Analyzing vowel distances

Therese Leinonen

LOT winter school, Groningen 21th January, 2009
Overview

• Segment distances in dialectometric research

• Acoustic properties of vowels

• Measuring vowel distances

• Perception of vowels

• Inducing sound segment distances using Pair HMM's
Using segment distances in the Levenshtein algorithm

• theorem: given segment distances, Levenshtein algorithm finds optimal alignment

• what are good segment distances?

• various feature systems: Vieregge-Cucchiarini, Almeida-Braun (Heeringa 2004)

• "acoustic" distance (Heeringa 2004)

• very limited improvement over binary segmental table
Phonetic Puzzle

Why is detailed phonetic information not helping?

• hypothesis 1: transcriptions are phonetically unreliable

• hypothesis 2: previous attempts were too ambitious, trying to characterize all distinctions

• hypothesis 3: we are past the size where fine discrimination matters

• others?
Acoustic properties of vowels

- The vocal tract is a resonator that resonates on given frequencies; by changing the size and shape of the tract (by moving the position of tongue, lips, yaw) we can adjust the resonant frequencies.

- The sound we produce with our vocal chords consists of a base tone and its harmonics.

- When some harmonic of the sound from the vocal chords matches or is close by a resonant frequency it will cause resonance.

- Formants = peaks in the frequency spectrum resulting from resonance in the vocal tract.

- Our perception of vowels is based on recognizing the formant frequencies characterizing each vowel.

- The first two formants (F1 and F2), corresponding well with vowel height and backness, are usually enough to distinguish vowels from each other.
spectrum

spectrogram

vowel

[i]  [æ]  [a]  [u]
Where symbols appear in pairs, the one to the right represents a rounded vowel.
Formant frequencies of Dutch vowels
Measuring the distance between vowels

Euclidean distance:

$$\sqrt{n \sum_{i=1}^{n} (p_i - q_i)^2}$$

Euclidean distance of the two first formants:

$$d_{ij} = \sqrt{(F_{1i} - F_{1j})^2 + (F_{2i} - F_{2j})^2}$$
Perception of vowels

- Frequency is measured in Hertz (cycles per second), however, human perception is roughly linear below 1000 Hz and roughly logarithmic above 1000 Hz.

- Bark and Mel are scales that are developed to correspond to perception.

- The formant frequencies depend on the size and shape of every speaker's vocal tract, as listeners we can normalize for this speaker-dependent variation, automatic procedures for doing the same are hard to find.

- The inventory of vowel phonemes in a language influences the perception of vowels.
Inducing sound segment distances using Pair HMM’s

(Wieling, Leinonen and Nerbonne 2007)

- Pair Hidden Markov Models (Pair HMM) were trained to align the pronunciation transcriptions of a large contemporary collection of Dutch dialect data (Goeman & Taeldeman, 1996)

- the PHMM give probabilities of two segments being aligned in the data set – these probabilities can be interpreted as segment distances

- we validated the substitution probabilities by acoustic measures (Euclidean distance of F1 and F2)

- acoustic data: pronunciation of Standard Dutch monophthongs by 50 male (Pols, Tromp and Plomp 1973) and 25 female speakers 25 female (Van Nierop, Pols and Plomp 1973) speakers
Transforming the data

- the occurrence frequency of the phonetic symbols influences substitution probability

- the substitution probabilities are divided by the product of the relative frequencies of the two phonetic symbols used in the substitution

- substitutions involving similar infrequent segments now get a much higher score than substitutions involving similar, but frequent segments – the logarithm of the score is used to bring the scores into a comparable scale
Transforming the data
Regression

\[ \text{acoustic distance} = 1.75 - 0.32 \times PHMM \]

\[ r = -0.72 \]
Dialectological results:

- dialect classification based on Pair HMM’s show only small differences to analyses based on the Levenshtein algorithm ($r = 0.89$)

- it would be valuable to test the method on dialect data for which perceptual distances are also available

Conclusions:

- alignments created by the Pair HMM are linguistically responsible

- the linguistic structure influences the range of linguistic variation

- similarity is a satisfying basis of comparison at local levels
References


