# Aggregate Analysis of Vowel Pronunciation in Swedish Dialects

## Leinonen, Therese, Vol 3, No 2 (2011)

Mei-Shin Wu[1]

[1]Computational Linguistics
University of Tbingen

27th of January, 2016



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


## Introduction

- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

## Data and Experimental Design

- Data
- Acoustic analysis
- Analysis of dialectal variation

## Result

- Geographic dialect continuum
- Dialect leveling

## Discussion And Conclusions

- The observation
- The pros and cons of aggregation analysis

# Outline


## Introduction

- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

## Data and Experimental Design

- Data
- Acoustic analysis
- Analysis of dialectal variation

## Result

- Geographic dialect continuum
- Dialect leveling

## Discussion And Conclusions

- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# The Goal

- Swedish dialects haven gone through massive **leveling** in the latter half of the 20th century.
- The aggregating methods are used for **identifying geographic dialect areas** and studying **ongoing changes** in Swedish dialects.



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


## Introduction

- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

## Data and Experimental Design

- Data
- Acoustic analysis
- Analysis of dialectal variation

## Result

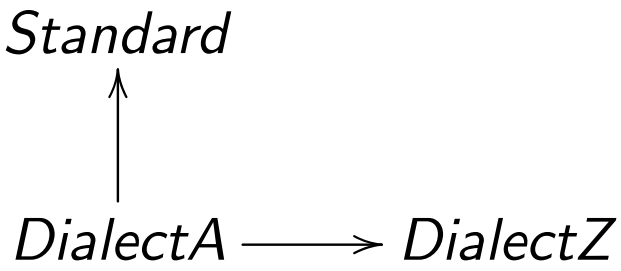
- Geographic dialect continuum
- Dialect leveling

## Discussion And Conclusions

- The observation
- The pros and cons of aggregation analysis

# Language leveling

- Dialect levelling is defined as the reduction of structural variation
- Dialect levelling makes $(a)$ individual dialects more homogeneous, and $(b)$ different dialects more similar and ,consequently, diasystems more homogeneous






Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


## Introduction

- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

## Data and Experimental Design

- Data
- Acoustic analysis
- Analysis of dialectal variation

## Result

- Geographic dialect continuum
- Dialect leveling

## Discussion And Conclusions

- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# The aggregating methods

- Feature-Based Variation Studies analyze variation in one element. While Aggregating methods consider **more than single feature**.
- Both rural Swedish dialects and regional varieties of Standard Swedish vary a lot when it comes to vowel pronunciation, and **vowels** have been **important for characterizing varieties of Swedish and classifying dialects**.
- are used for identifying **geographic dialect areas** and studying **ongoing change**.



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


## Introduction

- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

## Data and Experimental Design

- Data
- Acoustic analysis
- Analysis of dialectal variation

## Result

- Geographic dialect continuum
- Dialect leveling

## Discussion And Conclusions

- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# Swedish Dialects

- South Swedish dialects
- Gotaland dialects
- Svealand dialects
- Norrland dialects
- Gotland dialects
- Finland-Swedish dialects

Figure: The traditional Swedish dialect area according to Wessn




Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# Phonology data resource

- 98 rural sites
- from 1998 to 2001
- 1170 speakers
- 12 people for each location
- older group from 55 to 75 years
- younger group from 20 to 35 years
- 48 kHz sample rate and 16-bit amplitude resolution (recording)
- downsampled to 16 kHz/16 bit for analysis
- swedia.ling.gu.se

**Title of Project**
Swedia 2000 - Phonetics and phonology of Swedish dialects around the year 2000.

**Project Period**
01/01/1998 - 12/31/2008

**Subject areas**
Linguistics, Phonetics

**Head of research**
Eva Strangert, Department of language studies

# Phonology data resource-2

- Vowel segments were elicited with mono- or bi-syllabic words
- target vowels were surrounded by coronal consonants
- vowel recruitment criteria : long vowels, allophonic variants of /ɛː/, /øː/, vowels reflect the Historical development
- 19 words were used : dis[iː], disk[ɪ], dör[œː], dörr[œ], flytta[ʏ], lass[a], lat[aː], leta[eː], lett[e̞], lott[ɔ], lus[(ʉ)], lås/låt[oː], lär[æː], lös[ø], nät[ɛː], sot[uː], särk[æ], söt[øː], typ [yː]

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# Acoustic analysis

- All vowel segments were filtered with Bark filters.
- **principal component analysis(PCA)** : PC1 is related to **vowel height** and PC2 is related to **tongue advancement**
- males and females vocal tracts were normalized for by applying PCA separately to data from male and females speakers.



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# Analysis of dialectal variation

- multidimensional scaling (MDS) is a method for reducing complex distance data. Gap statistic is used for estimating the number of significant clusters.
- MDS is suitable for visualizing dialect continua.And cluster analysis detects dialect groups.
- This paper used Kruskal's non-metric MDS
- Euclidean distance represents the vowel quality.
- x, y are either two different sites or two different speaker groups.

$$distance(x,y) = \sqrt{\sum_{i=1}^{9}((PC1_{xi} - PC1_{yi})^2 + (PC2_{xi} - PC2_{yi})^2)}$$

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

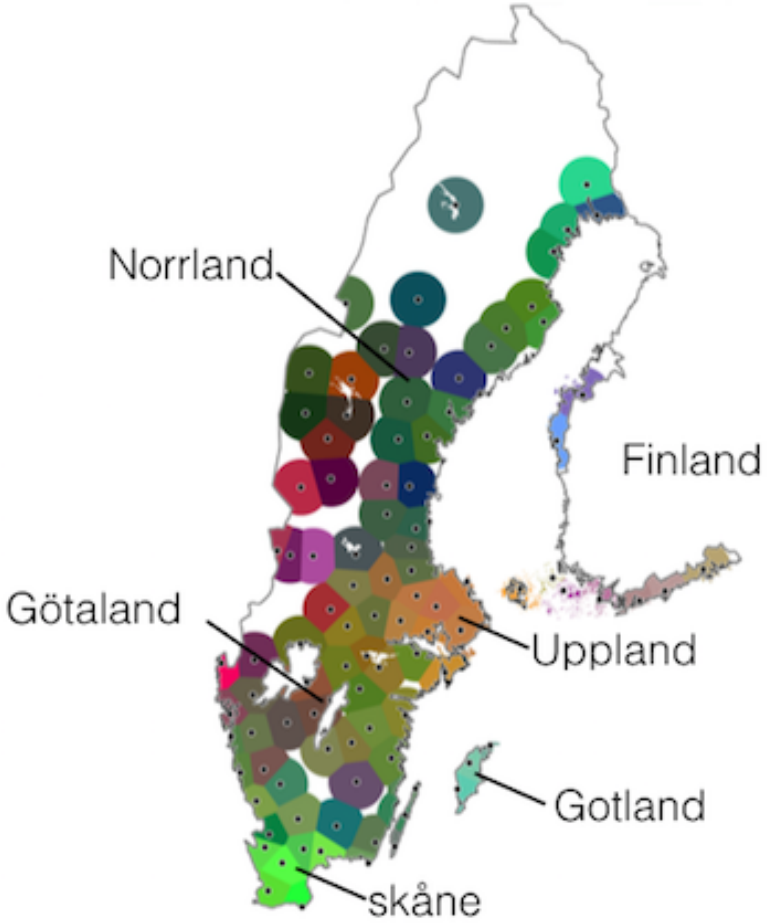
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# Geographic dialect continuum

- one dimension explains 81.4%, two dimension explains 93.6% and three dimension explains 96.3%
- the dimension cannot be interpreted as PC1 and PC2 fully.

# Geographic dialect continuum-2




- coherent area : Skåne, Gotland, Uppland
- incoherent area : Norrland, Götaland

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

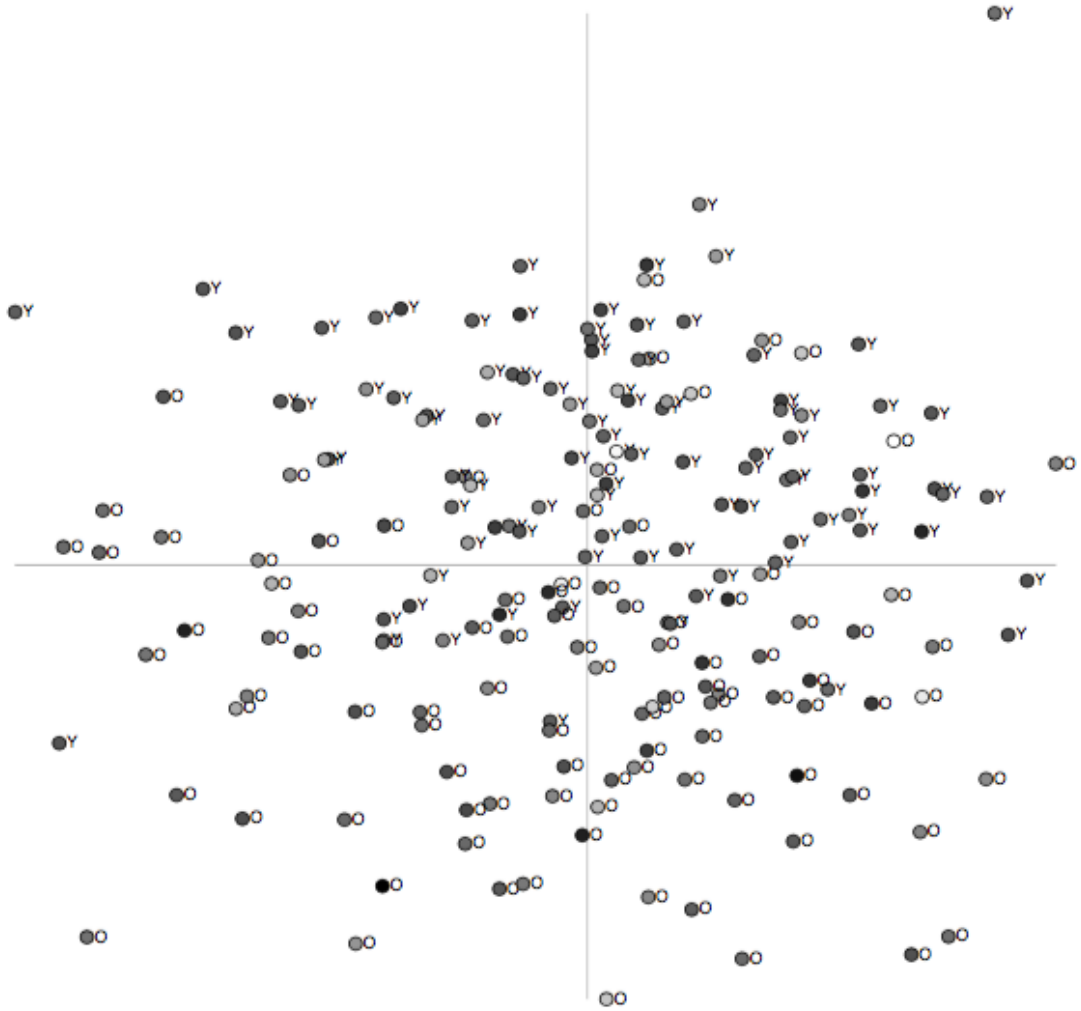
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

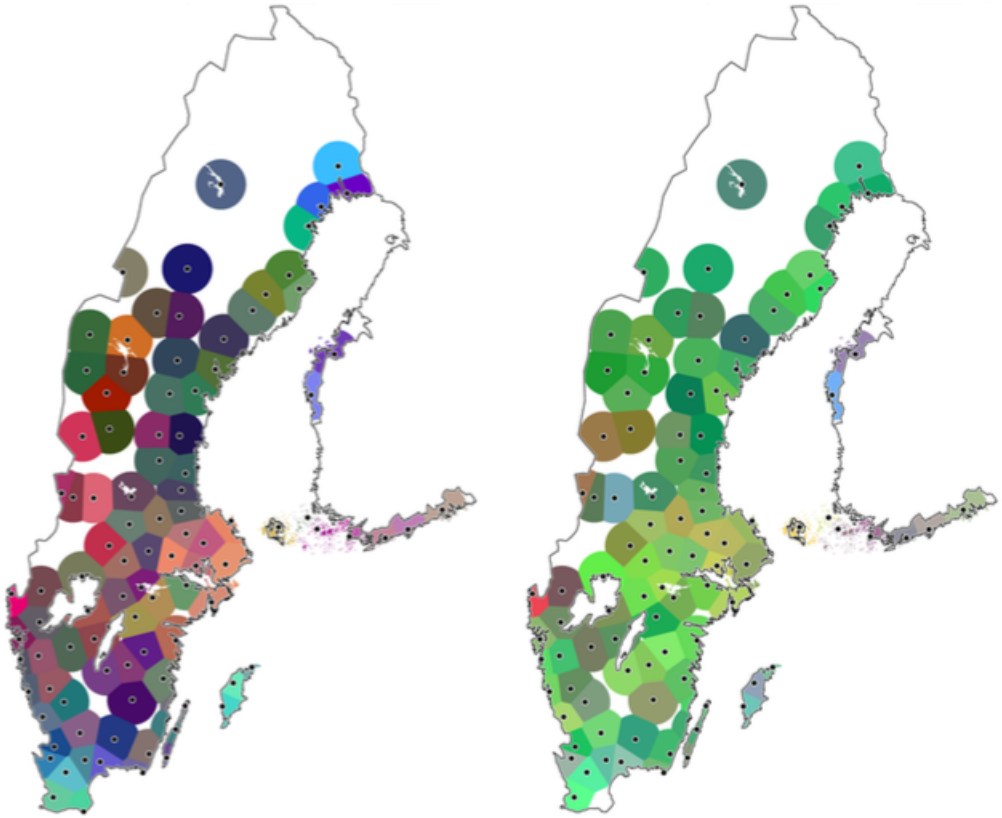
Reference

# Dialect leveling



- The extreme outlier is younger speakers of Löderup

# Dialect leveling-2

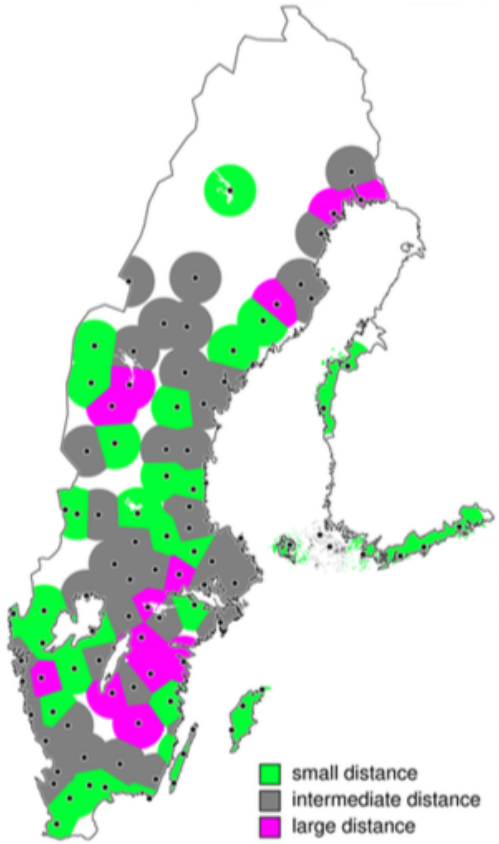



- A broad spectrum of colors is found in older group. While the younger group shows mainly green color
- A large scale of dialect leveling can be confirmed.
- The difference between younger speakers do exist! But the difference is so much smaller than between older speakers. Therefore, the color cannot reflect the differences precisely



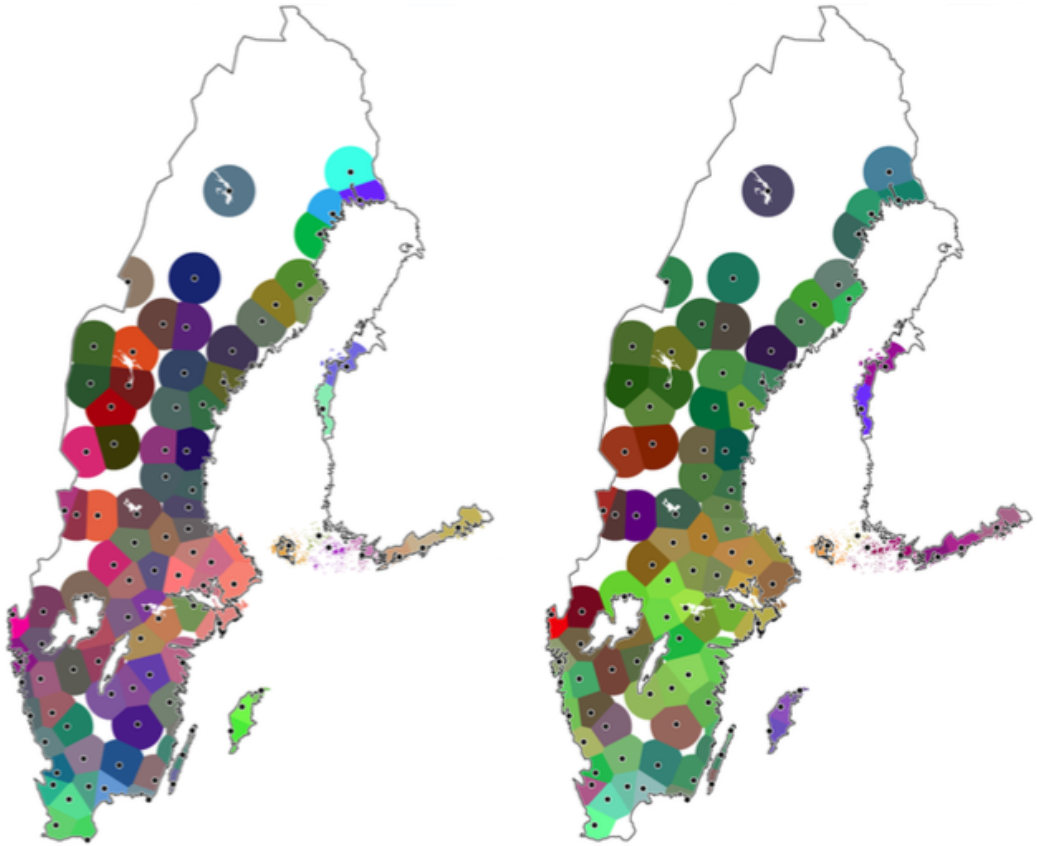
Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Dialect leveling-3




- the map shows the similarity distance.
- the sites close to biggest cities, such as Stockhom and Göteborg shows a large ongoing changes

# Dialect leveling-4




- In order to present the difference between younger groups, MDS were applied to two age groups seperatly.
- Two maps are independent.
- A more uniform spoken variety of Norrland is emerging

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# The observation

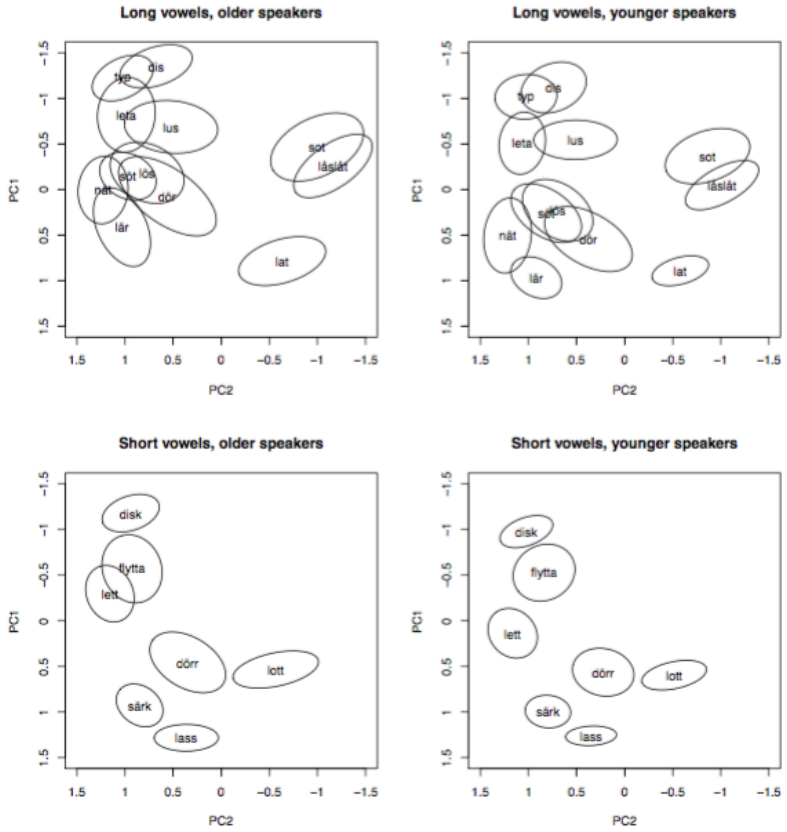
- The paper showed that the variation in vowel pronunciation across Swedish dialects is **continuous** and **abrupt dialect borders exist**. And MDS result **agree** with the traditional division of Swedish dialects by Wessén.
- The **dialect leveling** has been **confirmed statistically** by the aggregate analysis.



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# The observation-2

- A general lowering of front vowels going on
- The areas where the vowel shift is strongest are areas close to south-west of Stockholm and Göteborg
- The chain shift can be regarded as simplification of the vowel system : the vowel inventory becomes smaller.



# The observation-3

Figure: http://goo.gl/hbRDxa

- Not much change in vowel pronounciation : South Sweden (province Skåne), Gotland and the Swedish dialect area in Finland.
- Aggregation of linguistic data reveals global tendencies and overall relationships between varieties.

# Outline


Introduction
- The Goal
- Language leveling
- The aggregating methods
- Swedish Dialects

Data and Experimental Design
- Data
- Acoustic analysis
- Analysis of dialectal variation

Result
- Geographic dialect continuum
- Dialect leveling

Discussion And Conclusions
- The observation
- The pros and cons of aggregation analysis





Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects

Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation

Result
Geographic dialect continuum
Dialect leveling

Discussion And Conclusions
The observation
The pros and cons of aggregation analysis

Reference

# The pros and cons of aggregation analysis

The advantage

- Aggregation gives a more reliable signal of provenance than single features do.
- It allows us to examine the question whether dialectal variation can best be described in terms of areas of continua.

The limitation

- The linguist structure of the variation is not revealed
- The aggregate analysis does not reveal which linguistic features characterize the dialect areas.



Introduction
The Goal
Language leveling
The aggregating methods
Swedish Dialects
Data and Experimental Design
Data
Acoustic analysis
Analysis of dialectal variation
Result
Geographic dialect continuum
Dialect leveling
Discussion And Conclusions
The observation
The pros and cons of aggregation analysis
Reference

# Reference


- Leinonen, Therese. "Aggregate analysis of vowel pronunciation in Swedish dialects." Oslo Studies in Language 3.2 (2011).
- Hinskens, Frans. "Dialect levelling: A two-dimensional process." Folia linguistica 32.1-2 (1998): 35-52.
- Auer, Peter. "Dialect levelling and the standard varieties in Europe." Folia Linguistica 32.1-2 (1998): 1-9.
- Nerbonne, John. "DataDriven Dialectology." Language and Linguistics Compass 3.1 (2009): 175-198.
- multidimensional scaling http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/chapter3.pdf