



Why Measure Pronunciation Differences?

LOT Pronunciation Differences

- Dialectology & Variationist Linguistics
 - Aggregate analyses, maps of varietal differences
 - Identify characteristic differences (note that this is more demanding).
- Language Pathology
 - Speech of cochlear implant bearers (Sanders & Chin, 2009)
 - Stroke victims, ...
- Foreign accents
- Second language learning, second language pedagogy
- Pure phonetic, phonological interest (e.g. Obligatory Contour Principle: Frisch, Pierrehumbert & Broe, 2004)
- Quality of speech synthesis
- Better informed alignment procedures
 - Intimate connection alignment & distance (later today)
 - Enables identification of correspondences, e.g. interesting to historical linguistics



Measuring Pronunciation Differences

LOT Pronunciation Differences

- What to measure
 - Segments
 - Syllables
 - Words
 - Phrases
- How to measure
 - Nominal/Catgorical
 - Numerical

It's possible to use a nominal/categorical measure (Séguy, Goebel) but then more complex units are almost always different. Challenge: how to define a numerical measure?

We're particularly interested in techniques for measuring differences in the pronunciation of comparable material—the form most dialect atlases have.



Segment distances

LOT Pronunciation Differences

- Phones
 - Two segments are equal or different.
 - Distance between [ɪ] and [e] the same as between [ɪ] and [ɒ].
 - Rough, but easy to operationalize!
 - Rough measures reliable with large data sets.
- More sensitive measures later
- Challenge: How to lift segment distances to STRING DISTANCES.



String distances

- *Hamming distance* counts the positions where strings are different
- Example: *afternoon* is pronounced [ˈæftəˌnɪn] in Savannah and [ˌæftərˈnuːn] in Lancaster.
- The difference in string length is added to the number of different positions (if needed).

ˈæ	ə	f	t	ə	n	ˌɪn	n	
ˌæ	f	t	ə	r	n	ˈuː	n	
	1	1	1	1		1		$\Sigma = 5$

ˈæ	ə	f	t	ə	r	
ˈæ	f	t	ə	r		
	1	1	1	1	1	$\Sigma = 5$

- Notice treatment of stress
- Roughly same as “Manhattan distance”



String distances

LOT Pronunciation Differences

- *Levenshtein distance* calculates the (least) cost of changing one string into another
- Example: *afternoon* is pronounced [ˈæftənʌn] in Savannah and [ˌæftərˈnuːn] in Lancaster.

æftənʌn	delete ə	1
æftənʌn	insert r	1
æftərʌn	subst. ʌ/u	1
æftərnuːn		
<hr/>		3

- All operations cost *one* unit (initially)
- Problem: how to guarantee the *least* cost



Levenshtein algorithm(æəftən_ɪn,æftərnu:n)

LOT Pronunciation Differences

Create two-dimensional array

	æ	f	t	ə	r	n	u	n
æ								
ə								
f								
t								
ə								
n								
ɪ								
n								



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}ft\text{\e}n\text{\u}n, \text{\ae}\text{\o}t\text{\e}r\text{\n}u\text{\n}$)

		\ae	f	t	\e	r	n	u	n
\ae	0								
\e									
f									

- begin at upper left ($\Leftarrow 0$)

diag	above
left	$\min(\text{above} + \text{delete},$ $\text{diag} + \text{replace},$ $\text{left} + \text{insert})$

- to fill in a cell:



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}ft\text{\e}n\text{\u}n, \text{\ae}ft\text{\o}rn\text{\u}n$)

		\ae	f	t	\e	r	n	u	n
\ae	0	1	2	3	...				
\e	1	?							
f	2								
t	3								
	⋮								

Top horizontal row is always 1, 2, ... —cost of insertions

Left vertical column is always 1, 2, ... —cost of deletions

? is $\text{minimum}(\text{left} + \text{ins}, \text{above} + \text{del}, \text{diag} + \text{subst})$
 $\text{minimum}(1 + 1, 0 + 0, 1 + 1)$



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}\text{ft}\text{\e}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
	0	1	2	3	...				
\ae	1	0	$?_1$						
\e	2	$?_2$							
f	3								
t	:								
\e									
n									
\u									
n									



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}\text{ft}\text{\e}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
	0	1	2	3	...				
\ae	1	0	1						
\e	2	1	$?_1$						
f	3	$?_2$	$?_3$						
t	:								
\e									
n									
\u									
n									



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\ft}\text{\t}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{\ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
	0	1	2	3	...				
\ae	1	0	1						
\e	2	1	1						
f	3	2	1						
t	:								
\e									
n									
\tt									
n									



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\ft}\text{\t}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{\ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
	0	1	2	3	4	...			
\ae	1	0	1	2					
\e	2	1	1						
f	3	2	1						
t	4			1					
\e	:				1				
n									
\t									
n									



Algorithm

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}\text{ft}\text{\e}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
\ae	0	1	2	3	4	...			
\e	1	0	1	2					
f	2	1	1						
t	3	2	1	1					
\e	4			1					
n	:				1	2			
\u					2		2		
n								3	
n									3

- lower right corner contains Levenshtein distance, cost of least expensive set of transformations



Alignment

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}\text{ft}\text{\e}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{ft}\text{\e}\text{\r}\text{\n}\text{\u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
	0	1	2	3	...				
\ae	1	0	1	2					
\e	2	1	1						
f	3	2	1	1					
t	:			1					
\e					1	2			
n					2		2		
\u								3	
n									3



Alignment

LOT Pronunciation Differences

Levenshtein distance($\text{\ae}\text{\o}\text{ft}\text{\e}\text{\n}\text{\u}\text{\n}$, $\text{\ae}\text{ft}\text{\e}\text{\r}\text{\n}\text{u}\text{\n}$)

		\ae	f	t	\e	r	n	u	n
\ae	0	1	2	3	...				
\o	1	0	1	2					
f	2	1	1						
t	3	2	1	1					
\e	:			1					
n					1	2			
\u					2		2		
n								3	
n									3

\ae \o f t \e \emptyset n \u n
 \ae \emptyset f t \e r n u n



Dialect distances

LOT Pronunciation Differences

- Many sequence operations map [æəftənɪn] → [æftərnɪn]. Levenshtein distance = cost of cheapest mapping.
- Using w words the distance between two dialects is equal to the average of w Levenshtein distances.
 - automatically weights differences involving more frequent sounds more heavily
- All distances between n dialects are arranged in a $n \times n$ matrix for further analysis.



Pronunciation Difference Measurement

LOT Pronunciation Differences

- Rough (noted)
- Sensitive to stress (if mark occurs in transcription)
- Contextual effects *are* measured

t u \$ b æ t ə z *two batters*
t u \$ b æ r ø z

- Sources of contextual effects are *ignored* (in simple version)
if [t/r] occurs in same ratio in two pairs of varieties, distance will be the same—regardless of specifics of conditioning
- Longer words contribute more heavily to pronunciation difference since their string distances are longer



Short Words vs. Long Words

LOT Pronunciation Differences

- The simple Levenshtein distance can be normalized by the length of the word. For example, the sum of the operations is divided by the length of the longest alignment which gives the minimum cost. The longest alignment has the greatest number of matches.
- Example:

æ	ə	f	t	ə	∅	n	ʊ	n
æ	∅	f	t	ə	r	n	u	n
	1				1		1	

A total cost of 3 divided by a length of 8 gives a word distance of 0.38 or 38%.

- Using 125 words the distance between two dialects is equal to the average of 125 Levenshtein distances.
- All distances between n dialects are arranged in a $n \times n$ matrix.



More Sensitive Segment Distances

LOT Pronunciation Differences

- Levenshtein distance is also known as string distance and edit distance. Well-known example of DYNAMIC ALGORITHM (yet another name).
- Applications in other areas
 - software engineering** file differences
 - bioinformatics** differences between long strings of amino acids (DNA)
 - translation** aligning bilingual corpora
 - ethnology** tracking “folk processing” in bird calls
- Costs are often one for insertions and deletions, two for substitutions
 - seems wrong in assaying pronunciation dissimilarity



More Sensitive Segment Distances

LOT Pronunciation Differences

- Phones
 - Two segments are equal or different.
 - Rough, but simple!
- Features
 - Finer differentiation of segment distances.
 - Segment differences used to weight Levenshtein algorithm
 - Distance between [ɪ] and [e] smaller than between [ɪ] and [ɒ].
 - Distance between two bundles: sum of differences (simplest case).
 - * Heeringa (2004) experimented with Euclidean combination, a $(1 - r)$ measure ($r =$ Pearson's correlation coefficient) with little effect on aggregate distances.
- We might induce segment distances from frequencies of correspondences in alignments (Wieling lecture, Thurs.)



Dialect distances

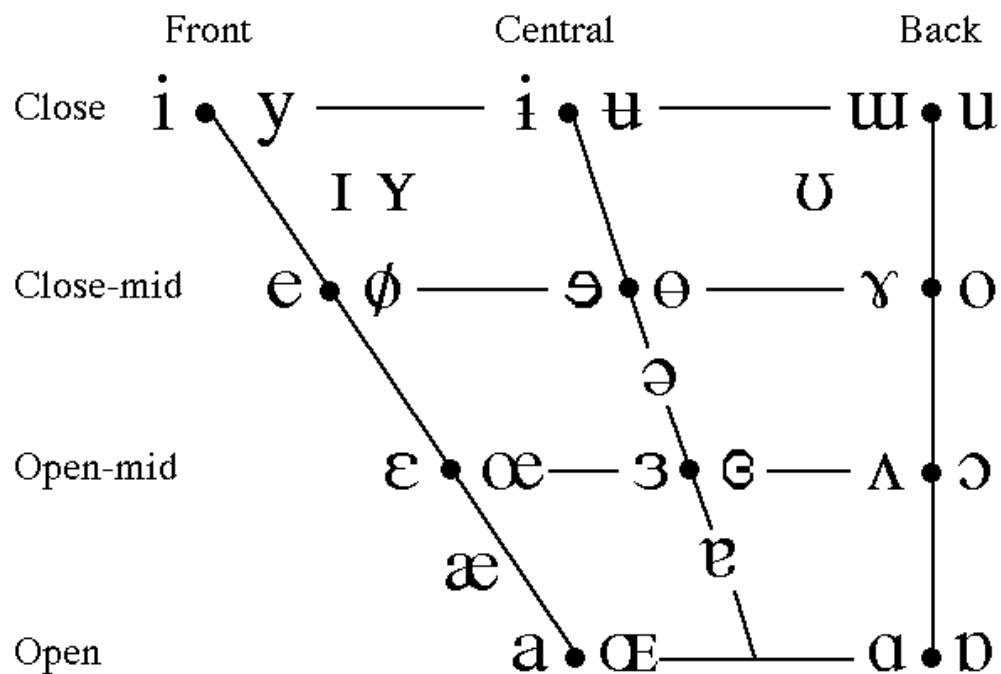
LOT Pronunciation Differences

- Refinement: feature bundle distances or acoustic distances as operation weights!
- We assure that the minimum cost is based on a alignment in which
 - a vowel matches with a vowel
 - a consonant matches with a consonant
 - the [j] or [w] matches with a vowel
 - the [i] or [u] matches with a consonant
 - the schwa matches with a sonorant
- We refer to this as the VC Levenshtein distance, or the variant that respects SYLLABICITY.



Discrete segment distances

LOT Pronunciation Differences



Vowel distances in the Almeida & Braun system: distances of 1 point:
ε vs. æ (height), ε vs. ɜ (advancement), ε vs. œ (round).



Feature-based Segment Distances

LOT Pronunciation Differences

Positive More sensitive distinctions

- In theory should yield no worse measurements

Negative

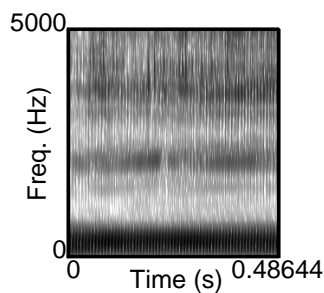
- Embarrassment of riches (many systems)
- Many more parameters (therefore weaker)
- Most feature systems developed to facilitate phonological description, not to provide foundation for description of dialectal similarity



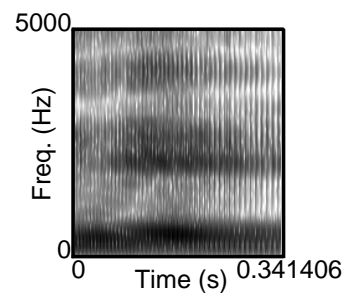
Acoustic segment distances

LOT Pronunciation Differences

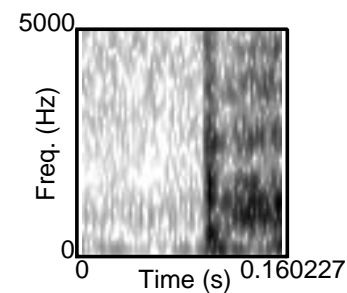
- Feature systems mostly not based on physical measurements.
- Samples of all IPA segments are found on the audio tape *The Sounds of the International Phonetic Alphabet* (1995).
- Calculate distances between the samples using their spectrograms or formant tracks.
- Intensity is processed, durations are made equal.



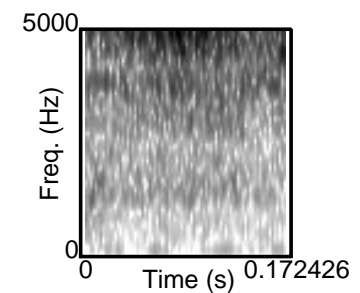
i



e



p

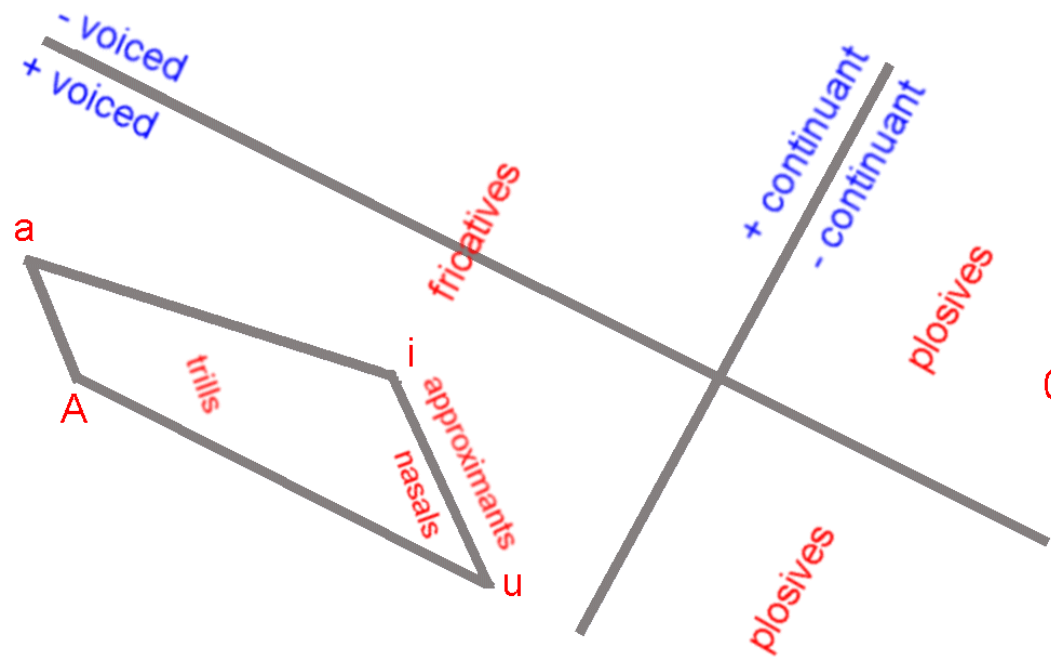


s



Acoustic segment distances (Heeringa)

LOT Pronunciation Differences

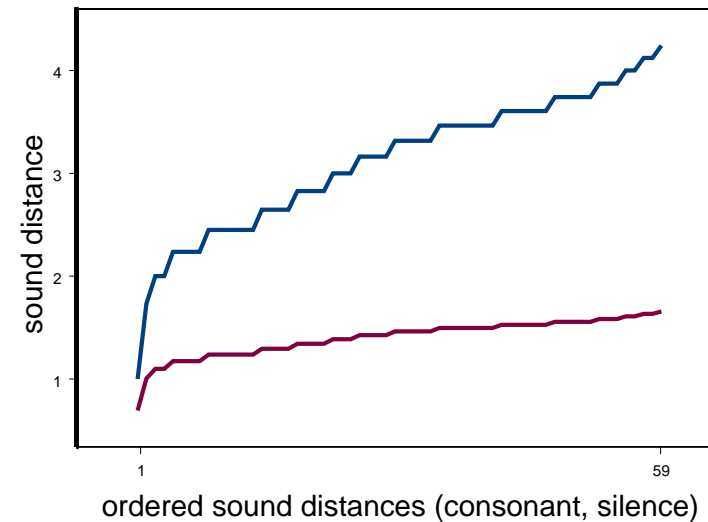
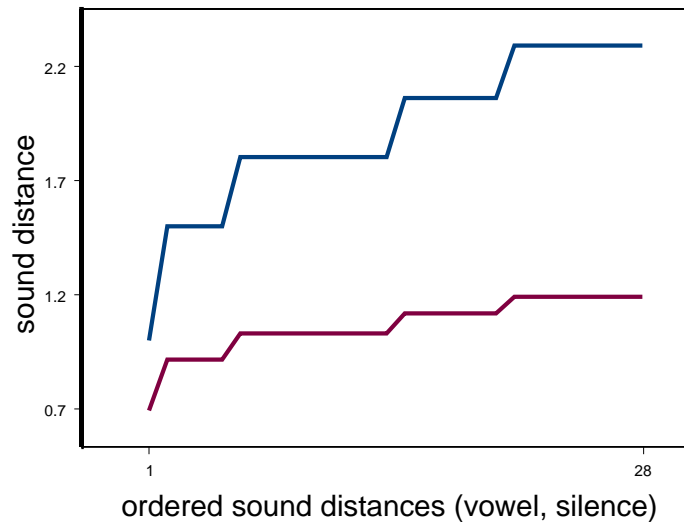


Distances among 88 segments (28 vowels, 59 consonants, silence) calculated using the Barkfilter and reduced from 88 dimensions to two dimensions with multidimensional scaling.



Linear and logarithmic segment distances

LOT Pronunciation Differences



Linear (upper) and logarithmic (lower) Almeida & Braun distances of 28 IPA vowels (left) and 59 IPA consonants (right) with respect to silence. Distances are sorted from low (left) to high (right). Greater distances are reduced more than smaller ones by using the logarithm.



Pronunciation Differences

LOT Pronunciation Differences

- Given database of pronunciations of comparable material, we can obtain various measures of pronunciation difference.
- Since we'll characterize the distance(s) numerically, we can analyze the results using numerical techniques.
- Levenshtein distance appropriate for dialect atlas material with comparable pronunciations indicated, but inappropriate for corpus material, i.e. material without indication of which pronunciations are to be compared.



Next Steps

LOT Pronunciation Differences

- LAMSAS pronunciations
 - No mapping to acoustic samples (too complex)
 - No logarithmic correction to flatten large differences
- Quality of results