

# **RuG/L04** **en webbapplikation**

Therese Leinonen



Nionde nordiska dialektologkonferensen

Uppsala, den 20 augusti 2010

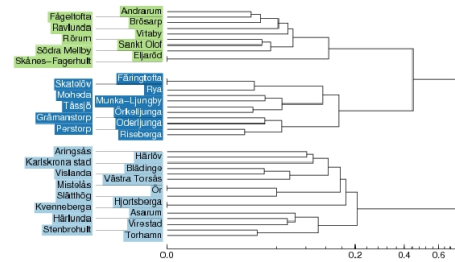
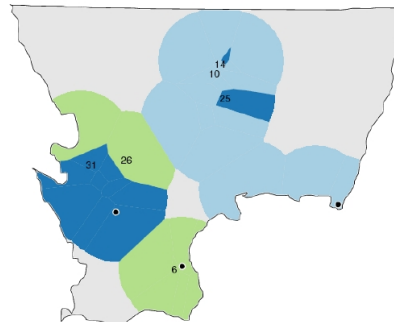
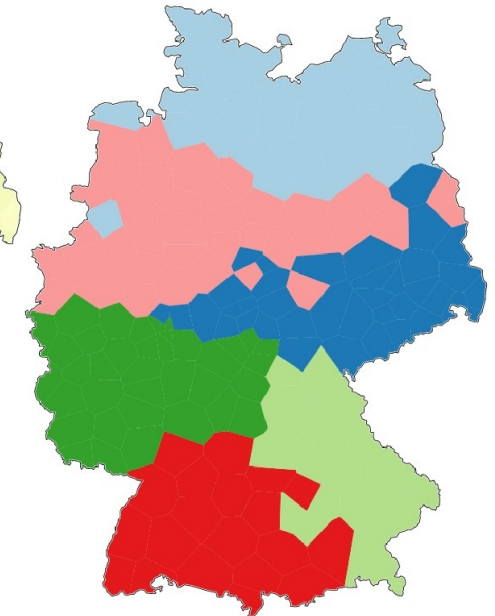
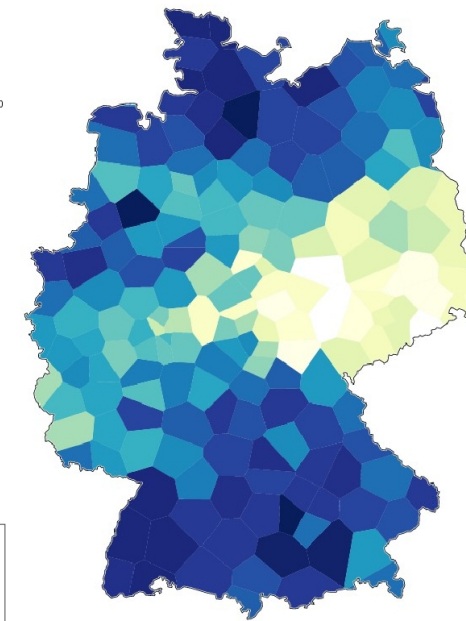
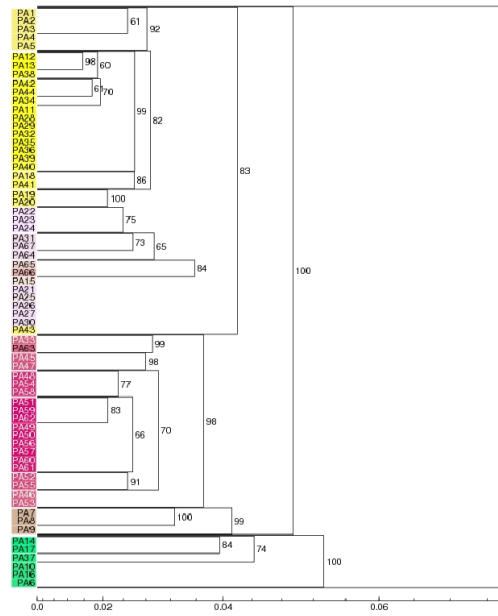
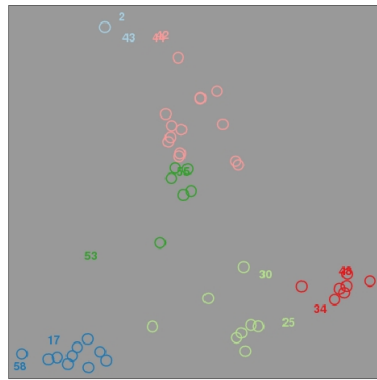
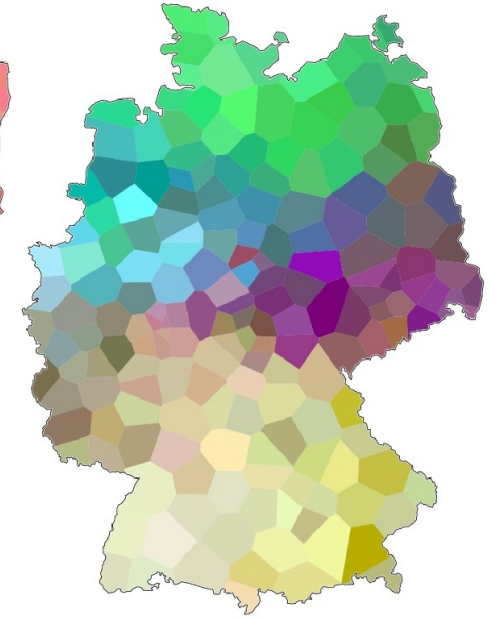
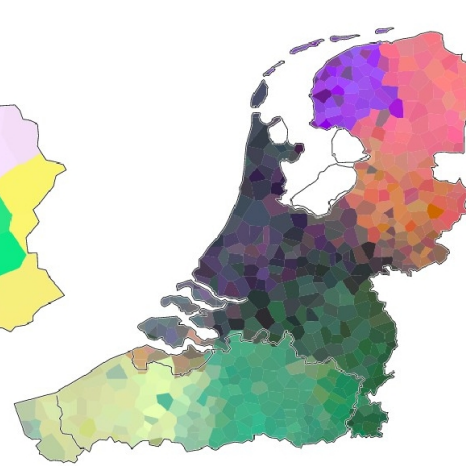
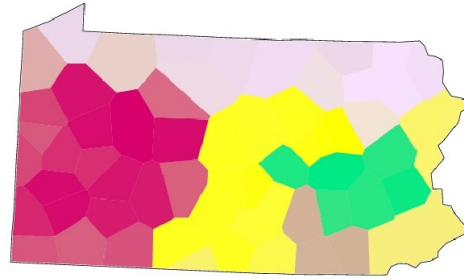
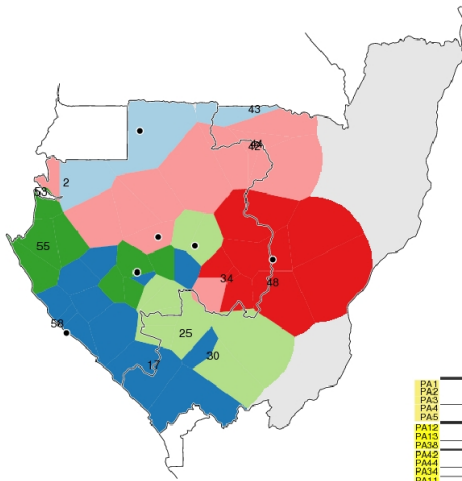
# Bakgrund

- RuG/L04: gratis programvara för dialektometri och kartografi
- utvecklad av Peter Kleiweg, universitet i Groningen
- existerar sedan 2001, har distribuerats gratis sedan 2004
- i den tidiga versionen av programvaran finns inget grafiskt användargränssnitt = komplicerat att använda för många potentiella användare (dialektologer, variationslingvister)
- Adept: CLARIN-NL finansierat projekt för att utveckla en webbapplikation av RuG/L04 programvaran

# Dialektometri

- dialektometri = mätning av dialekter
- syften: definiera dialektområden och beskriva dialektkontinuum
- datadrivet tillvägagångssätt
- vanliga metoder: multidimensionell skalering, faktoranalys, klusteranalys

# Output, exempel



# Dialektdata

- transkriberat dialektmaterial
- laddas upp i webbapplikationen i tabellform med Unicode teckenkodning (kan matas in t.ex. i Microsoft Excel)

Exempel:

	baka	bygga	golv
Andrarum	baga	bydja	gol
Aringsås	baka	bögga	gölv
Asarum	baka / böia	bygga	gål

# Dialektdata

Microsoft Excel - Book1

Type a question for help

File Edit View Insert Format Tools Data Window Help

Paste Special...

100%

10 B

G4 fx

	A	B	C	D	E	F	G	H	I	J	K
1		<b>baka</b>	<b>bygga</b>	<b>golv</b>							
2	<b>Andrarum</b>	baga	bydja	gol							
3	<b>Aringsås</b>	baka	bögga	gölv							
4	<b>Asarum</b>	baka / böia	bygga	gål							
5											

Sheet1 / Sheet2 / Sheet3

Ready

Microsoft Excel - Book1

File Edit View Insert Format Tools

Paste Special

G4 fx

	A	B	C
1		<b>baka</b>	<b>bygga</b>
2	<b>Andrarum</b>	baga	bydja
3	<b>Aringsås</b>	baka	bögga
4	<b>Asarum</b>	baka / böia	bygga
5			

Sheet1 / Sheet2 / Sheet3

Ready

**Save As**

Save in: data

My Recent Documents

Desktop

My Documents

My Computer

My Network Places

File name: Book1.xls

Save as type: Microsoft Office Excel Workbook (\*.xls)

Web Page (\*.htm; \*.html)

Template (\*.xlt)

Text (Tab delimited) (\*.txt)

Unicode Text (\*.txt)

Microsoft Excel 5.0/95 Workbook (\*.xls)

Microsoft Excel 97- Excel 2003 & 5.0/95 Workbook (\*.xls)

Save Cancel

# Dialektdata

Dialekt - Institutet för språk och folkminnen - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.sofi.se/4839

Google laserjet 1010 printable area Search W M hp laserjet 1010 printable area Sign in

RuG/L04 - webapp - project 12 / cluste... Dialekt - Institutet för språk och ...

**Institutet för språk och folkminnen**

Lyssna A- A+ Översikt Om webbplatsen Sök på denna webbplats

Startsida Våra tjänster Språkvård Minoritetsspråk **Dialekter** Namn Folkminnen

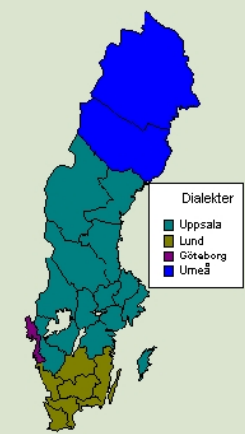
På gång Verksamhet Våra samlingar Lyssna på dialekter Litteratur Internationell samverkan Länkar In English

**Senaste nytt i databasen**

Nu är alla ord från Asarums socken i Blekinge inmatade i databasen.

**Geografiska ansvarsområden**

Institutets dialektverksamhet bedrivs i Göteborg, Lund, Umeå och Uppsala. Dessa orter har olika geografiska ansvarsområden. Se kartan nedan.



**Dialekter**

- Uppsala
- Lund
- Göteborg
- Umeå

**Syd- och västsvenska dialektord**

Stommen till Syd- och västsvensk dialektdata är nu byggd. Klicka på länken till höger för att komma direkt till databasen och söka bland vårt dialektmaterial.

Databasen är under uppbyggnad och innehåller än så länge bara en bråkdel av institutets samlingar. Den bygger på och återger de lapparkiv över sydsvenska dialekter (skånska, halländska, småländska, blekingska, och öländska), som finns på Dialekt- och ortnamnsarkivet i Lund, samt samlingar från Bohuslän, södra Dalsland och Göteborg på Dialekt-, ortnamns- och folkminnesarkivet i Göteborg.

Dialektorden återges med vanliga bokstäver. På de skannade arkivlapparna står uttalet ofta angivet med det s.k. landsmålsalfabetet. Detta är mer nyanserat men inte lämpligt att använda vid sökning. För den som vill läsa uttalet på lapparna kan en förklaring av landsmålsalfabetet vara till hjälp. Vid transkribering av lapparnas uttalsuppgifter till de dialektala formerna i databasen har alltså det vanliga svenska alfabetet använts. Det är ofta lätt att förstå vilket uttal som avses men följande är att tänka på:

- Bokstaven o betecknar ett o-ljud som i *sol* eller *ost* aldrig ett å-ljud som i *son* eller *gott*, vilka transkriberas "sån" resp. "gått". Om en pluraländelse ska återges med å-ljud, som ofta när standardspråket stavar med *or* (*dynor*) eller *er* (*töser*) blir det "år" (*dynår*, *tösår*).
- Som tecken för sje-ljudet skrivs "sj". Tje-ljudet skrivs "tj". Ng-ljudet återges som i vanlig skrift med *ng*.
- W återger ett uttal motsvarande engelskt *w*.

Tillbaka till startsidan

Uppdaterad 27 januari 2010

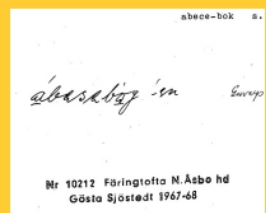
Ansvarig för sidan: Dialektavdelningen (DA)

Skriv ut

**Sök i databasen**

Sök i Syd- och västsvensk dialektdata

**Dialektmaterial**



Nr 10212 Färingtöfa N.Åsbo hd Gösta Sjöstedt 1967-68

Materialet i våra arkiv består av ordlappar som denna. Lappen visar uppslagsord, dialektal form och geografiskt ursprung. Det är detta material som dialektdatabasen bygger på.

**Nyutkommet**

Det har utkommit en ny ordbok över en svensk dialekt. Mandor Näslund har skrivit och gett ut andra bandet.

Find: printable Next Previous Highlight all Match case

Done

# Dialektdata

Syd- och västsvenska dialektord - sökresultat - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://130.235.143.139/dialekt/resultat.aspx?stickord=bygga&dialektord=bygga&ordklass=v&betydelse=bygga&exempel=0&illustration=0&sakupplysning=0&landskap=Blekinge&har. ☆ Google

laserjet 1010 printable area Search hp laserjet 1010 printable area Sign in

RuG/L04 - webbapp - project 12 / cluste... Dialekt - Institutet för språk och folkmin... Syd- och västsvenska dialektord...

<< Återgå till sökning  
<< Återgå till träfflista

|<< << >> >>|

Standardsvensk form	Dialektform	Betydelse	Landskap	Härad	Socken	E	I	S
bygga	bygga	bygga	Blekinge	Bräkne	Asarum	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bygga	bygga	Blekinge	Bräkne	Asarum	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bygga	bygga	Blekinge	Östra	Torhamn	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Blekinge	Östra	Torhamn	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Andrarum	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Brösarp	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Brösarp	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
bygga	byggja	bygga	Skåne	Albo	Fågelfohta	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Rörum	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Sankt Olof	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Sankt Olof	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Södra Mellby	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	byggja	bygga	Skåne	Albo	Södra Mellby	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Södra Mellby	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bydja	bygga	Skåne	Albo	Vitaby	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Aringsås	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Aringsås	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bygga	bygga	Småland	Allbo	Härlunda	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Härlöv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Härlöv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Slätthög	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bögga	bygga	Småland	Allbo	Västra Torsås	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bygga	bygga	bygga	Småland	Allbo	Västra Torsås	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

E: Exempel I: Illustration S: Sakupplysning

bygga

*bygga*, -'e, *bygde*, *bygt*, *bydja*  
*bygga*

*Asarum*

1929 M 586 Tage Ahlden *Bräkne*

Standardsvensk form: bygga  
Dialektal form: bygga  
Landsmålsform:  
Ordklass: v  
Betydelse: bygga  
Landskap: Blekinge  
Härad: Bräkne  
Socken: Asarum  
Övrigt:  
Bildlänk: [Större bild](#)  
Skriv ut bild: [Skriv ut](#)

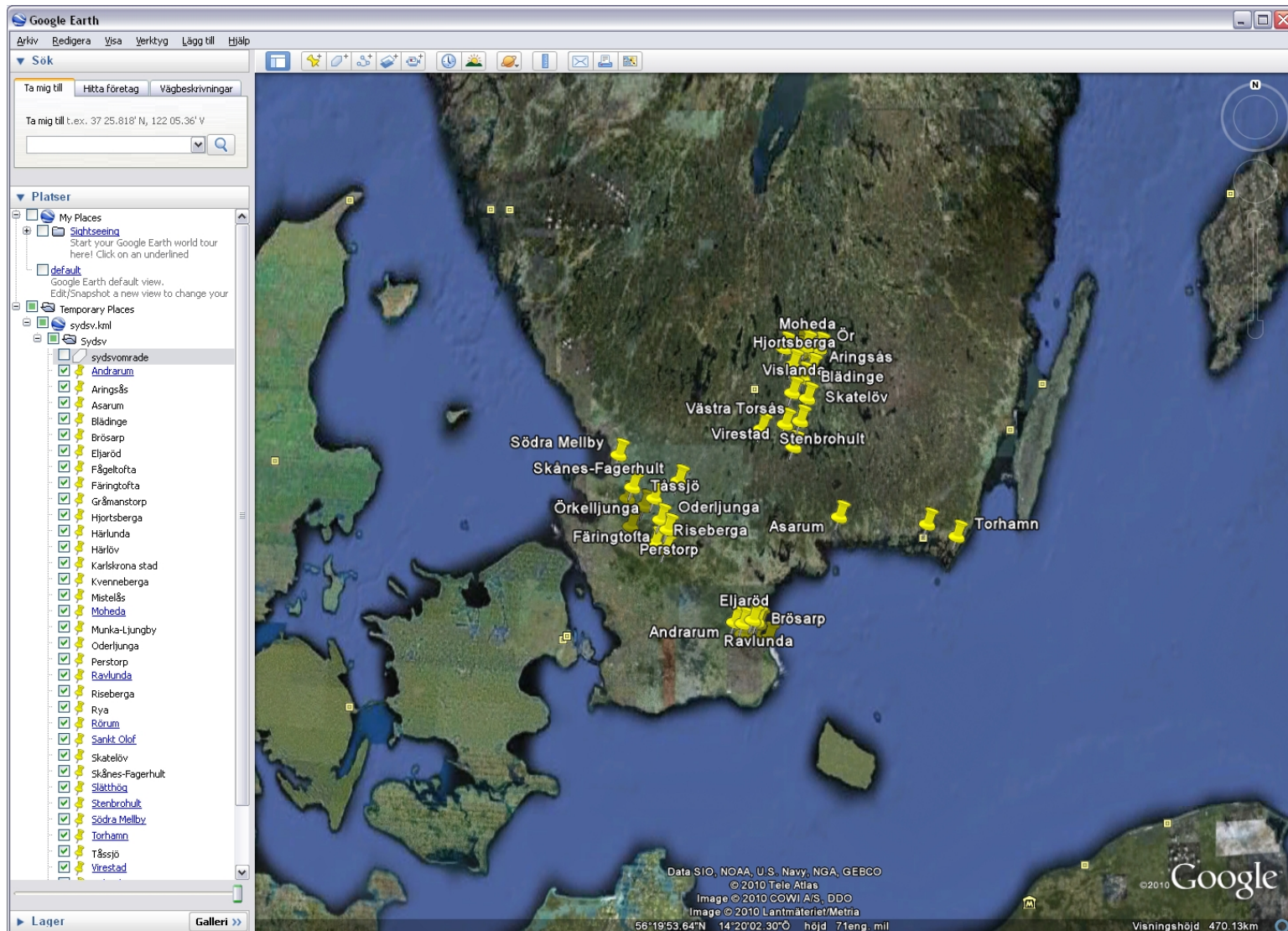
Find: printable Next Previous Highlight all Match case

Done



# Geografiska data

- importera geografiska data (insamlingsorter, områdesgränser) med hjälp av Google Earth (<http://earth.google.com/>)



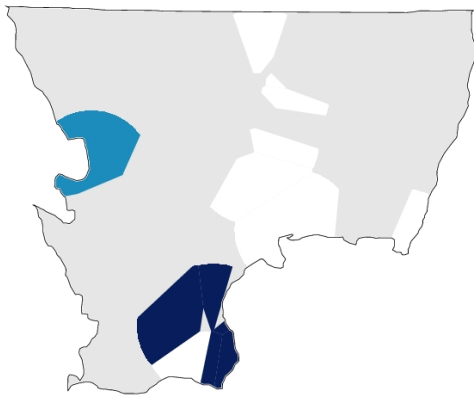
- 1 Andrarum
- 2 Aringsås
- 3 Asarum
- 4 Blädinge
- 5 Brösarp
- 6 Eljaröd
- 7 Fågeltofta
- 8 Färingtofta
- 9 Gråmanstorp
- 10 Hjortsberga
- 11 Härlunda
- 12 Härlöv
- 13 Karlskrona stad
- 14 Kvenneberga
- 15 Mistelås
- 16 Moheda
- 17 Munka-Ljungby
- 18 Oderljunga
- 19 Perstorp
- 20 Ravlunda
- 21 Riseberga
- 22 Rya
- 23 Rörum
- 24 Sankt Olof
- 25 Skatelöv
- 26 Skånes-Fagerhult
- 27 Slätthög
- 28 Stenbrohult
- 29 Södra Mellby
- 30 Torhamn
- 31 Tåssjö
- 32 Virestad
- 33 Vislanda
- 34 Vitaby
- 35 Västra Torsås
- 36 Ör
- 37 Örkelljunga



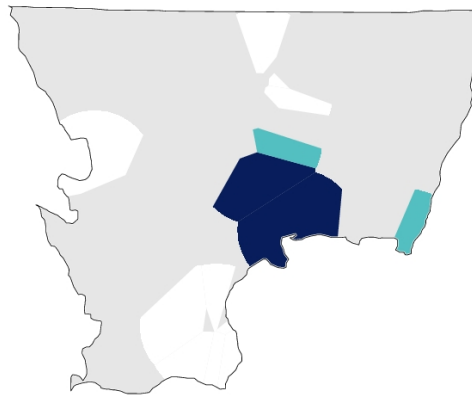
# Inspektera materialet

- sammanfattning av materialet (antal orter, antal ord, antal tecken etc.)
- distributionskartor:

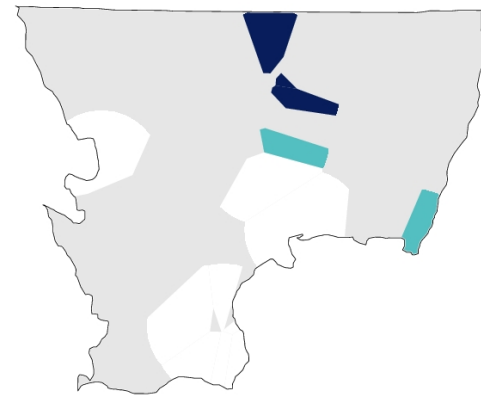
Exempel: bygga



bydja



bygga



bögga

# Levenshteinavstånd

- algoritm för att räkna ut det lingvistiska avståndet mellan dialekter utifrån transkriberade ord

Exempel:

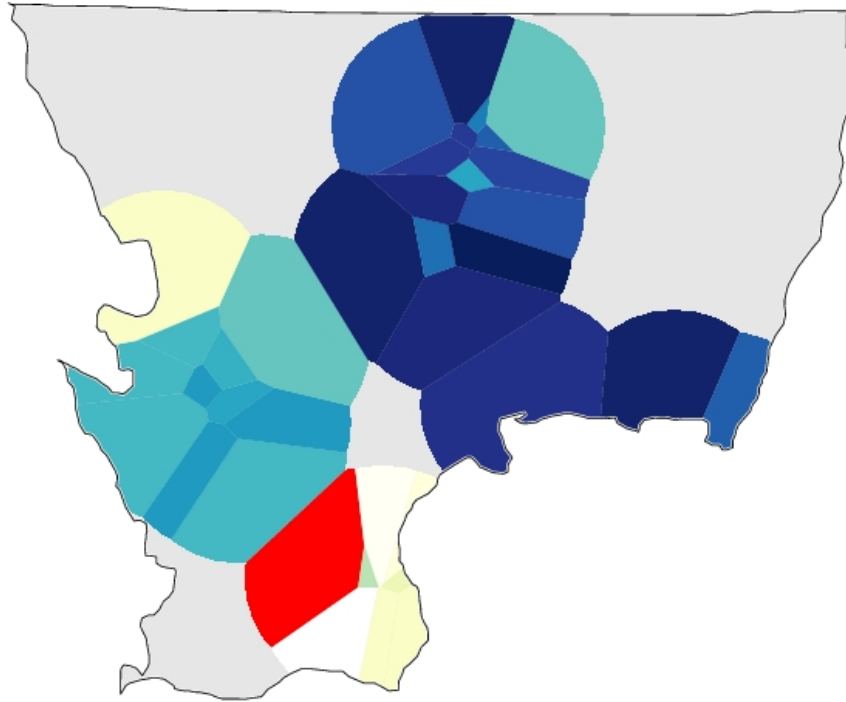
b	y	g	g	a
b	ö	g	g	a
0	1	1	1	1

b	y	g	g	a
b	y	d	j	a
0	0	1	2	2

b	ö	g	g	a
b	y	d	j	a
0	1	2	3	3

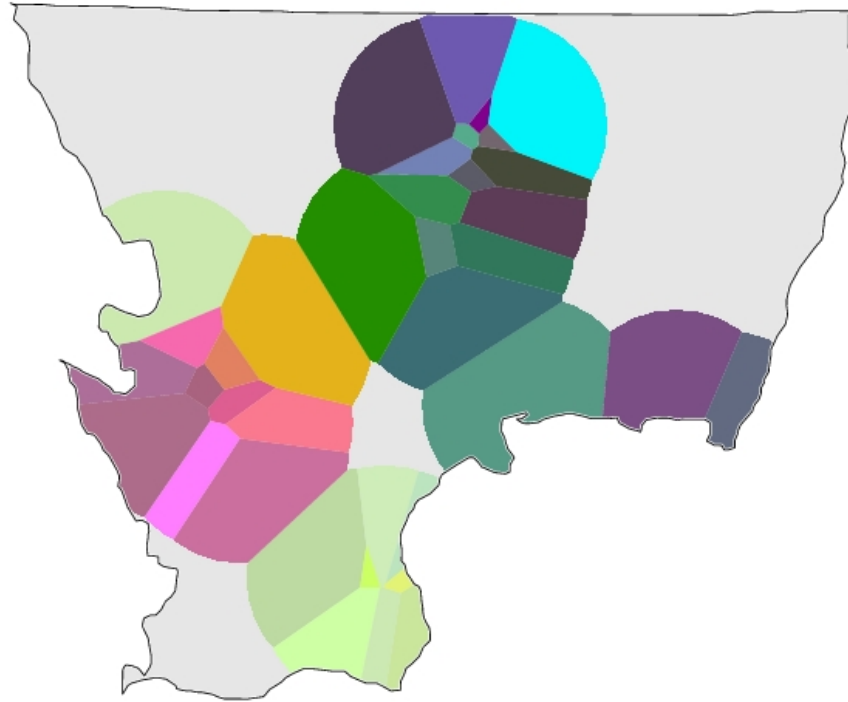
- avståndet räknas ut för alla dialektpar och alla ordpar
- avståndet mellan två dialekter är medeltalet av avståndet för alla ord som finns belagda i båda dialekterna

# Referenskartor



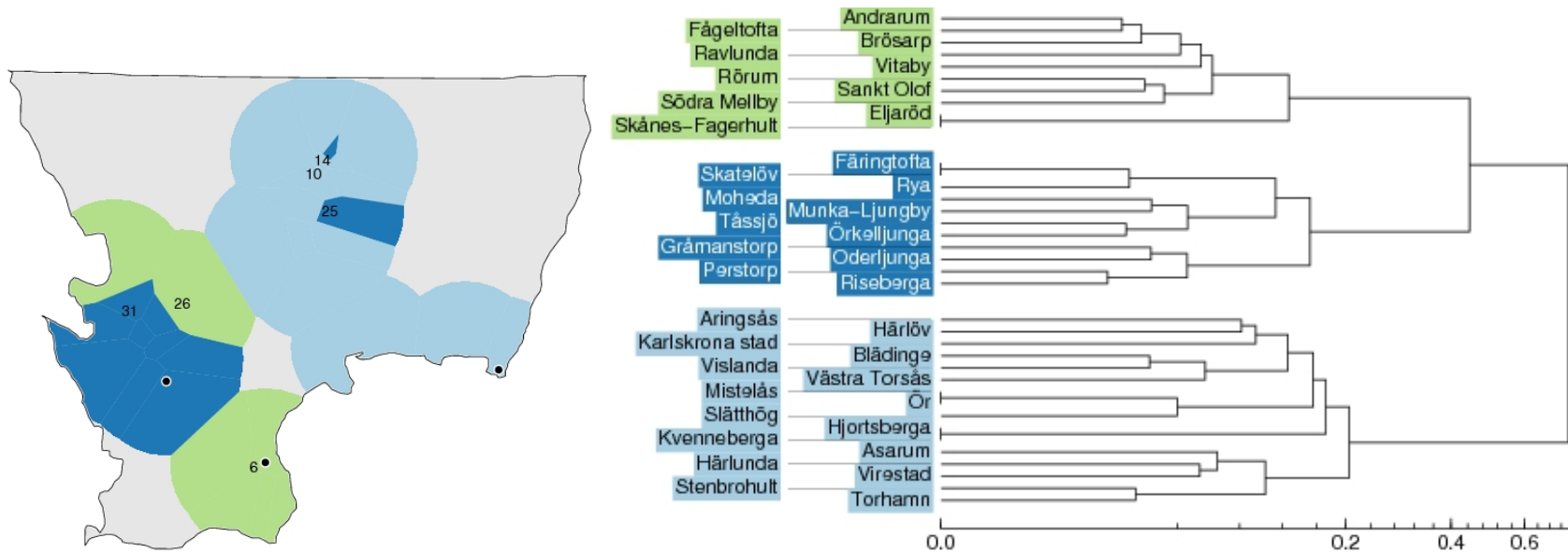
- lingvistiskt avstånd från en ort (röd: Andrarum) till alla andra orter (ju mörkare desto större avstånd)

# Statistiska analyser



- multidimensionell skalering: visar de lingvistiska relationerna mellan alla orter som ett kontinuum

# Statistiska analyser



- klusteranalys: grupperar orterna i grupper
- de dialekter som liknar varandra mest hamnar i samma grupp => dialektklassificering

Länk:

<http://www.let.rug.nl/~kleiweg/L04/webapp/>

OBS!!! Webbapplikationen är fortfarande **under uppbyggnad**. Det går att registrera sig som användare och prova sig fram, men alla funktioner är ännu inte utbyggda.