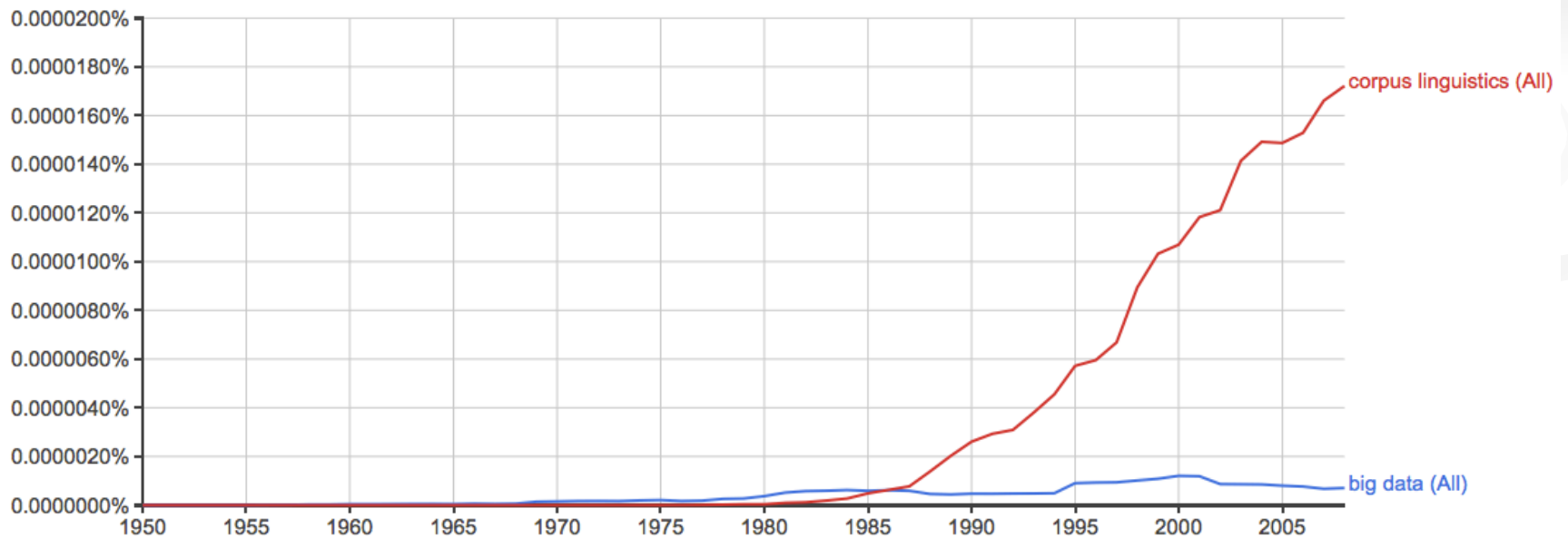


실험음성학회 12월 강독회

How to use Forced Alignment Systems

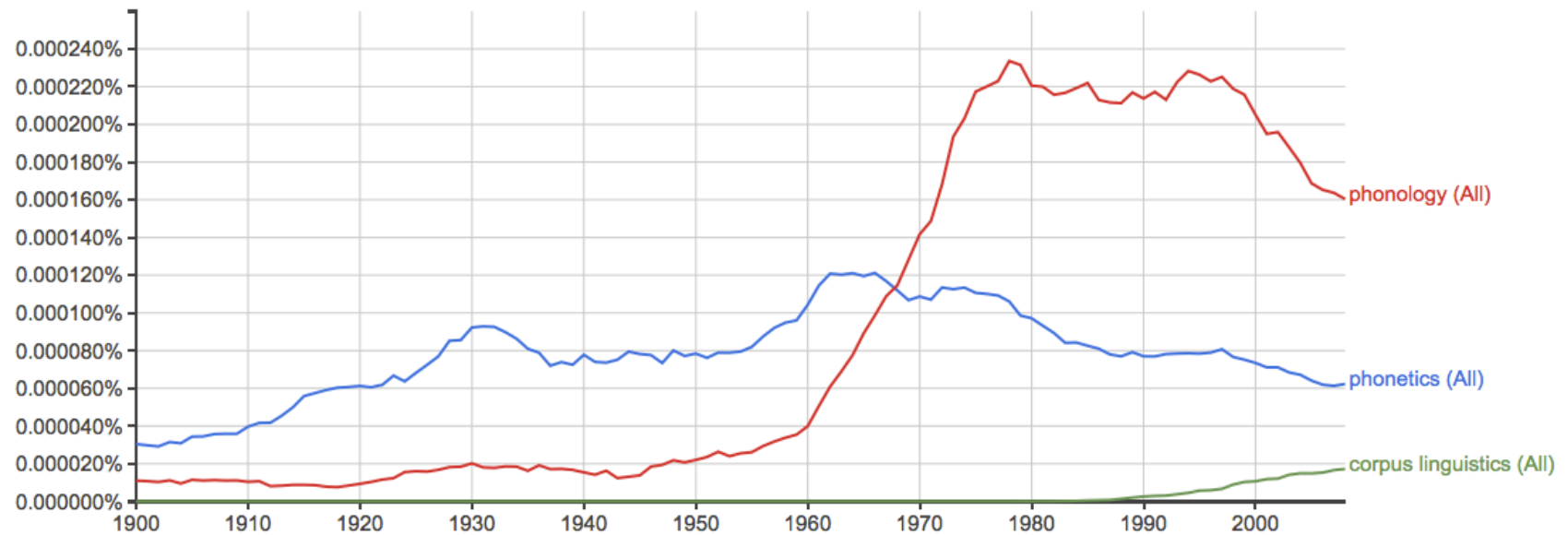
Tae-Jin Yoon
Department of English Language and
Literature
Cheonju University

The rise of corpus linguistics

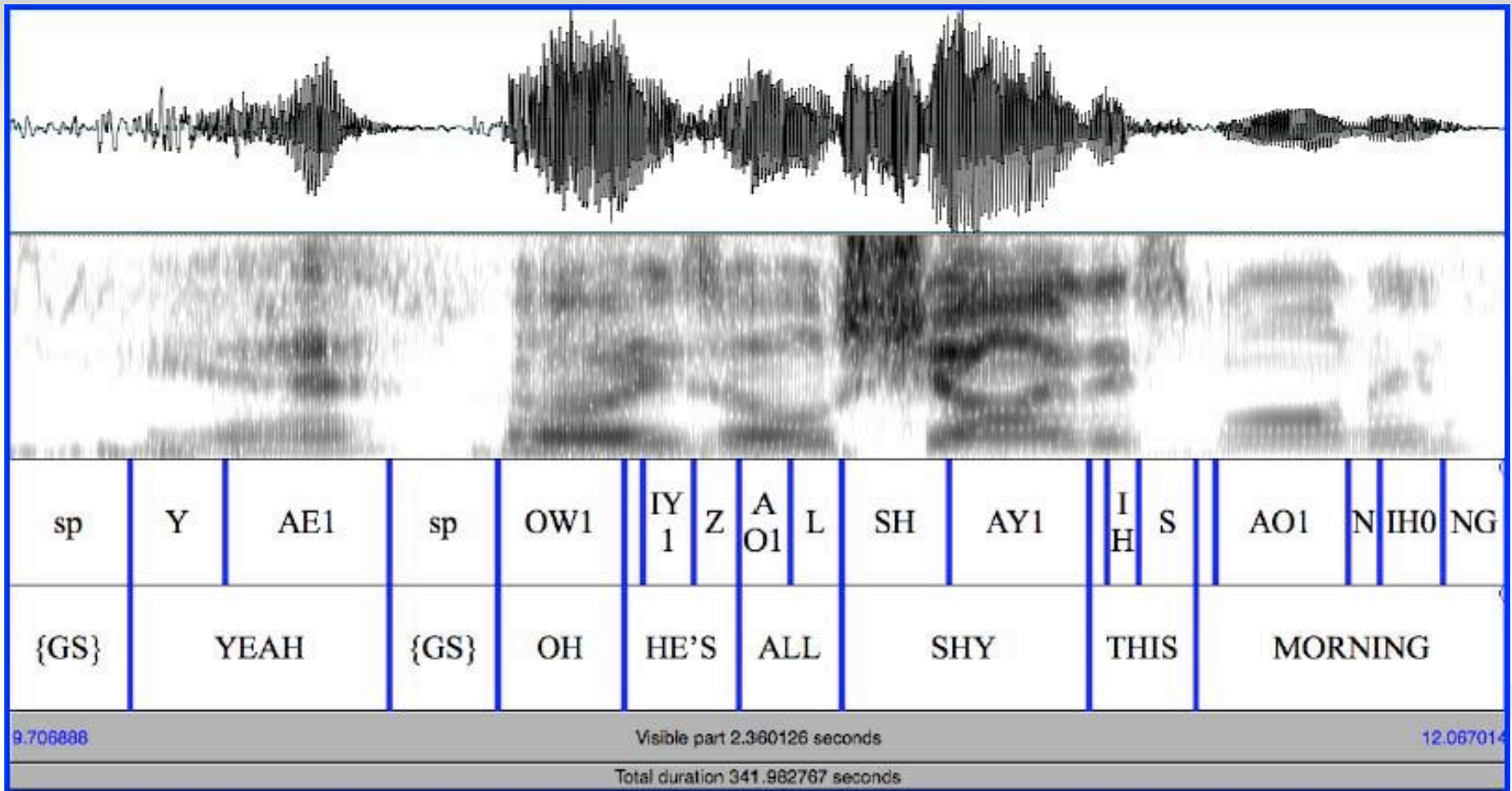


Google books Ngram Viewer

Phonetics, phonology and corpus linguistics



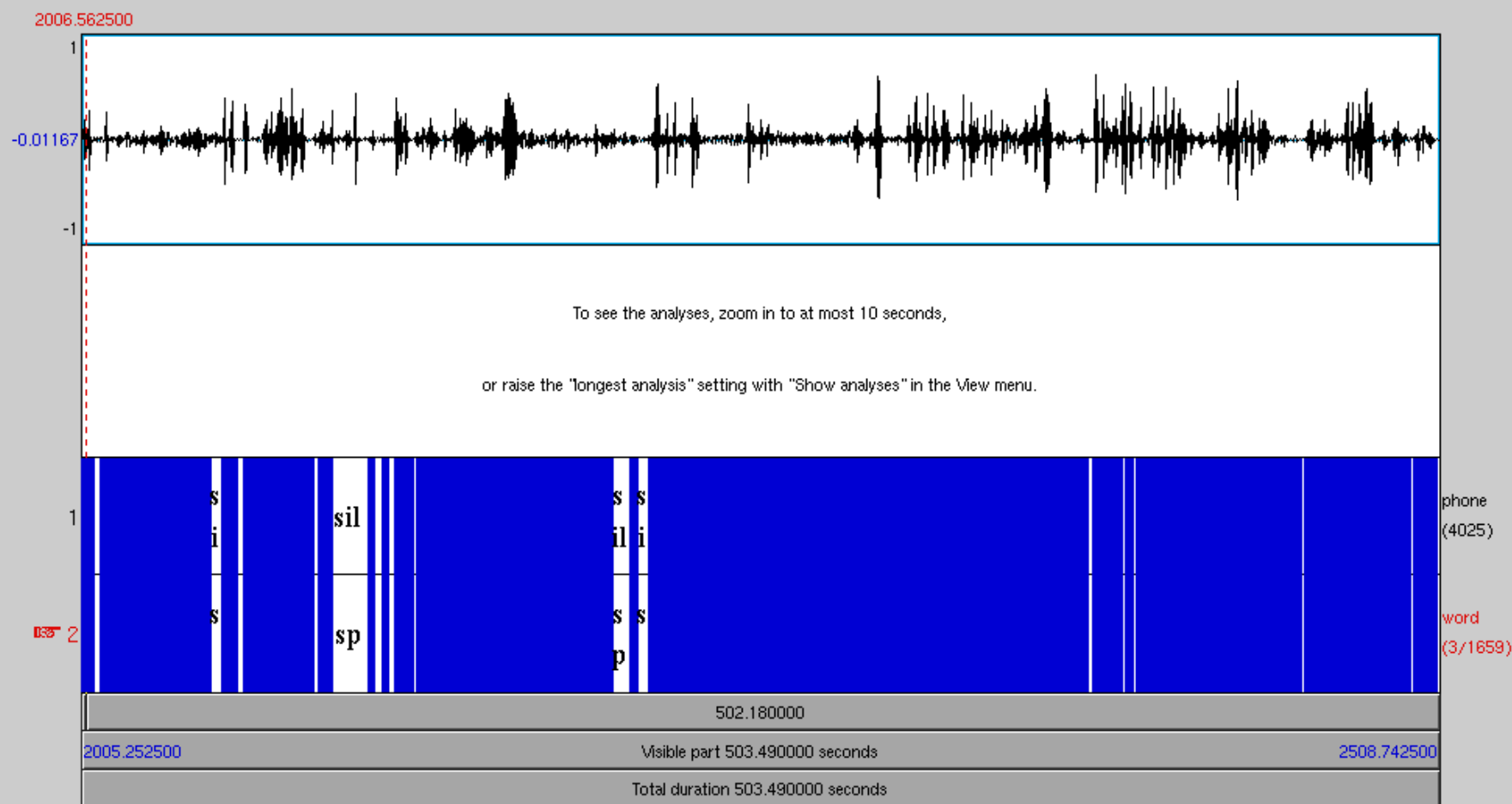
I have a sound file to be transcribed.
How do I transcribe?



x 21 million

How about this very long sound file?

→ Need for forced alignment



FAVE alignment

- An Online Interface to the Penn Forced Aligner
- <http://fave.ling.upenn.edu/FAAValign.html>
- An example file needed to be aligned



FAVE-align

An online interface for the Penn Forced Aligner

Upload files:

Please select the files to be aligned:

[?] Sound file:

[?] Transcription file:

Options:

☐ **-u:** Check transcription for unknown words[?]

☐ **-i:** Import dictionary transcriptions:[?]

Personal information:

First name:

Last name:

Institution:

Email address: *

confirm email address: *

[?] Comments:
How did you hear about this web site? Are you using it for research (which area/field?), or other purposes? Anything we can do to improve the interface?

* required

Align!

Depending on the size of your sound file and the speed of your internet connection, the file upload may take a while.

Webmaster: [Ingrid Rosenfelder](#)
Last modified: February 21, 2013

fave_test.wav
fave_test.lab

Pronunciation Dictionary

```

A42128 EY1 F A01 R T UW1 W AH1 N T UW1 EY1 T
AAA T R IH2 P AH0 L EY1
AABERG AA1 B ER0 G
AACHEN AA1 K AH0 N
AAKER AA1 K ER0
AALSETH AA1 L S EH0 TH
AAMODT AA1 M AH0 T
AANCOR AA1 N K A02 R
AARDEMA AA0 R D EH1 M AH0
AARDVARK AA1 R D V AA2 R K
AARON EH1 R AH0 N
AARON'S EH1 R AH0 N Z
AARONS EH1 R AH0 N Z

```

Phoneme	Example	Translation
AA	odd	AA D
AE	at	AE T
AH	hut	HH AH T
AO	ought	AO T
AW	cow	K AW
AY	hide	HH AY D
B	be	B IY
CH	cheese	CH IY Z
D	dee	D IY
DH	thee	DH IY
EH	Ed	EH D
ER	hurt	HH ER T
EY	ate	EY T
F	fee	F IY
G	green	G R IY N
HH	he	HH IY
IH	it	IH T
IY	eat	IY T
JH	gee	JH IY
K	key	K IY
L	lee	L IY
M	me	M IY
N	knee	N IY
NG	ping	P IH NG
OW	oat	OW T
OY	toy	T OY
P	pee	P IY
R	read	R IY D
S	sea	S IY
SH	she	SH IY
T	tea	T IY
TH	theta	TH EY T AH
UH	hood	HH UH D
UW	two	T UW
V	vee	V IY
W	we	W IY
Y	yield	Y IY L D
Z	zee	Z IY
ZH	seizure	S IY ZH ER

<http://www.speech.cs.cmu.edu/cgi-bin/cmudict>

C The ARPabet

Vowels			Consonants		
Phoneme	Example	Transcription	Phoneme	Example	Transcription
AA	bot	B AA T	B	be	B IY
AE	bat	B AE T	CH	cheese	CH IY Z
AH	but	B AH T	D	day	D EY
AO	bought	B AO T	DH	that	TH AE T
AW	bout	B AW T	F	fee	F IY
AY	bite	B AY T	G	go	G OW
EH	bet	B EH T	HH	he	HH IY
ER	bird	B ER D	JH	just	JH AH S T
EY	bait	B EY T	K	key	K IY
IH	bit	B IH T	L	late	L EY T
IY	beat	B IY T	M	me	M IY
OW	boat	B OW T	N	knee	N IY
OY	boy	B OY	NG	sing	S IH NG
UH	put	P UH T	P	pay	P EY
UW	boot	B UW T	R	read	R IY D
			S	sea	S IY
			SH	she	SH IY
			T	tea	T IY
			TH	thanks	TH AE NG K S
			V	vain	V EY N
			W	we	W IY
			Y	yes	Y EH S
			Z	zoo	Z UW
			ZH	pleasure	P L EH ZH ER

Example:

Stress is indicated by digits following the stressed vowels. There are three levels of stress:

Value	Level of stress
0	no stress
1	primary stress
2	secondary stress

in	AH0 N, IH1 N
the	DH AH0, DH AH1, DH IY0
dictionary	D IH1 K SH AH0 N EH2 R IY0
stress	S T R EH1 S
is	IH1 Z, AH0 Z
indicated	IH1 N D AH0 K EY2 T AH0 D
by	B AY1
digits	D IH1 JH AH0 T S
following	F AA1 L OW0 IH0 NG
stressed	S T R EH1 S T
vowels	V AW1 AH0 L Z

PH73-5-1-A-MCosta.txt

MC	Mary Costa	16.02	16.976	John F. Smith.
MC	Mary Costa	18.1	20.099	And the year was eighteen ninety-five
MC	Mary Costa	21.34	23.332	I never noticed before. Just when --
MC	Mary Costa	23.332	26.332	It's about, oh when we papered his house it was about
MC	Mary Costa	26.4	27.366	sixteen,
MC	Mary Costa	27.446	29.515	seventeen years ago when I papered the house.
MC	Mary Costa	29.57	33.274	I mean, we had it scraped. And that's when I noticed it.
MC	Mary Costa	35.32	35.766	Yeah.
MC	Mary Costa	38.702	43.063	Yeah, on Ninth Street, I know. I was born there, Eight Twenty-Four Christian Street.
MC	Mary Costa	48.103	49.678	Perry, yeah I know who you mean.
MC	Mary Costa	50.109	51.659	Yeah, ((Taker)), yeah I know.
MC	Mary Costa	52.395	52.929	I know.
MC	Mary Costa	53.52	55.08	We know the Perry's well.
MC	Mary Costa	55.537	57.619	Yeah, we know the Perry's well. Oh do you?
MC	Mary Costa	58.164	58.675	Oh yeah.
MC	Mary Costa	59.26	62.01	Well, Perry, ((Berry)), my sist' and all, we had him for --
MC	Mary Costa	62.2	62.718	(())
MC	Mary Costa	64.72	65.452	No.
MC	Mary Costa	65.5	67.492	Oh the -- No, he didn't die in jail.
MC	Mary Costa	67.532	69.907	Joe Perry. What's the matter with you, you made a mistake.
MC	Mary Costa	70.3	71.75	((I assure was a)) was a
MC	Mary Costa	72.13	72.874	somethin'
MC	Mary Costa	78.12	78.636	Yeah.
MC	Mary Costa	78.7	79.381	Yeah
MC	Mary Costa	80.87	82.131	((Geytan)), that's right.
MC	Mary Costa	82.34	83.42	Well, he died too.
MC	Mary Costa	85.05	86.145	Yeah I know. I heard he --
MC	Mary Costa	86.69	88.775	Well not as sudden as ((up at)) Darby.
MC	Mary Costa	89.99	93.443	The ((sons)) has a place up at Darby, ((Gate Dusserson))
MC	Mary Costa	97.32	99.518	Oh the daughter, his son. She still runs it there?
MC	Mary Costa	99.54	101.508	I haven't been around there so long, I don't know.
MC	Mary Costa	112.12	115.613	A f- fraternity like a brother. Like all of
MC	Mary Costa	117.67	120.279	Italy, you know. The country they come from.
MC	Mary Costa	120.43	121.037	((Gazent)).
MC	Mary Costa	122.42	126.636	Well it's cause I'm, I -- my mother and father, an', he comes from the same town.
MC	Mary Costa	127.71	130.48	It's called ((Terra Nova de Citre, Rovinchi, Gazent)).
MC	Mary Costa	131.01	131.44	That's it.

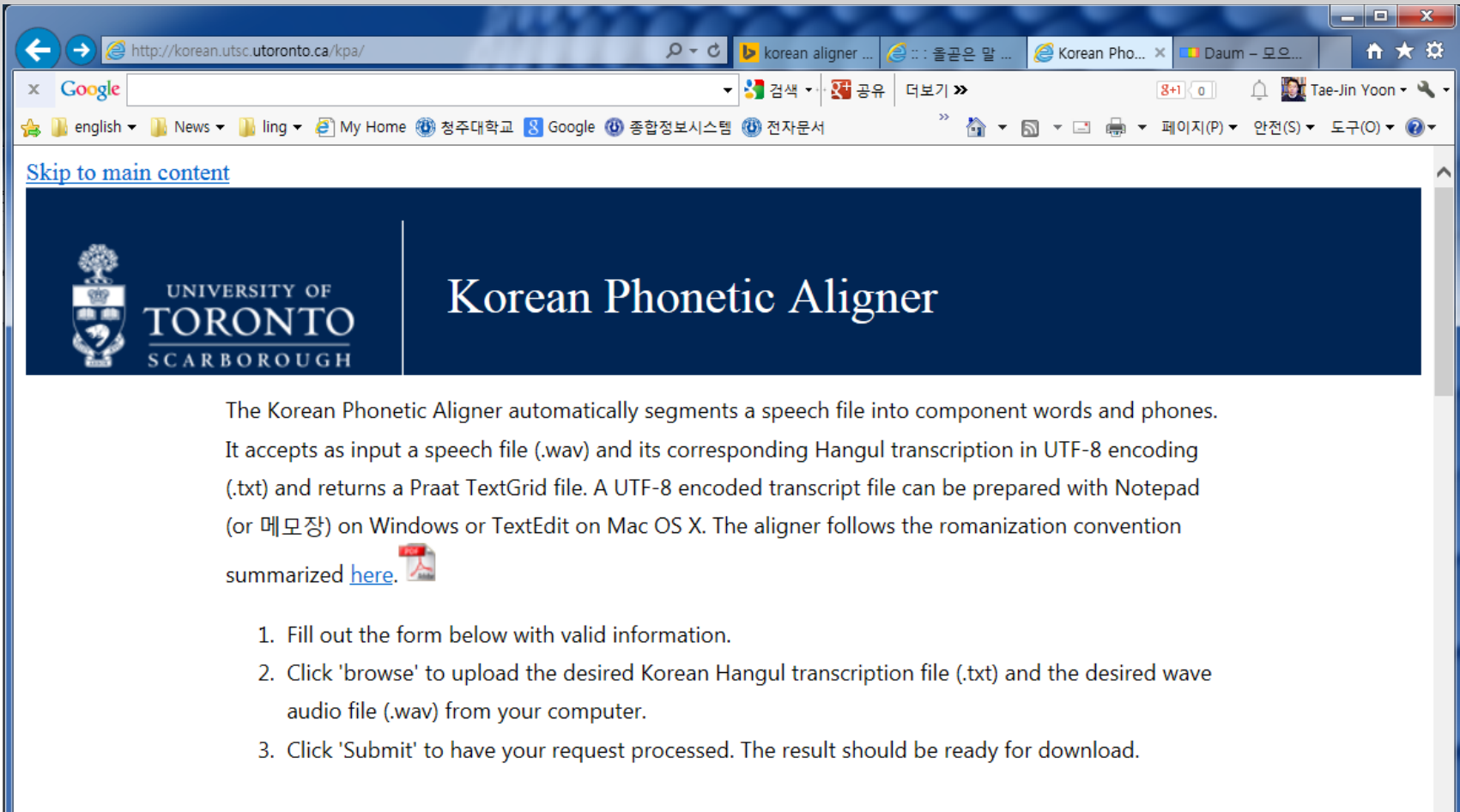
http://fave.ling.upenn.edu/downloads/Convert_To_FAVE-align_Input.praat

```
#####  
# This Praat script exports orthographic transcriptions in Praat  
## to a format suitable as input to the FAVE-align forced aligner  
## (http://fave.ling.upenn.edu/FAAValign.html).  
## The transcription will be converted to a 5-column tab-delimited .txt file  
## as outlined in the instructions on the FAVE web site.  
##  
## To run this program, select the TextGrid containing the transcriptions,  
## open this script, and select "Run" > "Run".  
##  
## This script was written by Ingrid Rosenfelder,  
## last modified October 31, 2011  
#####
```

```
## get TextGrid name
info$ = Info
filename$ = extractLine$(info$, "Object name: ")
outfile$ = filename$ + ".txt"
## ask the user before overwriting exiting file
if fileReadable(outfile$)
    pause 'filename$.txt already exists. Overwrite?'
    deleteFile(outfile$)
endif
## extract transcription info and write to file
n_tiers = Get number of tiers
for tier from 1 to n_tiers
    tiername$ = Get tier name... 'tier'
    n_intervals = Get number of intervals... 'tier'
    for interval from 1 to 'n_intervals'
        start = Get start point... 'tier' 'interval'
        end = Get end point... 'tier' 'interval'
        label$ = Get label of interval... 'tier' 'interval'
        if label$ <> ""
            fileappend'outfile$' 'tiername$' 'tab$' 'tiername$' 'tab$'
            ... 'start' 'tab$' 'end' 'tab$' 'label$' 'newline$'
        endif
    endfor
endfor
echo Written transcription in FAVE-align input format to file 'outfile$'.
```



Can I use the FAVE Aligner for Korean?


Korean Phonetic Aligner



The screenshot shows a web browser window with the URL <http://korean.utsc.utoronto.ca/kpa/>. The page features the University of Toronto Scarborough logo and the title "Korean Phonetic Aligner". Below the header, there is a paragraph describing the tool's functionality and a list of three steps for using it.

[Skip to main content](#)

 **Korean Phonetic Aligner**

The Korean Phonetic Aligner automatically segments a speech file into component words and phones. It accepts as input a speech file (.wav) and its corresponding Hangul transcription in UTF-8 encoding (.txt) and returns a Praat TextGrid file. A UTF-8 encoded transcript file can be prepared with Notepad (or 메모장) on Windows or TextEdit on Mac OS X. The aligner follows the romanization convention summarized [here](#). 

1. Fill out the form below with valid information.
2. Click 'browse' to upload the desired Korean Hangul transcription file (.txt) and the desired wave audio file (.wav) from your computer.
3. Click 'Submit' to have your request processed. The result should be ready for download.

<http://korean.utsc.utoronto.ca/kpa/>

1. Fill out the form below with valid information.
2. Click 'browse' to upload the desired Korean Hangul transcription file (.txt) and the desired wave audio file (.wav) from your computer.
3. Click 'Submit' to have your request processed. The result should be ready for download.

First Name *	<input type="text" value="TJ"/>
Last Name *	<input type="text" value="Yoon"/>
Email *	<input type="text" value="yoon.taejin@gmail.com"/>
Affiliation *	<input type="text" value="cju"/>
Upload transcription file (*.txt) *	<input type="text" value="C:\Users\Cju\Dropbox\실시간 음성 파일\"/> <input data-bbox="1136 562 1367 592" type="button" value="찾아보기..."/>
Upload Wav file (*.wav) *	<input type="text" value="C:\Users\Cju\Dropbox\실시간 음성 파일\"/> <input data-bbox="1136 642 1367 672" type="button" value="찾아보기..."/>
<input checked="" type="checkbox"/> I authorize the collection of the information provided in the form.	
<input data-bbox="285 732 367 762" type="button" value="Submit"/>	

한글_test1 - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

기차도 전기도 없었다.

Content

구성 ▼ 새 폴더

★ 즐겨찾기

Dropbox

다운로드

바탕 화면

최근 위치

iCloud 사진

라이브러리

문서

비디오

사진

음악



kdict1

파일 이름(N): 한글_test1.txt

파일 형식(T): 텍스트 문서(*.txt)

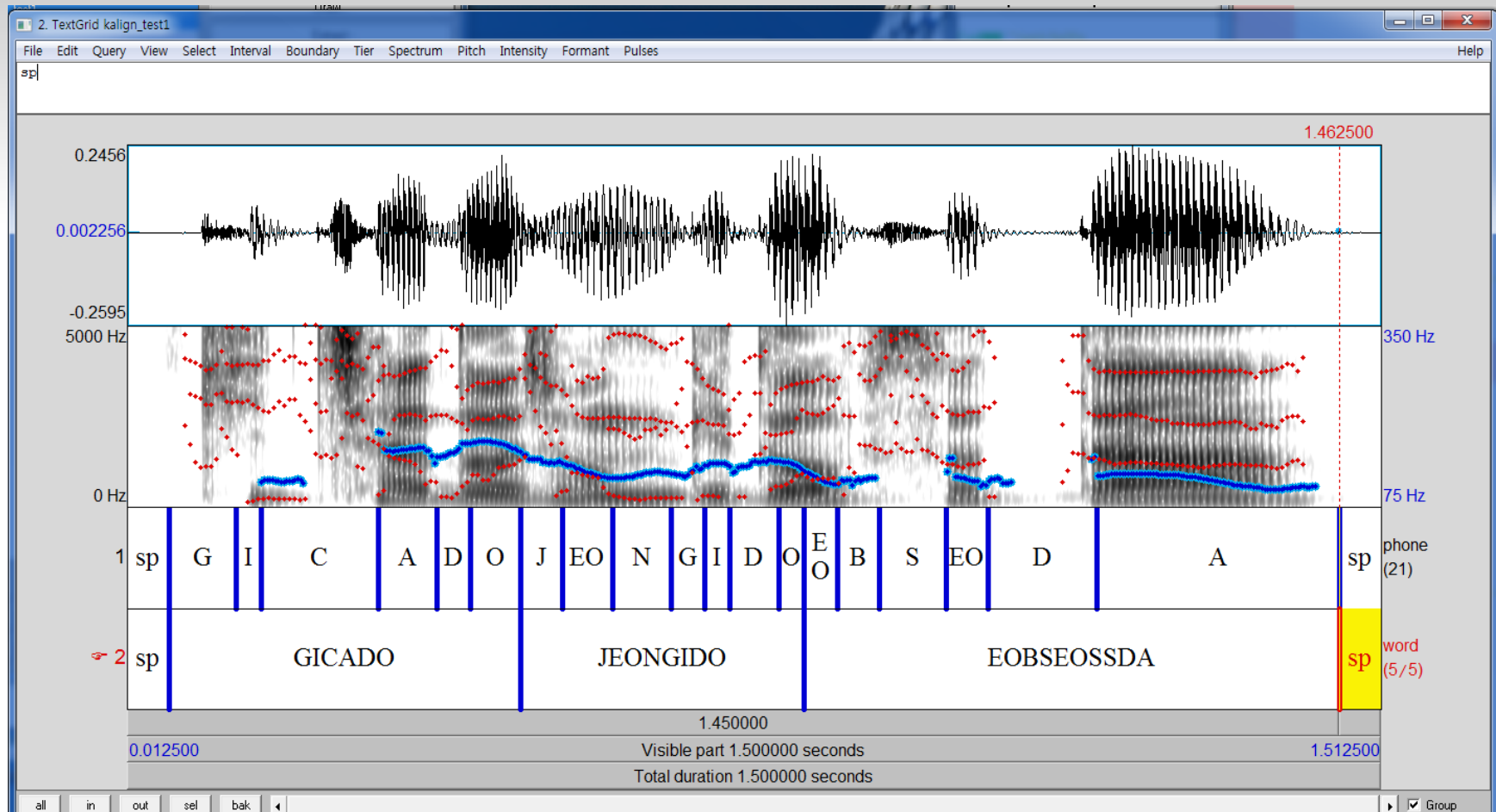
Save as .txt and utf-8

폴더 숨기기

인코딩(E): UTF-8

저장(S)

Result of Korean Phonetic Aligner



Romanization convention

Romanization convention used in the Korean Phonetic Aligner

<Vowels>

Transcription	Hangul	Phonetic
A	ㅏ	ɑ
AE	ㅐ	æ
E	ㅓ	e
EO	ㅜ	ə
EU	ㅡ	i
I	ㅣ	i
O	ㅗ	o
OE	ㅛ	we/ø
U	ㅜ	u
WA	ㅘ	wa
WAE	ㅙ	wæ
WE	ㅚ	we
WEO	ㅜ	wə
WI	ㅟ	wi/y
YA	ㅑ	ja
YAE	ㅓ	jæ
YE	ㅕ	je
YEO	ㅛ	jə
YI	ㅟ	hi
YO	ㅟ	jo
YU	ㅠ	ju

<Consonants>

Transcription	Hangul	Phonetic
B	ㅂ	p (b: between sonorants)
BB	ㅃ	p'
P	ㅍ	p ^h
D	ㄷ	t (d: between sonorants)
DD	ㄸ	t'
T	ㅌ	t ^h
G	ㄱ	k (g: between sonorants)
GG	ㄲ	k'
K	ㅋ	k ^h
J	ㅈ	ts (dz: between sonorants)
JJ	ㅉ	ts'
C	ㅊ	ts ^h
S	ㅅ	s (ʃ: before a high front vocoid)
SS	ㅆ	s' (ʃ': before a high front vocoid)
H	ㅎ	h
M	ㅁ	m
N	ㄴ	n
NG	ㅇ	ŋ
L(coda)/R (onset)	ㄹ	l (coda, geminate), r (intervocalic)
sp		Silent pause

Wow, it's great! Can I use the forced alignment for any types of sound files?

Challenges for alignments

- Transcription errors
- Long untranscribed portions
- Some transcribed regions with no audio (lost in copying)
- Broadcast recordings may include untranscribed commercials
- Transcripts generally edit out dysfluencies

Challenges for alignments

- Overlapping speakers
- Background noise/music/babble
- Variable signal loudness
- Reverberation
- Distortion
- Poor speaker vocal health/voice quality
- Unexpected accents: need multidialect pronouncing dictionary



I have 1000 files to process. Is there a way to process the multiple files much easily?

THERE ARE GEEKY WAYS!

Then, you need a Linux-like environment

- `./align.py filename.wav filename.lab filename.TextGrid`



```
Cju@ 00 ~/work/kp2fa
```

```
$ ./align.py test/fv01_t01_s01.wav test/fv01_t01_s01.lab test/fv01_t01_s01.TextGrid  
./tmp/sound.wav -> ./tmp/tmp.mfc  
Running HVite...
```

What do I need for the Linux-like environment?

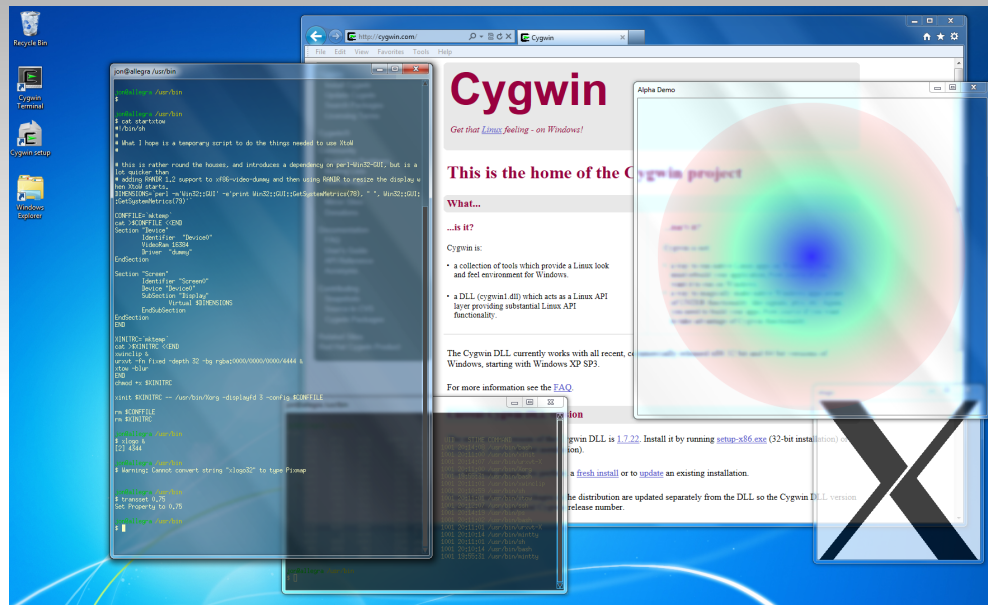
- Windows
 - Cygwin
 - HTK
- Mac OS X
 - XCode
 - HTK

On Windows: cygwin

- Cygwin (pronounced as SIG-win) = GNU + Cygnus + Windows.
- The mother site for Cygwin is www.cygwin.com.
- According to the developers, "Cygwin is a Open-source *Linux*-like environment for *Windows*."
- It consists of two parts:
 - A DLL (cygwin1.dll) which acts as a Linux API emulation layer providing substantial Linux API functionality.
 - A collection of tools which provide Linux look and feel."

YouTube Clip for installing Cygwin on Window 8

- <http://www.youtube.com/watch?v=hh-V6el8Oxk>



Cygwin설치하기 1편 및 2편

<http://detectiveboy.tistory.com/22>

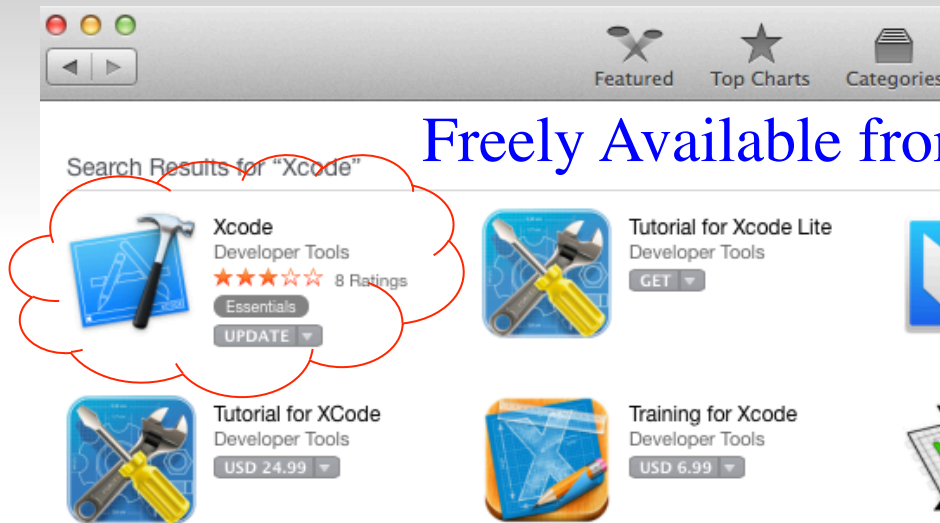
<http://detectiveboy.tistory.com/23>

On Mac OS X: terminal

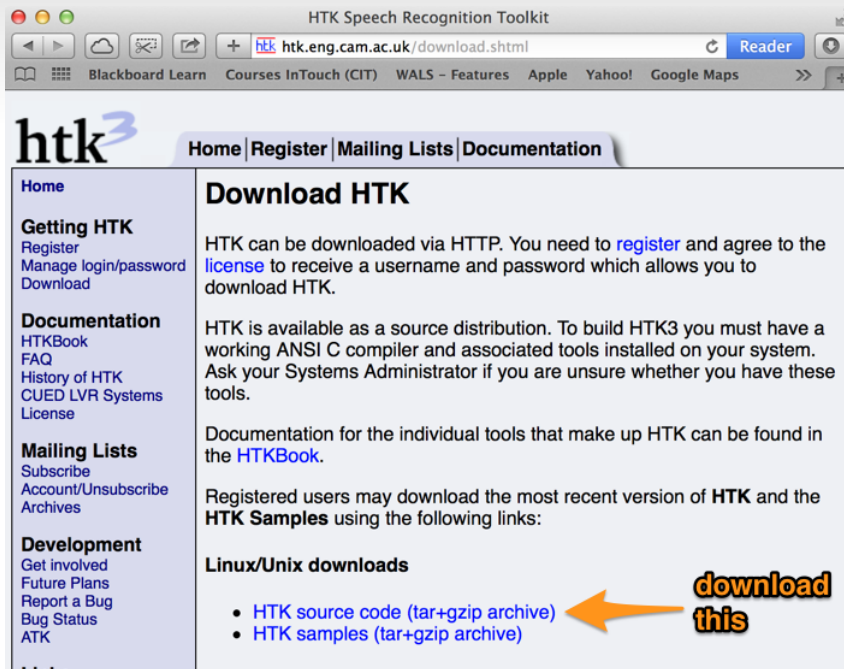


Mac OS X: Install Xcode

Freely Available from App Store



HTK (HMM Toolkit)



The screenshot shows the HTK website with the following content:

HTK Speech Recognition Toolkit
htk.eng.cam.ac.uk/download.shtml

Home | Register | Mailing Lists | Documentation

Getting HTK
Register
Manage login/password
Download

Documentation
HTKBook
FAQ
History of HTK
CUED LVR Systems
License

Mailing Lists
Subscribe
Account/Unsubscribe
Archives

Development
Get involved
Future Plans
Report a Bug
Bug Status
ATK

Download HTK
HTK can be downloaded via HTTP. You need to [register](#) and agree to the [license](#) to receive a username and password which allows you to download HTK.

HTK is available as a source distribution. To build HTK3 you must have a working ANSI C compiler and associated tools installed on your system. Ask your Systems Administrator if you are unsure whether you have these tools.

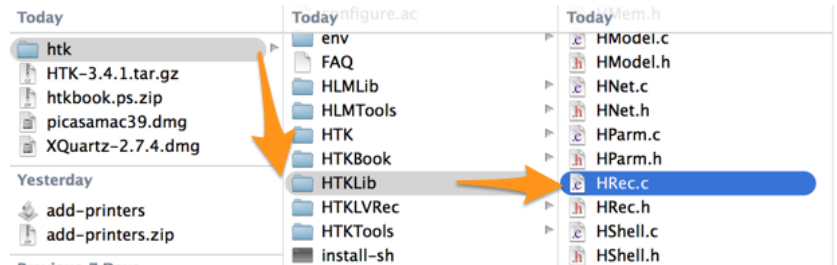
Documentation for the individual tools that make up HTK can be found in the [HTKBook](#).

Registered users may download the most recent version of **HTK** and the **HTK Samples** using the following links:

Linux/Unix downloads

- HTK source code (tar+gzip archive)
- HTK samples (tar+gzip archive)

download this (with an arrow pointing to the HTK source code link)



The screenshot shows a file manager window with the following structure:

Today

- htk
- HTK-3.4.1.tar.gz
- htkbook.ps.zip
- picasamac39.dmg
- XQuartz-2.7.4.dmg

Yesterday

- add-printers
- add-printers.zip

Today (in a subdirectory)

- env
- FAQ
- HLMLib
- HLMTTools
- HTK
- HTKBook
- HTKLib
- HTKLVRec
- HTKTools
- install-sh

Today (in a subdirectory)

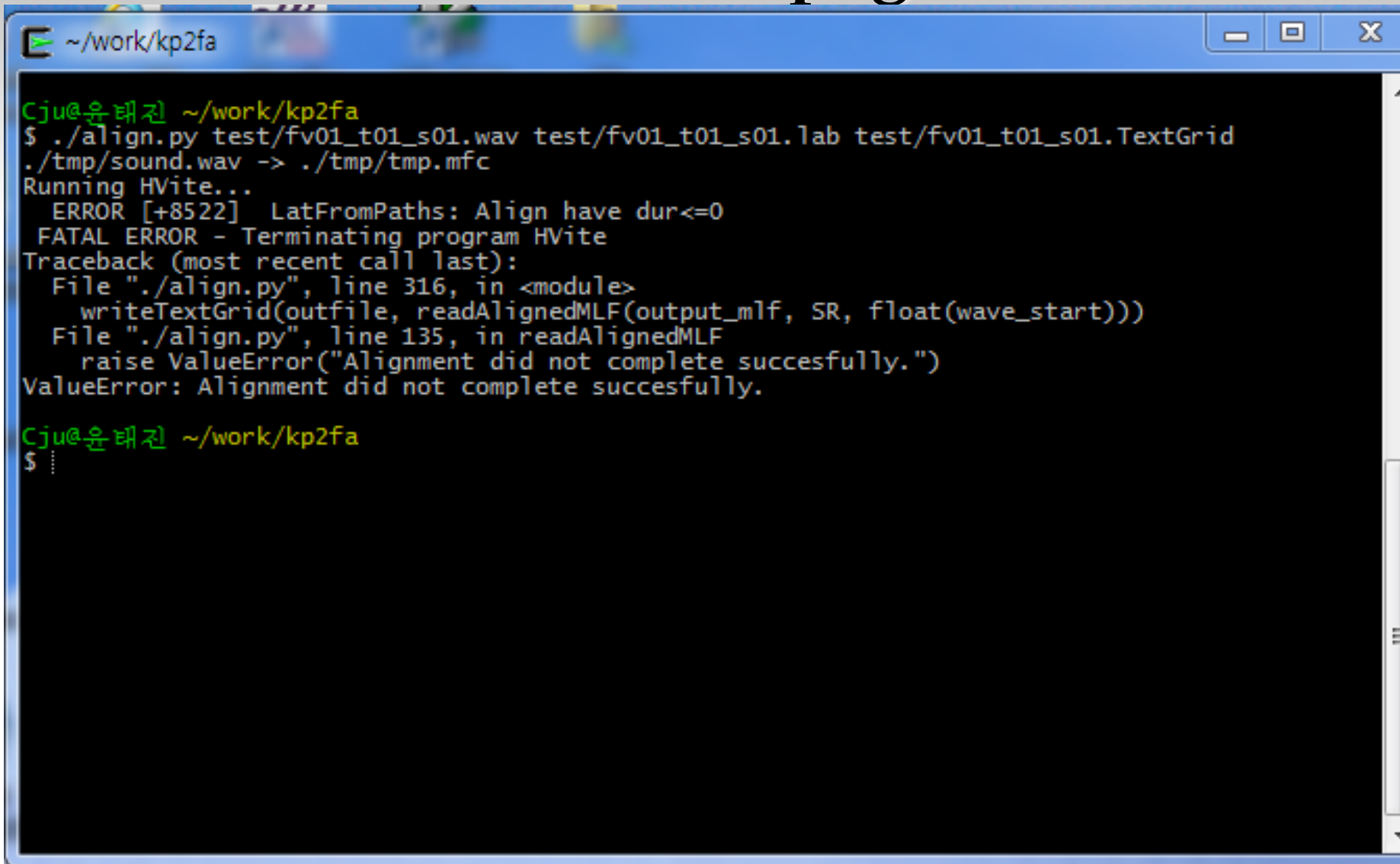
- HModel.c
- HModel.h
- HNet.c
- HNet.h
- HParm.c
- HParm.h
- HRec.c
- HRec.h
- HShell.c
- HShell.h

Arrows indicate the path from the 'htk' directory to the 'HTKLib' directory, and then to the 'HRec.c' file.

HTK Toolkit

- You'll need to register (free) with HTK to download it.
- Visit <http://htk.eng.cam.ac.uk/register.shtml> to register, and they will e-mail you a password.
- After that, visit <http://htk.eng.cam.ac.uk/download.shtml>, and download the HTK source code appropriate for your operating system.

If you have the following problem,
see the next page.

A screenshot of a terminal window with a blue title bar. The title bar contains a file explorer icon, the path ~/work/kp2fa, and standard window control buttons (minimize, maximize, close). The terminal has a black background with green text for the prompt and white text for the command and output. The prompt is Cju@윤태진 ~/work/kp2fa. The command executed is ./align.py test/fv01_t01_s01.wav test/fv01_t01_s01.lab test/fv01_t01_s01.TextGrid. The output shows the command ./tmp/sound.wav -> ./tmp/tmp.mfc, followed by Running HVite... and an error message: ERROR [+8522] LatFromPaths: Align have dur<=0. This is followed by a fatal error message: FATAL ERROR - Terminating program HVite. A traceback is shown, indicating the error occurred in ./align.py at line 316 in the writeTextGrid function and line 135 in the readAlignedMLF function. The final output is ValueError: Alignment did not complete successfully. The prompt Cju@윤태진 ~/work/kp2fa is shown again, followed by a single dollar sign \$ on the next line.

```
Cju@윤태진 ~/work/kp2fa
$ ./align.py test/fv01_t01_s01.wav test/fv01_t01_s01.lab test/fv01_t01_s01.TextGrid
./tmp/sound.wav -> ./tmp/tmp.mfc
Running HVite...
ERROR [+8522] LatFromPaths: Align have dur<=0
FATAL ERROR - Terminating program HVite
Traceback (most recent call last):
  File "./align.py", line 316, in <module>
    writeTextGrid(outfile, readAlignedMLF(output_mlf, SR, float(wave_start)))
  File "./align.py", line 135, in readAlignedMLF
    raise ValueError("Alignment did not complete succesfully.")
ValueError: Alignment did not complete succesfully.

Cju@윤태진 ~/work/kp2fa
$
```

Fixing the HTK Source

- In the htk directory, navigate to **htk/HTKLib**, and open the file HRec.c in a plain text editor. Find the line 1650 which reads

```
if (dur<=0 && labid != splabid) HError(8522,"LatFromPaths:  
Align have dur<=0 ");
```

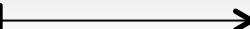
- labid in this line should be changed to labpr. It should look like this:

```
if (dur<=0 && labpr != splabid)  
HError(8522,"LatFromPaths: Align have dur<=0 ");
```

Now I have an HTK installed, how can I process multiple files?

■ Scripting in Bash shell

imagine you have
the data in the
following structure:



```
data/s01_01.wav  
data/s01_02.wav  
...  
data/s01_99.wav  
data/s02_01.wav  
data/s02_02.wav  
...  
data/s02_99.wav  
...  
data/s10_99.wav
```

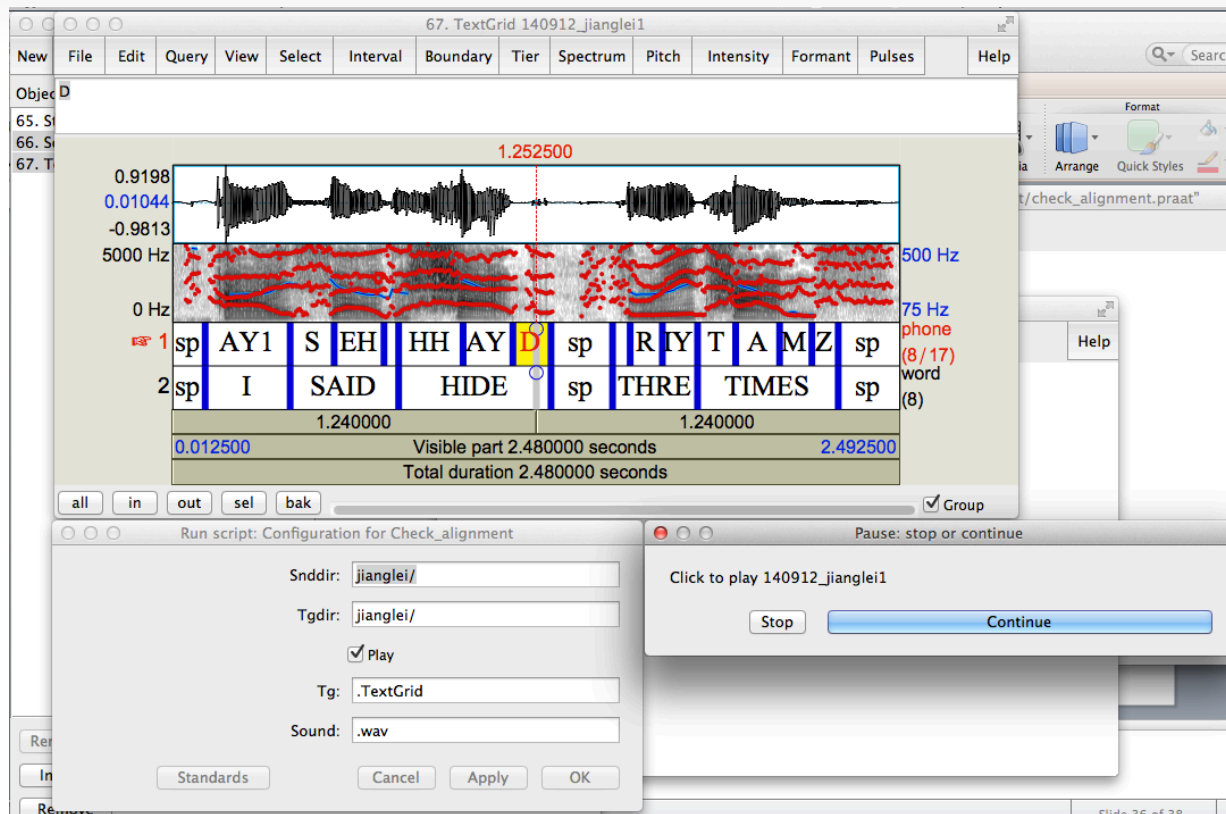
example: run.sh

```
#!/bin/sh
```

```
for SID in 0 1 2 3 4 5 6 7 8 9
do
    for CID in 0 1 2 3 4 5 6 7 8 9
    do
        if [ -e data/s${SID}_${CID}.wav ]; then
            echo 'Alignning the speech file of ' s${
{SID}_${CID}.wav
            ./align.py data/s${SID}_${CID}.wav data/s$
{SID}_${CID}.lab data/s${SID}_${CID}.TextGrid
        fi
    done
done
~
```

I want to check the aligned data

- `check_alignment.praat`



check_alignment.praat

```
form Configuration for Check_alignment
  sentence Snddir ./test/
  sentence Tgdir ./test/
  sentence Sound .wav
endform
clearinfo
Create Strings as file list... filelist 'snddir$'*'sound$'
num_files = Get number of strings

for j from 1 to num_files
  select Strings filelist
  fname$ = Get string... 'j'
  basename$ = fname$-sound$
  wavname$ = basename$+"sound$"
  tname$ = basename$ + ".TextGrid"
  printline We will label 'fname$'
  if fileReadable("'tgdir$'"+tname$)
    printline 'tname$' exists!
    Read from file... 'snddir$'"+fname$
    Read from file... 'tgdir$'"+tname$
    btime = Get start time
    etime = Get end time
    select Sound 'basename$'
    plus TextGrid 'basename$'
    View & Edit
      editor TextGrid 'basename$'
        pause Click to play 'basename$'
        Play... btime etime
        pause Click to close 'basename$'
        Close
      endeditor
    else
      printline 'tname$' does not exists!
    endif
  endfor
select Strings filelist
Remove
```