Statistical Acoustic-Phonetic Historical Linguistics: a short introduction

John Aston

John Coleman

UNIVERSITY OF CAMBRIDGE

Statistical Laboratory

UNIVERSITY OF OXFORD

Phonetics Laboratory

With thanks to: Davide Pigoli, Pantelis Hadjipantelis, Danny Yee, John Pybus, EPSRC and AHRC

Ancient Sounds

- What would comparative and historical phonology be like if we worked with sounds instead of symbols?
- Can quantitative methods give insights into language variation and change?
- Could we “bring back to life” the sounds of dead languages?

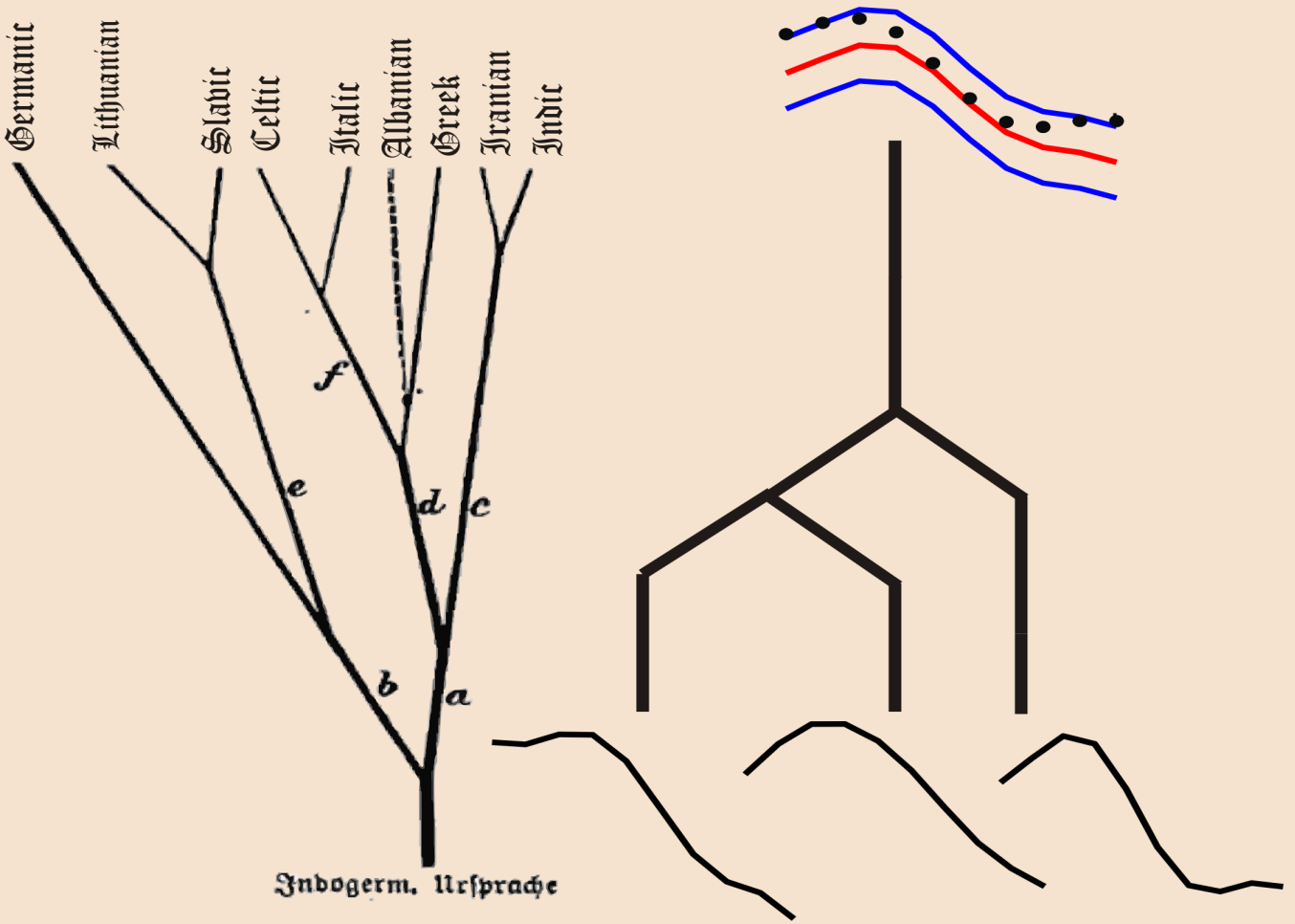
Organisation

I. Some basics

II. Spaces of variation

III. Variation through time

I. Some basics



After Schleicher (1860) *Deutsche Sprache,* and Aston, Buck, Coleman, Cotter, Jones, Macaulay, MacLeod, Moriarty, and Nevins (2011) Phylogenetic inference for function-valued traits: speech sound evolution. doi: 10.1016/j.tree.2011.10.001

Regular similarities ⇐ Shared ancestry

Dissimilarities ⇐ Historical divergence

	Latin		Italian	Spanish	Portuguese	French
1	*unus*	≈ ['u:nu-s/m]	['u:no]	['u:no]	[ũ(ŋ)]	[œ̃], [ɛ̃]
2	*duo*	≈ ['duo]	['du:e]	[dos]	[dois], [doiʃ]	[dø]
3	*tres*	≈ [tre:s]	[tre:]	[tres]	[tres], [treʃ]	[tʀwa]
4	*quattuor*	≈ ['kwatwor]	['kwat:ro]	['kwatro]	['kwatru]	[katʀ]
5	*quinque*	≈ ['kwi:ŋkwe]	['tʃiŋkwe]	['si:ŋko], ['θi:ŋko]	['sĩŋku]	[sɛ̃k]
6	*sex*	≈ [seks]	[sɛi]	[seis]	[sɐ̆is], [sɐ̆iʃ]	[sis]
7	*septem*	≈ ['septem]	['sɛt:e]	['sjete]	['sɛti]	[sɛt]
8	*octo*	≈ ['okto]	['ot:o]	['otʃo]	['oĭtu]	[ɥit]
9	*novem*	≈ ['novem]	['nɔve]	['nweve]	['nɔvɨ]	[nœf]
10	*decem*	≈ ['dekem]	['djɛtʃi]	['djɛs], ['djɛθ]	['dɛʃ]	[dis]

Dissimilarities ⇐ Historical divergence

Sound change rules describe such divergences, e.g.

[un] > [ũ] (Nasalization; fusional assimilation)

[ũ] > [œ̃] > [ɛ̃] (Nasal vowel lowering)

$[s] > [ʃ] \;/\; \left\{ \begin{matrix} i \\ e \end{matrix} \right\} \_\_\_$ (Postalveolarization)

	Latin		Italian	Spanish	Portuguese	French
1	*unus*	≈ ['u:nu-s/m]	['u:no]	['u:no]	[ũ(ŋ)]	[œ̃], [ɛ̃]
2	*duo*	≈ ['duo]	['du:e]	[dos]	[dois], [doiʃ]	[dø]
3	*tres*	≈ [tre:s]	[tre:]	[tres]	[tres], [treʃ]	[tʀwa]

But these () are not sounds!

	Latin	Italian	Spanish	Portuguese	French
1	[ˈuːnu-s/m]	[ˈuːno]	[ˈuːno]	[ũ(ŋ)]	[œ̃], [ɛ̃]
2	[ˈduo]	[ˈduːe]	[dos]	[dois], [doiʃ]	[dø]
3	[treːs]	[treː]	[tres]	[tres], [treʃ]	[tʀwa]
4	[ˈkwatwor]	[ˈkwatːro]	[ˈkwatro]	[ˈkwatru]	[katʀ]
5	[ˈkwiːŋkwe]	[ˈtʃiŋkwe]	[ˈsiːŋko], [ˈθiːŋko]	[ˈsĩŋku]	[sɛ̃k]
6	[seks]	[sɛi]	[seis]	[sɐ̆s], [sɐ̆ʃ]	[sis]
7	[ˈseptem]	[ˈsɛtːe]	[ˈsjete]	[ˈsɛti]	[sɛt]
8	[ˈokto]	[ˈotːo]	[ˈotʃo]	[ˈoĭtu]	[ɥit]
9	[ˈnovem]	[ˈnɔve]	[ˈnweve]	[ˈnɔvɨ]	[nœf]
10	[ˈdekem]	[ˈdjɛtʃi]	[ˈdjɛs], [ˈdjɛθ]	[ˈdɛʃ]	[dis]

But these () are not sounds!

	Latin	Italian	Spanish	Portuguese	French
1	['◆✚■◆✍▲✏❍]	[✝◆✚■❑]	[✝◆✚■❑]	[◆✈✝✉]	[œ], [✝]
2	['❄◆❑]	[✝❄◆✚❄]	[❄❑▲]	[❄❑❊▲], [❄❑❊✝]	[❄➽→]
3	[▼❒❄✚▲]	[▼❒❄✚]	[▼❒❄▲]	[▼❒❄▲], [▼❒❄✝]	[▼◗❁]
4	[✝✱◗❁▼◗❑❒]	[✝✱◗❁▼✚❒❑]	[✝✱◗❁▼❒❑]	[✝✱◗❁▼❒◆]	[✱❁▼✝]
5	[✝✱◗❊✚✝✱◗❄]	[✝▼✝❊✝✱◗❄]	[✝▲❊✚✝✱❑], [✝✪❊✚✝✱❑]	[✝▲✝✱◆]	[▲✝✱]
6	[▲❄✳▲]	[▲✝❊]	[▲❄❊▲]	[▲✝✝▲], [▲✝✝]	[▲❊▲]
7	[✝▲❄❒▼❄❍]	[✝▲✝▼✚❄]	[✝▲✲❄▼❄]	[✝▲✝▼❊]	[▲✝▼]
8	[✝❑✳▼❑]	[✝❑▼✚❑]	[✝❑▼✝❑]	[✝❑✝▼◆]	[✝❊▼]
9	[✝■❑❖❄❍]	[✝■✝❖❄]	[✝■◗❄❖❄]	[✝■✝❖✝]	[■œ❊]
10	[✝❄❄✳❄❍]	[✝❄✲✝▼✝❊]	[✝❄✲✝▲], [✝❄✲✪]	[✝❄✝]	[❄❊▲]

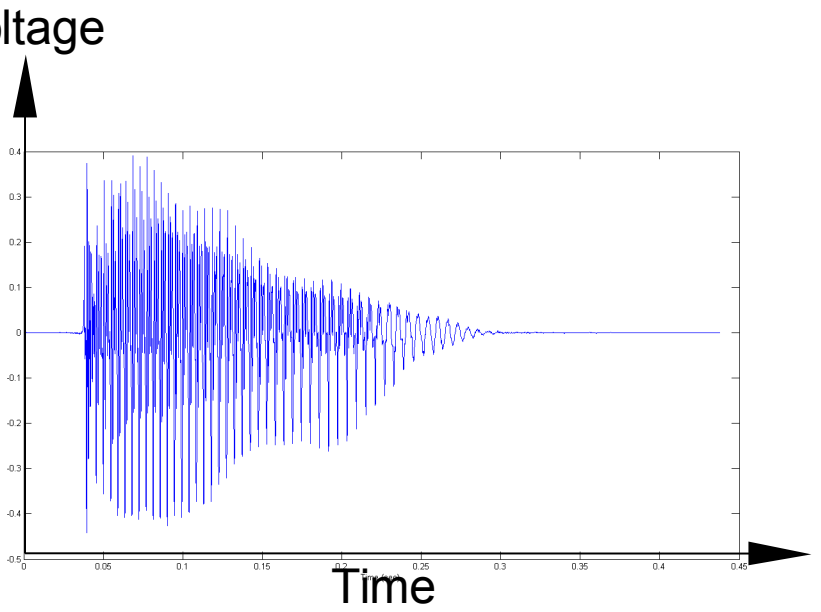
What are sounds?

What are sounds?

- Sound is waves of variation in air pressure (vibrations)
- The vibrations must be fast enough that our ears can detect it (between about 50 Hz and 18,000 Hz)
- We detect, measure and record sound waves using a *microphone*

What does a microphone do?

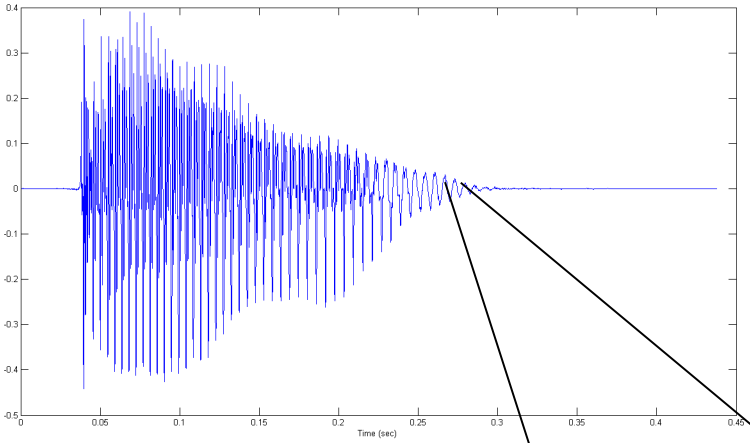
A microphone converts variations in air pressure to (corresponding) variations in voltage



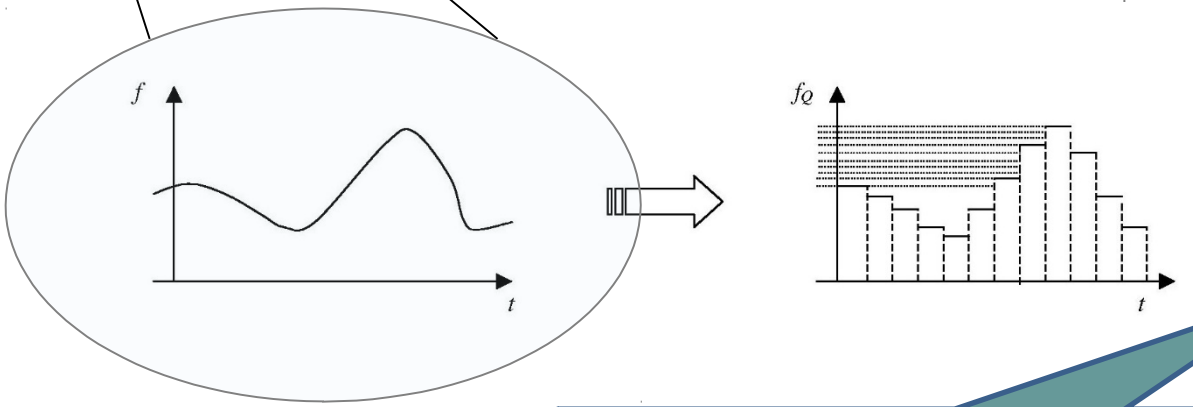
Digitization: turning sound into numbers

A computer's sound card contains an analogue-to-digital convertor, that measures the voltage tens of thousands of times per second and stores them as a (very long) list of numbers

Digitization: turning sound into numbers



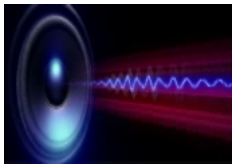
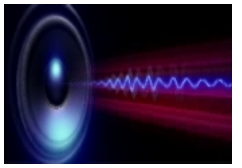
A computer's sound card contains an *analogue-to-digital convertor* that measures the voltage tens of thousands of times per second and stores them as a (very long) list of numbers

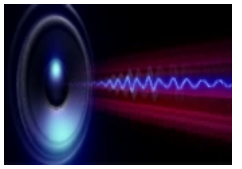


$$\begin{bmatrix} 104 \\ 32 \\ -81 \\ 23 \\ 98 \\ 200 \\ 80 \\ -3 \\ -105 \\ -75 \\ 12 \\ 56 \\ \vdots \\ \vdots \end{bmatrix}$$

This vector *is* a sound file, e.g. a wav file

Some things you can do with numbers

- Addition: sound 1 + sound 2 = 
  (*mixing*)
- Subtraction: noise + speech 
  (*noise cancellation*)

  – noise = 

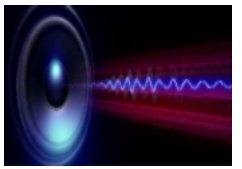
- Multiplication: $2 \times$ sound $=$ louder
  (*amplification*)
- Division: sound $\div 2 =$ quieter
  (*attenuation*)

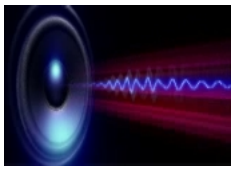
Some things we can do with numbers

How about something clever?

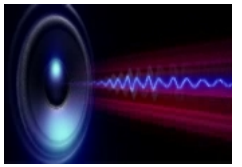
What sounds lie in between A and B?

- What is the average of [ũ] and [ɛ̃]?
  i.e. (sound 1 + sound 2) ÷ 2





- Is it [œ̃]?

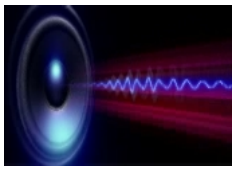
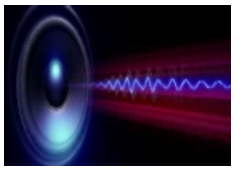


Some things we can do with numbers

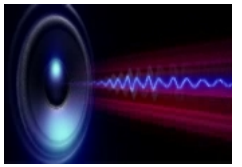
How about something clever?

What sounds lie in between A and B?

- What is the average of [ũ] and [ɛ̃]?
  i.e. (sound 1 + sound 2) ÷ 2

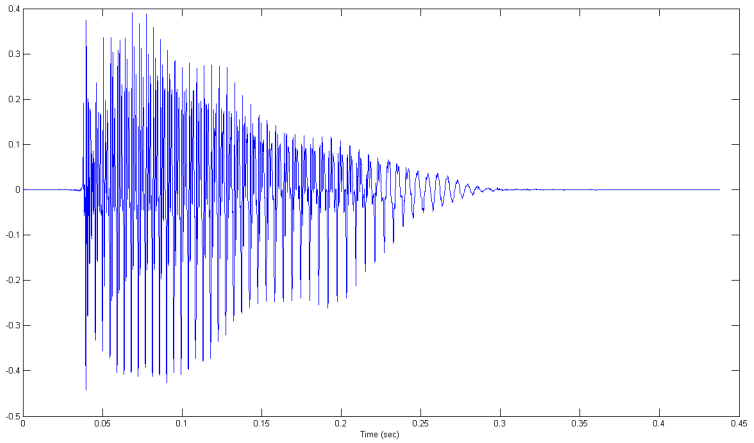



- Is it [œ̃]?



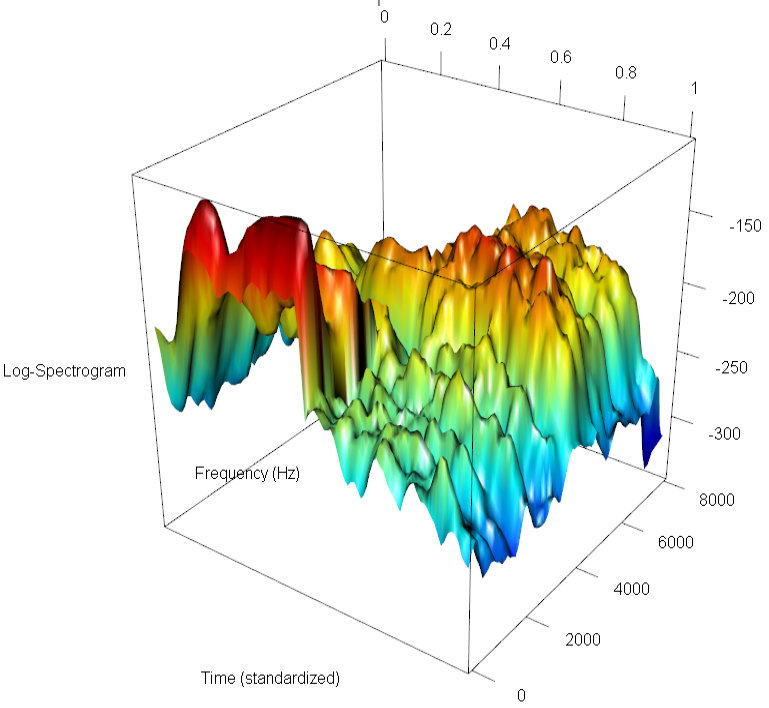
- No: to get something in between two sounds, we need to convert the 1-dimensional sound waves into 2-D surfaces: *spectrograms*

Spectrograms: sounds as *surfaces*

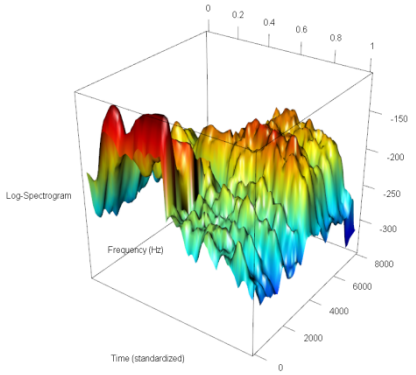


$$\begin{bmatrix} 104 \\ 32 \\ -81 \\ 23 \\ 98 \\ 200 \\ 80 \\ -3 \\ -105 \\ -75 \\ 12 \\ 56 \\ \vdots \\ \vdots \end{bmatrix}$$

$$\tilde{y}(w) = \log \int e^{i\omega t} y(t) dt$$



Averaging spectrograms



$$\begin{bmatrix} 14 & 72 & \cdots \\ 32 & 63 & \cdots \\ 81 & 89 & \cdots \\ 23 & 104 & \cdots \\ 98 & 64 & \cdots \\ 27 & 200 & \cdots \\ 80 & 3 & \cdots \\ 105 & 76 & \cdots \\ 7 & 18 & \cdots \\ 7 & 57 & \cdots \\ 12 & 160 & \cdots \\ 17 & 30 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} + \begin{bmatrix} 28 & 113 & \cdots \\ 587 & 613 & \cdots \\ 19 & 29 & \cdots \\ 28 & 84 & \cdots \\ 11 & 42 & \cdots \\ 121 & 110 & \cdots \\ 30 & 13 & \cdots \\ 10 & 16 & \cdots \\ 7 & 78 & \cdots \\ 17 & 103 & \cdots \\ 32 & 64 & \cdots \\ 51 & 12 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \div 2 = \begin{bmatrix} 21 & 93 & \cdots \\ 310 & 338 & \cdots \\ 50 & 59 & \cdots \\ 26 & 94 & \cdots \\ 55 & 53 & \cdots \\ 74 & 155 & \cdots \\ 55 & 8 & \cdots \\ 58 & 46 & \cdots \\ 7 & 48 & \cdots \\ 12 & 80 & \cdots \\ 22 & 112 & \cdots \\ 34 & 21 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

This surface is half-way between the other two

Sound change continua

Sound 1

spectrogram 1

100% spec 1 + 0% spec 2

90% spec 1 + 10% spec 2

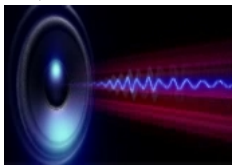
50% spec 1 + 50% spec 2

0% spec 1 + 100% spec 2

Interpolated spectrograms

Sound 2

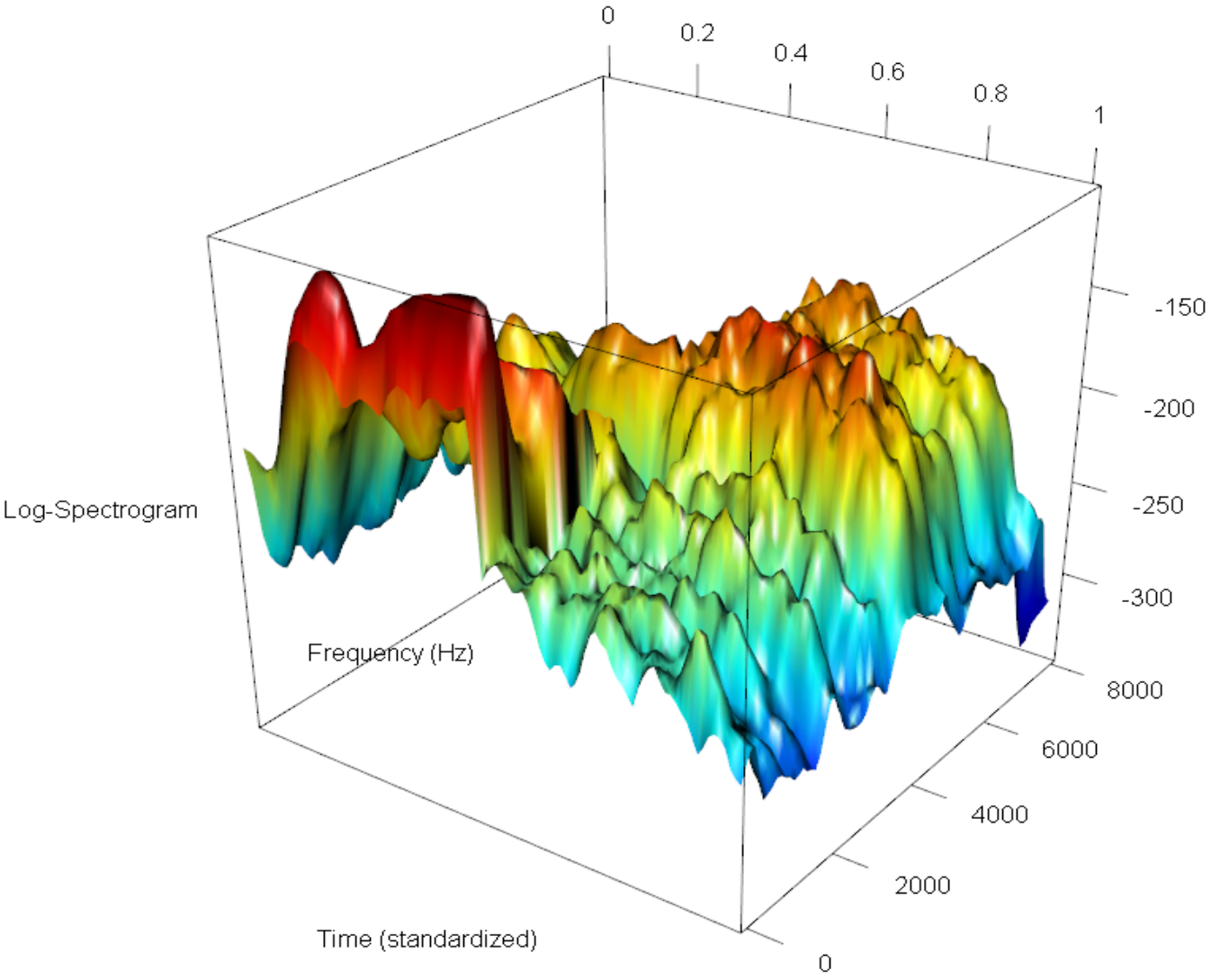
spectrogram 2



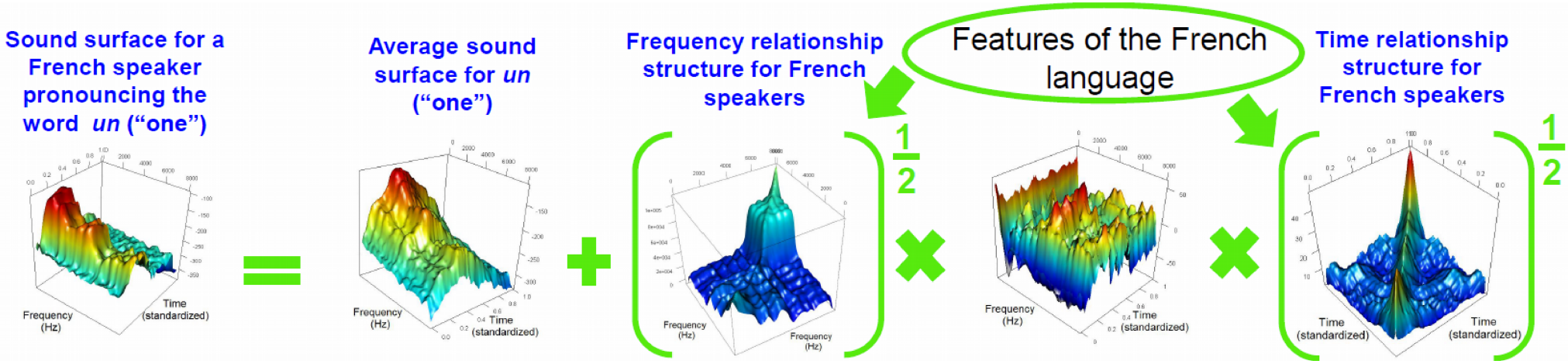
GENERATION OF SYNTHETIC SPEECH
D Moore, J Coleman, Eur. Patent 1,504,443

II. Spaces of variation

Sounds as data



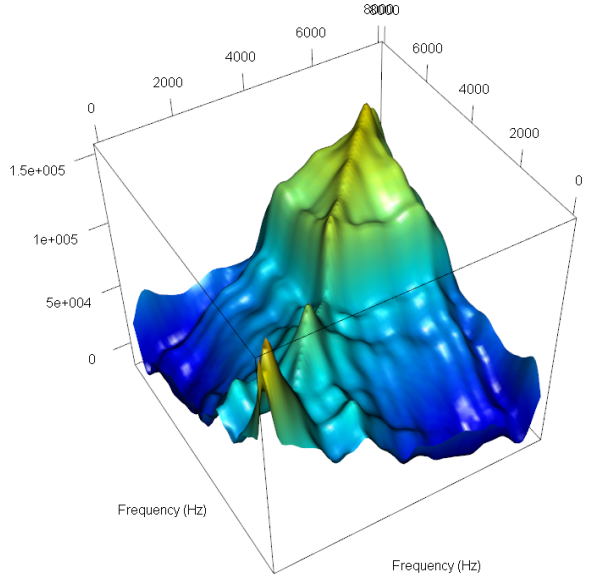
Decomposing acoustics of a word into its pronunciation, language, and speaker



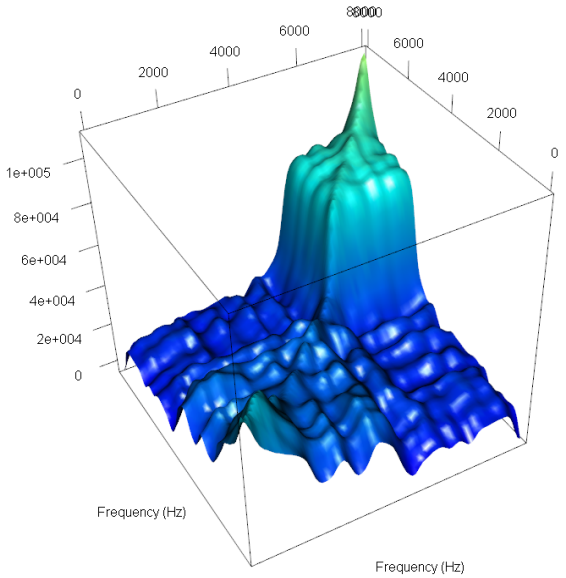
sound of word = pronunciation + lang. frequencies × speaker × lang. timing

Covariance surfaces of average spectrograms: “language” representations

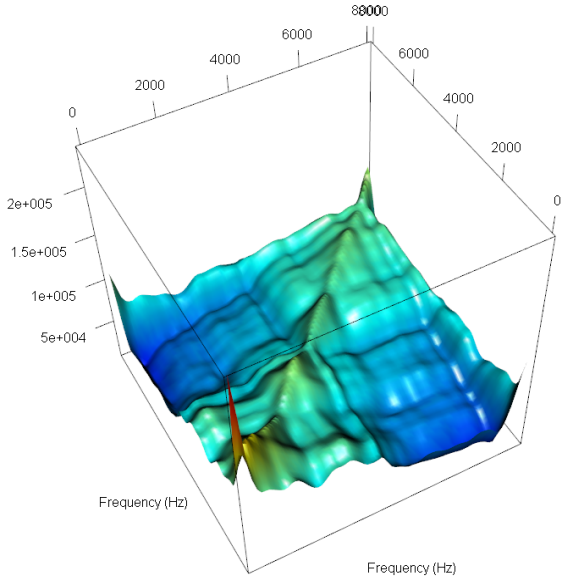
French



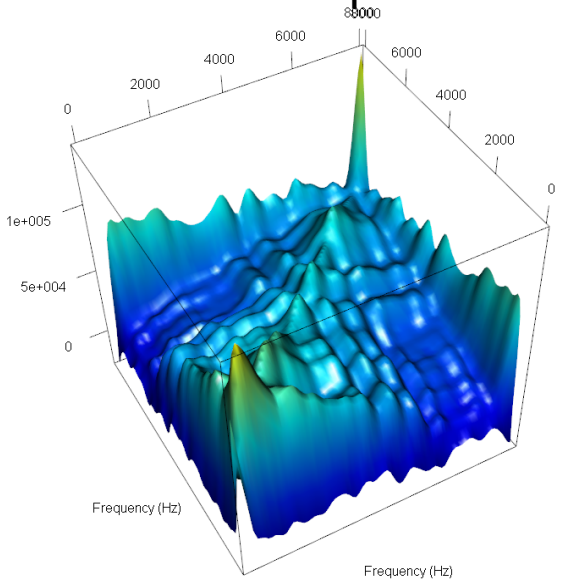
Portuguese



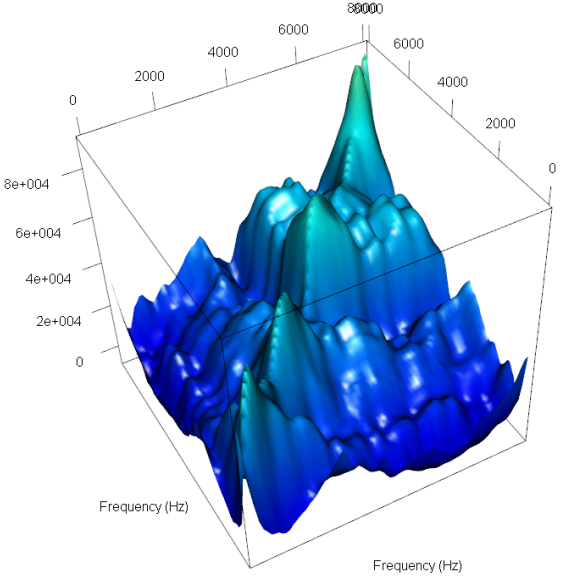
Italian



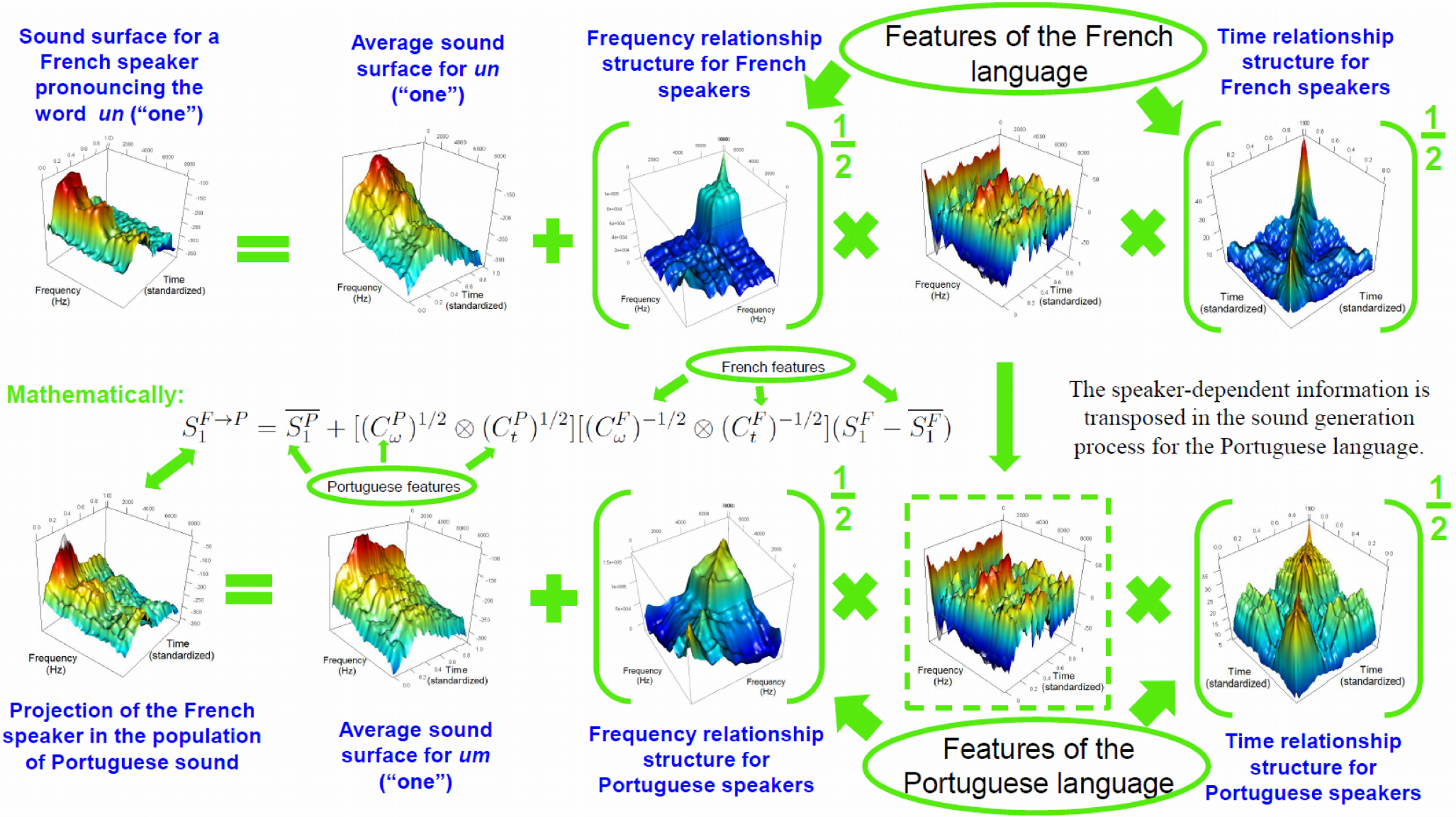
American Spanish



Iberian Spanish



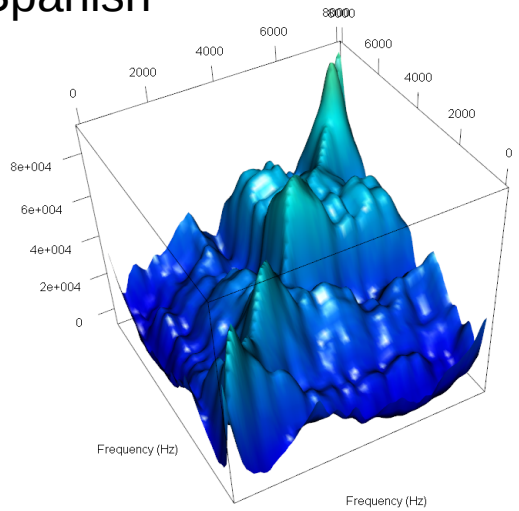
Transforming a word spoken by a speaker in one language into another language



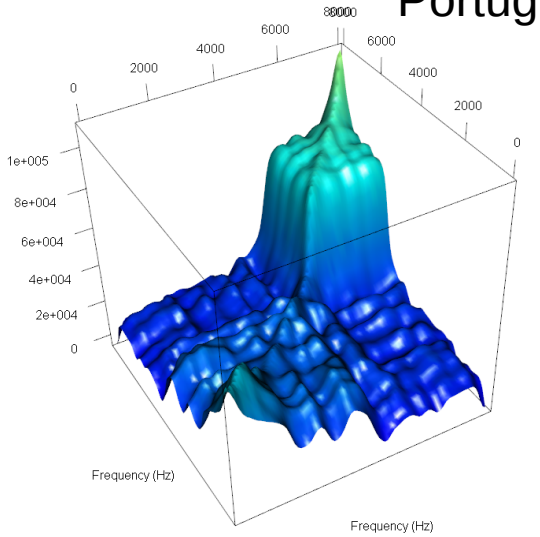
sound of word = pronunciation + lang. frequencies × speaker × lang. timing

Linguistic distances between sounds

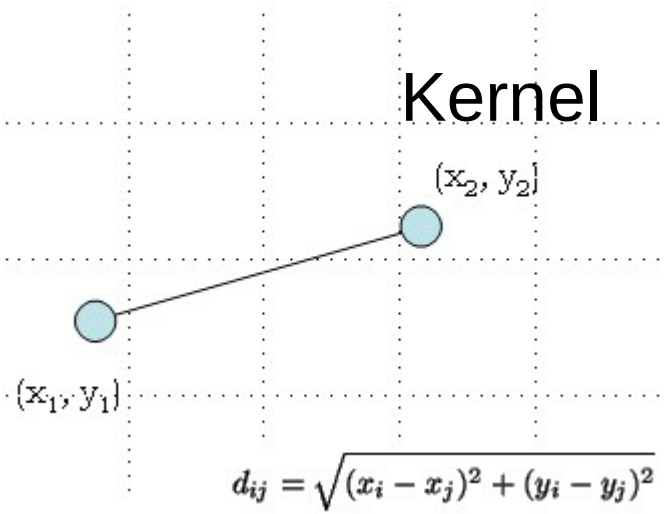
Iberian Spanish



Portuguese



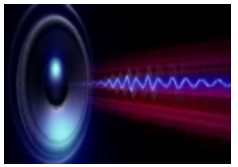
Kernel



Procrustes

Transforming a word spoken by a speaker in one language into another language

Example: French speaker saying “cinq” [sɛ̃k] transformed into French speaker saying Portuguese “cinco” [siŋko]



III. Variation through time

Problem: to model processes of change acoustically, we need recordings of older forms. We try 3 methods:

a) Use recordings from modern languages as proxies for ancestral recordings

b) Compute or edit intermediate (“hybrid”) forms from two modern recordings

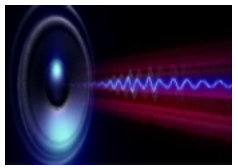
c) Unwind the sound changes and extrapolate backwards

a) Use modern recordings as proxies for ancestral recordings, e.g. “Latin” *un-us, -um*

[unʊ̀m]

Spectrogram 1

b) (Paradigmatic) Hybridization



[two]

LPC spectrogram 1

LPC spectrogram

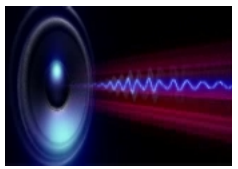
ectrogram

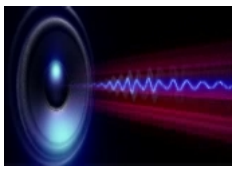
Interpolated LPC spectrograms

[*dwoh]

[dvah]

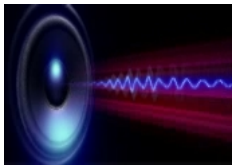
LPC spectrogram 2





b) (Syntagmatic) Hybridization

[*kwiŋkwe]



a)+b) proxy yielding a (testable) hybrid

[ũ]

LPC spectrogram 1

LPC

ctrogram

[œ̃]

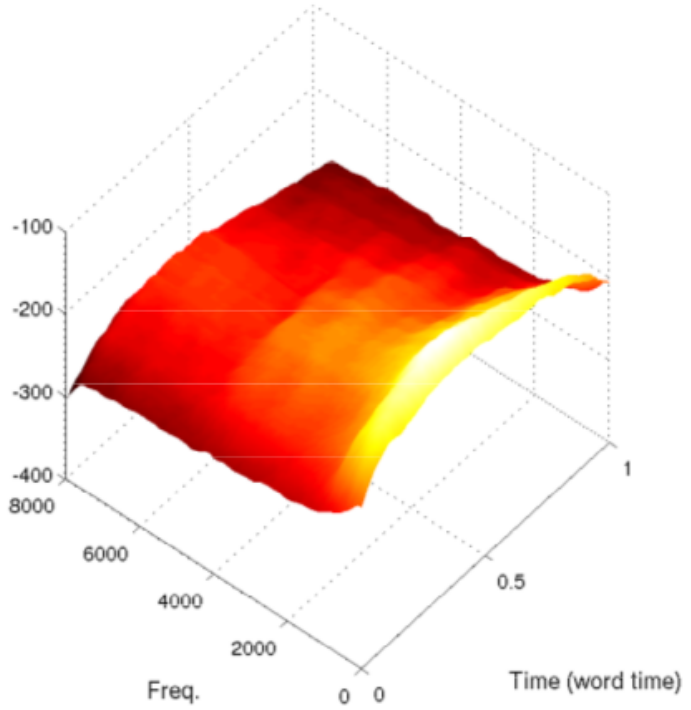
Interpolated LPC spectrograms

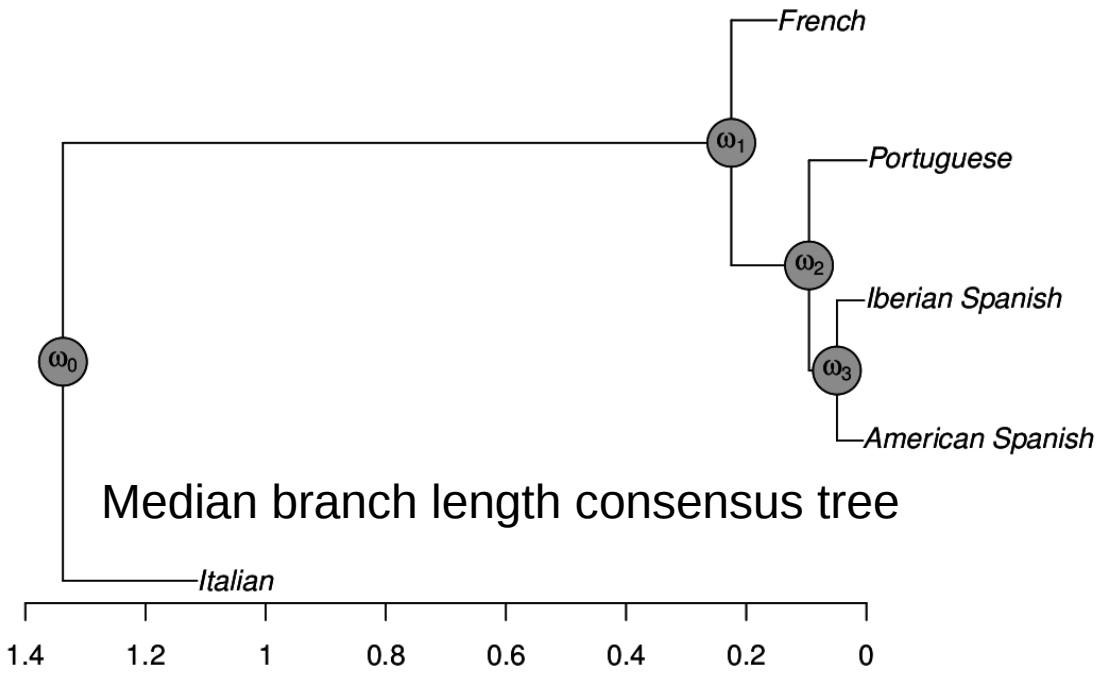
[ɛ̃]

LPC spectrogram 2

c) Compute most likely ancestor form by Phylogenetic Gaussian Process Regression

Hadjipantelis 2013 PhD *Functional Data Analysis in Phonetics*





Median branch length consensus tree

1. Factor out the phylogenetic variation from the non-phylogenetic variation

2. Turn the phylogenetic variation backwards to estimate ancestral parameters

3. Synthesize

Current and Future Directions

Exploring the direction and rate of language change

Current and Future Directions

Exploring the direction and rate of language change

Relating sound change to geographical distances

Current and Future Directions

Exploring the direction and rate of language change

Relating sound change to geographical distances

Broadening the approach to more languages and larger vocabulary

Current and Future Directions

Exploring the direction and rate of language change

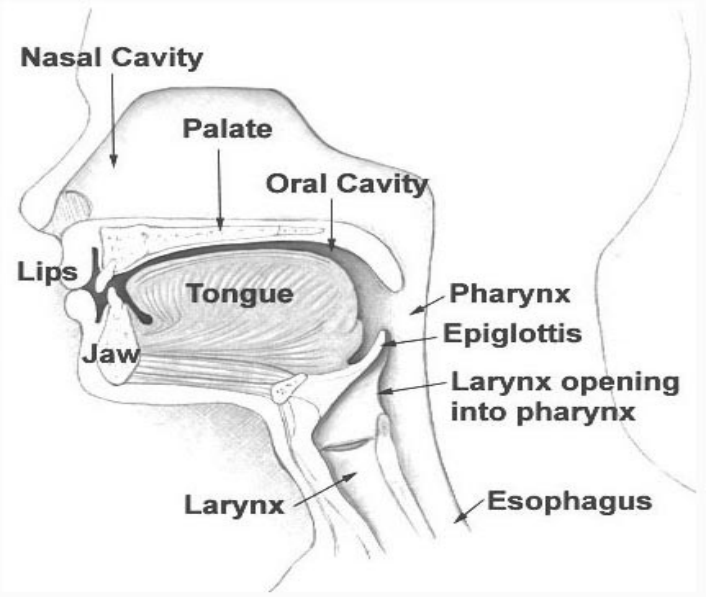
Relating sound change to geographical distances

Broadening the approach to more languages and larger vocabulary

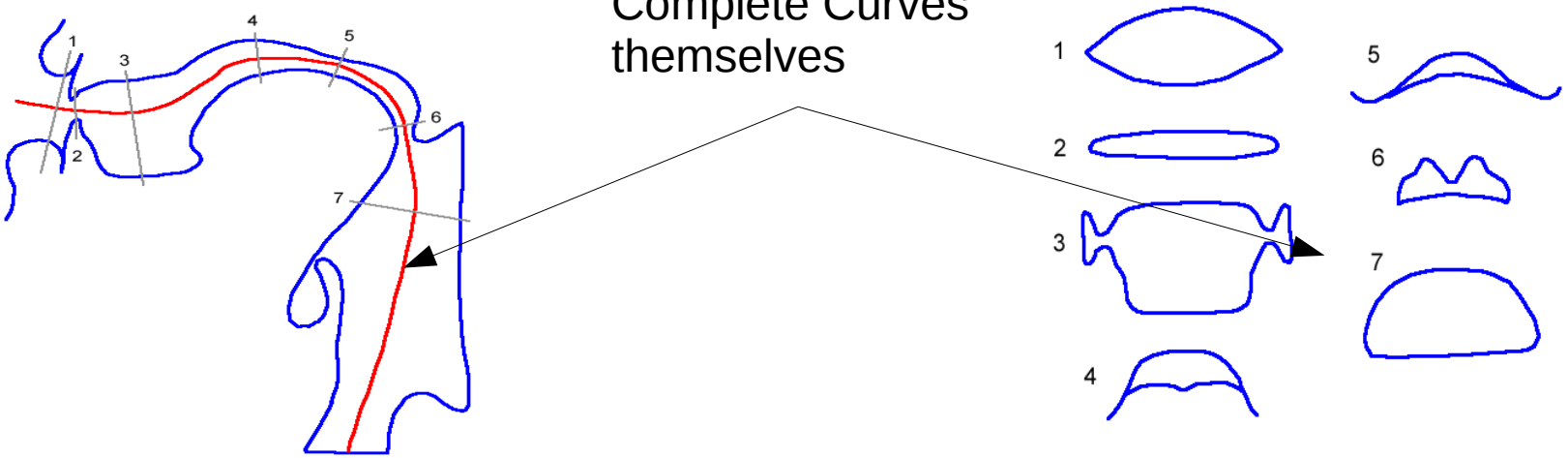
Spreading the word and getting more people involved!


That's all Folks!

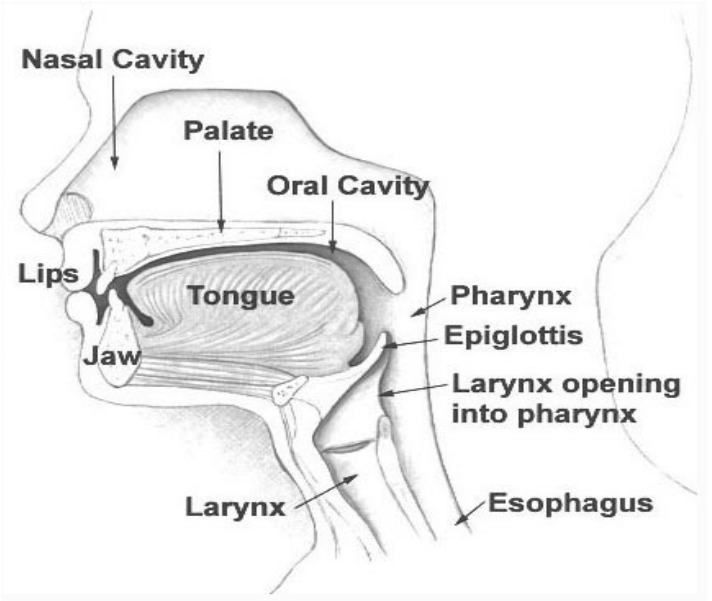
Functional Data Analysis



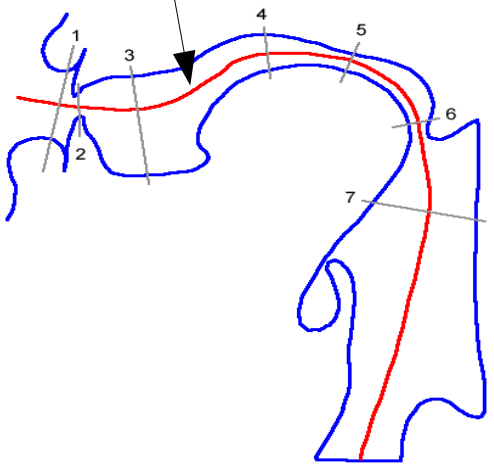
Functional Data:
Complete Curves themselves



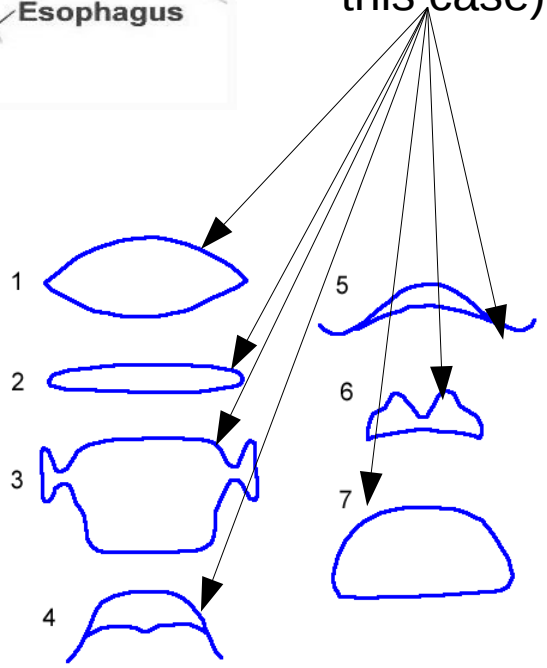
Functional Data Analysis



Univariate Data:
Length = 18cm



Multivariate Data:
Lengths of vocal tract
Cross-sections (7 in this case)



Sound change continuum: monophthongization

[due]

LPC spectrogram 1

Interpolated LPC spectrograms

[dø]

LPC spectrogram 2

Sound change continuum: postalveolarization

[tres]

LPC spectrogram 1

Interpolated LPC spectrograms

[treiʃ]

LPC spectrogram 2

Sound change continuum: postalveolarization

[seis]

LPC spectrogram 1

Interpolated LPC spectrograms

[seiʃ]

LPC spectrogram 2

Sound change continuum: diphthongization + uvularization

[tre]

LPC spectrogram 1

Interpolated LPC spectrograms

[tʁwa]

LPC spectrogram 2

Sound change continuum: monophthongization?

[seis]

LPC spectrogram 1

Interpolated LPC spectrograms

[si:s]

LPC spectrogram 2

Sound change continuum: postalveolarization + affrication

[set]

LPC spectrogram 1

Interpolated LPC spectrograms

[setʃ]

LPC spectrogram 2

Sound change continuum: postalveolarization + affrication

[ot:o]

LPC spectrogram 1

Interpolated LPC spectrograms

[otʃo]

LPC spectrogram 2

Sound change continuum: vowel raising

[des]

LPC spectrogram 1

Interpolated LPC spectrograms

[di:s]

LPC spectrogram 2