Atelier EMA

-

Part II: ema2wav & Praat

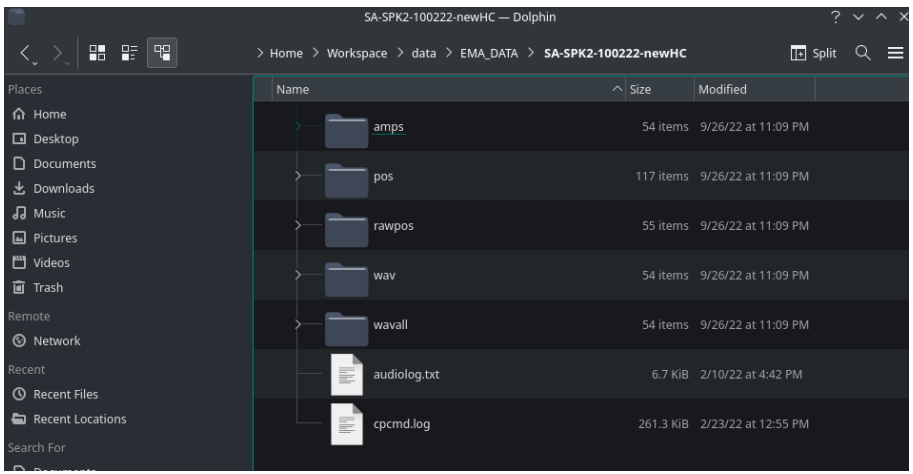
Philipp Buech

Laboratoire de Phonétique et Phonology - UMR 7018, CNRS/Sorbonne Nouvelle

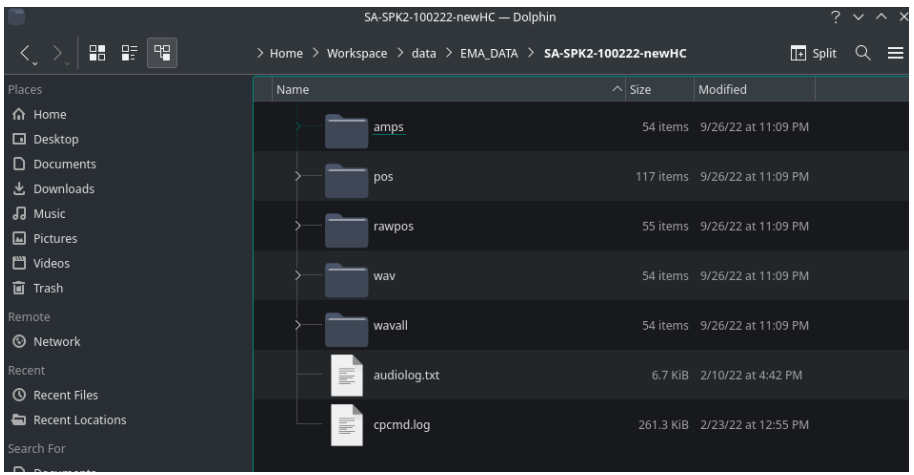
30/09/2022

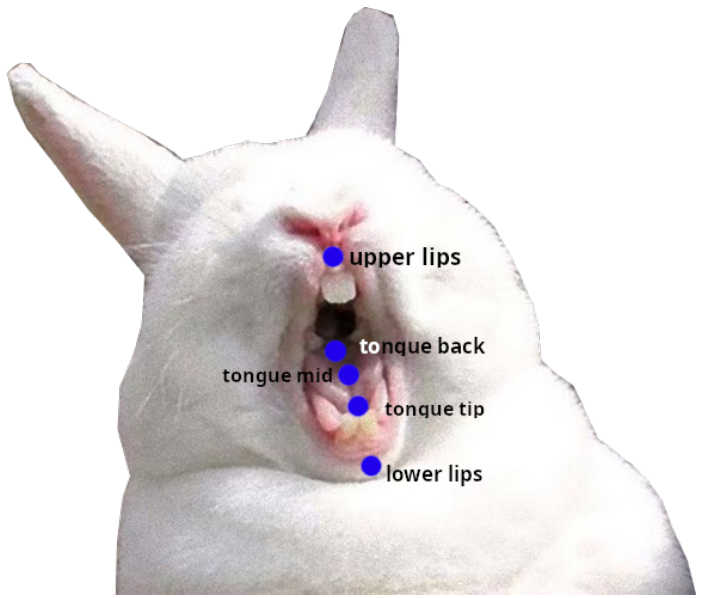
0 - Prerequisites

Data requirements



Data requirements





Software requirements

PRAAT

Software requirements

PRAAT

PYTHON

Software requirements

PRAAT

PYTHON

EMA2WAV

Python

- universal high-order programming language
- good readability
- free
- runs on:


Python - Anaconda (recommendation)

- **Anaconda distribution** for Windows, MacOS and Linux
- Python-centered, but supports also R
- easy handling of environments
- powerful package manager
- latest stable packages

ema2wav

- converts EMA data into multichannel WAV-files (& CSVs)
- Python-based, only free & open-source dependencies
- supports AG500/501 (Carstens Medizinelektronik GmbH)
- multiple options for derivations & calculations
- GUI, command line or python module
- work-in-progress
- .dmg (MacOS) & script (← always latest version)
- can be obtained **here**

1 - ema2wav & Praat

Program start

ema2wav can be started either

- using the **GUI**
  - via binary (.dmg)
  - via console
- using command-line (see documentation)
- by importing as a python module (custom script/notebook/**Google colab notebook**; see documentation)

Installation (Anaconda - script version)

- open Terminal/Console/Anaconda Prompt
- create conda environment:
  `conda create -name ema_env`
- activate the environment:
  `conda activate ema_env`
- install the dependencies:
  `pip install -r`
  `path/to/ema2wav/folder/src/requirements.txt`

Run ema2wav

- for Mac: run ema2wav_ app.app (see GitHub documentation)
- script version (GUI, from terminal, ema2wav src directory assumed):
  `python ema2wav_app.py`

User input

```
{
    "ema_device_info": "AG50x",
    "export_audio+ema": true,
    "export_to_csv": false,
    "export_raw_ema": false,
    "output_directory": "/path/to/your/output/folder/",
    "ema_input_directory": "/path/to/your/ema/folder/",
    "ema_samplerate": 1250,
    "ema_channels": 16,
    "audio_input_directory": "/path/to/your/wav/folder/",
    "audio_samplerate": 48000,
    "audio_channels": 1,
    "channel_allocation": {
        "TD": 3,
        "TT": 5
    },
    "parameters_of_interest": {
        "0_TT": "y",
        "1_TD": "y-vel",
        "2_TT": "tvel",
        "3_TD": "x"
    },
    "filter": null
}
```

User input - parts of the input file

```
{
    "ema_device_info": "AG50x",
    "export_audio+ema": true,
    "export_to_csv": false,
    "export_raw_ema": false,
    "output_directory": "/path/to/your/output/folder/",
    "ema_input_directory": "/path/to/your/ema/folder/",
    "audio_input_directory": "/path/to/your/wav/folder/",
    "ema_samplerate": 1250,
    "ema_channels": 16,
    "audio_samplerate": 48000,
    "audio_channels": 1,
    "channel_allocation": {
        "TD": 3,
        "TT": 5
    },
    "parameters_of_interest": {
        "0_TT": "y",
        "1_TD": "y-vel",
        "2_TT": "tvel",
        "3_TD": "x"
    },
    "filter": null
}
```

Device info (necessary)

output options (necessary)

folder paths (necessary)

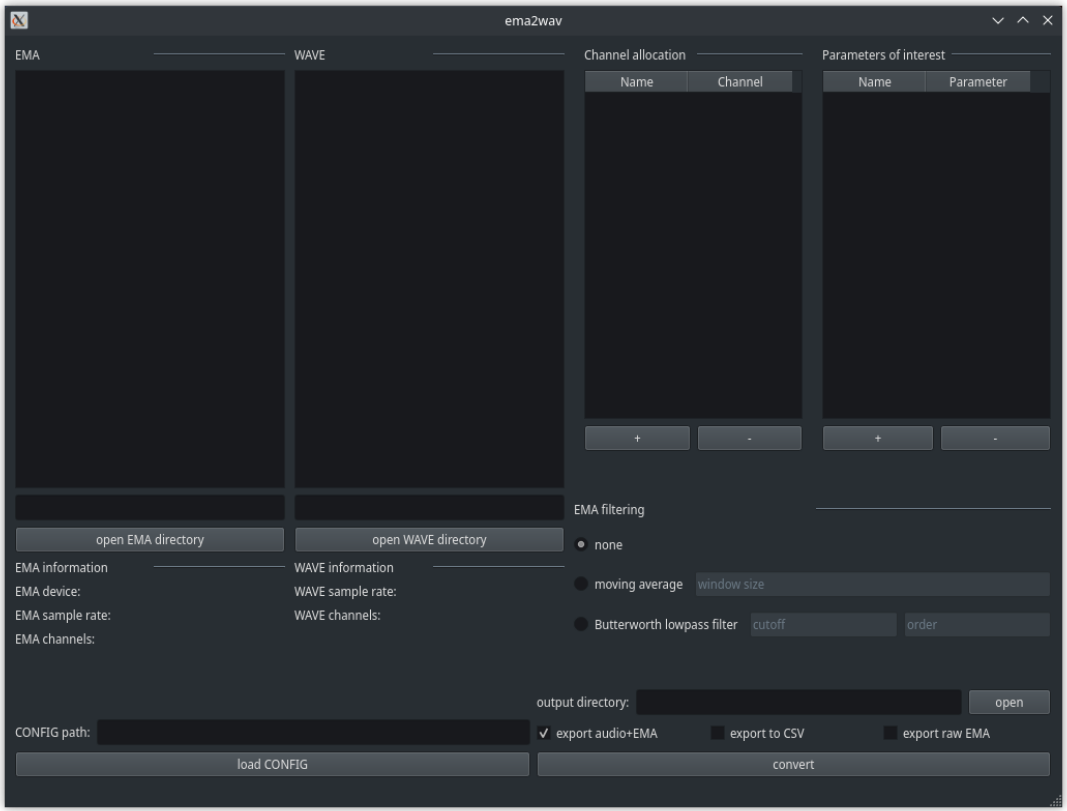
general information of the ema and audio files (optional)

channel information (necessary)

parameters of interest, consist of X_ + channel name & the parameter to extract (necessary)

filter information (necessary)

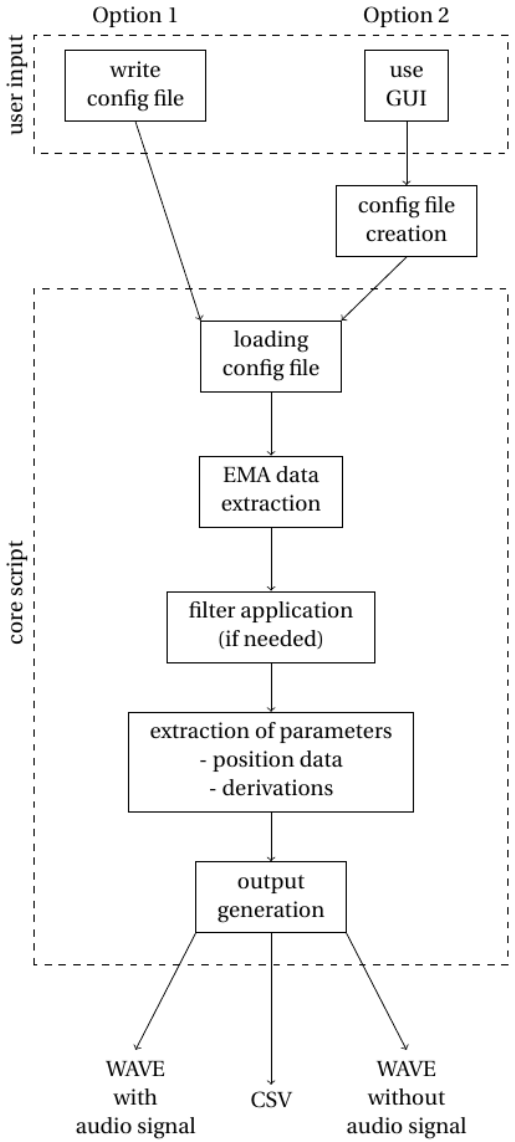
Start of the conversion process (GUI)



Start of the conversion process (manual)

- from terminal (ema2wav src directory assumed)
  ```
  python convert.py config.json
  ```
- as Python module:
  ```
  import ema2wav_core as ec
  config_file =
  "/path/to/your/config_file.json"
  ec.ema2wav_conversion(config_file)
  ```

Conversion process



Conversion process - reshaping of the data

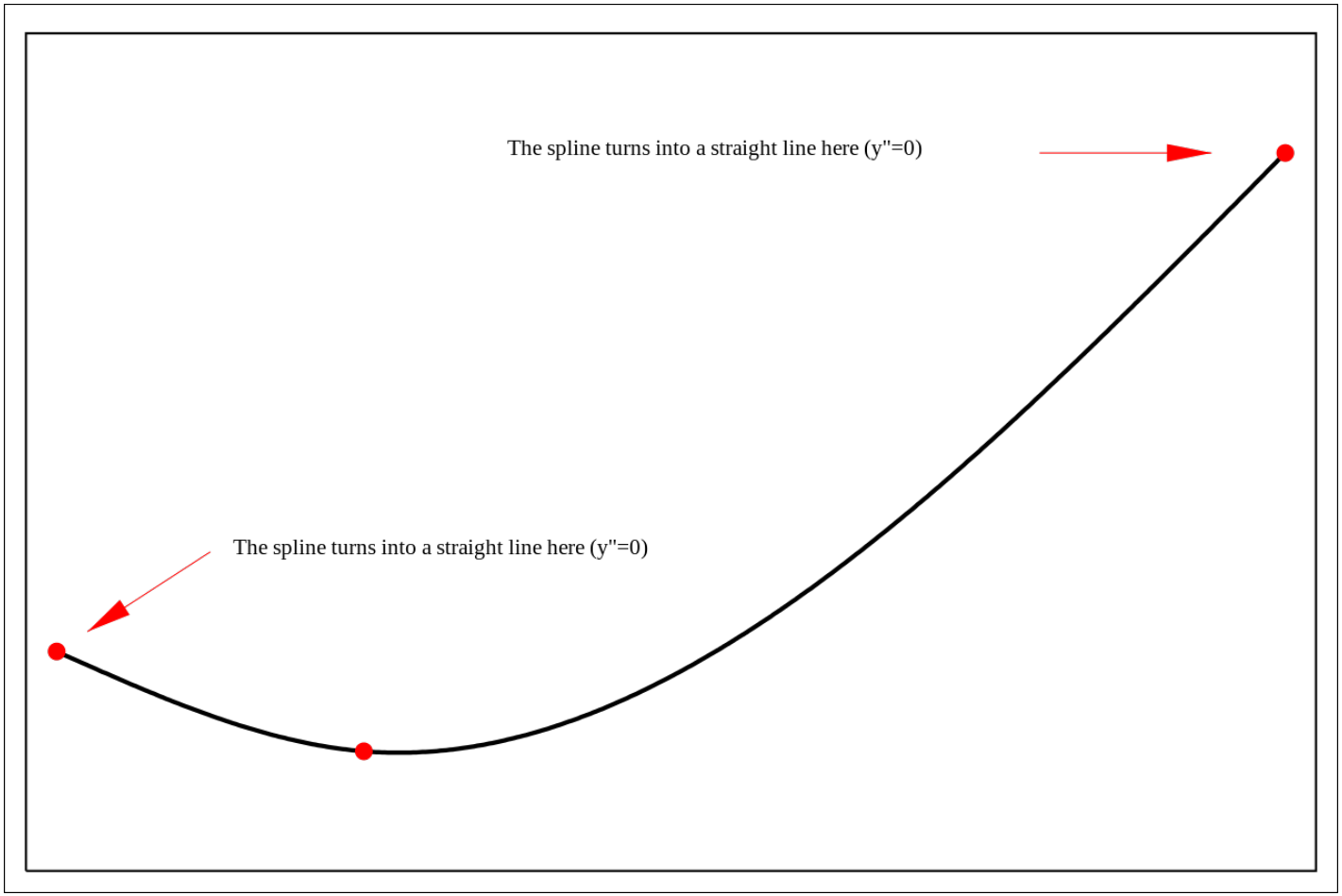
sample 1									sample 2					sample n
channel 1				channel 2				channel n	channel 1				channel n	⋯
x	y	z	⋯	x	y	z	⋯	⋯	x	y	z	⋯	⋯	⋯

sample n							
	x	y	z	phi	theta	rms	empty
channel 1	⋯	⋯	⋯	⋯	⋯	⋯	⋯
channel 2	⋯	⋯	⋯	⋯	⋯	⋯	⋯
channel n	⋯						

sample 2							
	x	y	z	phi	theta	rms	empty
channel 1	value	value	value	value	value	value	value
channel 2	value	value	value	value	value	value	value
channel n	⋯						

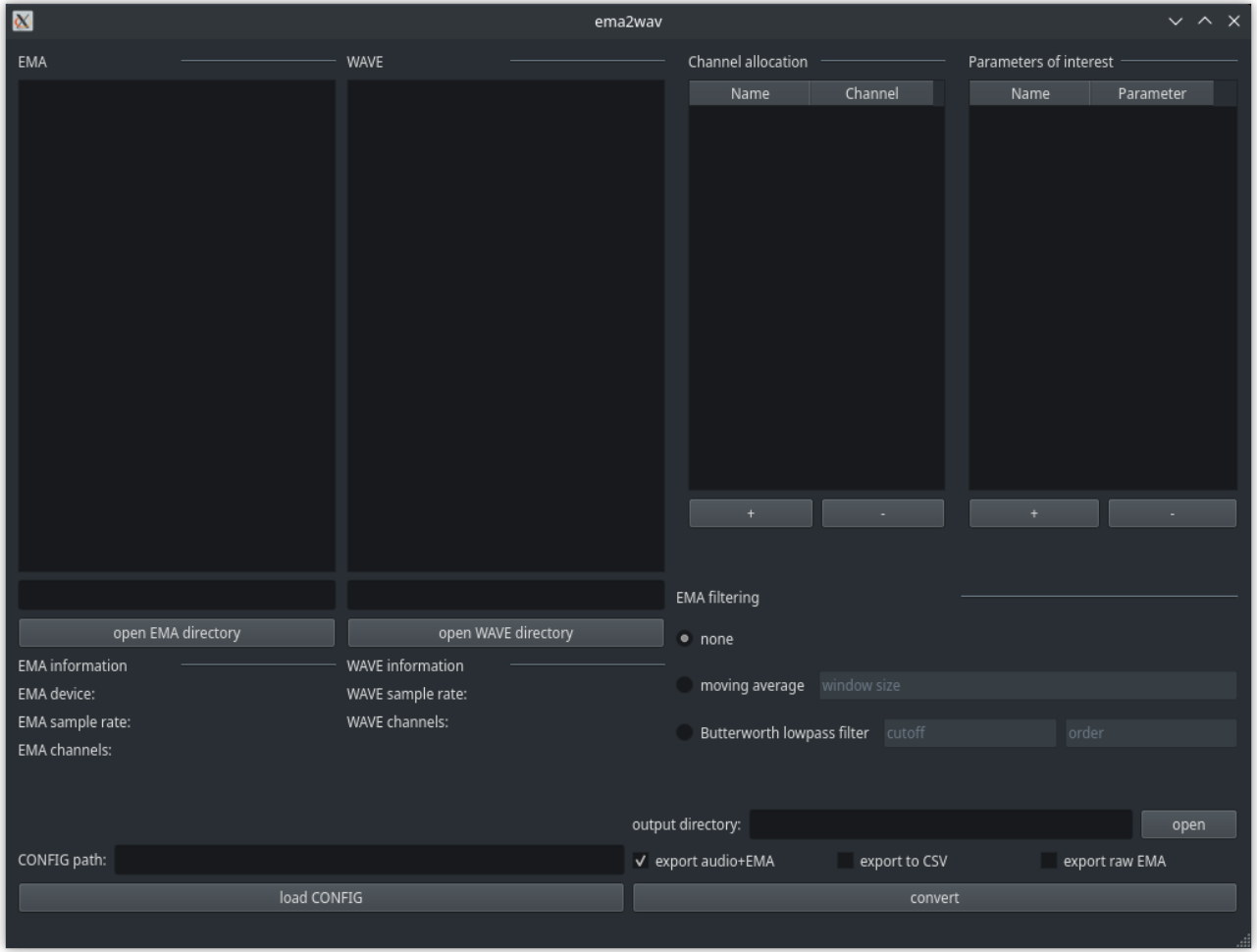
sample 1							
	x	y	z	phi	theta	rms	empty
channel 1	value	value	value	value	value	value	value
channel 2	value	value	value	value	value	value	value
channel n	⋯						

Conversion process - Interpolation



(https://en.wikipedia.org/wiki/Spline_interpolation/media/File:Cubic_splines_three_points.svg

Conversion process - GUI approach



Conversion process - GUI approach (2)

- open the folders containing .pos and .wav files
- channel allocation: enter name channel number
- parameters of interest: enter channel name parameters
- select filter (if necessary)
- open output folder
- start the conversion

Conversion process - GUI approach (3)

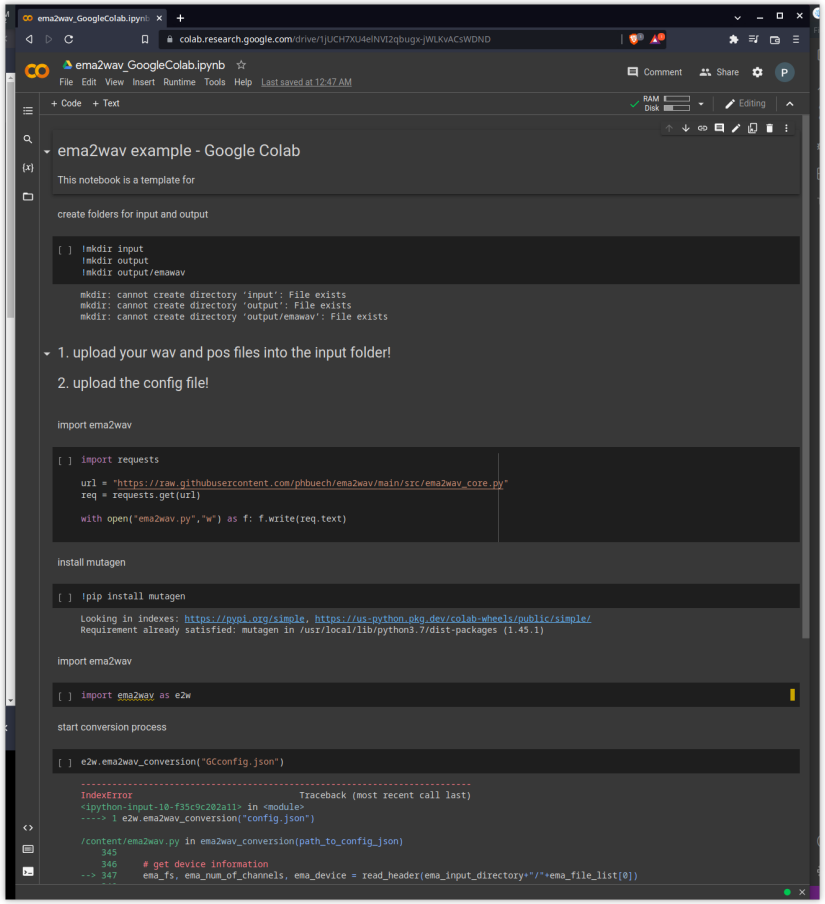
- config file as documentation
- can be used for replicating the conversion process (load CONFIG)

Conversion process - Google Colab approach

- Google Colab
- execute Jupyter Notebooks online
- free easy-to-use
- see example notebook

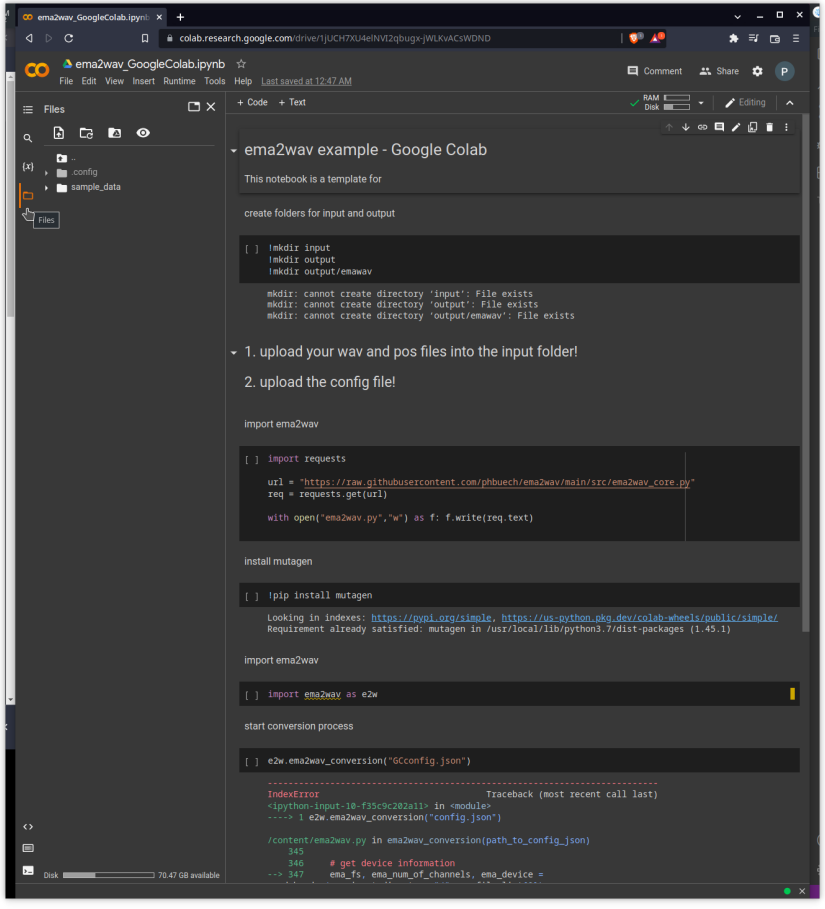
Conversion process - Google Colab (1)

- open the example notebook in Google Colab



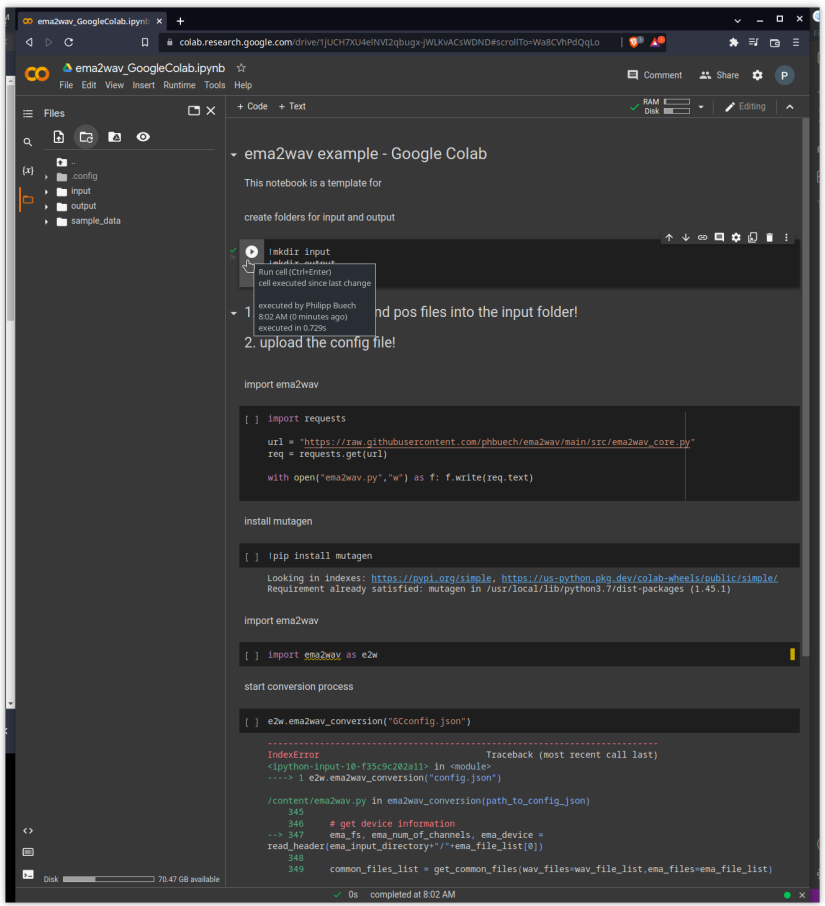
Conversion process - Google Colab (2)

- open the files folder



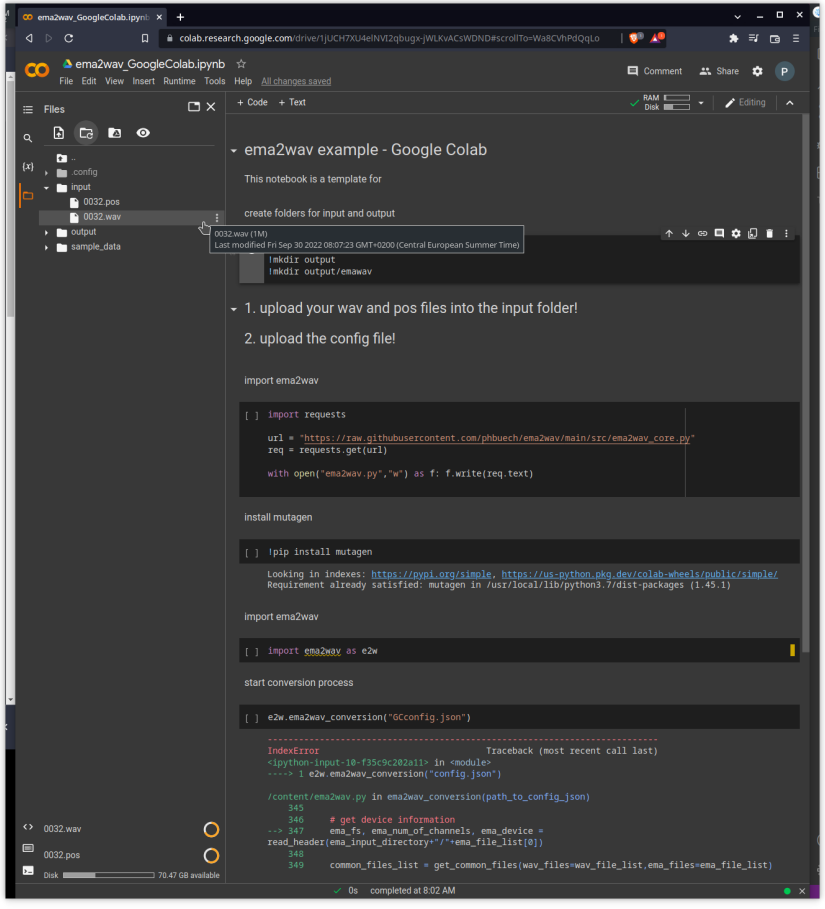
Conversion process - Google Colab (3)

- run the first code snippet



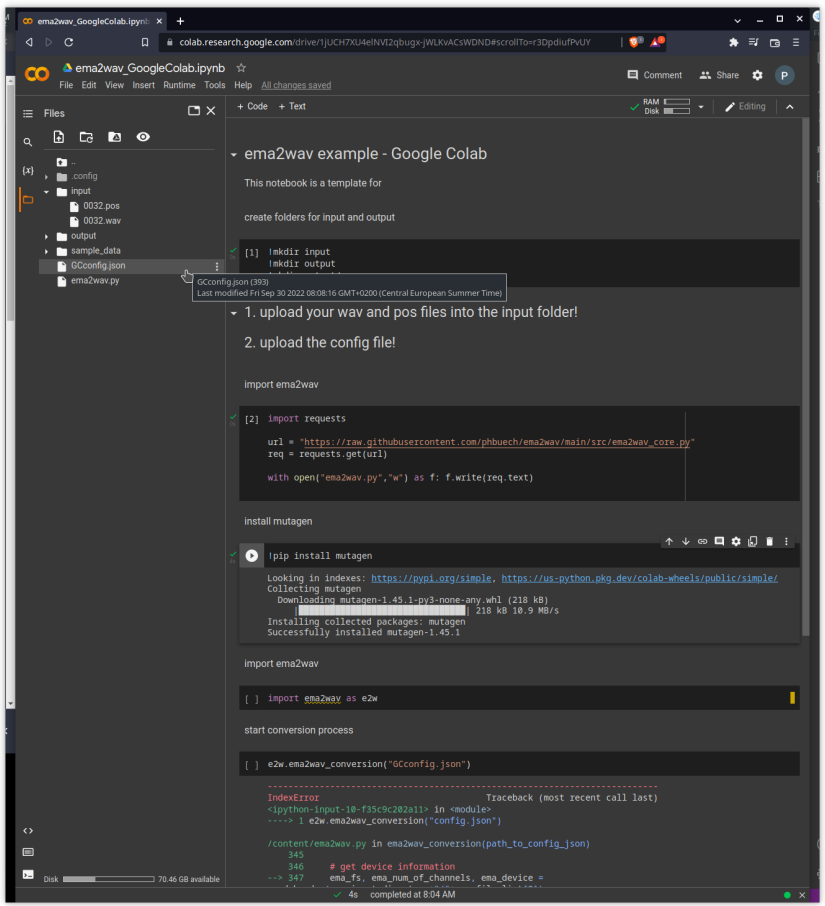
Conversion process - Google Colab (4)

- upload your .pos and .wav files



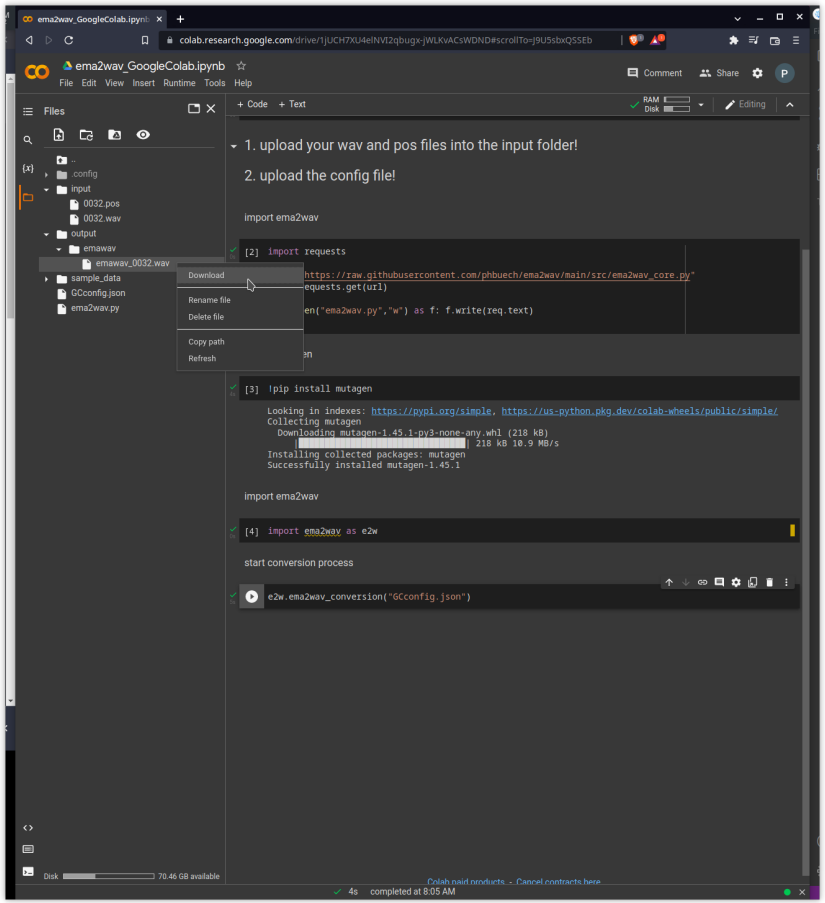
Conversion process - Google Colab (5)

- upload the GC config file



Conversion process - Google Colab (6)

- run the other code snippets and download converted files

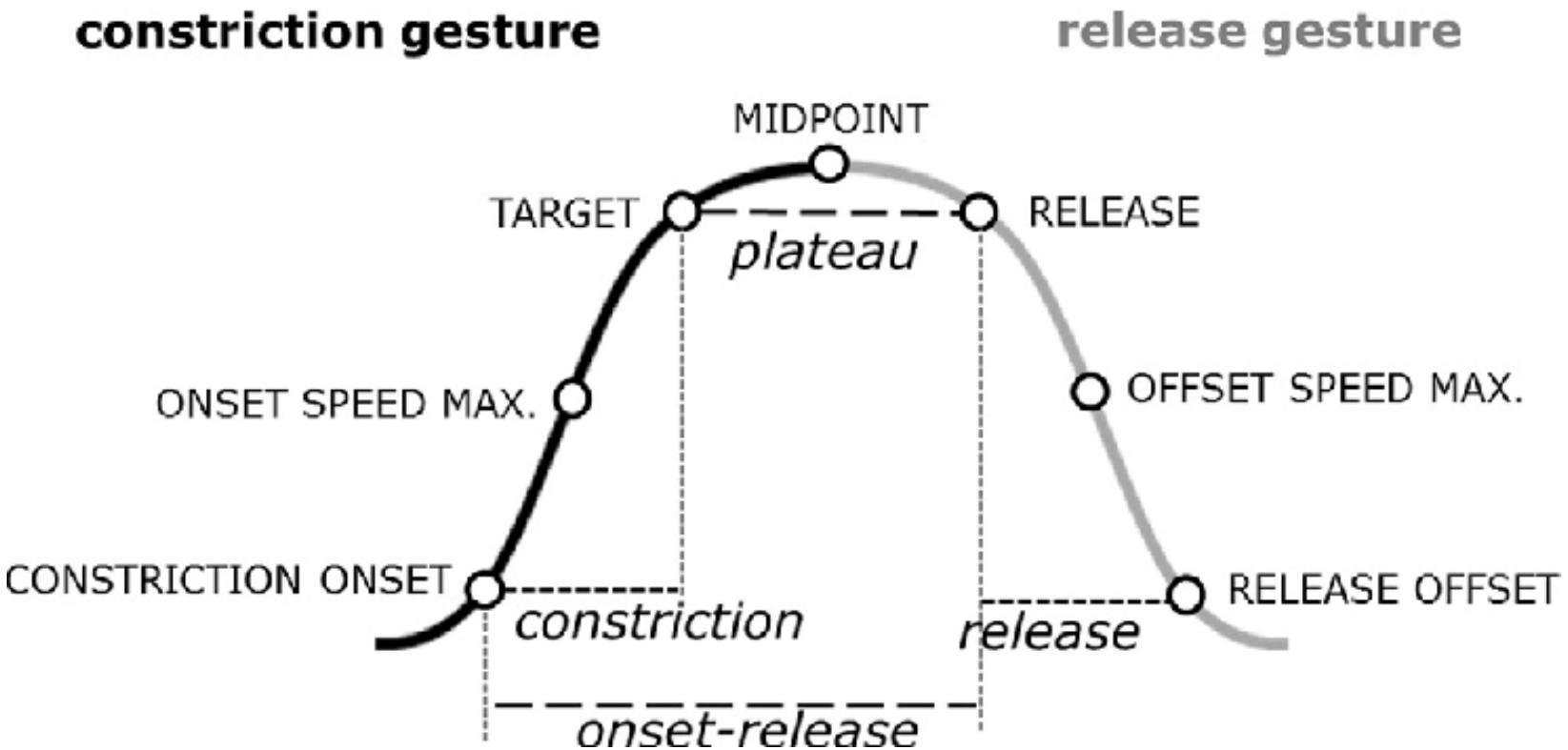


Annotation & Measurements in Praat

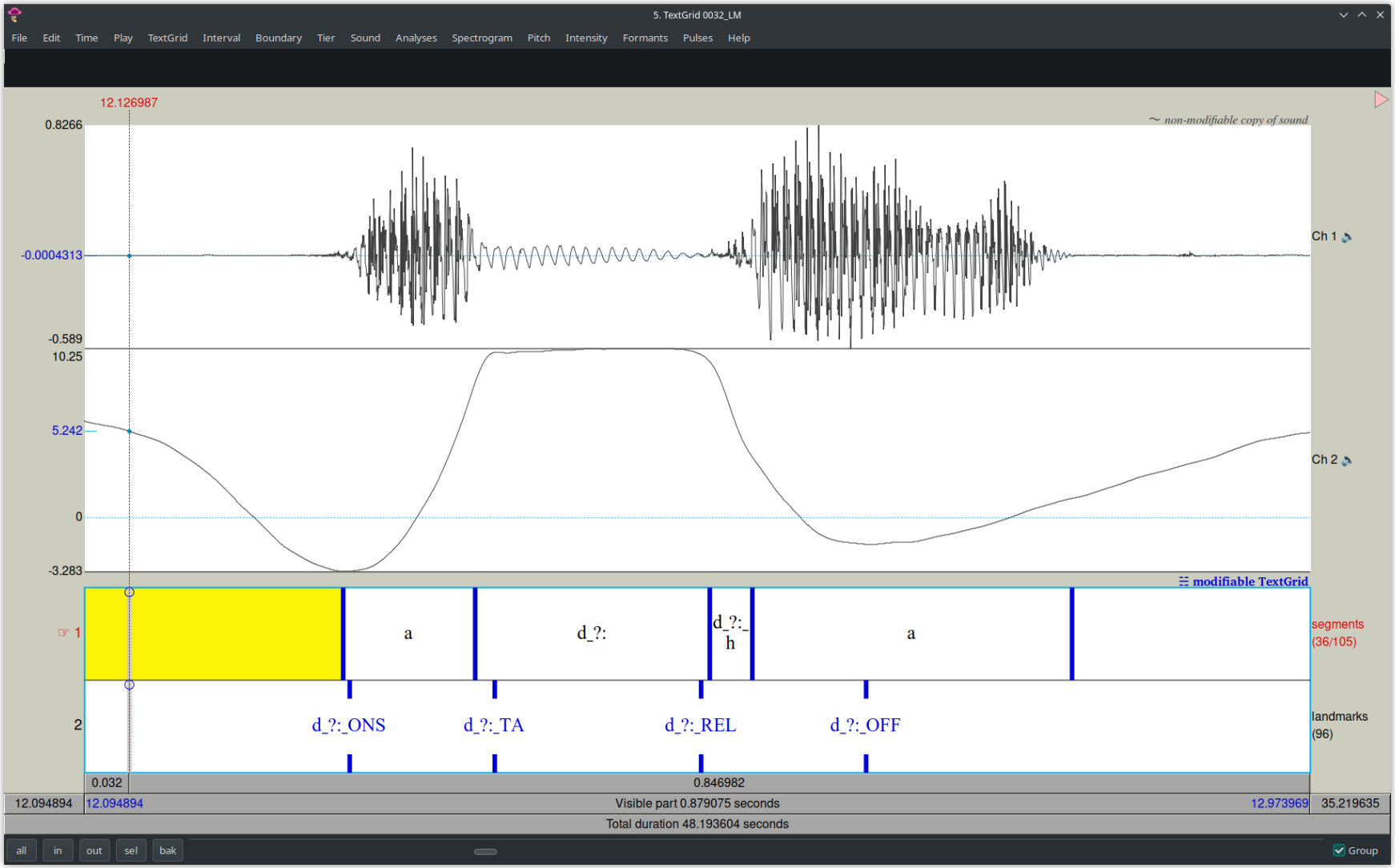
Praat tweaks

- Display default settings in Praat are not suitable for annotating EMA data
- Options for the best annotation experience:
  - disable the spectrogram
  - Change sound scaling in the editor window:
    (Sound > Sound scaling... > select ′by window and channel′)
  - Mute channels (if necessary):
    (Sound > Mute channels... > (ranges) > enter 2:X)

Landmark annotations



Gestural landmarks and intragestural intervals for a typical gestural complex (Tilsen, 2014)


5. TextGrid 0032_LM
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
12.126987
0.8266
-0.0004313
-0.589
non-modifiable copy of sound
Ch 1
10.25
5.242
0
-3.283
Ch 2
modifiable TextGrid
1
a
d_?:
d_?:_h
a
segments (36/105)
2
d_?:_ONS
d_?:_TA
d_?:_REL
d_?:_OFF
landmarks (96)
0.032
0.846982
12.094894
12.094894
Visible part 0.879075 seconds
12.973969
35.219635
Total duration 48.193604 seconds
all in out sel bak
Group

- Annotations of gestural landmarks using Point Tiers in Praat
- retrieve timing information as usual
- measure amplitude information by calling ′Get value at time′ applied to the Sound object
- see **measurement_example.praat**

Remarks

- never use the entire file for frequency/intensity/spectral measurements!
  Instead: extract the the audio track and use that for acoustic measurements
- make use of a good documentation of your files

3 - Hacks

Annotations

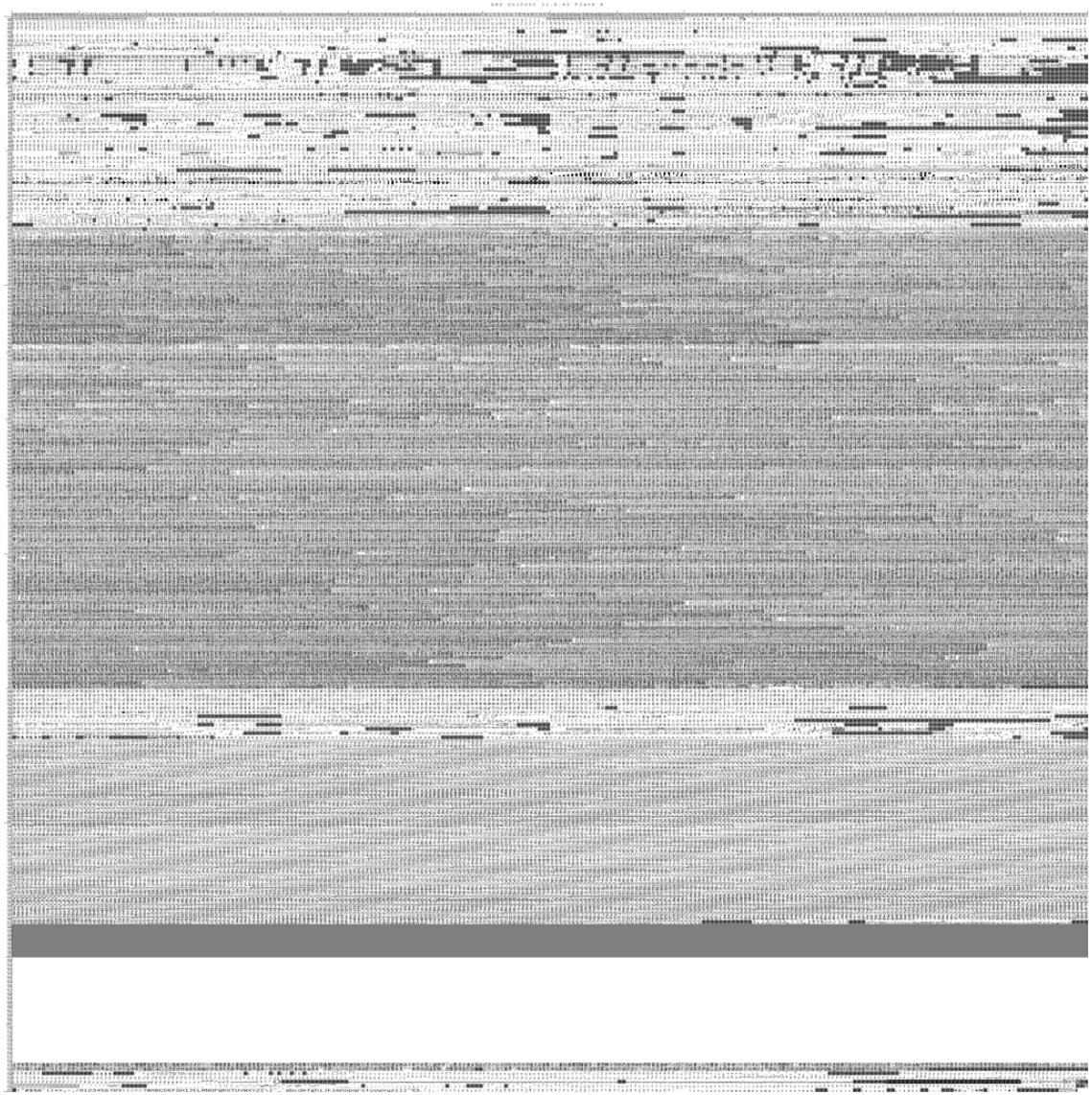
Why use IPA for annotations?

Encoding

- IPA symbols in Unicode (latin character, IPA extensions, Phonetic extensions)
- Whats the problem?
- different character encoding schemes:
  - ASCII
  - UTF-8
  - UTF-16
  - UTF-32

ASCII Table

Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char
0	0	0		32	20	40	[space]	64	40	100	@	96	60	140	`
1	1	1		33	21	41	!	65	41	101	A	97	61	141	a
2	2	2		34	22	42	"	66	42	102	B	98	62	142	b
3	3	3		35	23	43	#	67	43	103	C	99	63	143	c
4	4	4		36	24	44	$	68	44	104	D	100	64	144	d
5	5	5		37	25	45	%	69	45	105	E	101	65	145	e
6	6	6		38	26	46	&	70	46	106	F	102	66	146	f
7	7	7		39	27	47	'	71	47	107	G	103	67	147	g
8	8	10		40	28	50	(	72	48	110	H	104	68	150	h
9	9	11		41	29	51	)	73	49	111	I	105	69	151	i
10	A	12		42	2A	52	*	74	4A	112	J	106	6A	152	j
11	B	13		43	2B	53	+	75	4B	113	K	107	6B	153	k
12	C	14		44	2C	54	,	76	4C	114	L	108	6C	154	l
13	D	15		45	2D	55	-	77	4D	115	M	109	6D	155	m
14	E	16		46	2E	56	.	78	4E	116	N	110	6E	156	n
15	F	17		47	2F	57	/	79	4F	117	O	111	6F	157	o
16	10	20		48	30	60	0	80	50	120	P	112	70	160	p
17	11	21		49	31	61	1	81	51	121	Q	113	71	161	q
18	12	22		50	32	62	2	82	52	122	R	114	72	162	r
19	13	23		51	33	63	3	83	53	123	S	115	73	163	s
20	14	24		52	34	64	4	84	54	124	T	116	74	164	t
21	15	25		53	35	65	5	85	55	125	U	117	75	165	u
22	16	26		54	36	66	6	86	56	126	V	118	76	166	v
23	17	27		55	37	67	7	87	57	127	W	119	77	167	w
24	18	30		56	38	70	8	88	58	130	X	120	78	170	x
25	19	31		57	39	71	9	89	59	131	Y	121	79	171	y
26	1A	32		58	3A	72	:	90	5A	132	Z	122	7A	172	z
27	1B	33		59	3B	73	;	91	5B	133	[	123	7B	173	{
28	1C	34		60	3C	74	<	92	5C	134	\	124	7C	174	|
29	1D	35		61	3D	75	=	93	5D	135	]	125	7D	175	}
30	1E	36		62	3E	76	>	94	5E	136	^	126	7E	176	~
31	1F	37		63	3F	77	?	95	5F	137	_	127	7F	177	



character	encoding	bits
A	UTF-8	01000001
A	UTF-16	00000000 01000001
A	UTF-32	00000000 00000000 00000000 01000001
あ	UTF-8	11100011 10000001 10000010
あ	UTF-16	00110000 01000010
あ	UTF-32	00000000 00000000 00110000 01000010

Ç

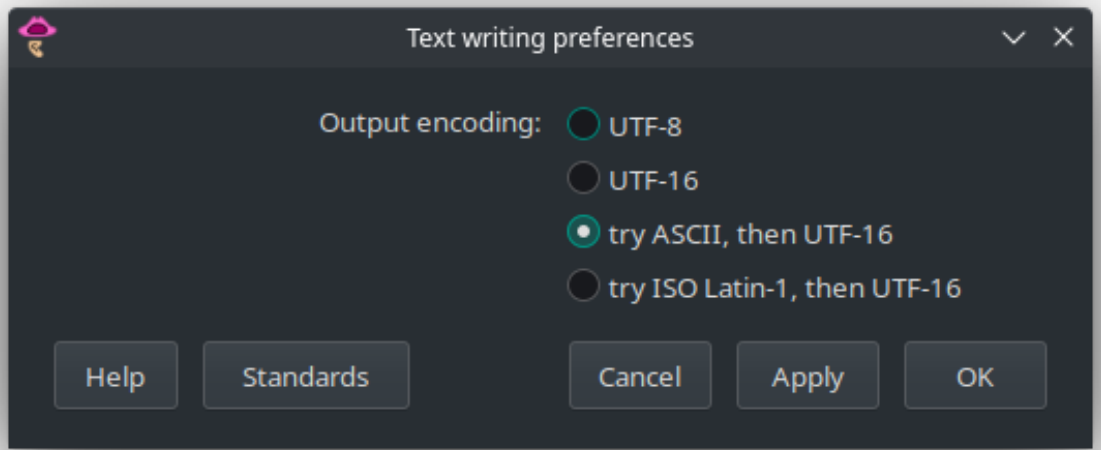
- How many glyphs?

ç

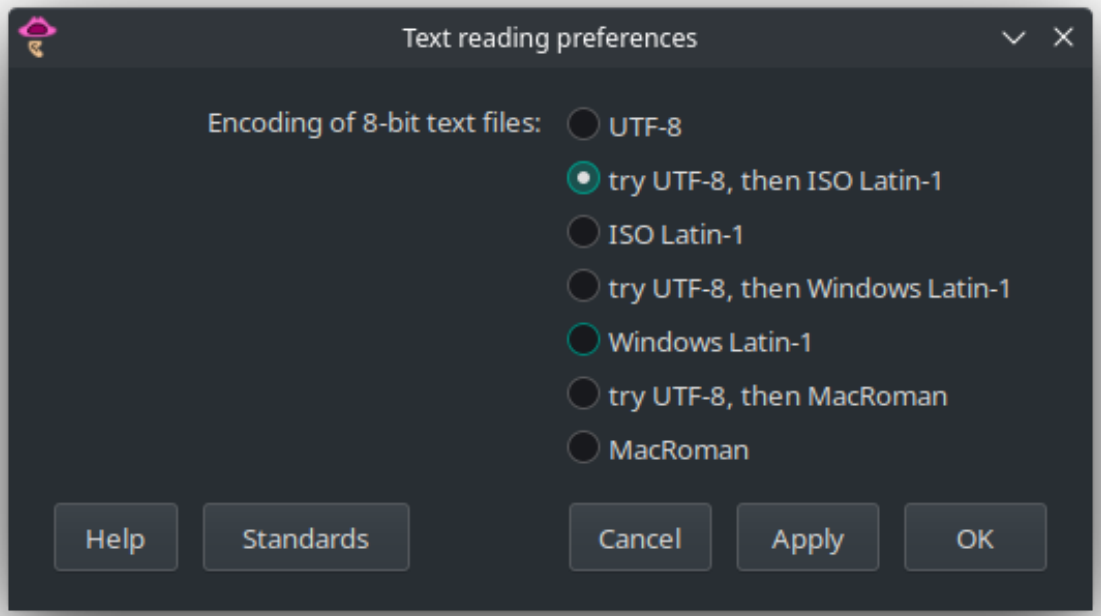
- How many glyphs?
- U+00E7
- c (U+0063) + ¸ (U+0328)

- ASCII and UTF-8 are compatible
- UTF-16 is not compatible with UTF-8
- but:
  - UTF-8 with or without BOM
  - UTF-16 Big Endian (Mac) vs UTF-16 Little Endian (Windows) (UNIX vs NUXI)
- Why is this important?
  - archiving of your data
  - data sharing with collaborators

Text encoding in Praat (text writing)



Text encoding in Praat (text reading)



Recommendations

- uniform encoding
- X-Sampa

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC) — WITH X-SAMPA EQUIVALENTS IN BLUE

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p p b b			t t d d		ʈ t` ɖ d`	c c ɟ J\	k k ɡ g	ɢ q G G\		ʔ ?
Nasal	m m	ɱ F		n n		ɳ n`	ɲ J	ŋ N	ɴ N\		
Trill	ʙ B\			r r					ʀ R\		
Tap or Flap				ɾ 4		ɽ r`					
Fricative	ɸ p\ β B	f f v v	θ T ð D	s s z z	ʃ S ʒ Z	ʂ s` ʐ z`	ç C ʝ j\	x x ɣ G	χ X ʁ R	ħ X\ ʕ ?\	h h ɦ h\
Lateral fricative				ɬ K	ɮ K\						
Approximant		ʋ P (or v\)		ɹ r\		ɻ r\`	j j	ɰ M\			
Lateral approximant				l l		ɭ l`	ʎ L	ʟ L\			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
ʘ O\ Bilabial	ɓ Bilabial b_<	’ _> as in:
ǀ |\ Dental	ɗ Dental/alveolar d_<	pʼ Bilabial p_>
! !\ (Post)alveolar	ʄ Palatal J_<	tʼ Dental/alveolar t_>
ǂ =\ Palatoalveolar	ɠ Velar g_<	kʼ Velar k_>
ǁ ||\ Alveolar lateral	ʛ Uvular G_<	sʼ Alveolar fricative s_>

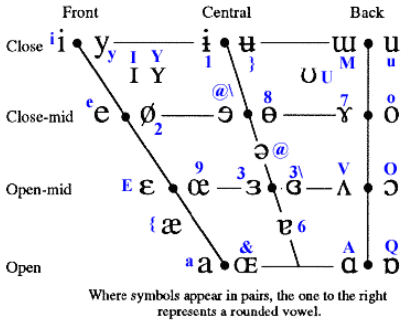
SUPRASEGMENTALS

ˈ "	Primary stress	%foUn@"tIS@n	ˌfoʊnəˈtɪʃən
%ˌ	Secondary stress		
ː :	Long	eː e:	
ˑ :\	Half-long	eˑ e:\	
˘ _X	Extra-short	ĕ e_X	
. .	Syllable break	ɹi.ækt	ɹi.{kt
| |	Minor (foot) group		
‖ ||	Major (intonation) group		
‿ -\	Linking (absence of a break)		

TONES & WORD ACCENTS

LEVEL		CONTOUR	
e̋ or 7	Extra high _T	ě or /	Rising _R
é	1 High _H	ê	\ Falling _F
ē	- Mid _M	e᷄	7 High rising _H_T
è	_ Low _L	e᷅	/ Low rising _B_L
ȅ	4 Extra low _B	e᷈	1 Rising-falling _R_F
↓ Downstep !		↗ Global rise	etc. <R>
↑ Upstep ^		↘ Global fall	<F>

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

DIACRITICS

X-SAMPA diacritics come after symbols, e.g. n_0
Diacritics may be placed above a symbol with a descender, e.g. n̊j

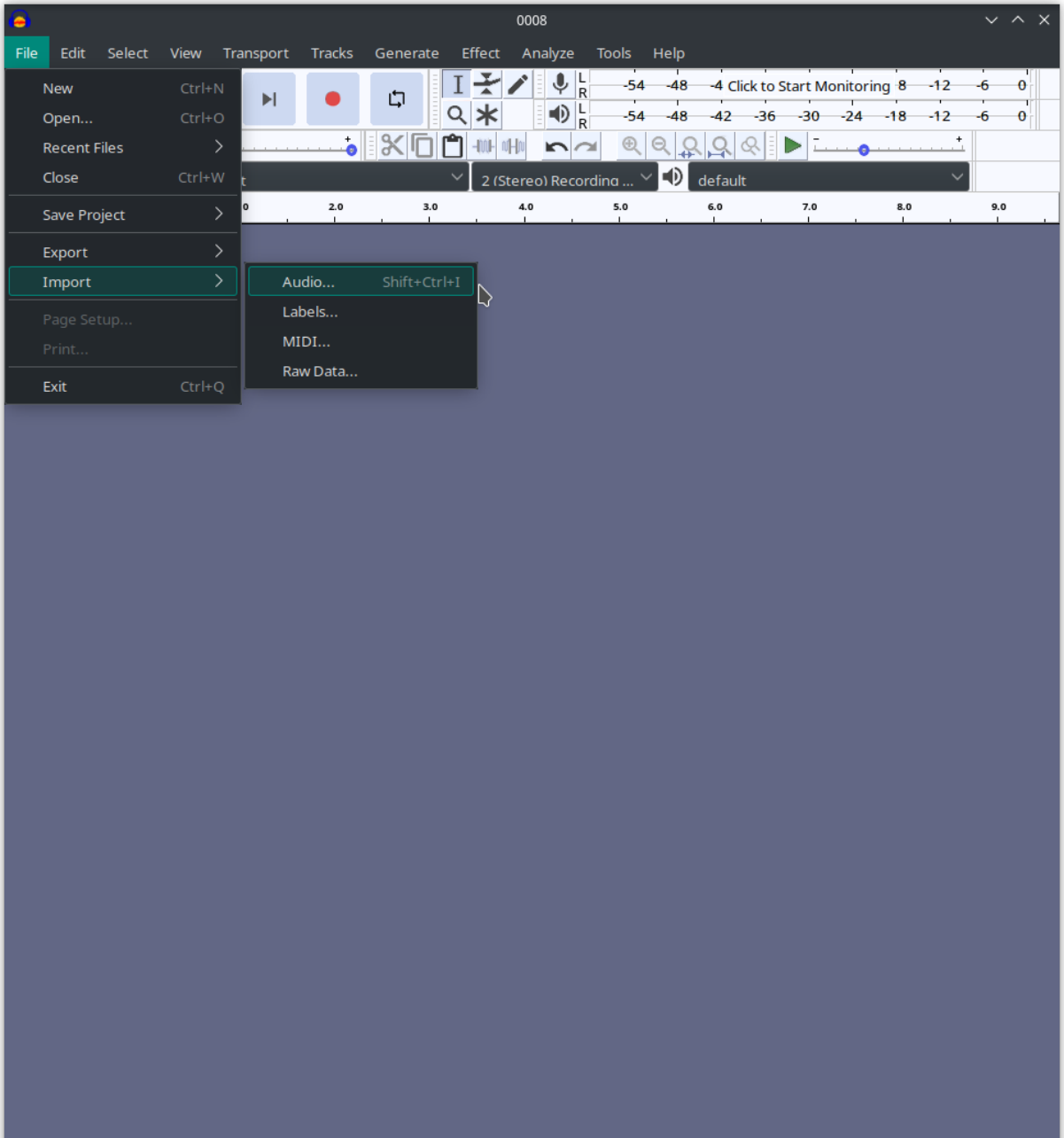
_0 ◌̥	Voiceless	n̥ d̥	_t ◌̤	Breathy voiced	b̤ a̤	_d ◌̪	Dental	t̪ d̪
_v ◌̬	Voiced	s̬ t̬	_k ◌̰	Creaky voiced	b̰ a̰	_a ◌̺	Apical	t̺ d̺
h _h	Aspirated	tʰ dʰ	_N ◌̼	Linguolabial	t̼ d̼	_m ◌̻	Laminal	t̻ d̻
_O ◌̹	More rounded	ɔ̹	w _w	Labialized	tʷ dʷ	~ (or _~)	Nasalized	e~ or ẽ
_c ◌̜	Less rounded	ɔ̜	j (or _j)	Palatalized	tʲ dʲ	n _n	Nasal release	dⁿ
_+ ◌̟	Advanced	u̟	Y _G	Velarized	tˠ dˠ	l _l	Lateral release	dˡ
_- ◌̠	Retracted	i̠	ʕ _?\	Pharyngealized	tˤ dˤ	~ _}	No audible release	d̚
_" ◌̈	Centralized	ë	~ _e	Velarized or pharyngealized	ɫ (or velarized l: 5)			
x _x ◌̽	Mid-centralized	e̽	_r ◌̝	Raised	e̝ (ɹ̝ = voiced alveolar fricative)			
= (or _=) ◌̩	Syllabic	ɹ̩	_o ◌̞	Lowered	e̞ (β̞ = voiced bilabial approximant)			

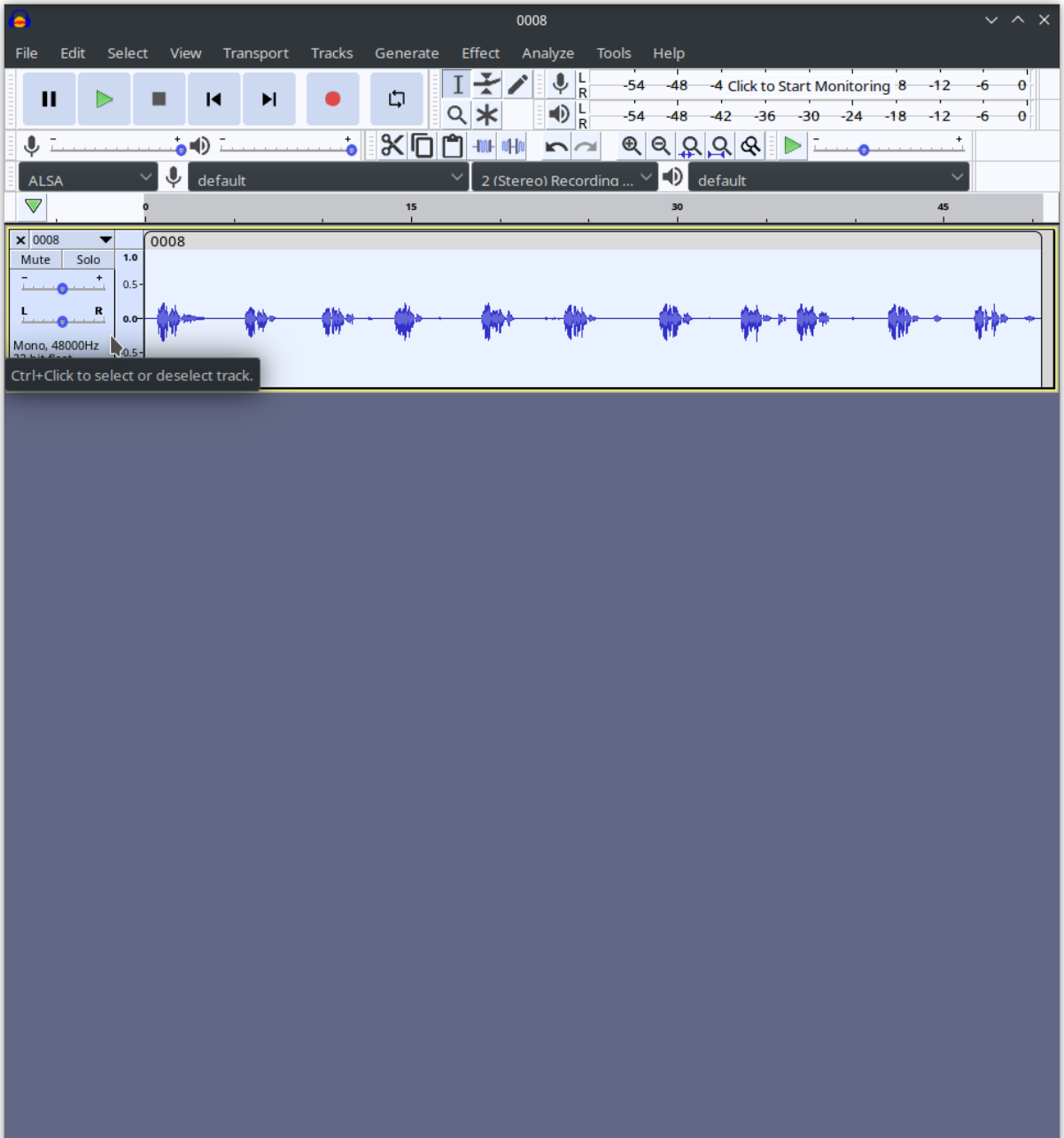
OTHER SYMBOLS

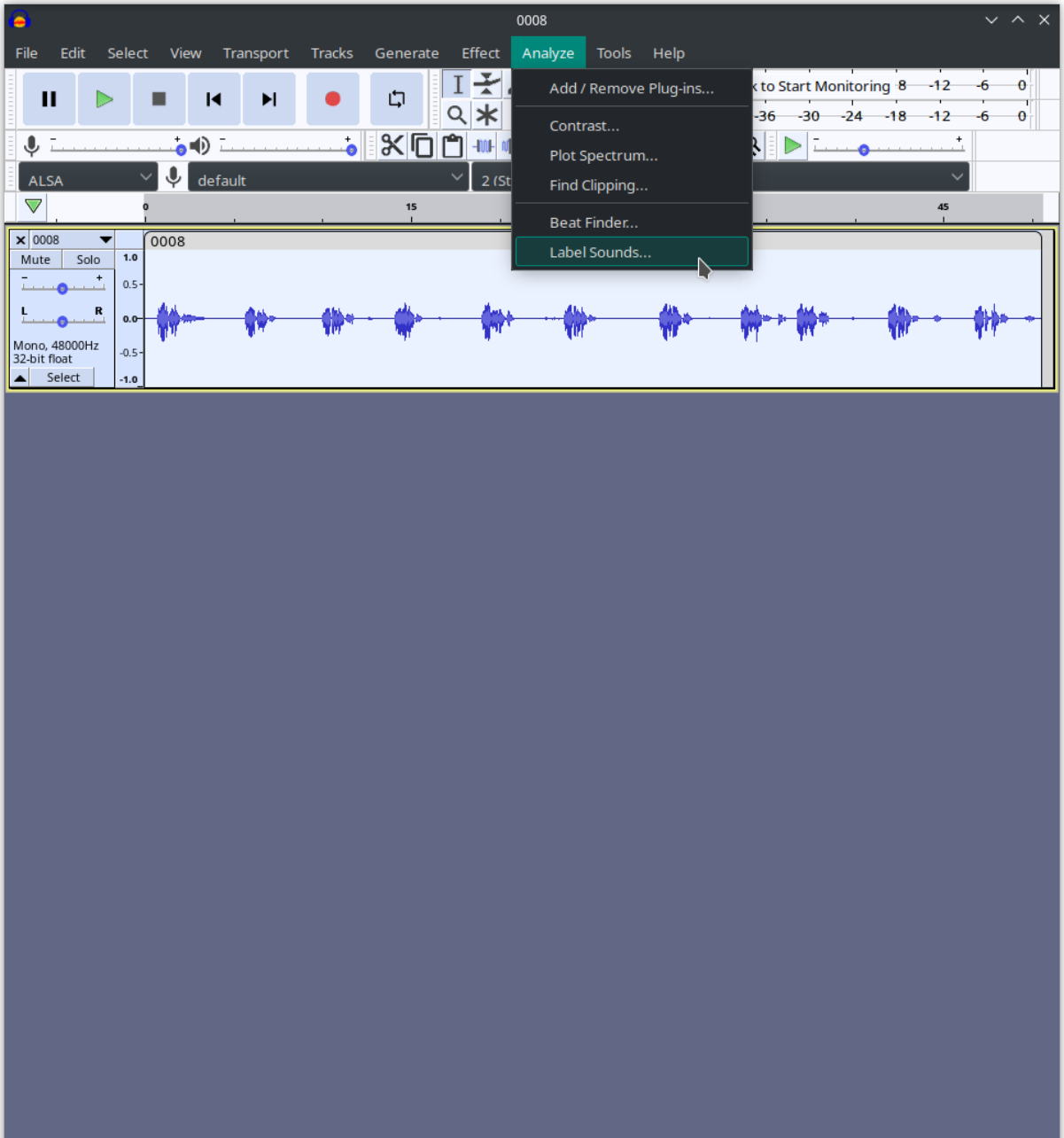
WM	Voiceless labial-velar fricative	s\ z\ ɕ ʑ	Alveolo-palatal fricatives
w W	Voiced labial-velar approximant	ɺ l\	Alveolar lateral flap
H ɥ	Voiced labial-palatal approximant	ʄ x\	Simultaneous ʃ and x
H\ H	Voiceless epiglottal fricative		
<\ ʢ	Voiced epiglottal fricative		

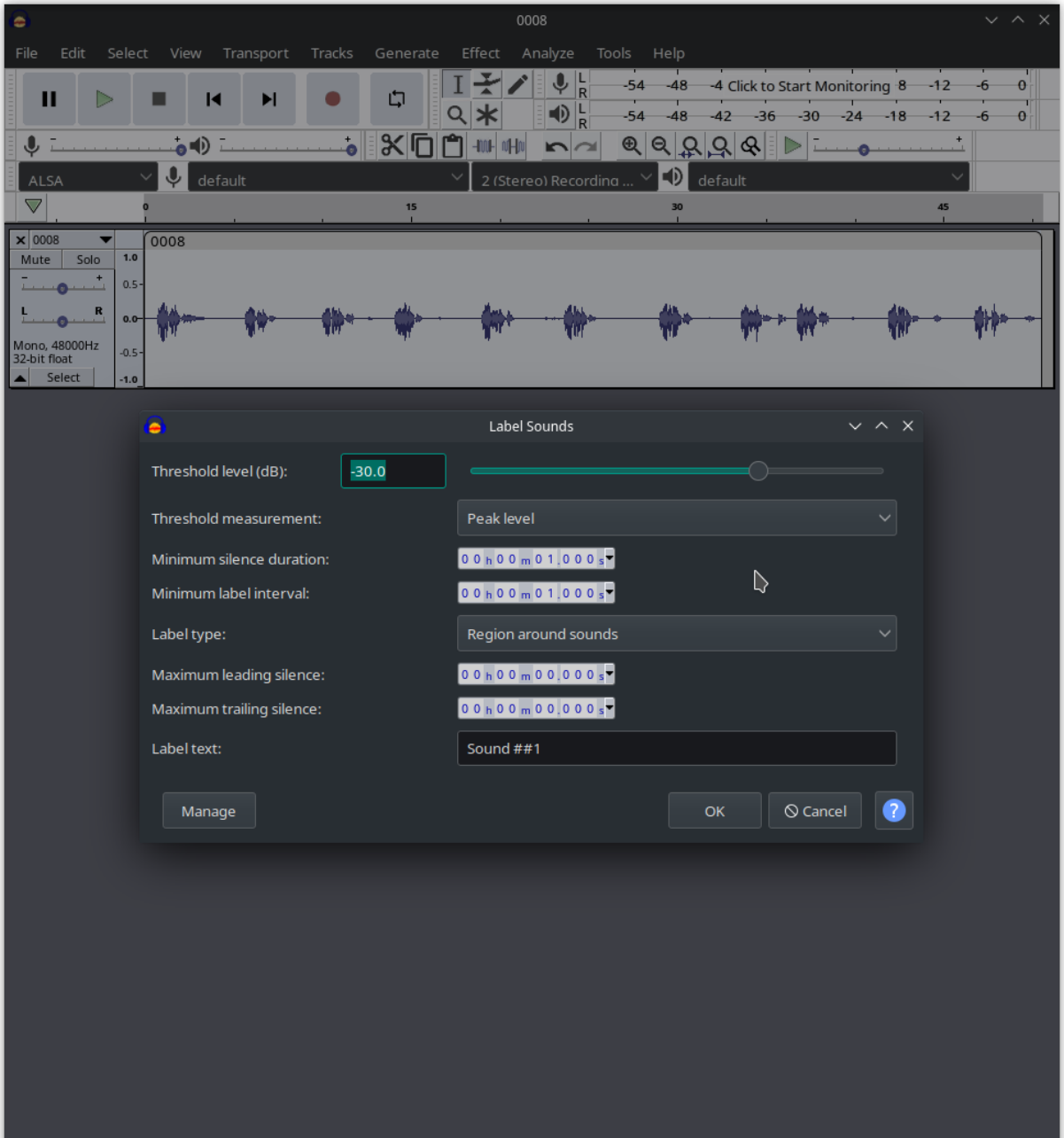
Affricates and double articulations can be represented by two symbols joined by a tie bar if

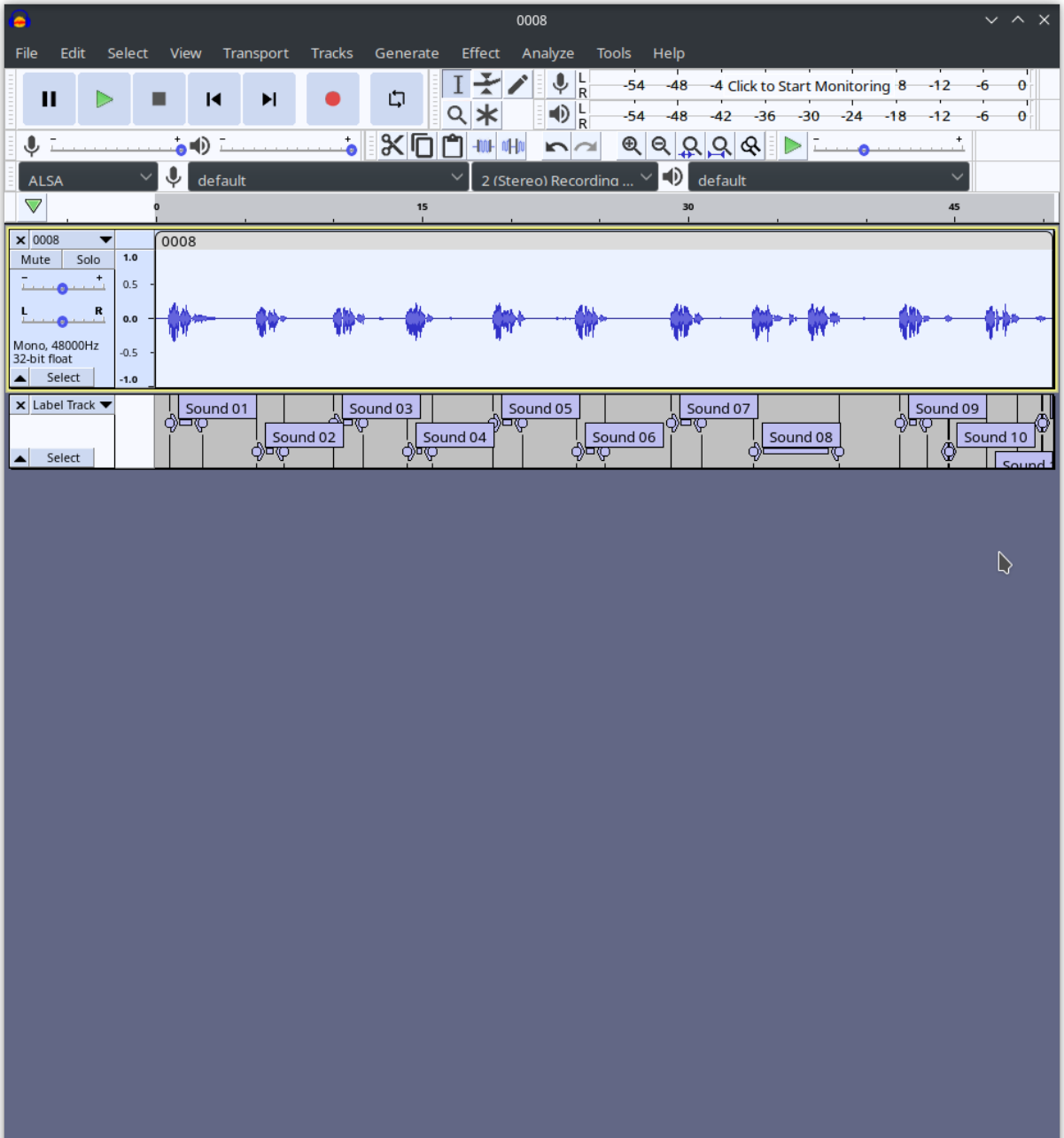
Hack for forced alignment

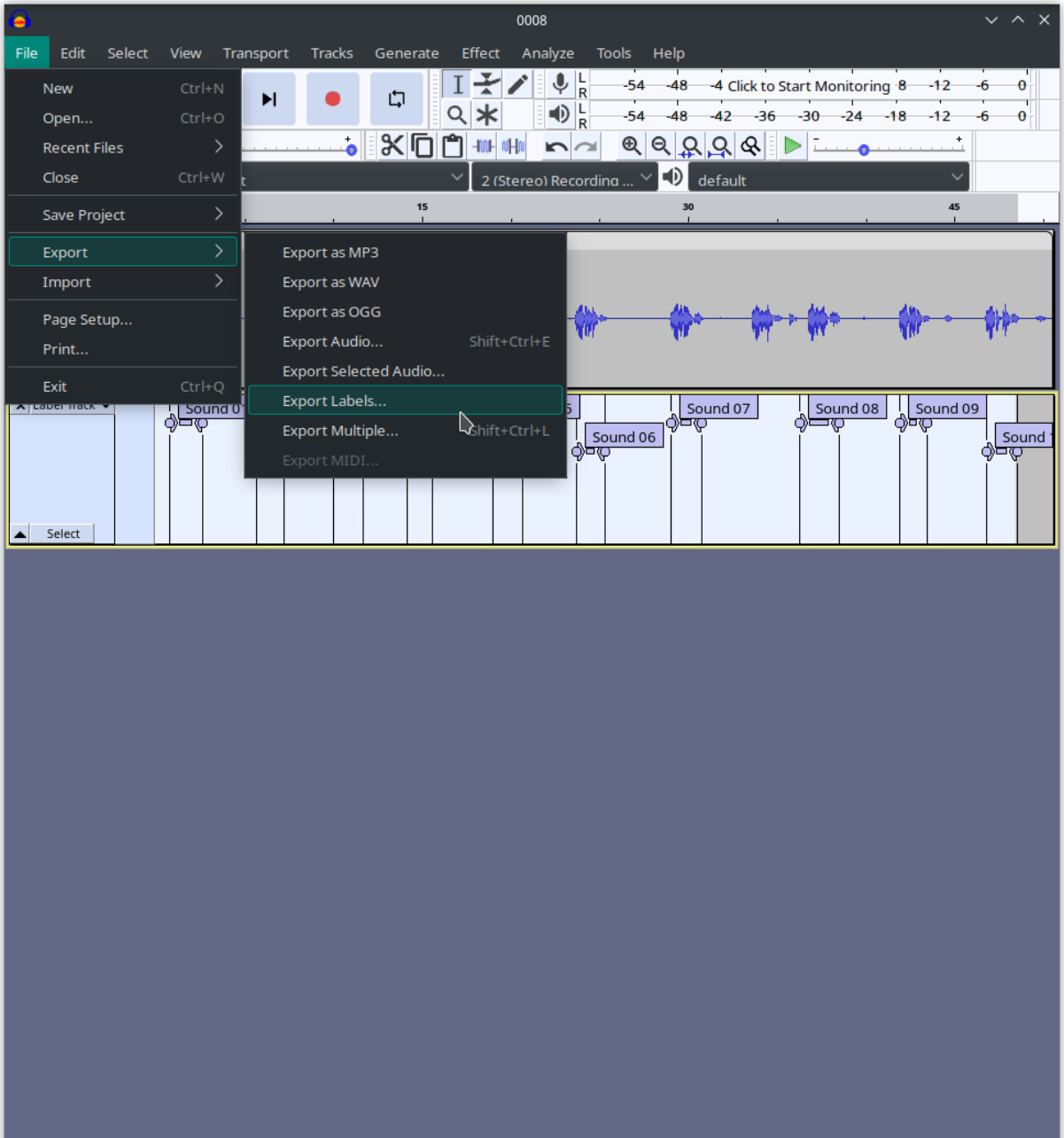

0008
File
Edit
Select
View
Transport
Tracks
Generate
Effect
Analyze
Tools
Help
New Ctrl+N
Open... Ctrl+O
Recent Files
Close Ctrl+W
Save Project
Export
Import
Page Setup...
Print...
Exit Ctrl+Q
Audio... Shift+Ctrl+I
Labels...
MIDI...
Raw Data...
Click to Start Monitoring
default

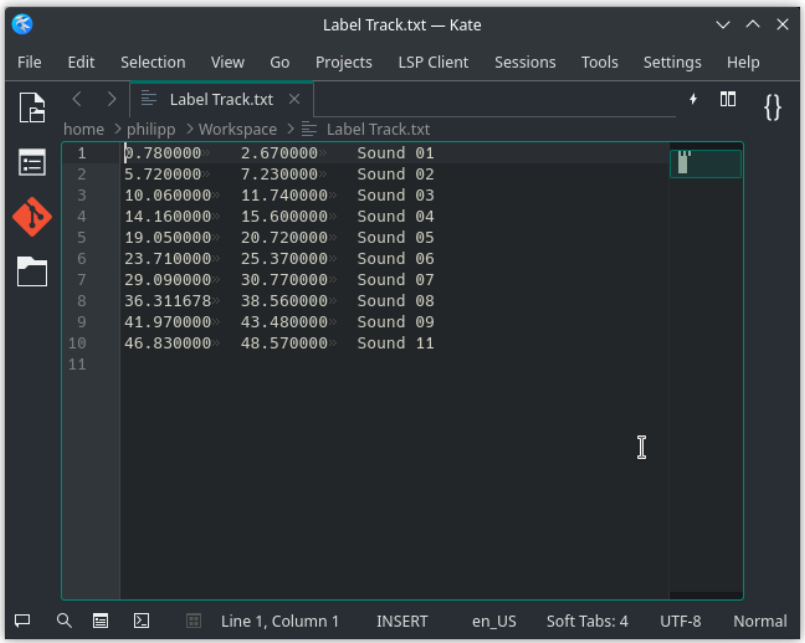

0008
File
Edit
Select
View
Transport
Tracks
Generate
Effect
Analyze
Tools
Help
Click to Start Monitoring
ALSA
default
2 (Stereo) Recording ...
default
0008
Mute
Solo
Mono, 48000Hz
Ctrl+Click to select or deselect track.


0008
File
Edit
Select
View
Transport
Tracks
Generate
Effect
Analyze
Tools
Help
Add / Remove Plug-ins...
Contrast...
Plot Spectrum...
Find Clipping...
Beat Finder...
Label Sounds...
to Start Monitoring
ALSA
default
0008
Mute
Solo
Mono, 48000Hz
32-bit float
Select


0008
File Edit Select View Transport Tracks Generate Effect Analyze Tools Help
Click to Start Monitoring
ALSA
default
2 (Stereo) Recording ...
default
0008
Mute Solo
Mono, 48000Hz
32-bit float
Select
Label Sounds
Threshold level (dB):
-30.0
Threshold measurement:
Peak level
Minimum silence duration:
Minimum label interval:
Label type:
Region around sounds
Maximum leading silence:
Maximum trailing silence:
Label text:
Sound ##1
Manage
OK
Cancel


0008
File
Edit
Select
View
Transport
Tracks
Generate
Effect
Analyze
Tools
Help
Click to Start Monitoring
ALSA
default
0008
Mute
Solo
Mono, 48000Hz
32-bit float
Select
Label Track
Sound 01
Sound 02
Sound 03
Sound 04
Sound 05
Sound 06
Sound 07
Sound 08
Sound 09
Sound 10


0008
File
Edit
Select
View
Transport
Tracks
Generate
Effect
Analyze
Tools
Help
New Ctrl+N
Open... Ctrl+O
Recent Files
Close Ctrl+W
Save Project
Export
Import
Page Setup...
Print...
Exit Ctrl+Q
Export as MP3
Export as WAV
Export as OGG
Export Audio... Shift+Ctrl+E
Export Selected Audio...
Export Labels...
Export Multiple... Shift+Ctrl+L
Export MIDI...
Click to Start Monitoring
default
Sound 06
Sound 07
Sound 08
Sound 09
Select


Label Track.txt — Kate
File Edit Selection View Go Projects LSP Client Sessions Tools Settings Help
Label Track.txt
home › philipp › Workspace › Label Track.txt
1 0.780000 2.670000 Sound 01
2 5.720000 7.230000 Sound 02
3 10.060000 11.740000 Sound 03
4 14.160000 15.600000 Sound 04
5 19.050000 20.720000 Sound 05
6 23.710000 25.370000 Sound 06
7 29.090000 30.770000 Sound 07
8 36.311678 38.560000 Sound 08
9 41.970000 43.480000 Sound 09
10 46.830000 48.570000 Sound 11
11
Line 1, Column 1 INSERT en_US Soft Tabs: 4 UTF-8 Normal

- use a script to extract the utterances of interest
- have one wav and one text file per utterance
- use the forced-alignment system of your choice