

Crossing the pond: Extending automatic alignment techniques to British English dialect data

Laurel MacKenzie and Danielle Turton

University of Manchester

New Ways of Analyzing Variation 42
20th October 2013



The University of Manchester

Introduction

- ▶ Increased recent use in sociolinguistic research of tools for automatic phonemicization and alignment of speech (e.g. Labov et al., 2013)
 - ▶ In a nutshell: convert an orthographic transcription into phonemes, automatically time-align words and phonemes to speech signal
 - ▶ But: reliance on a standard dictionary of reference pronunciations

Introduction

- ▶ Increased recent use in sociolinguistic research of tools for automatic phonemicization and alignment of speech (e.g. Labov et al., 2013)
 - ▶ In a nutshell: convert an orthographic transcription into phonemes, automatically time-align words and phonemes to speech signal
 - ▶ But: reliance on a standard dictionary of reference pronunciations

Introduction

- ▶ Increased recent use in sociolinguistic research of tools for automatic phonemicization and alignment of speech (e.g. Labov et al., 2013)
 - ▶ In a nutshell: convert an orthographic transcription into phonemes, automatically time-align words and phonemes to speech signal
 - ▶ But: reliance on a standard dictionary of reference pronunciations

Research questions

1. To what extent can these tools be used effectively with data from non-standard dialects?
2. What problem areas arise?
 - ▶ Current study: three aligners, four dialects

Research questions

1. To what extent can these tools be used effectively with data from non-standard dialects?
2. What problem areas arise?
 - ▶ Current study: three aligners, four dialects

Research questions

1. To what extent can these tools be used effectively with data from non-standard dialects?
2. What problem areas arise?
 - ▶ Current study: three aligners, four dialects

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

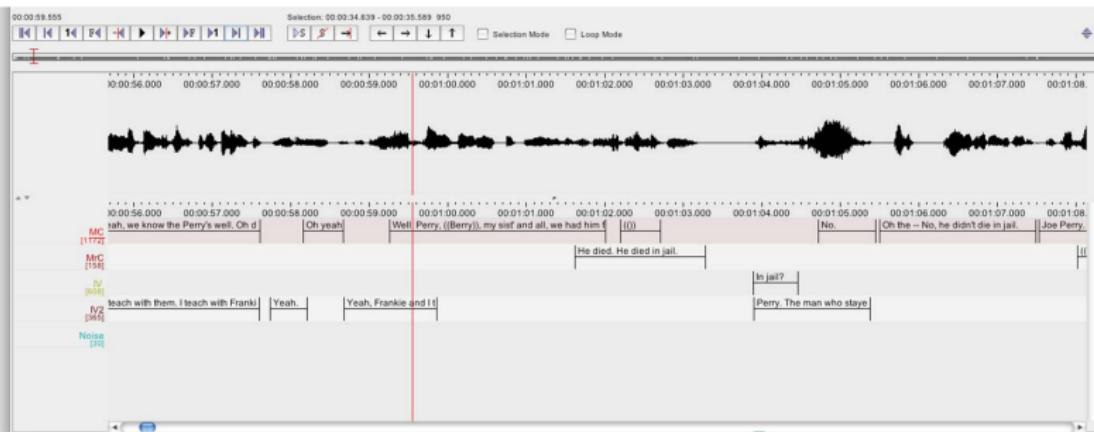
3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

Input

- ▶ Sound file
- ▶ Transcription



Input

- ▶ Sound file
- ▶ Transcription
- ▶ Pronouncing dictionary
 - ▶ Standard British: BEEP
 - ▶ Standard American:
CMUdict

```
TIMBLIN T IH1 M B L IH0 N
TIMBRE T IH1 M B ER0
TIMBROOK T IH1 M B R UH2 K
TIMBS T IH1 M Z
TIMBUKTU T IH2 M B AH0 K T UW1
TIME T AY1 M
TIME'S T AY1 M Z
TIMED T AY1 M D
TIMEFRAME T AY1 M F R EY2 M
TIMELESS T AY1 M L AH0 S
TIMELINE T AY1 M L AY0 N
TIMELINES T AY1 M L AY0 N Z
```

Input

- ▶ Sound file
- ▶ Transcription
- ▶ Pronouncing dictionary
 - ▶ Standard British: BEEP
 - ▶ Standard American:
CMUdict

```
TIMBLIN T IH1 M B L IH0 N
TIMBRE T IH1 M B ER0
TIMBROOK T IH1 M B R UH2 K
TIMBS T IH1 M Z
TIMBUKTU T IH2 M B AH0 K T UW1
TIME T AY1 M
TIME'S T AY1 M Z
TIMED T AY1 M D
TIMEFRAME T AY1 M F R EY2 M
TIMELESS T AY1 M L AH0 S
TIMELINE T AY1 M L AY0 N
TIMELINES T AY1 M L AY0 N Z
```

Output

- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction

time
↓

TIMBLIN	T IH1 M B L IH0 N
TIMBRE	T IH1 M B ER0
TIMBROOK	T IH1 M B R UH2 K
TIMBS	T IH1 M Z
TIMBUKTU	T IH2 M B AH0 K T UW1
TIME	T AY1 M
TIME'S	T AY1 M Z
TIMED	T AY1 M D
TIMEFRAME	T AY1 M F R EY2 M
TIMELESS	T AY1 M L AH0 S
TIMELINE	T AY1 M L AY0 N
TIMELINES	T AY1 M L AY0 N Z

Output

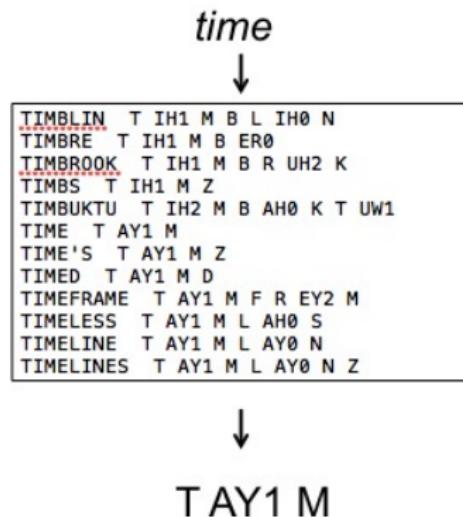
- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction

time
↓

TIMBLIN	T IH1 M B L IH0 N
TIMBRE	T IH1 M B ER0
TIMBROOK	T IH1 M B R UH2 K
TIMBS	T IH1 M Z
TIMBUKTU	T IH2 M B AH0 K T UW1
TIME	T AY1 M
TIME'S	T AY1 M Z
TIMED	T AY1 M D
TIMEFRAME	T AY1 M F R EY2 M
TIMELESS	T AY1 M L AH0 S
TIMELINE	T AY1 M L AY0 N
TIMELINES	T AY1 M L AY0 N Z

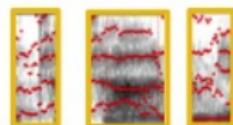
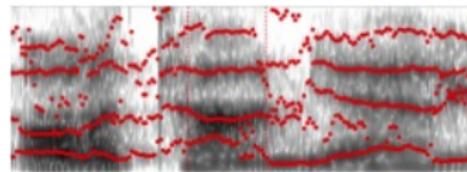
Output

- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction



Output

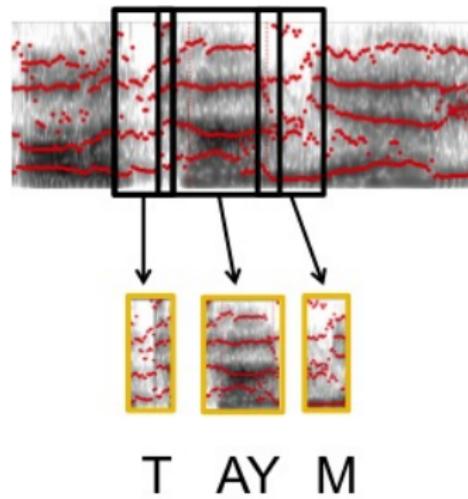
- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction



T AY M

Output

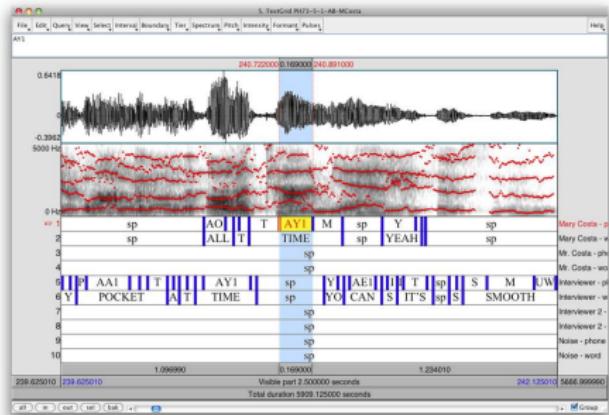
- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction



Output

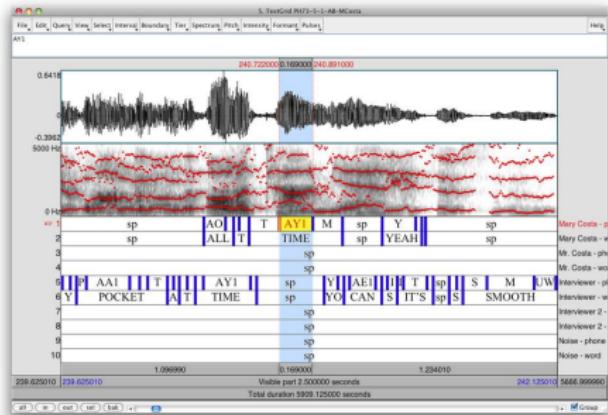
- ▶ Phonemicisation of transcription via pronouncing dictionary
 - ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
 - ▶ Production of aligned phonemic transcription

→ Functions as input for scripts for hand-coding and vowel extraction



Output

- ▶ Phonemicisation of transcription via pronouncing dictionary
- ▶ Matching of speech stream chunks to 'gold standard' acoustic models of phonemes
- ▶ Production of aligned phonemic transcription
- Functions as input for scripts for hand-coding and vowel extraction



1 Introduction

2 Methodology

- Components of forced alignment
- **Different aligners**
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

Aligners under study

- ▶ FAVE (Forced Alignment and Vowel Extraction; Rosenfelder et al. 2011)
- ▶ PLA (ProsodyLab-Aligner; Gorman et al. 2011)
- ▶ SPPAS (SPeech Phonetization Alignment and Syllabification; Bigi 2012)

Aligners under study

- ▶ FAVE (Forced Alignment and Vowel Extraction; Rosenfelder et al. 2011)
- ▶ PLA (ProsodyLab-Aligner; Gorman et al. 2011)
- ▶ SPPAS (SPeech Phonetization Alignment and Syllabification; Bigi 2012)

Aligners under study

- ▶ FAVE (Forced Alignment and Vowel Extraction; Rosenfelder et al. 2011)
- ▶ PLA (ProsodyLab-Aligner; Gorman et al. 2011)
- ▶ SPPAS (SPeech Phonetization Alignment and Syllabification; Bigi 2012)

Loci of variation

There are a number of parameters on which the aligners differ from one another, making cross-aligner comparison non-trivial.

Loci of variation

1. Format of input

- ▶ FAVE: long sound file
- ▶ PLA, SPPAS: breath group sized-sound files

Loci of variation

1. Format of input

- ▶ FAVE: long sound file
- ▶ PLA, SPPAS: breath group sized-sound files

Loci of variation

1. Format of input

- ▶ FAVE: long sound file
- ▶ PLA, SPPAS: breath group sized-sound files

Loci of variation

1. Format of input

- ▶ FAVE: long sound file
- ▶ PLA, SPPAS: breath group sized-sound files

PLA was designed for use with short experimental stimuli.
Accordingly, it is not well equipped to deal with the noise and
overlapping speech common to sociolinguistic interviews.

Loci of variation

2. Format of dictionary

- ▶ FAVE, PLA: ARPAbet
- ▶ SPPAS: X-SAMPA

```
TIMBLIN T IH1 M B L IH0 N
TIMBRE T IH1 M B ER0
TIMBROOK T IH1 M B R UH2 K
TIMBS T IH1 M Z
TIMBUKTU T IH2 M B AH0 K T UW1
TIME T AY1 M
TIME'S T AY1 M Z
TIMED T AY1 M D
TIMEFRAME T AY1 M F R EY2 M
TIMELESS T AY1 M L AH0 S
TIMELINE T AY1 M L AY0 N
TIMELINES T AY1 M L AY0 N Z
```

Loci of variation

2. Format of dictionary

- ▶ FAVE, PLA: ARPAbet
- ▶ SPPAS: X-SAMPA

ticking	"tI.kɪN
tight	"talt
tight	"tal?
til	"tɪl
till	"tɪl
timing	"tal.mɪN
tired	"tal.@d
toilets	"tOI.l@ts
toilets	"tOI.llts
Topman	"tQp.m{n
torture	"tO:.tS@
Toyota	tOI."@U.t@
Toyota	tOI."@U.?@
track	"tr{k
tracks	"tr{ks
traders	"treld@z
trading	"treld.iN
trail	"trell
trainers	"treln@z

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it has no acoustic models for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for

	British dictionary, ARPA (IPA)	FAVE-friendly ARPA (IPA)
START	AA (a:)	→ AA (a)
THOUGHT	AO (ɔ:)	→ AO (ɔ)
LOT	OH (ɒ)	→ N/A (our solution: AA (a))

Loci of variation

3. Acoustic model

- ▶ FAVE's 'gold standard' acoustic models are based on speech from a corpus of Supreme Court Justices
 - ▶ As such, its models contain "General American" formant structures
 - ▶ And it **has no acoustic models** for phonemes that do not exist in General American
 - ▶ Accordingly, transcriptions must be coerced into the vowel classes that it does have models for
- ▶ By contrast, other aligners create their own acoustic models based on your data

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

Data sources

Four speakers, representing four different dialects of UK English:

Data sources

Four speakers, representing four different dialects of UK English:

- ▶ **David Attenborough,**
nature documentary narrator: RP
- ▶ **Victoria “Posh Spice” Beckham,**
fashion designer & former
member of the Spice Girls: Essex
- ▶ **Melanie “Sporty Spice” Chisholm,**
performer & former member
of the Spice Girls: Liverpool
- ▶ **Holly,** university student: Manchester



Data sources

Four speakers, representing four different dialects of UK English:

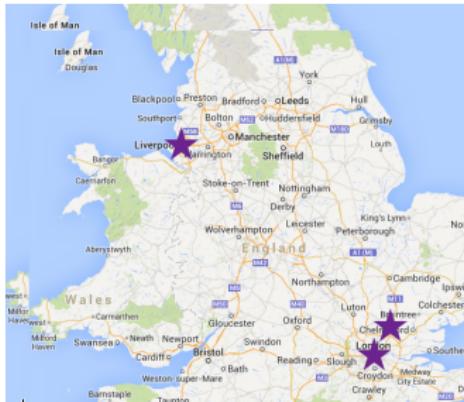
- ▶ **David Attenborough**,
nature documentary narrator: RP
- ▶ **Victoria “Posh Spice” Beckham**,
fashion designer & former
member of the Spice Girls: Essex
- ▶ **Melanie “Sporty Spice” Chisholm**,
performer & former member
of the Spice Girls: Liverpool
- ▶ **Holly**, university student: Manchester



Data sources

Four speakers, representing four different dialects of UK English:

- ▶ **David Attenborough,**
nature documentary narrator: RP
- ▶ **Victoria “Posh Spice” Beckham,**
fashion designer & former
member of the Spice Girls: Essex
- ▶ **Melanie “Sporty Spice” Chisholm,**
performer & former member
of the Spice Girls: Liverpool
- ▶ **Holly, university student: Manchester**



Data sources

Four speakers, representing four different dialects of UK English:

- ▶ **David Attenborough**,
nature documentary narrator: RP
- ▶ **Victoria “Posh Spice” Beckham**,
fashion designer & former
member of the Spice Girls: Essex
- ▶ **Melanie “Sporty Spice” Chisholm**,
performer & former member
of the Spice Girls: Liverpool
- ▶ **Holly**, university student: Manchester



Evaluating aligner performance

We identified five vowels which differentiate the dialects under study:

- ▶ FOOT /ʊ/, STRUT /ʌ/ ▶ BATH /ɑ:/, TRAP /æ/
 - RP & Essex have phonemic splits; Manchester & Liverpool have one phoneme for each
- ▶ GOOSE /u:/
 - involved in fronting

Evaluating aligner performance

We identified five vowels which differentiate the dialects under study:

- ▶ FOOT /ʊ/, STRUT /ʌ/ ▶ BATH /ɑ:/, TRAP /æ/
 - RP & Essex have phonemic splits; Manchester & Liverpool have one phoneme for each
- ▶ GOOSE /u:/
 - involved in fronting

Evaluating aligner performance

We identified five vowels which differentiate the dialects under study:

- ▶ FOOT /ʊ/, STRUT /ʌ/ ▶ BATH /ɑ:/, TRAP /æ/
 - RP & Essex have phonemic splits; Manchester & Liverpool have one phoneme for each
- ▶ GOOSE /u:/
 - involved in fronting

Evaluating aligner performance

We identified five vowels which differentiate the dialects under study:

- ▶ FOOT /ʊ/, STRUT /ʌ/ ▶ BATH /ɑ:/, TRAP /æ/
 - RP & Essex have phonemic splits; Manchester & Liverpool have one phoneme for each
- ▶ GOOSE /u:/
 - involved in fronting

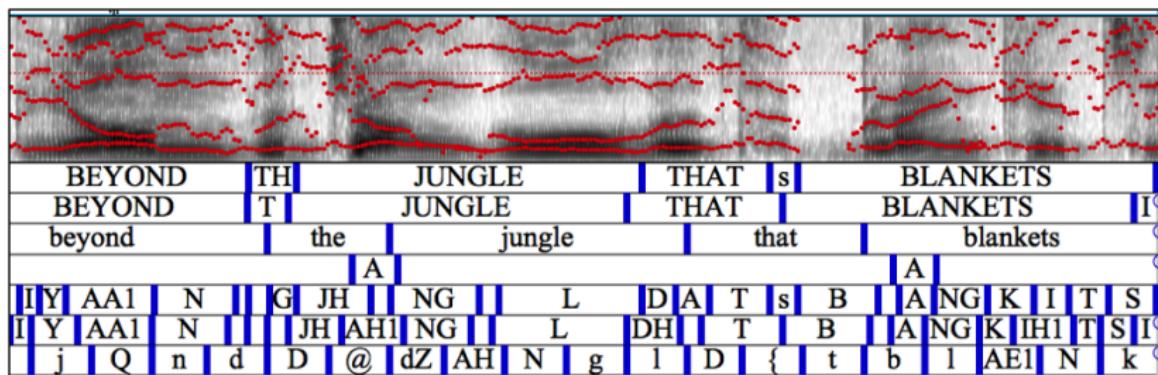
Evaluating aligner performance

We identified five vowels which differentiate the dialects under study:

- ▶ FOOT /ʊ/, STRUT /ʌ/ ▶ BATH /ɑ:/, TRAP /æ/
 - RP & Essex have phonemic splits; Manchester & Liverpool have one phoneme for each
- ▶ GOOSE /u:/
 - involved in fronting

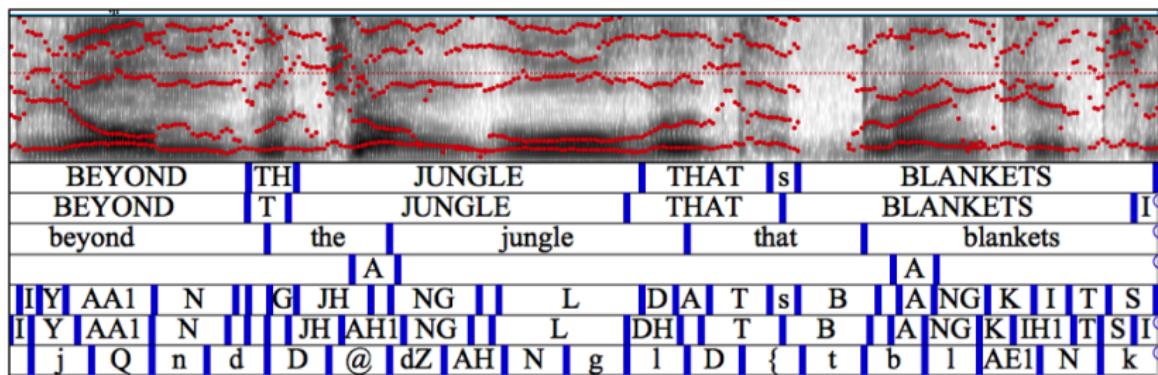
Evaluating aligner performance

- ▶ 1188 tokens of the vowels of interest were hand-annotated
- ▶ These “gold standard” annotations were then compared to the aligners’ placement of the same vowels



Evaluating aligner performance

- ▶ 1188 tokens of the vowels of interest were hand-annotated
- ▶ These “gold standard” annotations were then compared to the aligners’ placement of the same vowels



- ▶ Two measurements per vowel per aligner: onset displacement from gold standard, offset displacement from gold standard

Table of Contents

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

Overall performance

How well do the aligners' vowel placements match the humans'?

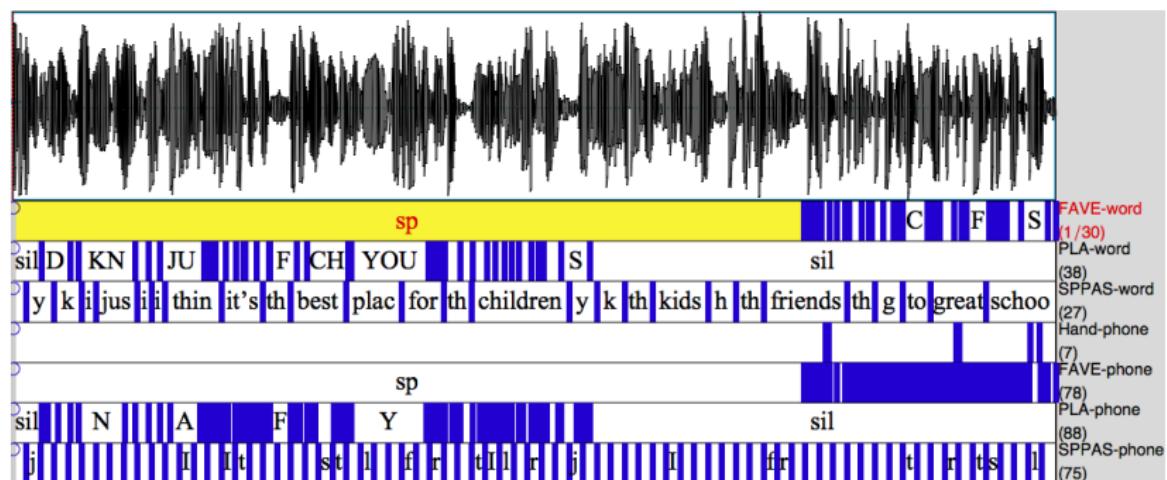
→ Measurement: vowel boundary displacement (in seconds)

	Median		Mean		Max	
	Onset	Offset	Onset	Offset	Onset	Offset
FAVE	0.009	0.009	0.019	0.021	0.583	0.588
PLA	0.015	0.019	0.267	0.252	55.473	55.488
SPPAS	0.150	0.144	0.504	0.480	68.903	67.408

Table 1: Vowel boundary displacement in seconds for three aligners

Sources of error

Substantial errors may occur when the aligner misidentifies noise or speech from an interlocutor as speech from the target speaker:



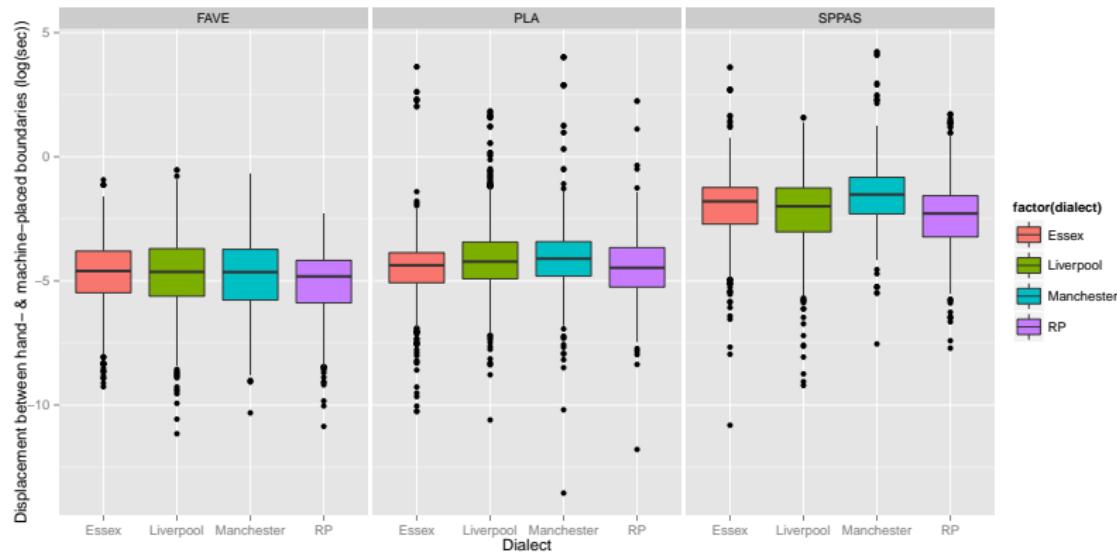
Performance on standard dialect

But, restrict the data to David Attenborough and the aligners improve:

	PLA		SPPAS	
	RP	Other dialects	RP	Other dialects
median	0.011	0.014	0.102	0.164
mean	0.067	0.312	0.281	0.554
max	9.436	55.490	5.576	68.900

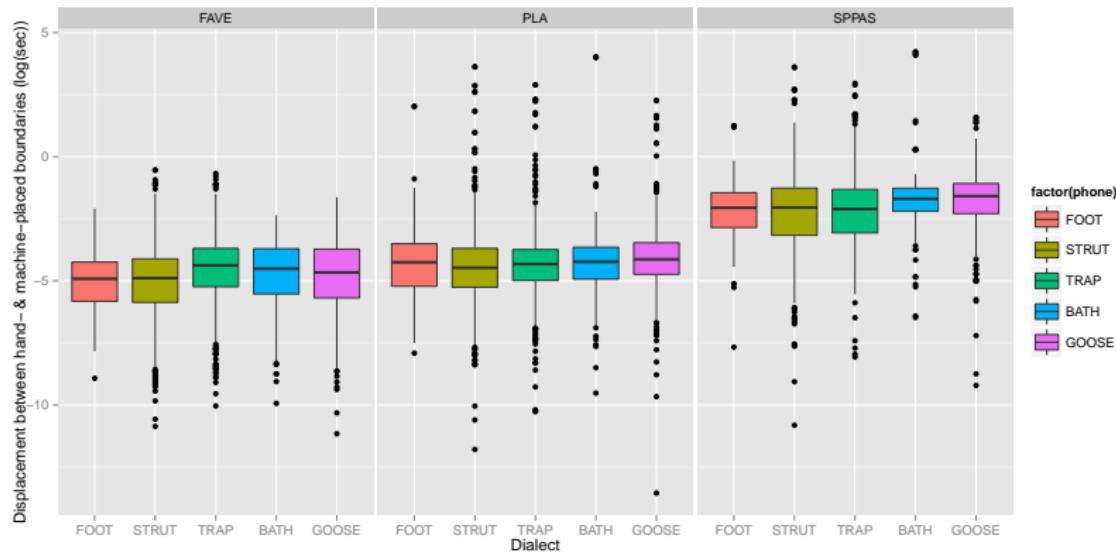
Table 2: Vowel displacement by dialect

Performance by dialect



Significant interaction between PLA & Manchester ($\beta = 0.722, p = 7.15 \times 10^{-4}$) and SPPAS & Manchester ($\beta = 1.034, p = 1.78 \times 10^{-6}$)

Performance by phoneme



Significant interaction between PLA & BATH ($\beta = 1.33, p = 6.80 \times 10^{-4}$) and SPPAS & BATH ($\beta = 1.66, p = 2.40 \times 10^{-5}$)

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- **Qualitative findings**
- User-friendliness

4 Conclusion

Coping with non-standard dialect features

Deletion processes

- ▶ h-dropping
- ▶ rhoticity (and linking /r/)
- ▶ l-vocalisation

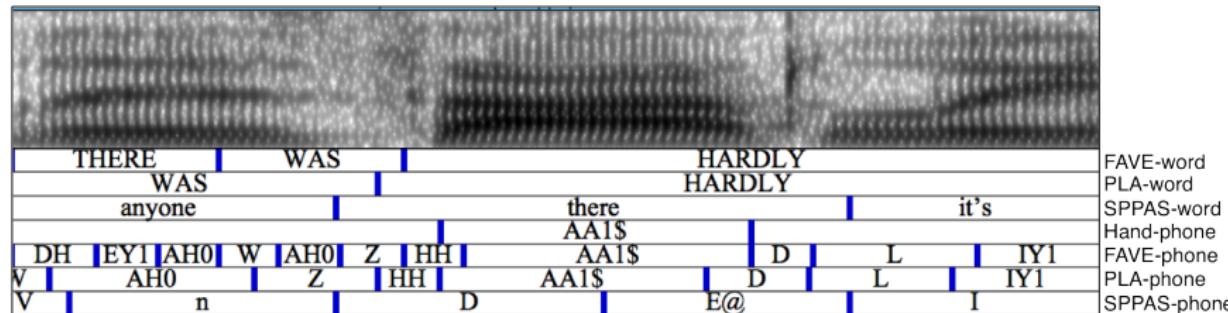
Lenition processes

- ▶ t-glottalling
- ▶ stop lenition

Unexpected segments

- ▶ post-nasal stops
- ▶ intrusive /r/

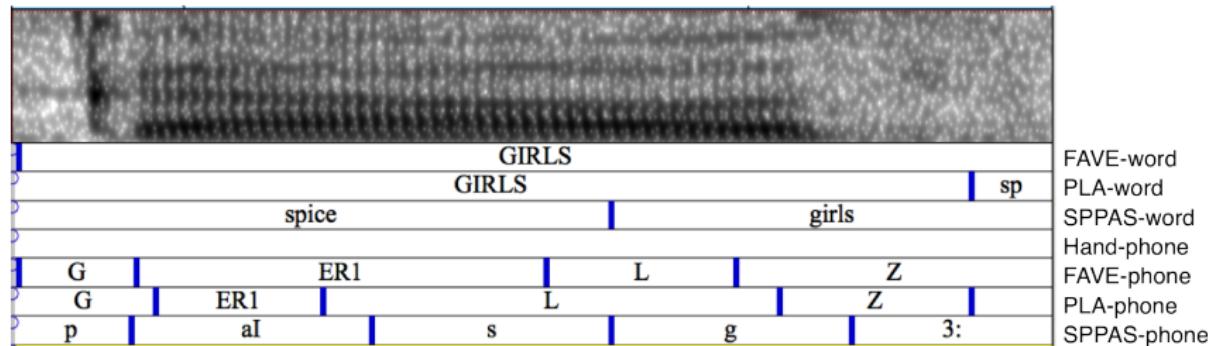
h-dropping



Holly, Manchester: *There was hardly anyone, hardly*

- ▶ Human hears no /h/, but aligners have to find a place for one
- ▶ If dictionary had alternate pronunciations with deleted /h/s, this could be avoided

/l-vocalisation

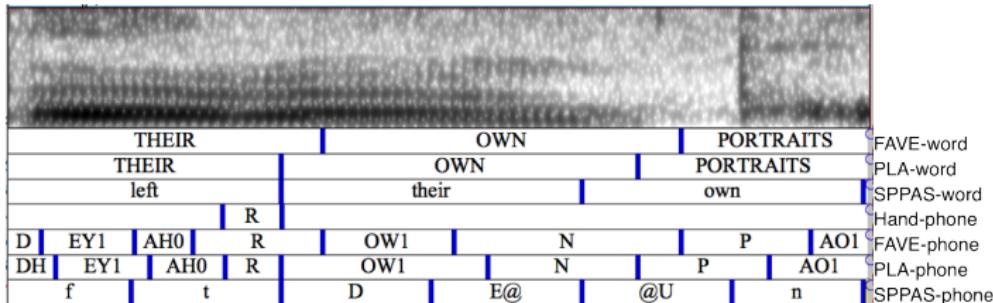


Victoria Beckham, Essex: *girls*

- ▶ Human hears vocalized /l/; aligners disagree on where to place boundary between vowel and /l/

Linking /r/

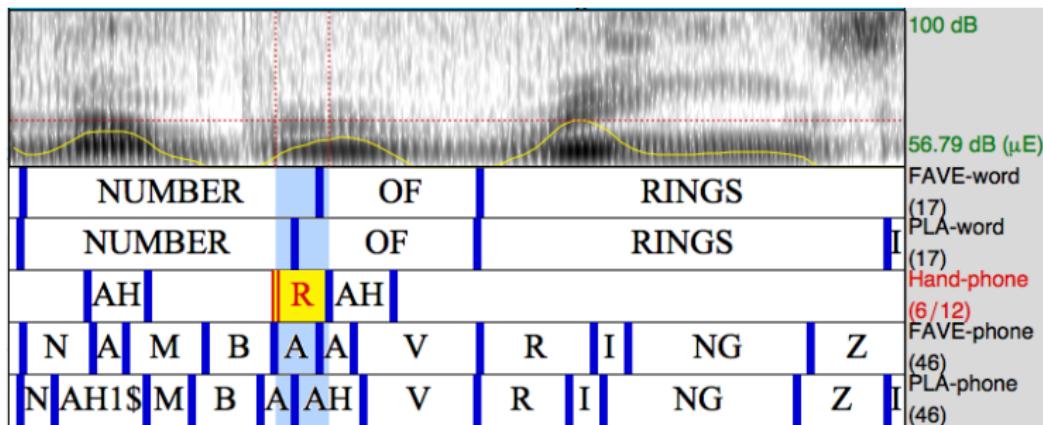
- ▶ The BEEP dictionary has both rhotic and non-rhotic pronunciation for words with post-vocalic /r/.
- ▶ This is useful, as it means the aligners can look out for /r/s in linking position, e.g. in the phrase below, *their own portraits*.



David Attenborough, RP: *their own*

Linking /r/

However, if the /r/ has a particularly weak realisation, the aligner won't include it.

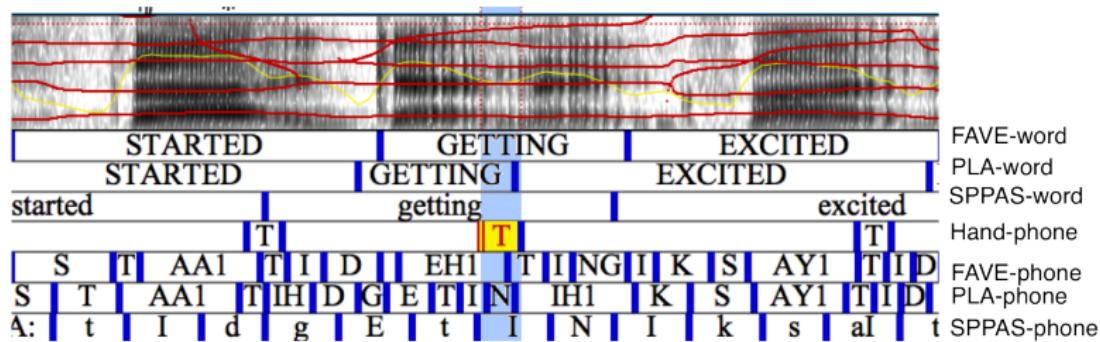


David Attenborough, RP: *number of rings, number of*

The very faint, yet audible, /r/ is not detected by the speech recognition software, which is expecting a more consonantal /r/.

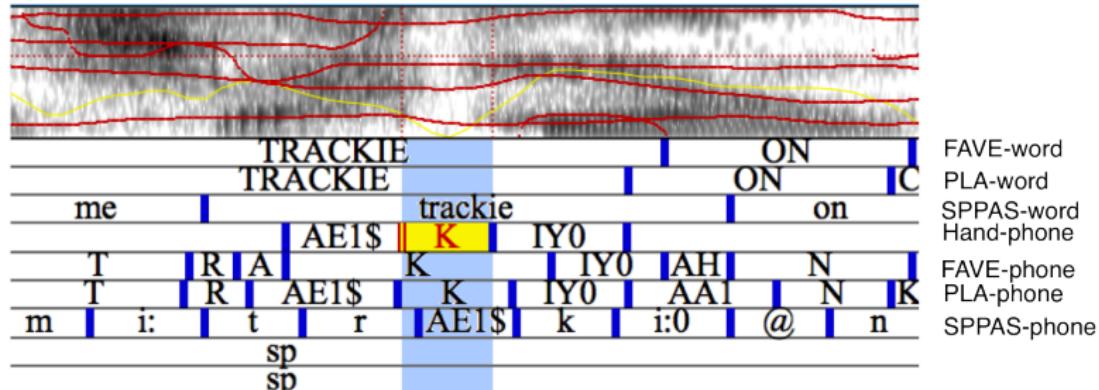
t-glottalling

Varying degrees of success with intervocalic glottalling.



Holly, Manchester: *When he started getting excited*

Liverpool stop lenition

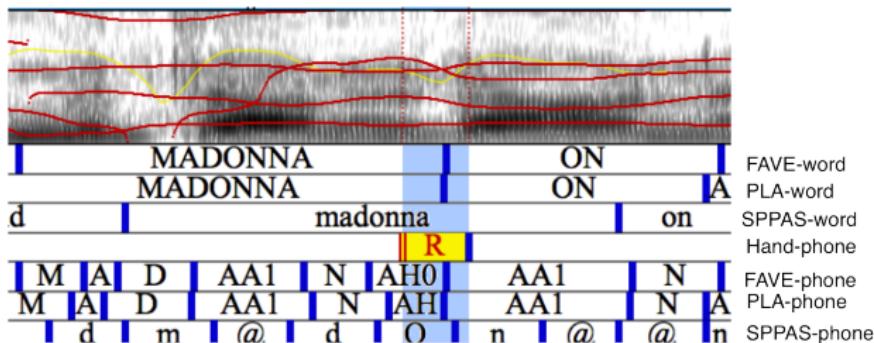


Mel C, Liverpool: *when I've got my trackie on, trackie*

- ▶ Human hears /k/ as fricativised [χ], but aligners don't recognise this as the stop phoneme.

Intrusive /r/

None of the dictionaries are equipped to deal with intrusive /r/, where the /r/ was not present etymologically and therefore is not reflected in the spelling.

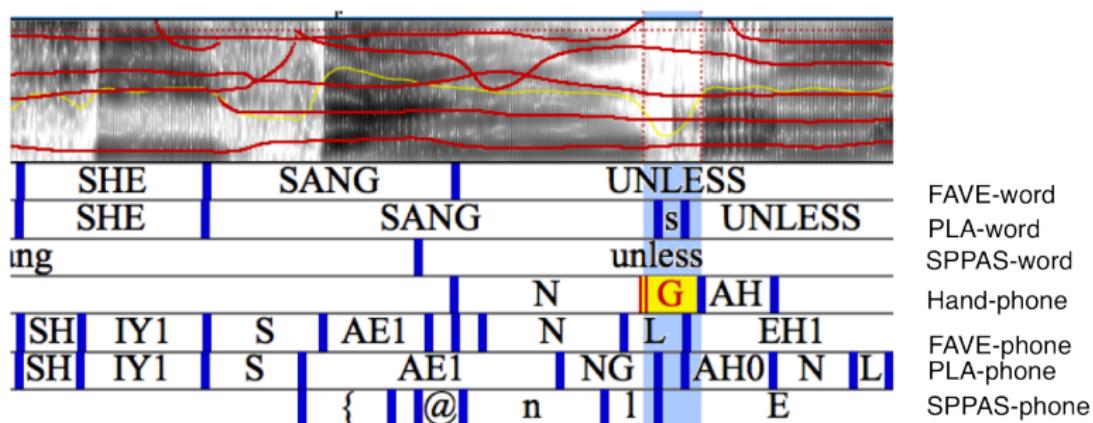


Mel C, Liverpool: *stick the old Madonna on, Madonna on*

- ▶ The /r/ is not accounted for by any of the aligners.
- ▶ Perhaps transcribers could take such tokens into account and include some indication in the orthography e.g. (r)

Post-nasal stops

Dialects in the North-West of England retain the stop after velar nasals in words such as *sang, king*.



Holly, Manchester: *I don't think she sang, sang*

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

User-friendliness

- ▶ No aligner is trivial to install locally: requires HTK Speech Recognition Toolkit, SoX (sound processing program) . . .
 - ▶ Instructions for local installation of FAVE only available for Macs

User-friendliness

- ▶ No aligner is trivial to install locally: requires HTK Speech Recognition Toolkit, SoX (sound processing program) . . .
 - ▶ Instructions for local installation of FAVE only available for Macs

User-friendliness

- ▶ No aligner is trivial to install locally: requires HTK Speech Recognition Toolkit, SoX (sound processing program) . . .
 - ▶ Instructions for local installation of FAVE only available for Macs
- ▶ FAVE has a user-friendly online interface, but this is not terribly customizable (cannot be customized to permanently use British pronouncing dictionary)

Home | FAVE-align | FAVE-extract | How to use FAVE-align | How to use FAVE-extract | About

 **FAVE-align**
An online interface for the Penn Forced Aligner

Upload files:
Please select the files to be aligned:
? Sound file: No file chosen
? Transcription file: No file chosen

Options:
 -u: Check transcription for unknown words ?
 -i: Import dictionary transcriptions: ? No file chosen

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?

User-friendliness

- ▶ No aligner is trivial to install locally: requires HTK Speech Recognition Toolkit, SoX (sound processing program). . .
 - ▶ Instructions for local installation of FAVE only available for Macs
- ▶ FAVE has a user-friendly online interface, but this is not terribly customizable (cannot be customized to permanently use British pronouncing dictionary)
- ▶ PLA requires substantial pre-processing (breaking sound & transcription files into breath group-sized chunks)

Table of Contents

1 Introduction

2 Methodology

- Components of forced alignment
- Different aligners
- The present study

3 Findings

- Quantitative findings
- Qualitative findings
- User-friendliness

4 Conclusion

Conclusion

- ▶ In general, we were impressed at how well the aligners performed
- ▶ State-of-the-art forced alignment techniques are a viable option for non-standard sociolinguistic data

Our recommendations

- ▶ Take the time to install FAVE locally, then use it with a standard pronouncing dictionary
 - !! The installation instructions on FAVE's website aren't up to date with the latest Mac OS. Contact us if you need help
- ▶ Coercion to American vowel classes will require hand-recoding if your vowel of interest is affected, which is a pain...
- ▶ ...but PLA, which is equipped to deal with this, just can't handle the messiness of sociolinguistic interviews
 - That said, if you have non-messy data (e.g. word lists), PLA is a good alternative, if you don't mind the pre-processing

Our recommendations

- ▶ Take the time to install FAVE locally, then use it with a standard pronouncing dictionary
 - !! The installation instructions on FAVE's website aren't up to date with the latest Mac OS. Contact us if you need help
- ▶ Coercion to American vowel classes will require hand-recoding if your vowel of interest is affected, which is a pain...
- ▶ ...but PLA, which is equipped to deal with this, just can't handle the messiness of sociolinguistic interviews
 - That said, if you have non-messy data (e.g. word lists), PLA is a good alternative, if you don't mind the pre-processing

Our recommendations

- ▶ Take the time to install FAVE locally, then use it with a standard pronouncing dictionary
 - !! The installation instructions on FAVE's website aren't up to date with the latest Mac OS. Contact us if you need help
- ▶ Coercion to American vowel classes will require hand-recoding if your vowel of interest is affected, which is a pain...
- ▶ ...but PLA, which is equipped to deal with this, just can't handle the messiness of sociolinguistic interviews
 - That said, if you have non-messy data (e.g. word lists), PLA is a good alternative, if you don't mind the pre-processing

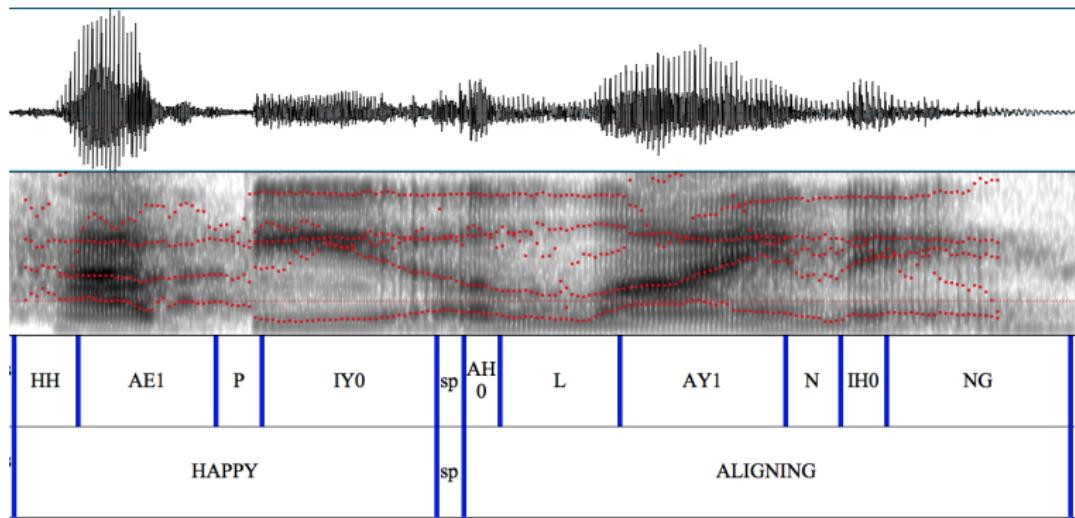
Our recommendations

- ▶ Take the time to install FAVE locally, then use it with a standard pronouncing dictionary
 - !! The installation instructions on FAVE's website aren't up to date with the latest Mac OS. Contact us if you need help
- ▶ Coercion to American vowel classes will require hand-recoding if your vowel of interest is affected, which is a pain...
- ▶ ...but PLA, which is equipped to deal with this, just can't handle the messiness of sociolinguistic interviews
 - That said, if you have non-messy data (e.g. word lists), PLA is a good alternative, if you don't mind the pre-processing

Our recommendations

- ▶ Take the time to install FAVE locally, then use it with a standard pronouncing dictionary
 - !! The installation instructions on FAVE's website aren't up to date with the latest Mac OS. Contact us if you need help
- ▶ Coercion to American vowel classes will require hand-recoding if your vowel of interest is affected, which is a pain...
- ▶ ...but PLA, which is equipped to deal with this, just can't handle the messiness of sociolinguistic interviews
 - That said, if you have non-messy data (e.g. word lists), PLA is a good alternative, if you don't mind the pre-processing

Happy Aligning!



Acknowledgements

With special thanks to the members of Autumn 2012's LELA30981
Language Change Across the Lifespan for providing their
transcriptions, our speakers, and you!

References

- Gorman, Kyle, Jonathan Howell, and Michael Wagner. 2011. Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39:192–193.
- Harrington, Jonathan, Felicitas Kleber, and Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123:2825.
- Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: linear incrementation, reversal, and reanalysis. *Language* 89:30–65.
- Robinson, Anthony. 1994. The British English Example Pronunciation dictionary, v0. 1. *Cambridge University* .
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, and Jiahong Yuan. 2011. FAVE (forced alignment and vowel extraction) program suite. URL <http://fave.ling.upenn.edu> .
- Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.