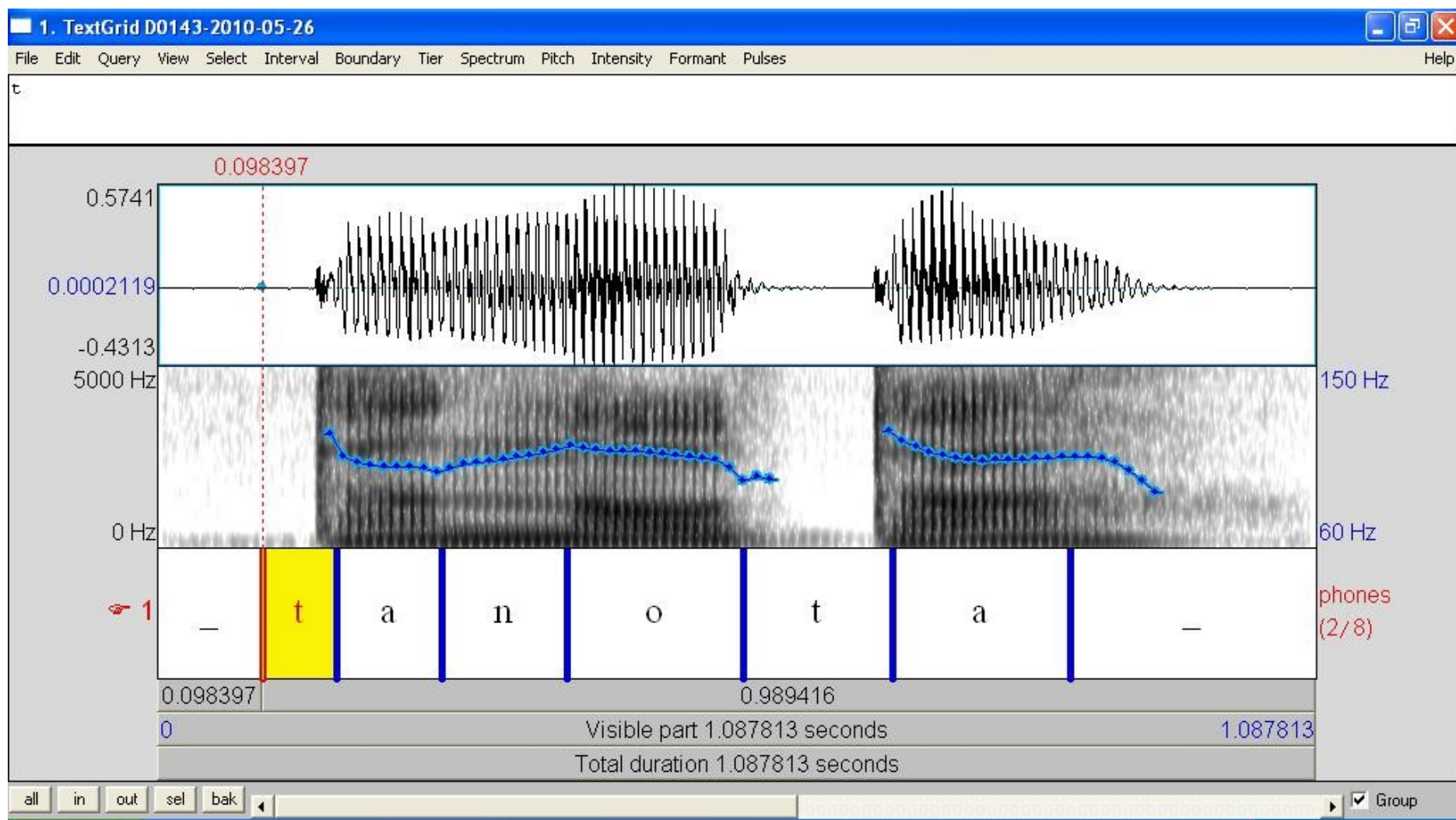
After downloading this file, double-click it and click OK to install Praat on your desktop. If this installation does not work, you may use the ZIP archive instead (after downloading, double-click to open):

[praat5023_win.zip](#) (May 9, 2008; 1.9 MB)

(if you have Windows 95/98/NT4/ME, try [praat4601_win98.zip](#))

Either way, a file called **Praat** or **Praat.exe** will appear. This is the Praat program. You can put it anywhere on your hard disk (so that you can use Praat even if your system administrator does not allow you to install other programs!).

Anotacja



Spektrogram

Oscylogram

Anotacja

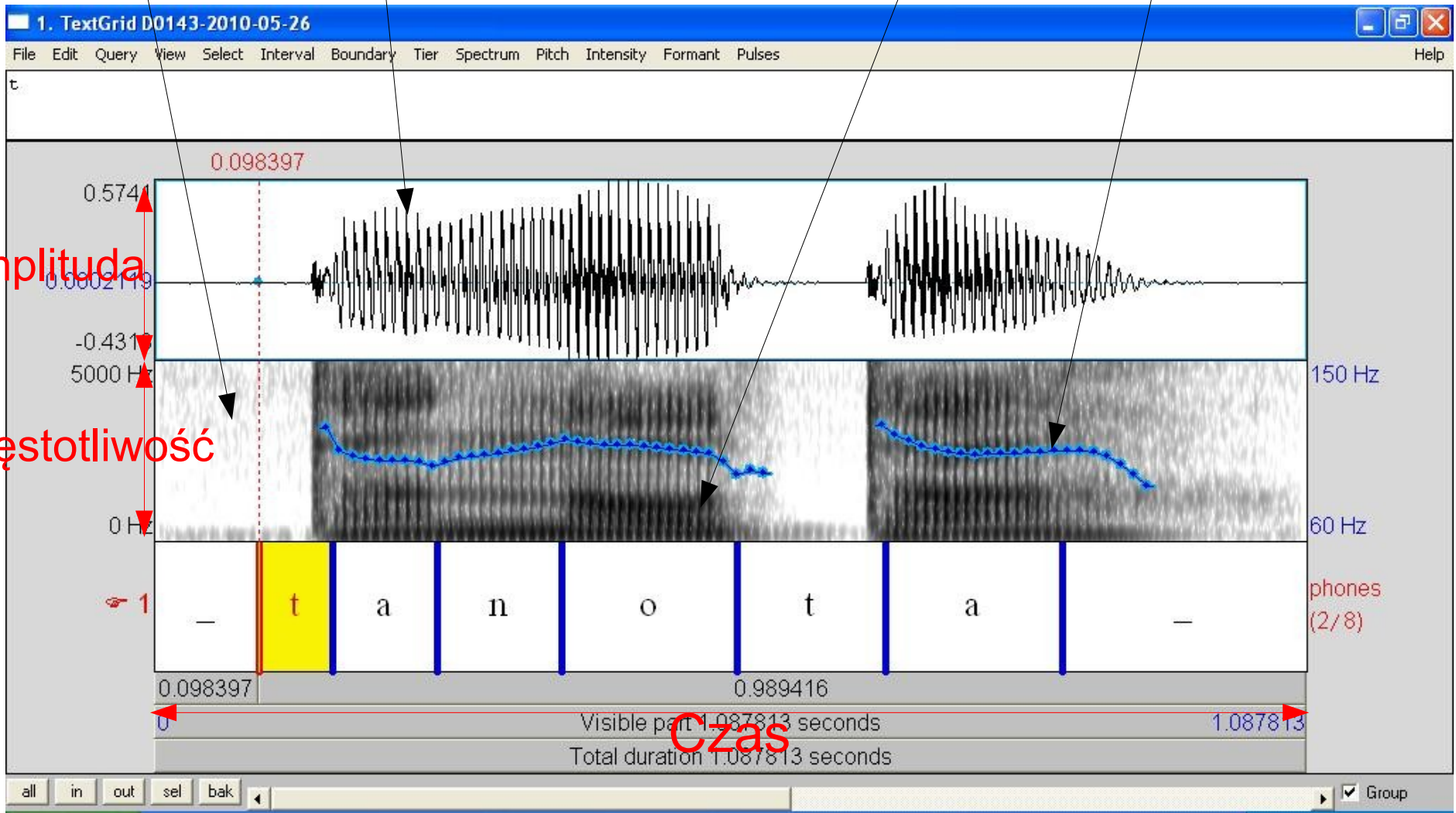
Częstotliwość podstawowa F0

Formanty

Amplituda

Częstotliwość

Czas



Spółgłoski

	Dwuwargowe	Wargowo- zębowe	Zębowe	Zadziąsłowe	Przednio- podniebienne	Miękko- podniebienne
Zwarte	p b		t d			k g
Zwarto- szczelinowe			ts dz	tʃ dʒ	tɕ dʑ	
Szczelinowe		f v	s z	ʃ ʒ	ɕ ʑ	x
Nosowe	m		n		ɲ	ŋ
Półotwarte	w		ɫ	l r	j	

Samogłoski i ich formanty

- Formant to pasmo częstotliwości, uwypuklone w barwie dźwięku.

Vowel formant centers

Vowel	IPA	Formant f_1	Formant f_2
u	u	320 Hz	800 Hz
o	o	500 Hz	1000 Hz
ɑ	ɑ	700 Hz	1150 Hz
a	a	1000 Hz	1400 Hz
ø	ø	500 Hz	1500 Hz
y	y	320 Hz	1650 Hz
æ	ɛ	700 Hz	1800 Hz
e	e	500 Hz	2300 Hz
i	i	320 Hz	2500 Hz

Praat: ćwiczenie

- Otwórz program PRAAT
- Załaduj plik:
 - bachan.speechlabs.pl > MA & PhD Theses > English original recording
 - http://bachan.speechlabs.pl/files/ai_eng_eng_m_read_na_selection.wav
- Wykonaj anotację IPA w PRAATcie

Praat: ćwiczenie

- Teraz wykonaj anotację korzystając z IPA specjalnie stworzonego dla PRAATa
 - Help > Search Praat manual > "Phonetic symbols"
- Edytuj plik TextGrid i porównaj kodowanie anotacji

Jak te informacje są przechowywane?

Format pliku TextGrid

```
File type = "ooTextFile"
Object class = "TextGrid"

xmin = 0
xmax = 0.2929375
tiers? <exists>
size = 1
item []:
  item [1]:
    class = "IntervalTier"
    name = "Phones"
    xmin = 0
    xmax = 0.2929375
    intervals: size = 4
```

```
intervals [1]:
  xmin = 0
  xmax = 0.0718
  text = "t_h"
intervals [2]:
  xmin = 0.0718
  xmax = 0.2323
  text = "aI"
intervals [3]:
  xmin = 0.2323
  xmax = 0.25579
  text = "g"
intervals [4]:
  xmin = 0.2557
  xmax = 0.2929
  text = "r\`="
```

Jak ułatwić sobie pracę?

Jak ułatwić sobie pracę?

SAMPA

computer readable phonetic alphabet

SAMPA
"**s{mpA:**

speech assessment methods
phonetic alphabet

<i>PL1 Polish SAMPA</i>	<i>Orthography</i>	<i>Phonemic transcription</i>	<i>PL1 Polish SAMPA</i>	<i>Orthography</i>	<i>Phonemic transcription</i>
p	<u>pik</u>	<u>pik</u>	i	kit	kit
b	bit	bit	I	<u>typ</u>	<u>tIp</u>
t	test	test	e	test	test
d	<u>dym</u>	<u>dIm</u>	a	pat	pat
k	<u>kat</u>	<u>kat</u>	o	pot	pot
g	gen	gen	u	<u>puk</u>	<u>puk</u>
f	fan	fan	e~	<u>geś</u>	<u>ge~s'</u>
v	<u>wilk</u>	<u>vilk</u>	o~	<u>was</u>	<u>vo~s</u>
s	<u>syk</u>	<u>sIk</u>	m	<u>mysz</u>	<u>mIS</u>
z	<u>zbir</u>	<u>zbir</u>	n	<u>nasz</u>	<u>naS</u>
S	<u>зык</u>	<u>SIk</u>	n'	<u>koń</u>	<u>kon'</u>
Z	<u>zyto</u>	<u>ZIto</u>	N	<u>pek</u>	<u>peNk</u>
s'	<u>świt</u>	<u>s'fit</u>	l	<u>luk</u>	<u>luk</u>
z'	<u>źle</u>	<u>z'le</u>	r	<u>ryk</u>	<u>rIk</u>
x	hymn	<u>xImn</u>	w	<u>lyk</u>	<u>wIk</u>
ts	<u>cyk</u>	<u>tsIk</u>	j	<u>jak</u>	<u>jak</u>
dz	<u>dzwon</u>	<u>dzvon</u>			
tS	<u>czyn</u>	<u>tSIn</u>			
dZ	<u>dżem</u>	<u>dZem</u>			
ts'	<u>ćma</u>	<u>ts'ma</u>			
dz'	<u>dźwig</u>	<u>dz'vik</u>			

Praat: ćwiczenie

- Wykonaj anotację w SAMPIE dla pliku
<http://bachan.speechlabs.pl/files/A0310.wav>

Synteza mowy

Synteza mowy

- System przetwarzania tekstu pisanego w mowę
- Text-to-Speech (TTS)
- TTS powinien być w stanie przeczytać każdy tekst, ale w praktyce nie jest to takie proste do zrealizowania

Synteza mowy

- Parametryczna synteza mowy
 - synteza formantowa
 - synteza artykulacyjna
- Konkatenacyjna synteza mowy
 - synteza difonowa
 - synteza trifonowa
 - unit selection

Synteza mowy

- Parametryczna synteza mowy
 - synteza formantowa
 - synteza artykulacyjna
- Konkatenacyjna synteza mowy
 - synteza difonowa
 - synteza trifonowa
 - unit selection

Konkatenacyjna synteza mowy

- Konkatenacyjna synteza mowy łączy mniejsze jednostki nagranej mowy (difony, trifony, sylaby, wyrazy) w większą całość (wyrazy, zdania)
- System jest oparty na bazie nagrań mowy. Baza posegmentowana jest na mniejsze elementy (głoski, difony, wyrazy), z których później “skleja się” wypowiedzi.

Co to jest Close Copy Speech Synthesis?

Close Copy Speech Synthesis

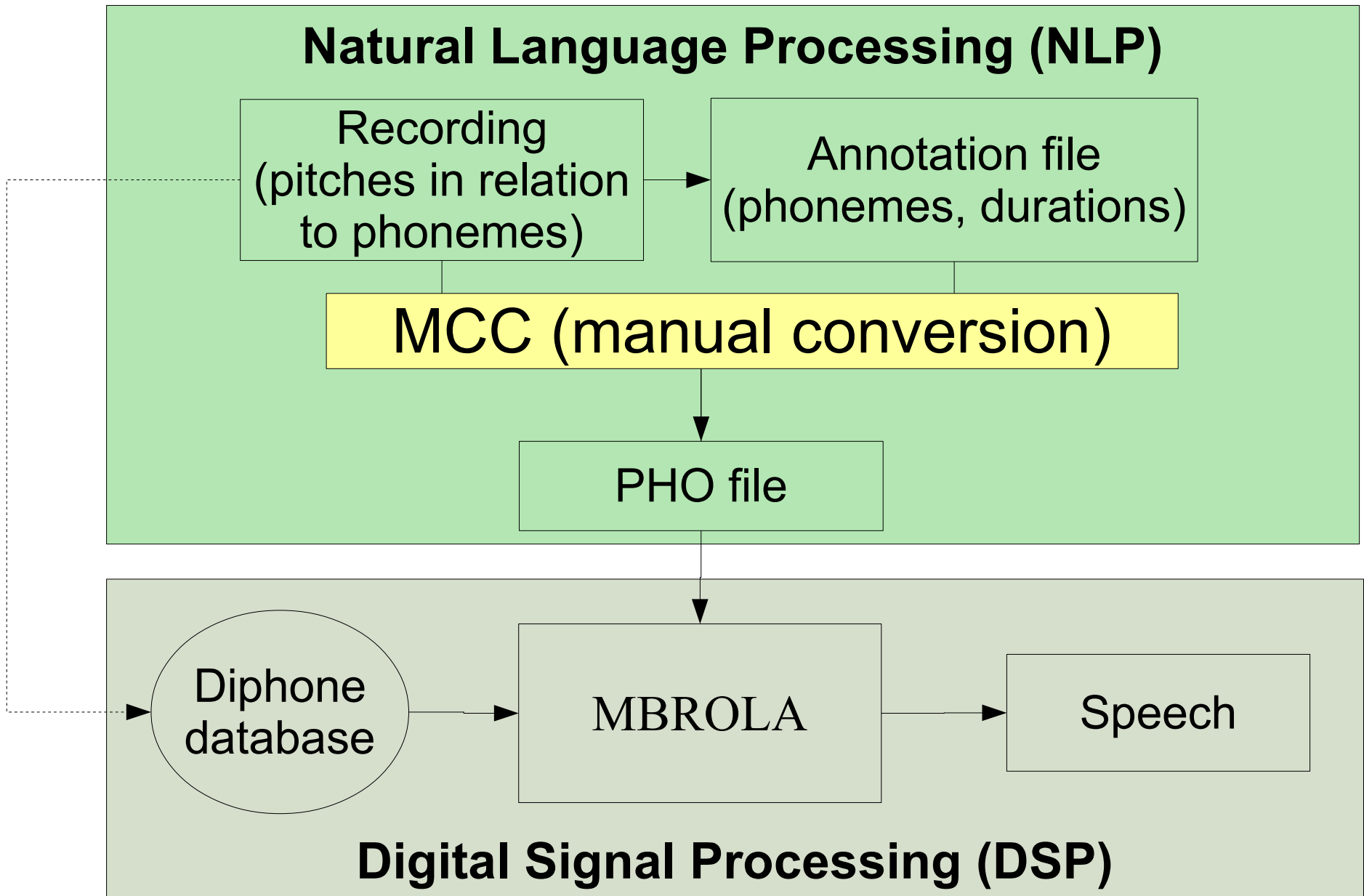
The CCS synthesis system produces a sound which “repeats an utterance produced by a human speaker with a synthetic voice, while keeping the original prosody” (Dutoit, 1996).

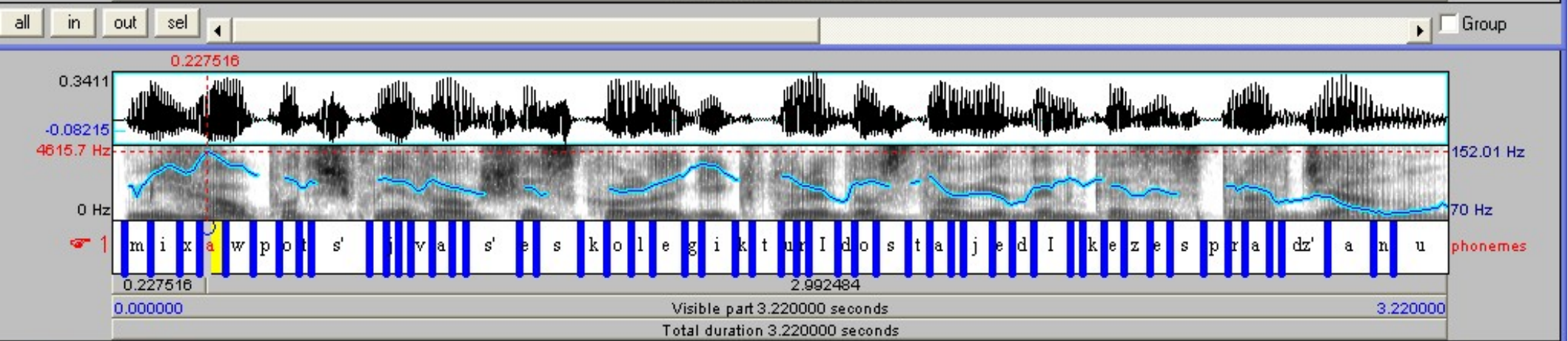
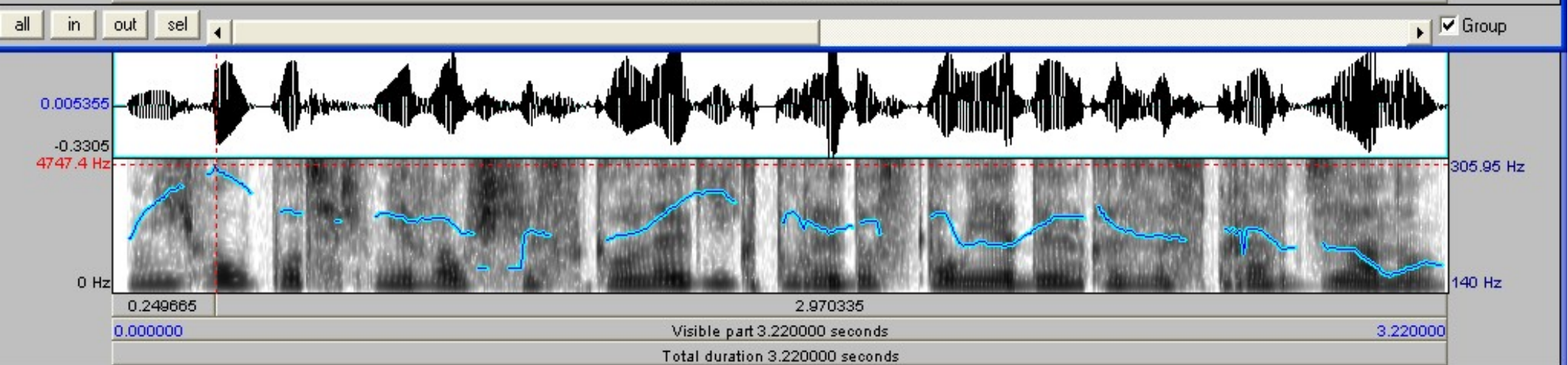
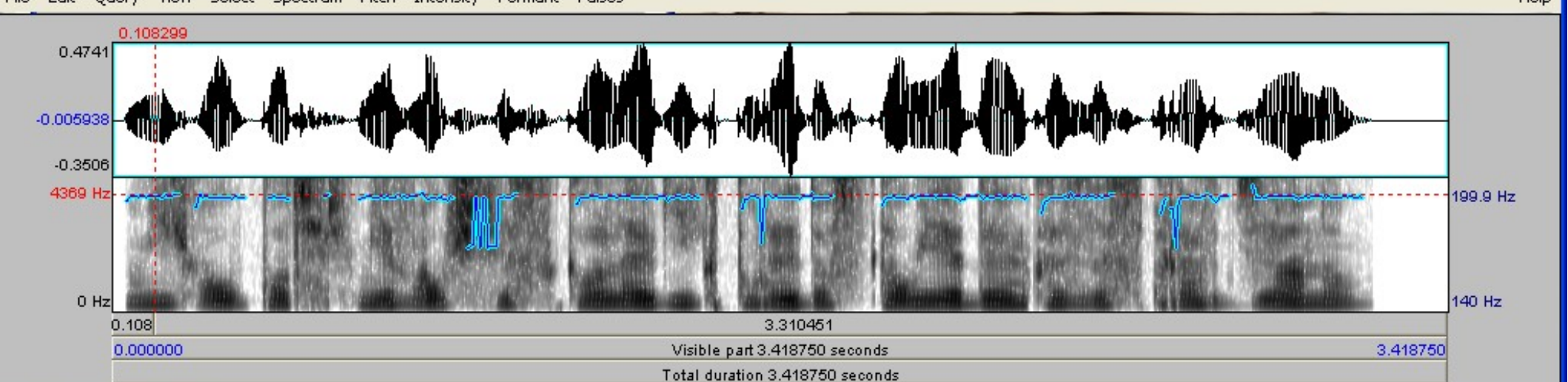
Manual Close Copy Speech (MCCS) Resynthesis

Komponenty MCCS

- Wejście: Mowa
 - nagrania mowy
 - anotacja nagrań mowy
- Syntezator mowy (tu: MBROLA)
 - baza difonów (tzw. głos MBROLI)
 - silnik syntezy

MCCS synthesis





Synteza MCCS

- Monotonny

Monotone

- Synteza MCCS

MCCS

- Oryginał

Original recording

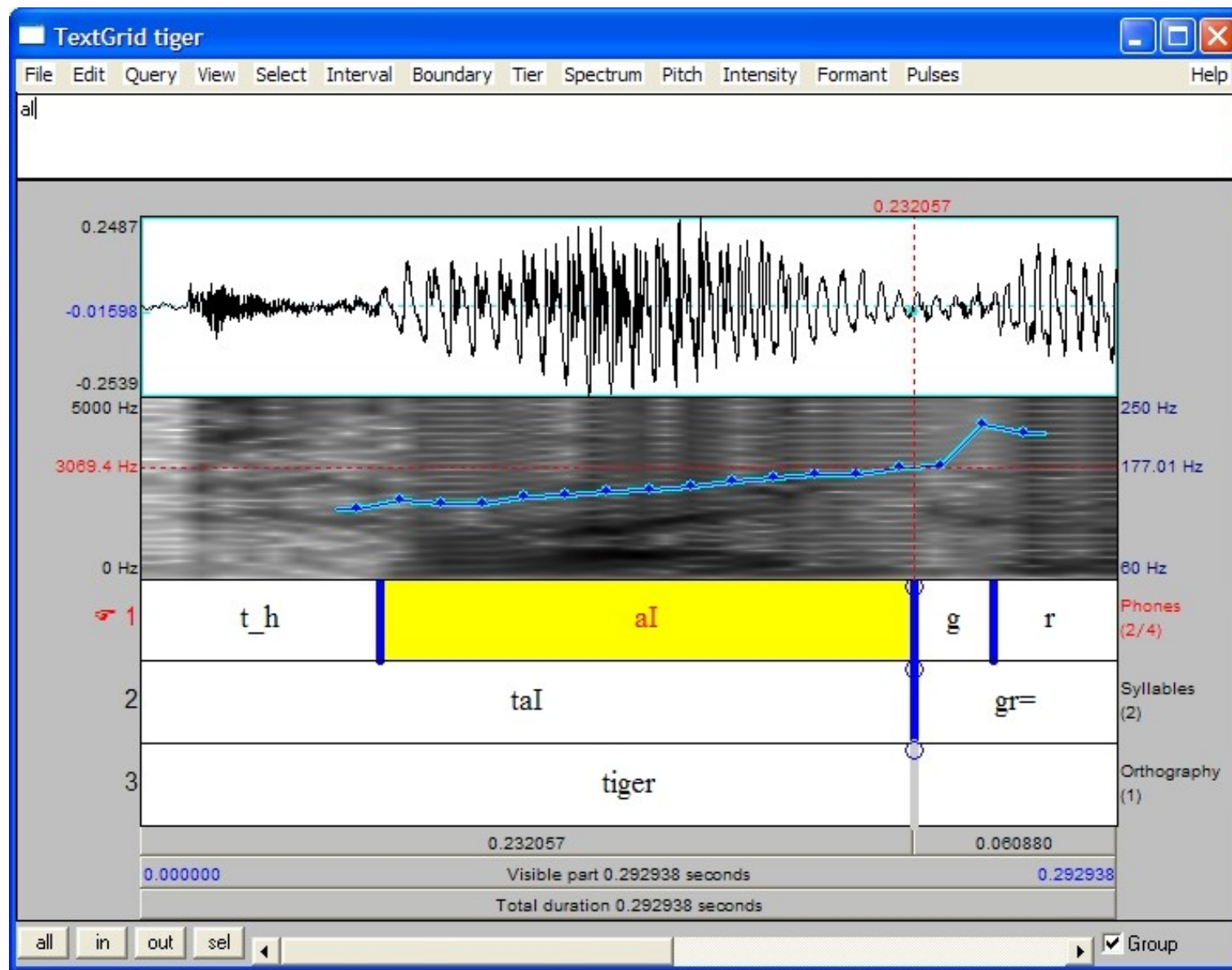
Zastowowania MCCS

- Tworzenie bodźców na potrzeby testów percepcyjnych mowy
- Narzędzie w nauczniu fonetyki
- Eksperymentowanie z prozodią
- Sprawdzanie anotacji

Anotacja mowy w Praacie

Anotacja dla syntezy CCS

REQUIRED
FOR
SPEECH
RE-SYNTHESIS



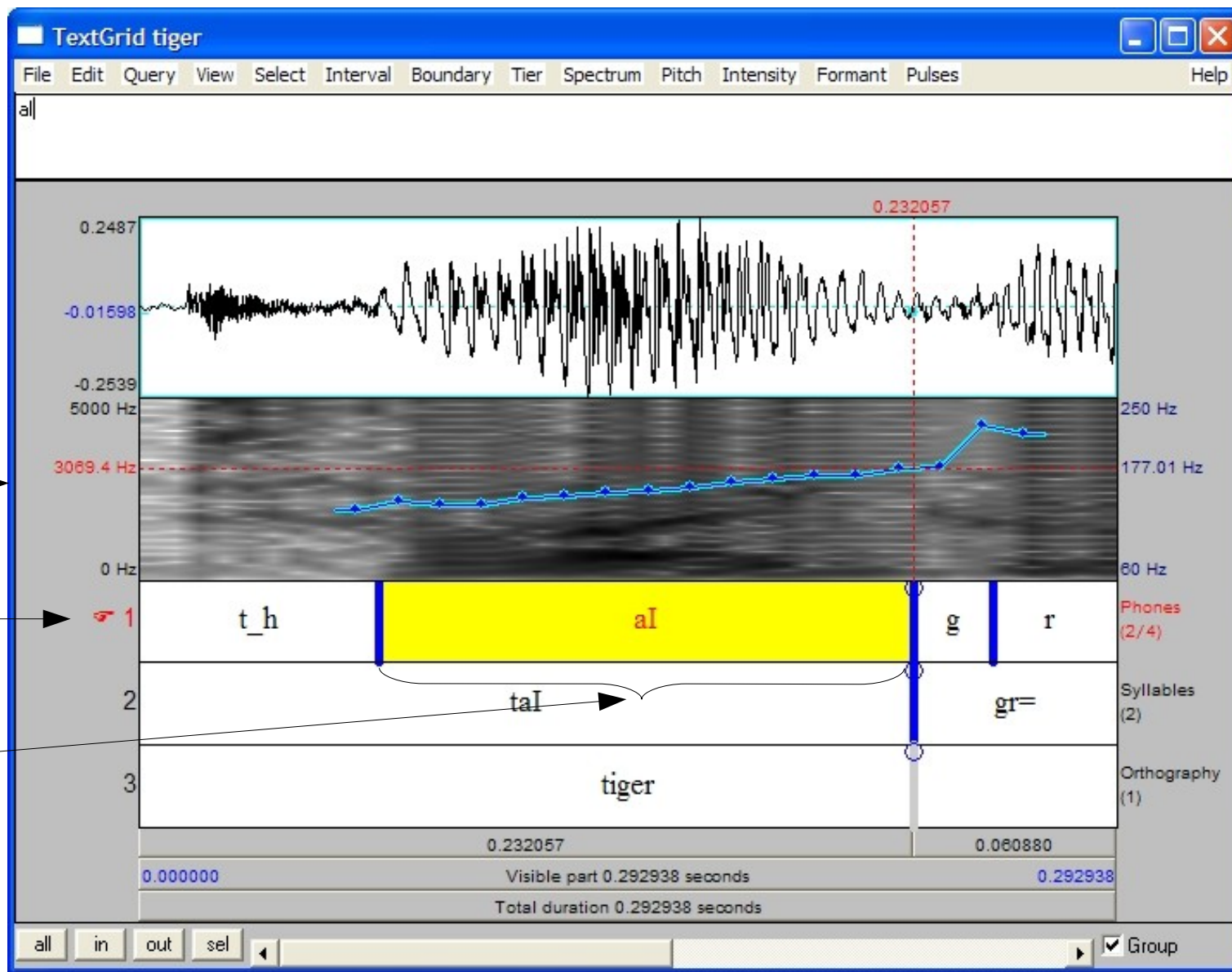
Anotacja dla syntezy CCS

REQUIRED
FOR
SPEECH
RE-SYNTHESIS

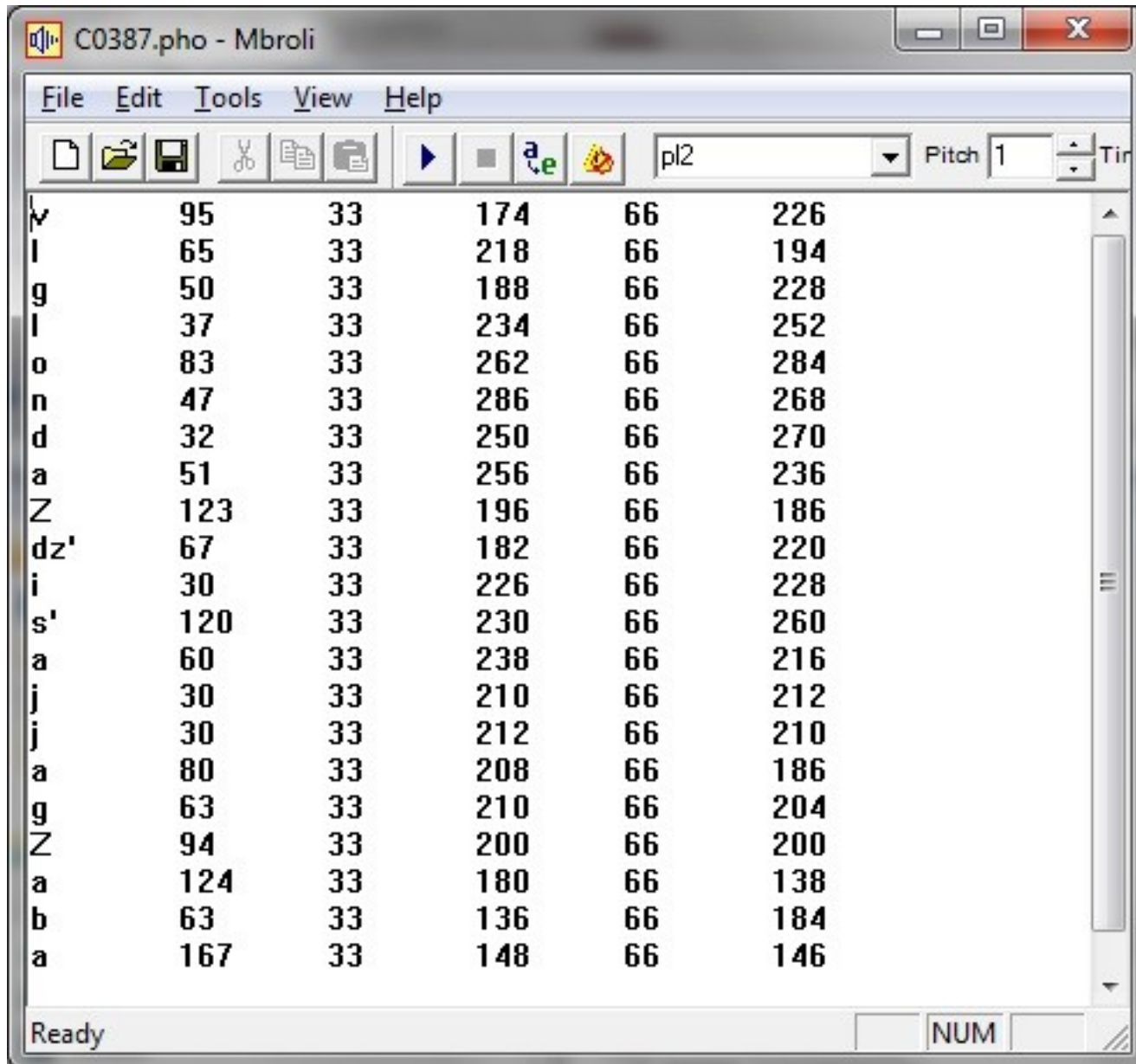
Pitch

Phones

Durations



Format pliku PHO



The screenshot shows a window titled "C0387.pho - Mbroli" with a menu bar (File, Edit, Tools, View, Help) and a toolbar. The main area contains a table of phonetic data. The status bar at the bottom shows "Ready" and "NUM".

v	95	33	174	66	226
l	65	33	218	66	194
g	50	33	188	66	228
l	37	33	234	66	252
o	83	33	262	66	284
n	47	33	286	66	268
d	32	33	250	66	270
a	51	33	256	66	236
Z	123	33	196	66	186
dz'	67	33	182	66	220
i	30	33	226	66	228
s'	120	33	230	66	260
a	60	33	238	66	216
j	30	33	210	66	212
j	30	33	212	66	210
a	80	33	208	66	186
g	63	33	210	66	204
Z	94	33	200	66	200
a	124	33	180	66	138
b	63	33	136	66	184
a	167	33	148	66	146

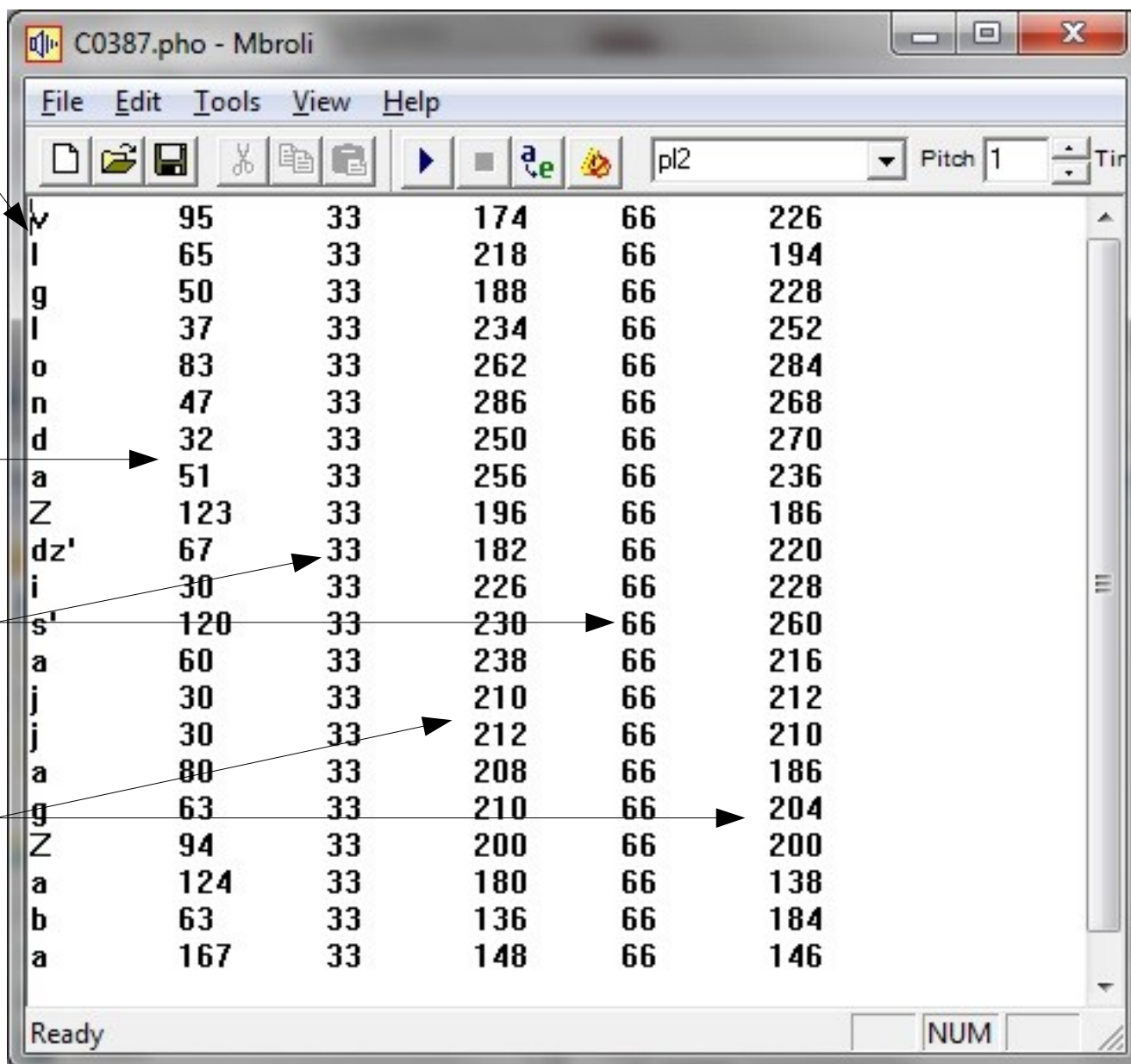
Format pliku PHO

Głoski:
SAMPA

Iloczas

Pozycja F0

Wartość F0



The screenshot shows a window titled "C0387.pho - Mbroli" with a menu bar (File, Edit, Tools, View, Help) and a toolbar. The main area contains a table of phonetic data. Annotations with arrows point to specific parts of the table: "Głoski: SAMPA" points to the first column, "Iloczas" points to the second column, "Pozycja F0" points to the third column, and "Wartość F0" points to the fourth column.

v	95	33	174	66	226
l	65	33	218	66	194
g	50	33	188	66	228
l	37	33	234	66	252
o	83	33	262	66	284
n	47	33	286	66	268
d	32	33	250	66	270
a	51	33	256	66	236
Z	123	33	196	66	186
dz'	67	33	182	66	220
i	30	33	226	66	228
s'	120	33	230	66	260
a	60	33	238	66	216
j	30	33	210	66	212
j	30	33	212	66	210
a	80	33	208	66	186
g	63	33	210	66	204
Z	94	33	200	66	200
a	124	33	180	66	138
b	63	33	136	66	184
a	167	33	148	66	146

Start z MBROLA

- bachan.speechlabs.pl/files/2008-06-03_JB_OprUzytkowe_MBROLAdownload_installation.pdf
- Wykonaj syntezę MCCS
 - Przygotuj plik PHO dla poanotowanego pliku dla języka polskiego lub angielskiego

Zadanie domowe

- Przeczytaj „Pismo” dra [Tomasza Wicherkiewicza](#)
- <http://pl.languagesindanger.eu/book-of-knowledge/writing/>

Do zobaczenia za tydzień!