Counting Back Through History

Extrapolating Historical Phonetic Forms With Computational Methods

Joseph Rhyne & Betsy Miller

SECOL 2017, 10 March

Overview

- Traditional historical methods of sound reconstruction & acoustic reconstruction
- Data source: Slavic languages
- Our process following Coleman et al. 2015
- Results
- Potential applications and further exploration

Context

Historical Linguistics: Comparative Method (Trask 2000:64-67)

1. Establish genetic relationship prima facie
   - Fairly easy to do for closely related languages, e.g. Romance
2. Identify cognate sets through systematic correspondences of sounds in words of similar meaning
3. Set up proto-forms from the correspondence sets
   - Allows for reconstruction of the target proto-language
   - Allows for detection of sound changes between mother and daughter languages

Comparative Method: ‘hundred’

- Reconstructed proto-form: PIE *ḱm̥tóm

- Process based on textual representations
  - Phonetic qualities are extrapolated

Language	Word
Latin	*centum*
Greek	*hekaton*
Tocharian B	*kante*
Old Irish	*cét*
Middle Welsh	*cant*
Gothic	*hund*
Sanskrit	*śatám*
Avestan	*satəm*
Lithuanian	*šim̃tas*
Old Church Slavic	*sŭto*

Acoustic Modeling Of Sound Change

- Comparative Method (CM):
  - Typically models sound change: x>y
  - Leaves out intermediate stages
- All sound change starts with articulation (Lindblom 1963,Labov 1994):
  - Undershoot: PIE *ḱ→ Skt. ś
  - Redundancy deletion: [ãn]→[ã] in French
- Acoustic modeling uses attested methods from other fields:
  - Speech synthesis techniques (Moore & Coleman 2005)
  - Functional data in biology and mathematics (Meyer & Kirkpatrick 2005)

Possibilities for acoustic reconstruction

A. Use modern recordings that resemble what we think historical pronunciations sounded like

B. Splice together forms from modern recordings

C. Use statistical regression over phylogenetic tree to extrapolate back to ancestral forms from modern languages

   - ❖We use a simplified version of this approach

Acoustic Reconstruction Methods

Follow the general outline of Coleman et al. (2015):

1. Gather recordings of words from speakers in different languages
2. Extract acoustic parameters for numerical transformations
3. Extrapolate back to ancestral forms through transformations of the extracted parameters
4. Resynthesize transformed parameters into speech

Project Goals

- Create acoustic reconstructions to see if *hearing* historical forms is possible
- Improve upon traditional historical methods for reconstruction using acoustic analysis with current technology
- Extend proposed methods to untested Slavic data
- Propose further applications for these methods

Data

Slavic Languages

- Sub-Branch of Indo-European
- Has three branches:
  - South Slavic:
    - Western South Slavic: Serbo-Croatian, Slovenian
    - Eastern South Slavic: Bulgarian, Macedonian
  - East Slavic: Russian, Ukrainian, Belorussian, Rusyn
  - West Slavic:
    - Lekhitic: Polish, Kashubian
    - Czecho-Slovak: Czech, Slovak

Why Slavic?

- Acoustic reconstruction methods have been applied solely to Romance languages (Coleman et al. 2013, Pigoli et al. 2015)
  - Romance easiest to work with because of attestation, written records of the common stage (i.e. Latin)
- Expanding methods to another set of data allows for additional testing
  - The stage of common development for Slavic is unattested
- Ultimate goal: synthesize Proto-Indo-European words
  - Need to look at all Indo-European branches to reconstruct PIE

Common Slavic

- The hypothesized proto-language/common stage of Slavic
  - Reconstructed through the comparative method
  - Shares many features with Old Church Slavonic
- Period of shared development that lasted until about 1200 CE for what would become the modern Slavic languages
- This stage is unattested

Data

- Targets for reconstruction: spoken forms of numbers 1-10
- 5 Slavic Languages: Russian, Czech, Croatian, Polish, Bulgarian
  - Covers the different branches of Slavic
  - 4 tokens per number per language (200 total)
- Sounds samples gathered from Internet
  - Grammar websites
  - Corpora (Pelcra Spelling and NUmbers Voice database)
- Recordings converted from .mp3 to .wav with a sample rate of 11,025 Hz

Collected Tokens

	Russian	Bulgarian	Croatian	Czech	Polish
One	odín	edín	jedan	jeden	jeden
Two	dva	dve	dva	dva	dwa
Three	tri	tri	tri	tři	trzy
Four	četýre	čétiri	četiri	čtyři	cztery
Five	pjatʹ	pet	pet	pět	pięć
Six	šestʹ	šest	šest	šest	sześć
Seven	semʹ	sedem	sedam	sedm	siedem
Eight	vósemʹ	ósem	osam	osm	osiem
Nine	devjatʹ	devet	devet	devět	dziewięć
Ten	désjatʹ	déset	deset	deset	dziesięć

Methods

Methods

- Follow the general outline of Coleman et al. (2015)
- Same general functions recreated using PRAAT and R
  - Source code was not available
- R used to do data manipulations
  - R packages: phonTools, seewave, TuneR, simecol
- PRAAT used to combine sound files together

Methods: Functional Data

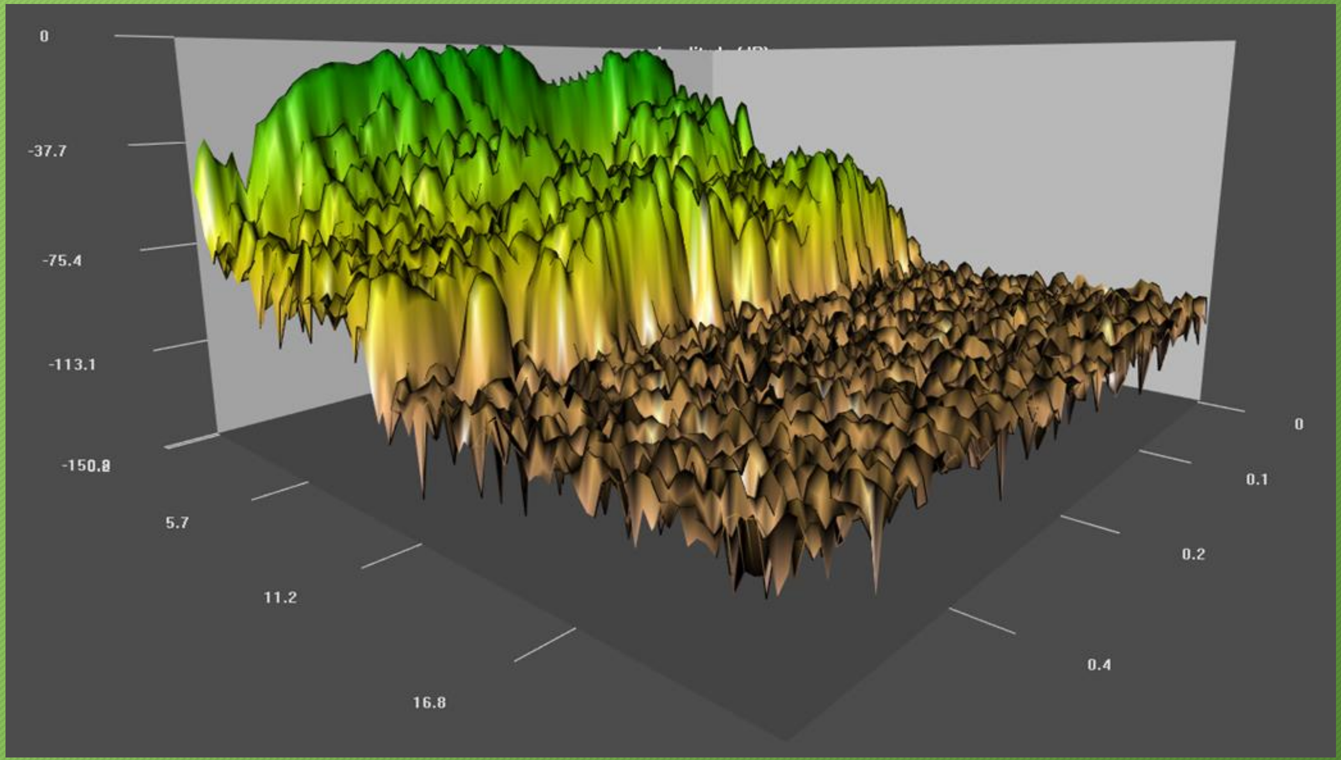
- Model sound as Functional Data (see Horvath & Kokoszka 2012, Ramsay & Silverman 2005)
- Data are represented as continuous mathematical functions
  - Standard statistical methods used for univariate and multivariate data have been extended to functional data
  - Used frequently in mathematics, statistics, machine-learning and other fields
  - Use smoothness and regularity of the functions to allow statistical analysis
- Spectrograms come from recordings and can estimate covariance operators
- Data taken from these surfaces (e.g. F_0) allow comparisons between languages

Methods: Log-Spectrograms

- Use log-spectrograms to determine the average of two sounds
  - Can’t simply mix sounds together
- Spectrograms can be viewed as functional data
- Spectrograms can be averaged
  - For comparison purposes
  - For other mathematical and statistical tasks

Methods: “Averaged” Spectrogram

- Averaged log-spectrogram for ‘one’ in Slavic
- Created from 20 total tokens



Methods: Estimate acoustic parameters

- Extract the acoustic parameters from spectrograms
- Within 5ms frames:
  - Estimate voicing
  - Estimate F_0
  - Estimate noise source parameters
- Create a snapshot of a speaker/language
- Future goal: deconstruct sound into speaker and language-specific components

Methods: Linear Predictive Coding (LPC)

- Used widely in speech synthesis, speech recognition
- Premise: speech sample can be approximated as a linear combination of past samples
  - Speech modeled as a linear, time-varying system
  - LPC provides an estimate of the characteristics that make up speech, removing the effects of formants and leaving just a buzz (intensity and frequency)
- Easy to convert back to synthetic speech

Methods: Interpolation Of Acoustic Parameters

- Estimated acoustic and LPC parameters combined into source + spectral parameter matrices
  - Mathematical representations of acoustic data allow for manipulations
- Simple linear interpolation between Ancestral form (A) and Modern recording (M) (Coleman et al. 2015)
  - $M = A + k\delta_g$
    - k= number of generations, δ_g = quantum of change/generation
  - Intermediate matrices interpolated; yield a continua of sound change
  - Sound files synthesized from intermediate steps
  - Culminates in reconstruction

Methods: Review

- Extract acoustic parameters from spectrograms in matrix format
  - Get a snapshot of the language by averaging the matrices
- Compare languages through linear interpolation
  - Create continua through intermediate forms resulting in a reconstructed form
- Re-synthesize transformed parameters into audible sounds

Results

Results

	Russian	Bulgarian	Croatian	Czech	Polish	PSl/CS
One	odín	edín	jedan	jeden	jeden	*(j)edinъ
Two	dva	dve	dva	dva	dwa	*dъva
Three	tri	tri	tri	tři	trzy	*trьje
Four	četýre	čétiri	četiri	čtyři	cztery	*četyre
Five	pjat′	pet	pet	pět	pięć	*pętь
Six	šest′	šest	šest	šest	sześć	*šestь
Seven	sem′	sedem	sedam	sedm	siedem	*sedmь
Eight	vósem′	ósem	osam	osm	osiem	*osmь
Nine	devjat′	devet	devet	devět	dziewięć	*devętь
Ten	désjat′	déset	deset	deset	dziesięć	*desętь

Results: Failures

- Synthesis not perfect: failed to capture the original jer vowels of Common Slavic
  - Weak vs. strong: strong jers became other vowels, while weak jers were lost
  - See remnants of them in Russian palatalization
  - Need to look to other words for these sounds
  - Add them to the number sounds, through manipulation of matrices, not through splicing
- More advanced techniques can be used

Results: Successes

- Overall synthesis is successful
  - Audible forms are reconstructed, and we can compare them to the textual attestations
- Synthesis can be combined with phylogenetic data for better results (Aston et al. 2011; Shiers et al. 2014)
- Compare reconstructed acoustic forms to using modern languages as proxies, such as through splicing together sound files

Applications and Explorations

Applications and Further Exploration

- Eventual acoustic reconstruction of proto-languages
  - Need to look at more branches of IE to reconstruct PIE
  - Only Romance (and now Slavic) have been processed
- Refining these methods for future historical exploration is crucial
  - Historical data of the future will be acoustic in addition to textual
  - Technology changes may prevent accessing data we are creating now
- Compare synthesized interpolants to attested intermediate stages, like Old East Slavic

Applications and Further Exploration

- Speech synthesis: translate recordings from one language into another, preserving speakers' voice characteristics
  - Use distances between covariance structure to predict how a speaker might sound in another language
  - With enough data, capturing what each language sounds like may be possible
- Modify synthesized speech to sound like a specific speaker
  - Commercial applications such as video games, movies, voice recognition, personal assistants (Siri), etc.

Selected References

Coleman, J., J. Aston and D. Pigole. 2015. Reconstructing the sounds of words from the past. In The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK: the University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0296; also 2013 (slide 11)

Lindblom, Bjorn. 1963. Spectrographic study of vowel reduction. Journal of the Acoustical Society of America 35, 1773-1781.

Labov, William. 1994. *Principles of Linguistic Chance, Internal Factors.* Oxford: Blackwell.

Moore, D., Coleman, J. 2005. Generation of synthetic speech. US Patent Application 20050171777.

Meyer, K. and Kirkpatrick, M. (2005) Up hill, down dale: quantitative genetics of curvaceous traits. Philos. Trans. R. Soc. B 360, 1443–1455

Pigoli D., Aston, J.A.D., Dryden, I.L. and Secchi, P. (2014) “Distances and Inference for Covariance Operators”, Biometrika, 101, 409–422.

Horvath, L. and Kokoszka, P. 2012. Inference for Functional Data with Applications. Springer, New York.

Ramsay, J. O. and Silverman, B. W. 2005. Functional data analysis. Springer Series in Statistics. Springer, New York, second edition.

Aston, J. et al. (The Functional Phylogenies Group). 2011. Phylogenetic inference for function-valued traits: speech sound evolution. Trends in Ecology and Evolution 27 (3), 160–166.

Shiers, Nathaniel, John A. D. Aston, Jim QSmith, and John S. Coleman. 2016. Gaussian Tree Constraints Applied to Acoustic Linguistic Functional Data. Journal of Multivariate Analysis Vol. 154, 199-215..

Acknowledgements

- Dr. Peggy Renwick, UGA for inspiration and guidance
- The Phonetics & Phonology Reading group at UGA for hosting practice talks and providing feedback
  - Dr. Peggy Renwick, Joey Stanley, Michael Olsen for specific suggestions for improvements to this project

- Thank you!

Questions/comments: joseph.rhyne@uga.edu, semnewman@uga.edu