



UiO : Det humanistiske fakultet

Unn Røyneland

Sosiolingvistik

Höstsymposium Åbo

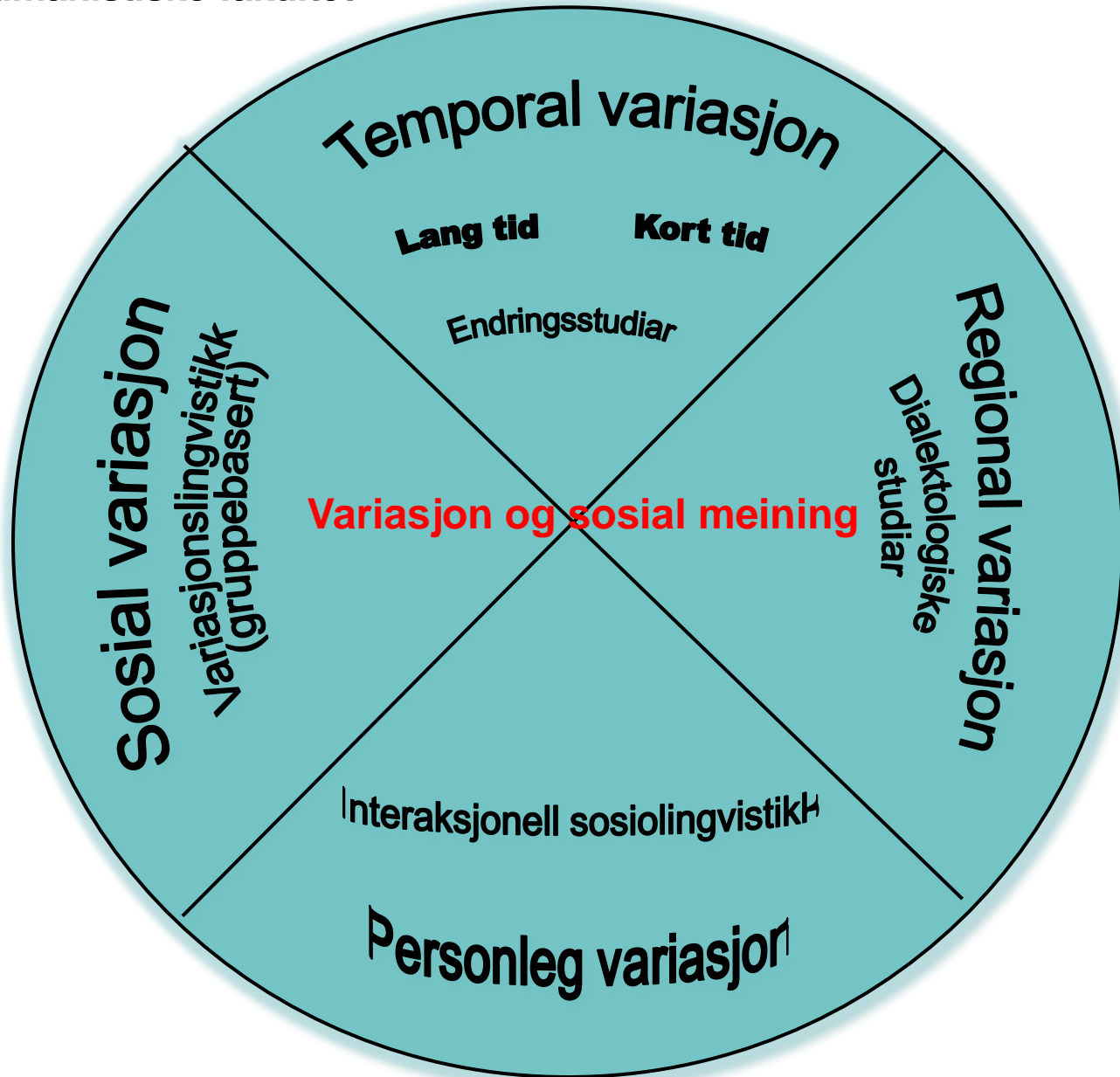
5.-6. oktober 2012



Litt oppvarming

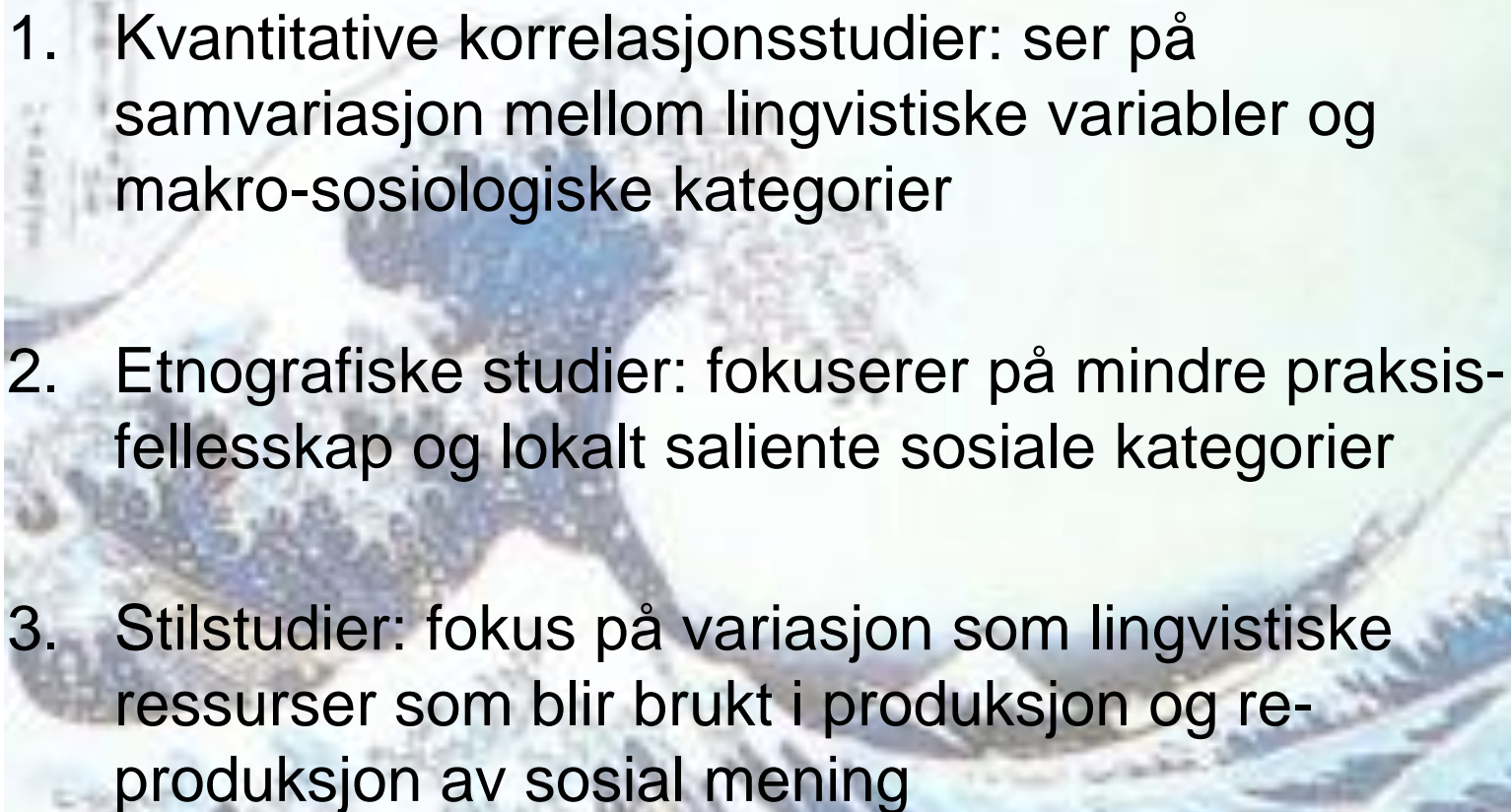


- Hva er sosiolingvistikk?
- Variasjon!



Variasjonsstudium – tre bølger

Penny Eckert (2012). Three Waves of Variation Study

1. Kvantitative korrelasjonsstudier: ser på samvariasjon mellom lingvistiske variabler og makro-sosiologiske kategorier
 2. Etnografiske studier: fokuserer på mindre praksis-fellesskap og lokalt saliente sosiale kategorier
 3. Stilstudier: fokus på variasjon som lingvistiske ressurser som blir brukt i produksjon og re-produksjon av sosial mening
- 

Endring i synet på variasjon

- Fra variasjon sett som *refleksjon* av det sosiale rommet til variasjon som *ressurs* for å konstruere sosial mening
- Fra studium av struktur til studium av praksis
- Ulik metodologisk og analytisk praksis

Men alle er interessert i:

- Variasjon
- ... og da opererer man med variabler dvs.
 - data som kan ta ulike verdier
 - tellbare (målbare) data
- ... og da *kan* man benytte seg av statistikk

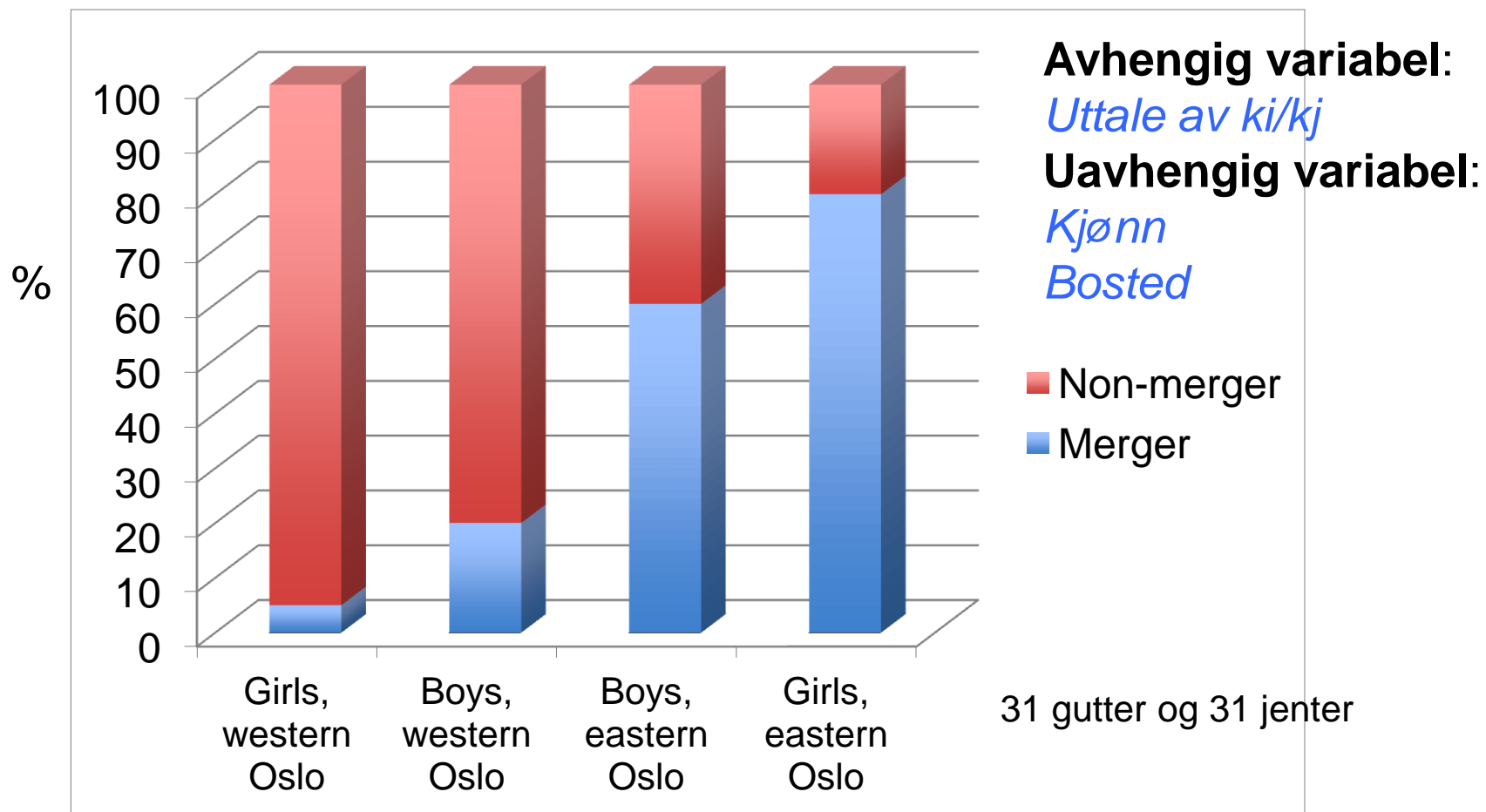
STATISTISKE METODER SOSIOLINGVISTIKKEN

Sosiolingvistiske analysemetoder

- Variasjonslingvistikken
 - Bivariat statistikk – samvariasjon mellom språklige og sosiale makrovariabler
 - Avhengige variabler
 - Uavhengige variabler
 - Signifikanstesting
 - Chi-kvadrat
 - T-test

/ʃøre/ eller /çøre/? (NoTa-Oslo)

N=1421, ny uttale: /ʃ/ 555



Sosiolingvistiske analysemetoder

- Multivariat statistikk
 - Faktoranalyse
 - Mulippel regressjon analyse
 - Principle component analysis
 - ANOVA
 - OSV.
 - Korrespondanseanalyse

Hvordan kan statistikk hjelpe meg?

- gir muligheter for datareduksjon (oppdage og sammenfatte utviklingstendenser, finne fellestrekk ved et datasett)
- gir muligheter for å trekke slutninger (inferential statistics)
- gir muligheter for å oppdage sammenhenger
- gir muligheter for å utforske prosesser som handler om sannsynlighet

Hva er jeg ute etter?

- Hvilke forskingsspørsmål har jeg?
- Teste hypoteser?

eller

- Finne mønster?

Hvilke type data har jeg?

- Er observasjonene:
 - Kategoriske
 - Kontinuerlige
 - Representativt utvalgt
 - Normalfordelte
 - Osv..
- Ulike analysemetoder stiller ulike krav til data
- Hva slags statistisk analyse passer dataene/observasjonene mine til?

Spør en:

Statistiker



Hva gjør statistikk for oss?

- I think that one of the main points to keep in mind about drawing conclusions from data is that it is all guessing. Really. But what we are trying to do with statistical summaries and hypothesis testing is to quantify just how reliable our guesses are.

Keith Johnson: *Quantitative methods in linguistics*, 2008

KORRESPONDANSEANALYSE

Korrespondanseanalyse – historikk

- Utviklet og brukes primært innenfor fransk analysemetodologi
- Jean-Paul Benzécri, 1960-tallet – utg.pkt i Karl Pearson's arbeid
 - Mål: lage ein induktiv og eksplorerende teknikk for analyse av lingvistiske data, i opposisjon til mer deduksjonistisk orientert lingvistikk – ”distribusjonell språkteori” (Benzécri 1976)

Korrespondanseanalyse – historikk

- Lenge relativt ukjent i nordisk og anglo-amerikansk sosialvitenskap
- 1980-tallet: markedsanalyse
- Pierre Bourdieu: *La Distinction*, (1979)
- Statistikerne Michael Greenacre og Jörg Blasius (1994, 1998)
- Statistiker og filolog Daniel Apollon (1990)

Hva er korrespondanseanalyse

- Multivariat analysemetode
- Avdekke de underliggende strukturene i komplekse datamatriser
- Flerdimensjonalt punktdiagram – *korrespondansekart*
- Oppsummere mest mulig av variansen i et datamaterialet langs så få akser som mulig

Hva er korrespondanseanalyse?

- Deskriptiv og eksplorerende analysemetode
- Modellen skal følge data, ikke data modellen
- Primært hypotese-*genererende* og ikke hypotese-*testende*
- Avdekke systemet av relasjoner mellom ulike variabler

Ulike typer korrespondanseanalyser

- Enkel korrespondanseanalyse
 - Krysstabeller / frekvenstabeller

Fritidsaktivitet	Alder		Kjønn		Utdanning	
	Ung	Voksen	K	M	Høy	Lav
Kjøre skuter						
Snowboard						

- Multippel korrespondanseanalyse
 - Multivariate tabeller / disjunktive tabeller (Burt-matrix)

UiO : Det humanistiske fakultet

ALTERNATIV2: INNHOLDER FREKVENS AV LINGVISTISKE MARKØRER

individ-id	S-PAL-LOKAL	S-PAL-NIVELLERT	N-PAL-LOKAL	N-PAL-NIVELLERT	L-PAL-LOKAL	L-PAL-NIVELLERT	MON-LOKAL	MON-NIVELLERT	NEG-LOKAL	NEG-NIVELLERT	TRYKK-LOKAL	TRYKK-NIVELLERT	JAM-LOKAL	JAM-NIVELLERT
R-Ailin	0	59	54	68	41	38	0	7	34	9	50	12	4	30
R-Bente	0	38	42	34	13	30	1	2	41	2	56	3	9	5
R-Carola	15	41	78	57	34	58	4	1	41	2	50	0	5	7
R-Iris	1	38	18	99	5	49	0	8	9	36	5	69	1	18
R-Linn	0	40	49	26	16	29	3	1	47	0	21	11	7	10
R-Chris	60	9	39	72	27	42	5	0	48	0	82	13	12	5
R-Eirik	68	1	90	21	47	8	7	0	40	2	55	1	5	9
R-Magne	2	30	20	89	8	52	1	1	18	20	44	24	8	4
R-Sondre	0	27	4	65	9	54	0	2	21	3	24	7	0	4
R-Truls	0	19	2	67	3	26	0	4	40	5	29	4	0	3
R-Bjørn	1	54	2	115	10	52	1	0	0	47	33	12	10	5
R-Finn	61	1	109	7	86	5	13	0	45	0	34	1	25	3
R-Hogne	7	0	21	5	15	0	1	0	18	0	14	0	6	0
R-Trym	18	1	37	0	19	0	1	0	20	0	21	2	9	0
R-Tore	14	0	47	7	15	1	4	1	15	0	10	1	6	1
R-Ragnar	17	12	95	18	57	15	10	0	48	0	83	14	14	0
R-Stein	36	3	80	29	59	5	2	0	43	0	57	3	20	0
R-Tom	1	50	32	130	16	62	5	5	25	12	15	98	10	13
R-Anna	32	0	22	24	29	16	5	1	30	3	13	0	6	2
R-Eli	68	1	42	21	36	13	5	0	40	0	29	2	10	1
R-Gunn	36	0	125	3	91	1	6	0	44	0	38	0	28	0
T-Åsne	*	*	0	69	0	53	5	7	0	49	30	9	1	4
T-Ane	*	*	0	34	2	39	8	0	0	42	12	4	1	3
T-Laila	*	*	0	51	0	51	0	0	0	48	4	22	0	11
T-Ina	*	*	0	77	0	58	6	10	0	50	29	11	1	18
T-Mia	*	*	0	83	0	65	0	6	0	49	41	3	2	17
T-Gunnar	*	*	0	39	0	46	0	5	0	44	23	5	1	10
T-Ola	*	*	0	84	0	78	0	1	0	49	44	2	1	12
T-Kåre	*	*	0	42	0	60	0	3	0	48	4	40	0	11
T-Arne	*	*	0	22	0	31	2	0	0	36	25	9	0	4
T-Hans	*	*	0	13	0	19	1	3	0	13	9	1	0	3
T-Varg	*	*	3	22	4	14	0	3	0	27	9	0	1	0
T-Bo	*	*	1	40	0	25	3	2	0	27	12	2	0	5
T-Alva	*	*	18	58	1	57	6	0	48	0	27	8	0	4
T-Ask	*	*	0	62	0	62	8	1	2	33	26	13	0	10
T-Oskar	*	*	0	55	2	45	1	0	0	38	20	4	0	4
T-Njål	*	*	7	71	1	52	6	6	0	47	40	1	0	14

Dialektnivellering i Norge

Røynealand 2005. *Dialektnivellering, ungdom og identitet*. Oslo: UniPub.

TYNSET

RØROS





Variabler

- 37 respondenter
- 15 språklige variabler: fonologiske, morfo-fonologiske, morfologiske og leksikalske
- 12 sosiale variabler:
 - Kjønn jente/gutt
 - Alder ung/voksen
 - Skolretning allmennfag/yrkesfag
 - Foreldrebakgrunn begge/ein/ingen frå heimlassen
 - Sosial bakgrunn akademisk/ikkje-akademisk
 - Lokalisering sentrum/utkant
 - Gruppetilhørighet bus-råner/soss-skater/skule-idrett/skule-kultur
 - Framtidsplan bli/flytte/flytte tilbake
 - Lokal orientering pluss/nøytral/minus
 - Urban orientering pluss/nøytral/minus
 - Holdning til hjemstedet pluss/nøytral/minus
 - Holdning til egen dialekt pluss/nøytral/minus

Tredelt analyse

- Variabelanalyse
 - Detaljert analyse av de 15 språklige variablene: variantfordeling i forhold til sted, alder og på individnivå
 - Bivariat statistikk
- Korrespondanseanalyse
 - Multivariat, eksplorerende statistisk analyse
 - To- el. tredimensjonale kart over alle informantene, de språklige og de sosiale variablene
- Informantportrett
 - Hermeneutisk orientert analyse av noen av ungdommene på basis av intervjuene

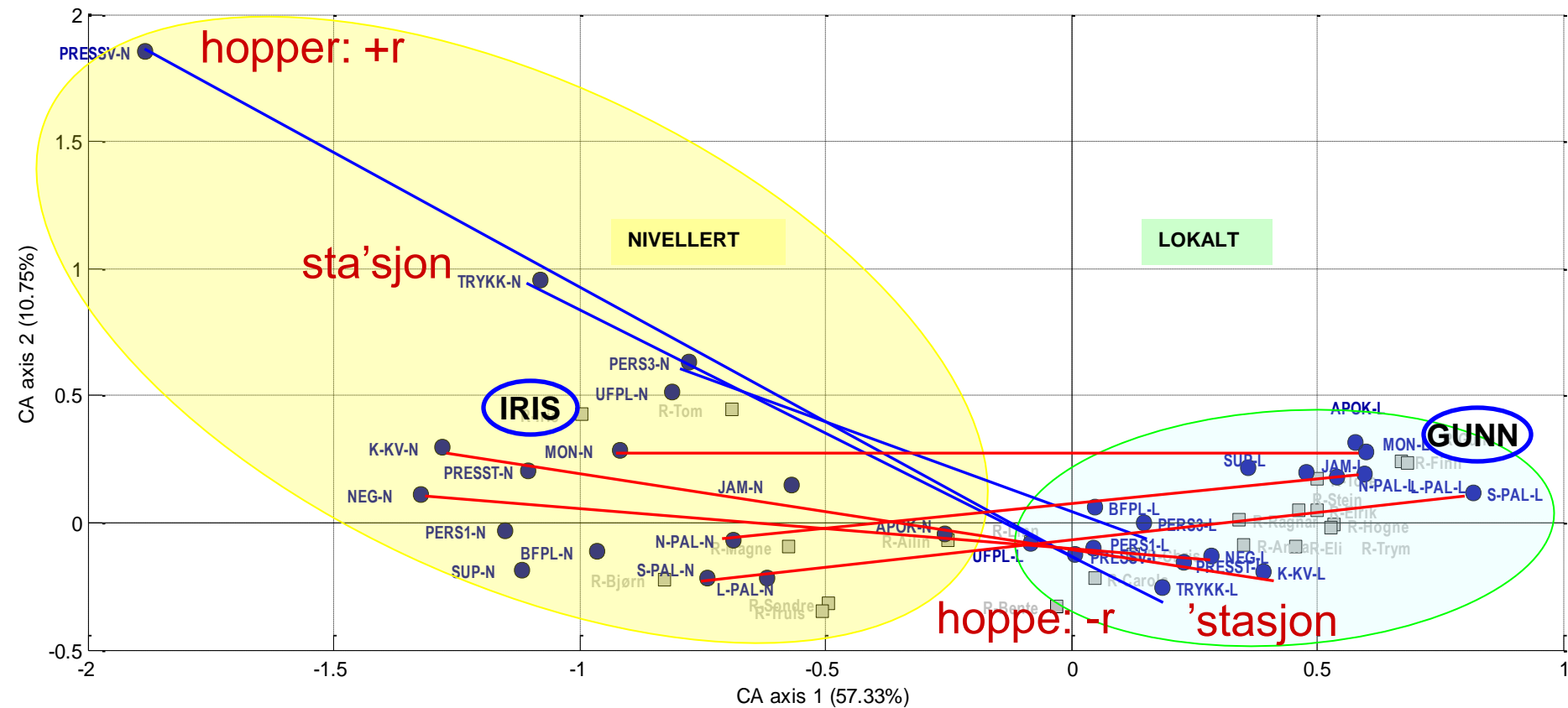
UiO : Det humanistiske fakultet

ALTERNATIV2: INNHOLDER FREKVENS AV LINGVISTISKE MARKØRER

individ-id	S-PAL-LOKAL	S-PAL-NIVELLERT	N-PAL-LOKAL	N-PAL-NIVELLERT	L-PAL-LOKAL	L-PAL-NIVELLERT	MON-LOKAL	MON-NIVELLERT	NEG-LOKAL	NEG-NIVELLERT	TRYKK-LOKAL	TRYKK-NIVELLERT	JAM-LOKAL	JAM-NIVELLERT
R-Ailin	0	59	54	68	41	38	0	7	34	9	50	12	4	30
R-Bente	0	38	42	34	13	30	1	2	41	2	56	3	9	5
R-Carola	15	41	78	57	34	58	4	1	41	2	50	0	5	7
R-Iris	1	38	18	99	5	49	0	8	9	36	5	69	1	18
R-Linn	0	40	49	26	16	29	3	1	47	0	21	11	7	10
R-Chris	60	9	39	72	27	42	5	0	48	0	82	13	12	5
R-Eirik	68	1	90	21	47	8	7	0	40	2	55	1	5	9
R-Magne	2	30	20	89	8	52	1	1	18	20	44	24	8	4
R-Sondre	0	27	4	65	9	54	0	2	21	3	24	7	0	4
R-Truls	0	19	2	67	3	26	0	4	40	5	29	4	0	3
R-Bjørn	1	54	2	115	10	52	1	0	0	47	33	12	10	5
R-Finn	61	1	109	7	86	5	13	0	45	0	34	1	25	3
R-Hogne	7	0	21	5	15	0	1	0	18	0	14	0	6	0
R-Trym	18	1	37	0	19	0	1	0	20	0	21	2	9	0
R-Tore	14	0	47	7	15	1	4	1	15	0	10	1	6	1
R-Ragnar	17	12	95	18	57	15	10	0	48	0	83	14	14	0
R-Stein	36	3	80	29	59	5	2	0	43	0	57	3	20	0
R-Tom	1	50	32	130	16	62	5	5	25	12	15	98	10	13
R-Anna	32	0	22	24	29	16	5	1	30	3	13	0	6	2
R-Eli	68	1	42	21	36	13	5	0	40	0	29	2	10	1
R-Gunn	36	0	125	3	91	1	6	0	44	0	38	0	28	0
T-Åsne	*	*	0	69	0	53	5	7	0	49	30	9	1	4
T-Ane	*	*	0	34	2	39	8	0	0	42	12	4	1	3
T-Laila	*	*	0	51	0	51	0	0	0	48	4	22	0	11
T-Ina	*	*	0	77	0	58	6	10	0	50	29	11	1	18
T-Mia	*	*	0	83	0	65	0	6	0	49	41	3	2	17
T-Gunnar	*	*	0	39	0	46	0	5	0	44	23	5	1	10
T-Ola	*	*	0	84	0	78	0	1	0	49	44	2	1	12
T-Kåre	*	*	0	42	0	60	0	3	0	48	4	40	0	11
T-Arne	*	*	0	22	0	31	2	0	0	36	25	9	0	4
T-Hans	*	*	0	13	0	19	1	3	0	13	9	1	0	3
T-Varg	*	*	3	22	4	14	0	3	0	27	9	0	1	0
T-Bo	*	*	1	40	0	25	3	2	0	27	12	2	0	5
T-Alva	*	*	18	58	1	57	6	0	48	0	27	8	0	4
T-Ask	*	*	0	62	0	62	8	1	2	33	26	13	0	10
T-Oskar	*	*	0	55	2	45	1	0	0	38	20	4	0	4
T-Njål	*	*	7	71	1	52	6	6	0	47	40	1	0	14

RØROS

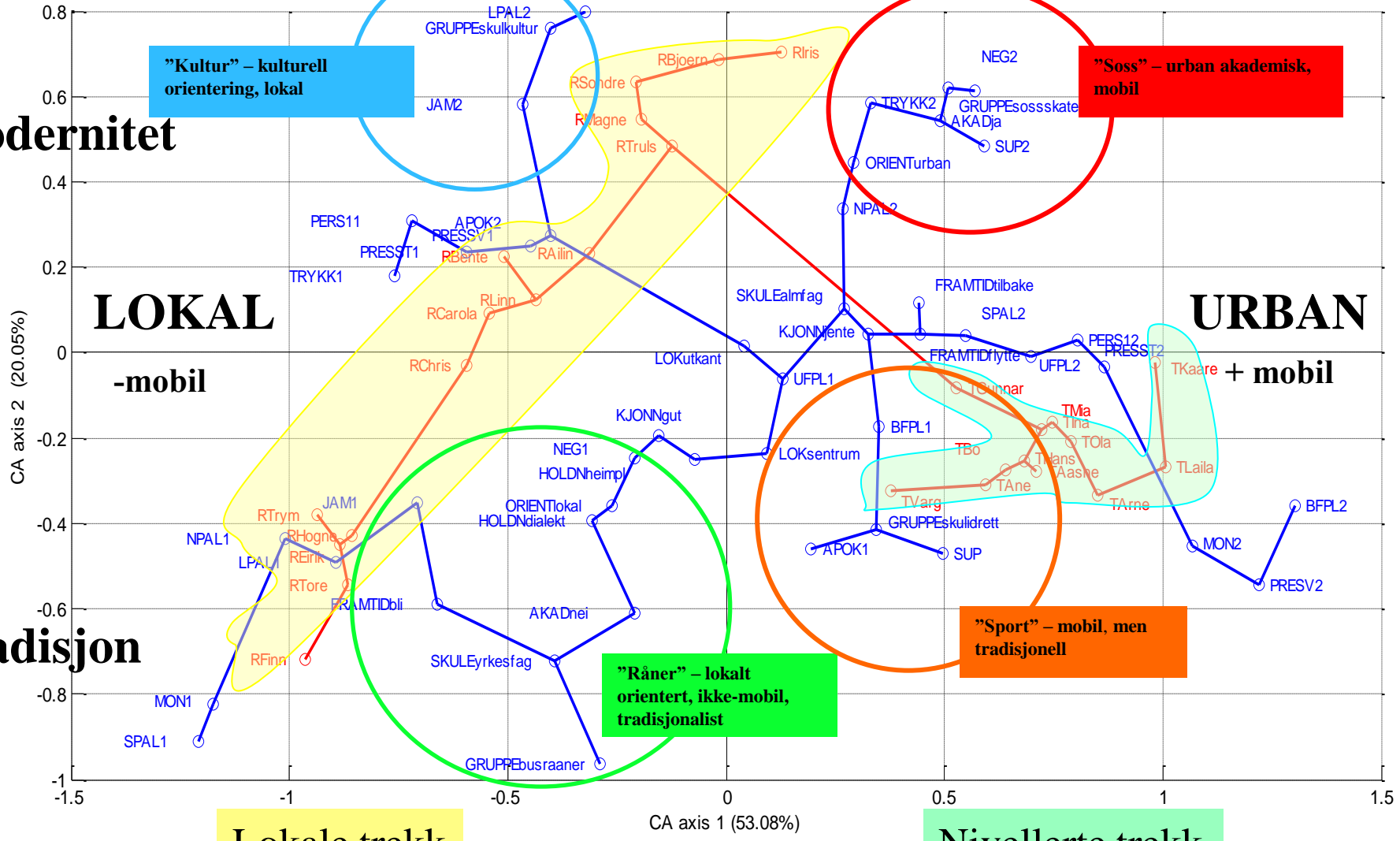
Correspondence Analysis unn_roros_b_arefono.dat



Tolking av korrespondansekartet

- Ut fra den relative posisjonen mellom punkt kan resultatet tolkes som klynger av egenskaper i rommet
 - Nærhet mellom punkt avslører gjensidig affinitet og kan bety at disse punktene utgjør et praksisfelt
 - Avstand mellom punkt viser at de sjelden korresponderer
 - Ekstremt stor avstand mellom punkt viser at disse egenskapene genererer den viktigste eller mest åpenbare variasjonen i datasettet
 - Punkt ved origo ligg nær gjennomsnittet for hele datasettet og har en profil som ligner marginalfordelingene for hele materialet

Correspondence Analysis unn_uge.dat



Relasjonistisk analysemetode

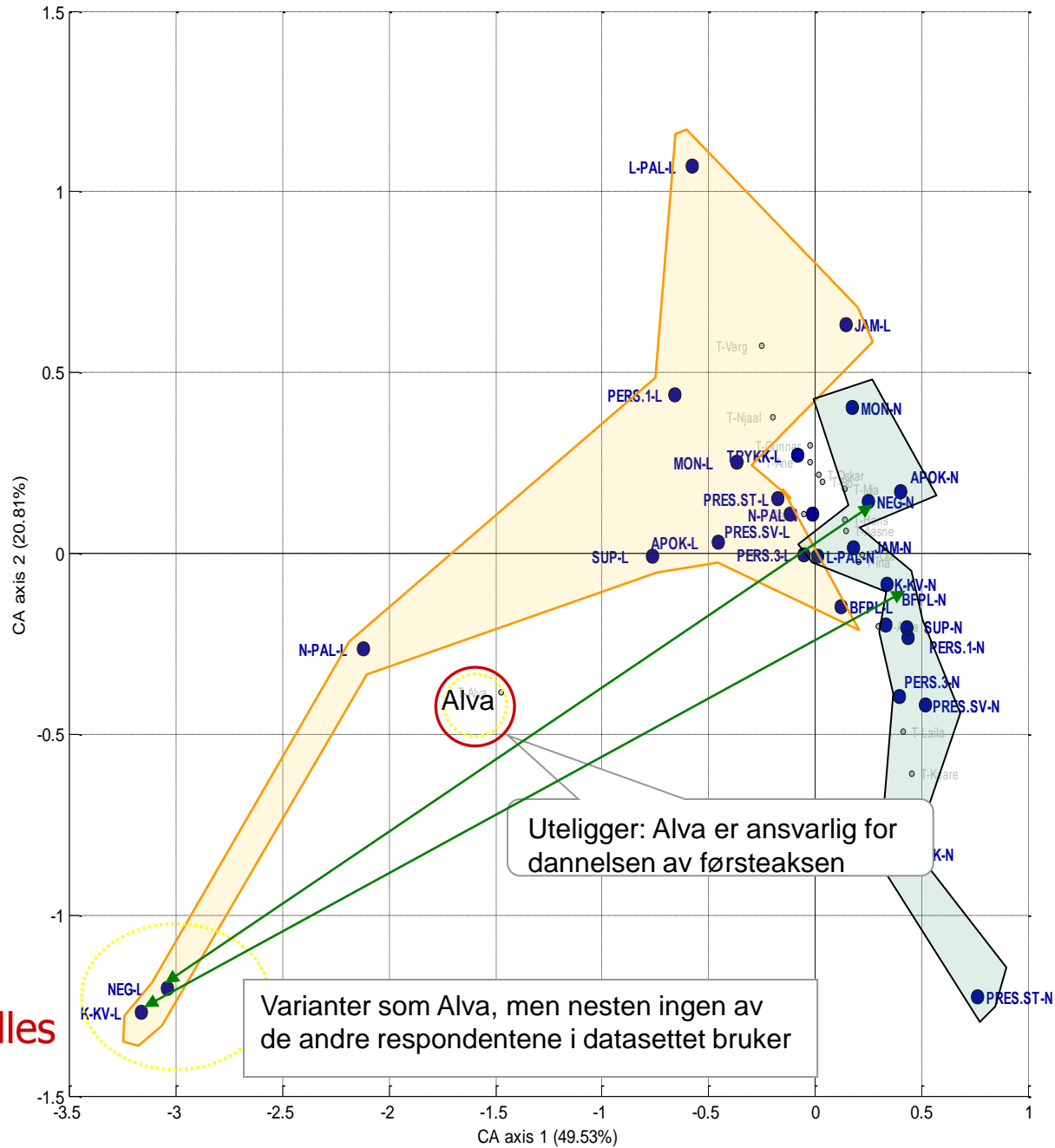
- [...] various statistical techniques contain implicit social philosophies that need to be made explicit. When you perform a regression analysis, a path analysis, or a factorial analysis, you need to know what social philosophy you are bringing in, and more especially what philosophy of causality, action, the mode of existence of social things, and so on
- [...]. I make a lot of use of the analysis of correspondences, because I think it's an essentially *relational* technique whose philosophy entirely corresponds to what social reality is in my view. It's a technique that "thinks" in terms of *relationships*, as I try to do with the idea of the field. (Bourdieu & Krais 1991:254).

Relasjonistisk sosiologi

- Sosiale fenomen eksisterer i et system av *relasjoner*
- Plasseringa av et punkt i korrespondanseanalysen er gitt ut fra punktet sin relasjon til alle andre punkt i matrisen
- Ifølge Bourdieu er det å knytte ett isolert punkt i det sosiale rommet til ett punkt i det språklige rommet, å kortslutte analysen fordi det reduserer heterogene data til endimensjonale variabler som skjuler viktige multidimensjonale sammenhenger.
- Systemet av relasjoner mellom sosiale posisjoner må relateres til systemet av relasjoner mellom språklige trekk, for at *mønsteret av relasjoner* kan tre fram.

Fordeler ved korrespondanseanalyse

- Data-drevet og basert på induksjon
- Eksplorerende og dynamisk
 - Legge til eller fjerne variabler i løpet av analysen for å utforske ulike aspekt og teste analysen
 - Fjerne upålitelige data
 - Fjerne data med lav svar-frekvens
 - Fjerne “uteliggere” dvs. variabler som strukturerer datasettet i ekstrem grad



Hvorfor korrespondanseanalyse?

- Gir et omfattende bilde av komplekse datamateriale
- Opner opp for “felt”-tenkning – lingvistiske og sosiale praksisfelt
- Er svært fleksibel
- Tillater ulike typer data
- God metode for å finne mønstre i kvalitative data
- Hypotese- og teori-genererende
- Resultatet blir presentert i relativt forståelige og lett tilgjengelige grafiske bilder



UiO : **Det humanistiske fakultet**

Unn Røyneland

Takk for oppmerksomheten!



Litteratur

Apollon, Daniel 2003. Kurs 12.-13. juni 2003. Ppt-fil: Innføring i praktisk korrespondanseanalyse.

<http://huminf.uib.no/~daniel/coran/Korrespondanseanalyse.ppt>.

— 1990. Dataanalytiske metoder i filologien. I Haugen, O. E. & E. Thomassen (red.) *Den filologiske vitenskap*. Oslo: Solum, 181-208.

Benzécri, Jean-Paul 1976. *Histoire et préhistoire de l'analyse des données*. Paris: Dunod.

Bourdieu, Pierre 1979. *La Distinction: Critique sociale du jugement*. Paris: Editions de Minuit.

Broady, Donald 1991. *Sociologi och epistemologi: om Pierre Bourdieus författarskap och den historiska epistemologin*. 2. korr. oppl. Stockholm: HLS förlag.

Clausen, Sten-Erik 1998. *Applied Correspondence Analysis: An Introduction*. Sage University Paper series on Quantitative Applications in the Social Sciences, 121. Thousand Oaks, CA: Sage.

Litteratur

- Greenacre, Michael J. & J. Blasius (red.) 1998. *Visualization of categorical data*. San Diego, Calif.: Academic Press.
- (red.) 1994. *Correspondence Analysis in the Social Sciences: Recent Developments and Applications*. London: Academic Press.
- Eckert, Penny 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Variation. *Annual Review of Anthropology*, 41, 87-100.
- Hjellbrekke, Johs. 1999. *Innføring i korrespondanseanalyse*. Bergen: Fagbokforlaget.
- Johnson, Keith 2008. *Quantitative Methods in Linguistics*. Oxford: Blackwell.
- Rosenlund, Lennart 1992. *Korrespondanseanalyse. Metode for kvantitativ behandling av kvalitative data*. Stavanger: Arbeidspapirer fra Høgskolesenteret i Rogaland.
- Røyneland, Unn 2004. *Dialektnivellering, ungdom og identitet. Ein komparativ studie av språkleg variasjon og endring i to tilgrensande dialektområde, Røros og Tynset*. Oslo: Acta Humaniora or as pdf-files: <http://folk.uio.no/unnr/avhandling/>