

Therese Leinonen

Talet lever!

Fyra studier i svenskt talspråk i Finland

Studier i
nordisk
filologi 86

SKRIFTER UTGIVNA AV SVENSKA
LITTERATURSÄLLSKAPET I FINLAND NR 796

STUDIER I NORDISK FILOLOGI 86

STUDIER I NORDISK FILOLOGI
UTGIVNA GENOM CAMILLA WIDE
ÅTTIOSJÄTTE BANDET

Therese Leinonen

TALET LEVER!

Fyra studier i svenskt talspråk i Finland

Svenska litteratursällskapet i Finland

Helsingfors 2015

Utgiven med stöd av Ingrid, Margit och Henrik Höijers donationsfond II
inom Svenska litteratursällskapet i Finland

© Författaren och Svenska litteratursällskapet i Finland 2015

Detta verk är licensierat under Creative Commons
Erkännande-Ickekommersiell-IngaBearbetningar 4.0 Internationell
(CC BY-NC-ND 4.0).

Redaktionens adress:
Svenska litteratursällskapet i Finland
PB 158
00171 Helsingfors

www.sls.fi

Omslag: Peter Sandberg
Typografi: Antti Pokela
Ombrytning: Mary Meinander
Typsnitt: Arno och Myriad

ISBN 978-951-583-292-4 (tryckt utgåva)
ISBN 978-951-583-446-1 (pdf), URN:NBN:fi:sls-978-951-583-446-1
ISSN 0356-0376 Studier i nordisk filologi
ISSN 0039-6842 Skrifter utgivna av Svenska litteratursällskapet i Finland
UDK 80

INNEHÅLL

Förord	9
Inledning	11
Materialet	13
Jämförelse av språkdrag i delmaterialen	16
<i>di ~ dåm</i>	16
<i>i</i> -uttal i <i>se</i> och <i>tre</i>	17
Infinitivmärket	19
Vasasvenskan mellan dialekt och standard	21
DEL I: ÅBOLÄNDSKA PÅ 2000-TALET	23
1 Introduktion	25
1.1 Åboland	26
1.2 Tidigare forskning	31
1.3 Syfte	36
1.4 Material och metoder	39
1.4.1 Informanter	41
1.4.2 Val av språkliga variabler och excerpering	42
1.4.3 Materialredovisning	44
1.4.4 Kvantitativa analysmetoder	46
1.5 Dialekternas ställning i Åboland	48
1.5.1 Dialekt och standardspråk	48
1.5.2 Stora skillnader inom Houtskär	51
2 Ljudlära	53
2.1 Vokaler	53
2.1.1 Primära diftonger	54
2.1.2 Diftongering av långa vokaler	58

2.1.3 <i>a</i> och <i>ä</i> före <i>ld</i> , <i>mb</i> , <i>nd</i> och <i>ng</i>	61
2.1.4 Bevarat <i>a</i>	62
2.1.5 Kort <i>u</i>	64
2.2 Konsonanter	66
2.2.1 <i>ng</i> -ljudet	67
2.2.2 Bevarade konsonantkombinationer: <i>ld</i> , <i>mb</i>	71
2.2.3 Inskottskonsonanter: <i>ld</i> , <i>mb</i> , <i>nd</i>	74
2.2.4 <i>dj</i> -	75
2.2.5 Initialt <i>g</i> -, <i>k</i> - och <i>sk</i> -	76
2.2.6 Norrländsk förmjukning	77
2.3 Kvantitet	79
2.3.1 Kortstavighet	80
2.3.2 Överlång stavelse	85
3 Formlära och formord	87
3.1 Substantiv	87
3.2 Adjektiv	93
3.2.1 <i>-er</i>	94
3.2.2 <i>-og</i>	96
3.3 Pronomen	97
3.3.1 <i>het</i>	97
3.3.2 <i>han</i> och <i>hon</i> som objektsformer	101
3.3.3 <i>token/tolin</i>	102
3.4 Verb	103
3.4.1 Infinitivmärket	105
3.4.2 Infinitiv	112
3.4.3 Preteritum	119
3.5 Negationen	122
4 Kvantitativ analys	125
4.1 Jämförelse av variablerna	128
4.2 Jämförelse av informanterna	131
4.3 Principalkomponentanalys	132
5 Slutdiskussion	137

**DEL II: FRÅN *DI* OCH *DEM* TILL *DÅM*
– VARIATION OCH FÖRÄNDRING I FYRA
STÄDER I SVENSKFINLAND** 145

1	Introduktion	147
2	Material	151
3	Grammatisk indelning	153
4	Sociala variabler	156
5	Språkliga variabler	157
	5.1 Grammatisk kategori och betoning	158
	5.2 Ordföljd	159
6	Intervjuarna och deras inflytande på informanterna	161
7	Slutsatser	162

**DEL III: FONETISK KVANTITET
I FINLANDSSVENSKA STADSMÅL** 165

1	Introduktion	167
2	Tidigare forskning	168
3	Material	174
4	Förhållandet V:/C	176
	4.1 Språkliga variabler	177
	4.2 Sociala variabler	178
	4.3 Taltempo	180
	4.4 Slumpintercept	182
5	Distinktionen V:C–VC:	187
6	Diskussion	190

**DEL IV: ORDKLASSTAGGNING AV
FINLANDSSVENSKT TALSPRÅK** 195

1	Introduktion	197
2	Automatisk ordklasstagning	199
3	Arbetsmoment och hjälpmedel	200
	3.1 För- och efterbehandling av materialet	201
	3.2 SUC	201
	3.3 Tagguppsättningen	203
	3.4 Automatisk taggning och handkorrigering	203

4 Taggningsresultat	203
4.1 Korsvalidering	204
4.2 Inlärningskurvan	206
4.3 Svårigheter med tagguppsättningen	209
5 Slutsatser	210
Figur- och tabellförteckning	212
Figurer	212
Tabeller	213
Referenser	214
Arkivsamlingar	214
Litteratur	214
Bilaga A: Regressionsanalyser, åboländska	224
1 Ljudlära	224
1.1 <i>ei</i> -diftongen	224
1.2 Kort <i>u</i>	225
1.3 <i>ng</i> -ljudet	226
2 Formlära och formord	227
2.1 Infinitivmärket	227
2.2 Infinitiv	228
2.3 Preteritum	229

Förord

Artiklarna i denna volym är resultatet av forskning som jag har bedrivit som anställd vid Svenska litteratursällskapet i Finland (SLS). Genom anställningen som forskare vid SLS har jag haft möjlighet att bedriva forskning på heltid vilket har varit ett stort privilegium.

Åren 2011–2012 arbetade jag inom projektet *Svenskan i Finland – i dag och i går* där mitt uppdrag var att ta fram nya forskningsresultat för volymerna I:1 *Dialekter och småstadsspråk* och I:2 *Gruppspråk, samspråk, två språk*. Det fanns helt enkelt luckor i kunskapen om svenskan i Finland i tal, och mitt uppdrag var att fylla några av dem. Under år 2011 arbetade jag med de åboländska dialekterna, vilket resulterade i del I. Under år 2012 bestod uppdraget i att beskriva variation i städerna Helsingfors, Åbo, Mariehamn och Vasa, vilket gav upphov till del II och del III. Ann-Marie Ivars, Camilla Wide och Marika Tandefelt tilldelade mig breda, konkreta forskningsuppdrag inom projektet samtidigt som de gav mig mycket fria händer i valet av analysmetoder och språkliga variabler. De har stöttat mig under arbetets gång och läst och gett värdefulla synpunkter på olika versioner av manuset till denna bok vilket jag är mycket tacksam över.

Under några månader våren 2013 var jag anställd vid Språkarkivet vid SLS för att inleda arbetet med den lingvistiska annoteringen av talspråkskorpusen Talko. Syftet var att utvärdera automatiska ordklasstaggare och påbörja ordklassuppmärkningen av korpusmaterialet. I arbetet hade jag mycket stor nytta av de erfarenheter och tekniska lösningar som forskarna vid Tekstlaboratoriet vid Universitetet i Oslo byggt upp i arbetet med Nordic Dialect Corpus och som de generöst delade med sig av. Del IV beskriver detta arbete.

De fyra delarna i boken binds samman av att talspråksmaterialet som alla fyra delar bygger på är insamlat i projektet *Spara det finlands-svenska talet* vid SLS. Jag har haft förmånen att arbeta med ett omfattande språkmaterial som bearbetats på olika sätt av flera händer före mig. På så vis kunde jag direkt sätta igång med de språkliga analyserna av materialet. Jag vill därför tacka alla dem som arbetat med ljudenliga och ortografiska transkriptioner av *Spara talet*-materialet och särskilt språkarkivarierna vid SLS som jag har haft ett givande och uppbyggande samarbete med i arbetet med de olika delstudierna. Tack Ann-Sofie Grönroos, Janina Öhman, Lisa Södergård och Pamela Gustavsson!

Jag vill dessutom tacka SLS förlag som tagit in boken i serien Studier i nordisk filologi och som bidragit till att kvaliteten på texten har höjts genom två anonyma granskares konstruktiva kommentarer och genom förlagsredaktören Jonas Lillqvists noggranna arbete.

Sist men inte minst vill jag tacka alla informanter vars röster och berättelser sparats i *Spara talet*-projektet samt de inspelningsassistenter som gjorde inspelningar på olika håll i Svenskfinland! Talet är sparat men lever vidare – både ute i samhället, där det fortsätter att talas och förändras, och i arkivet där det bearbetas och förses med värdefull tilläggsinformation och görs tillgängligt för forskare.

Therese Leinonen

Inledning

Varje levande språk utvecklas och förändras ständigt av sina talare. Så även svenskan i Finland. För att dokumentera de olika finlandssvenska talspråken i början av 2000-talet genomfördes projektet *Spara det finlandssvenska talet* (kort *Spara talet*) åren 2005–2008. I projektet, som utfördes vid Svenska litteratursällskapet i Finland, gjordes inspelade intervjuer med både äldre och yngre finlandssvenskar från alla delar av språkområdet. Sammanlagt intervjuades närmare tusen personer. Genom att materialet är insamlat under en begränsad tidsrymd och på ett enhetligt sätt i hela språkområdet ger det möjlighet till direkt jämförbara studier av talspråket i de finlandssvenska regionerna. Och det är precis det som är syftet med denna volym.

Boken omfattar fyra delstudier, som alla bygger på material samlat i *Spara talet*-projektet. Volymen inleds med en längre studie av talspråket på olika håll i den åboländska skärgården (del I), följd av studier där språket i de fyra finlandssvenska s.k. regionhuvudstäderna Helsingfors, Åbo, Mariehamn och Vasa jämförs (del II och del III). Kartan i figur 1 visar var regionerna och regionhuvudstäderna ligger. Eftersom materialen är hämtade från ett och samma insamlingsprojekt går det också att jämföra resultaten från de olika studierna med varandra i de fall då samma språkdrag har analyserats. Den sista artikeln (del IV) handlar inte så mycket om språket i sig, utan om vilka tekniska möjligheter det finns för automatisk analys av talspråk. Det visar sig att det med relativt enkla medel går att anpassa lösningar utvecklade för svenskt skriftspråk till att passa även för finlandssvenskt talspråk. På så vis kan man utnyttja resurser från större språk för att skapa forskningsinfrastruktur även för mindre språk.



Figur 1. Karta som visar de finlandssvenska regionerna med regionhuvudstäderna utmärkta.

Studierna i del I–III följer den sociolingvistiska forskningstraditionen. I denna forskningstradition, som bygger på Labovs tidiga studier från New York (Labov 1972), studeras språkdrag som uppvisar variation inom en språkgemenskap, och man analyserar hur sociala faktorer som ålder, kön och socioekonomisk status inverkar på den språkliga variationen. I de tidigaste sociolingvistiska studierna av svenskan i Finland var det Helsingforssvenskan som stod i fokus. Reuter (1982) studerade fonetisk kvantitet i Helsingforssvenskan medan Nyholm (1976 och 1984) analyserade ett flertal fonologiska och morfologiska variabler och Saari (1975) bedrev syntaktiska studier. I ett senare projekt användes sociolingvistiska metoder i studiet av språket i fyra mindre städer

(Ekenäs, Jakobstad, Kristinestad och Lovisa; Ivars 1996; Gullmets-Wik 2004). Efter de tidiga sociolingvistiska studierna har studiet av Helsingforssvenskan i stor utsträckning följt andra forskningsinriktningar (t.ex. i projektet *Svenska samtal i Helsingfors*; Lehti-Eklund 1998). Samtidigt saknas jämförbara studier av svenskan i Helsingfors och andra städer i Finland, och vi vet också att språket i Helsingfors och på andra håll har fortsatt att förändras. Artiklarna här utgör alltså ett slags *Hur gick det sen?* Artikeln om kvantitet (del III) bygger vidare på Reuters (1982) resultat, medan artikeln om *di* och *dom* (del II) behandlar ett av de drag som Nyholm (1984) studerade. Dialekterna i Åboland har tidigare beskrivits framför allt av Zilliacus (1992) som arbetade enligt traditionella dialektologiska metoder. Språket Zilliacus beskrev var de dialekter som talades under första hälften av 1900-talet. I artikeln i del I används variationslingvistiska metoder för att beskriva hur de åboländska dialekterna har utvecklats.

Flera av resultaten i denna volym beskrivs också i bokserien *Svenskan i Finland – i dag och i går* som riktar sig till en bredare läsekrets. Om de åboländska dialekterna kan man läsa i volym I:1 *Dialekter och småstadsspråk* (Ivars 2015, kap. 2 och 5) och variation i språket i regionhuvudstäderna behandlas i volym I:2 *Gruppspråk, samspråk, två språk* (Tandefelt 2015, kap. 3).

Materialet

Spara talet-projektet finns beskrivet på annat håll (se t.ex. Ivars & Södergård 2007 och 2009 samt Södergård 2009). Därför beskrivs här bara kort de delar av materialet som berör artiklarna i just denna volym.

Intervjuerna är cirka en timme långa och genomfördes i regel i informanternas hem. De spelades in med mikrofoner av typ mygga och Hi-MD-spelare.¹ I intervjuerna behandlas ämnen som hemort, skolgång, arbete och fritidsintressen. I flera fall handlar en del av intervjun också om attityder till eget språkbruk och andras och om den egna dialekten. I de flesta fall skedde intervjuerna mellan en intervjuare och en informant, men i vissa fall har två informanter intervjuats till-

1. Spelaren var inställd på 44,1 kHz samplingsfrekvens och 16 bitars bitdjup. Inspelningarna är arkiverade som BWF-filer.

sammans. Intervjuerna är i de allra flesta fall uppvuxna på den ort där intervjuerna gjordes och talar det lokala språket.

Vid Språkarkivet vid SLS arbetar man med utvecklandet av den finlandssvenska talspråkskorpusen Talko som ska göra talspråksmaterial från hela det svenska språkområdet i Finland tillgängligt för språkvetenskaplig forskning. Material från *Spara talet*-projektet utgör stommen i korpusen. I Talko kommer omkring en tredjedel av de informanter som spelades in inom projektet att ingå. Ur de cirka en timme långa intervjuerna har ett tjugo minuter långt klipp valts ut. Klippet har transkriberats med en grov ljudskrift och transkriptionen har därefter ”översatts” till standardsvensk ortografi. Dialekttranskriptionen har alltså inte översatts till ren standardsvenska, utan det handlar om en ord-för-ord-översättning eller återgivning med standardiserad ortografi. Materialet är transkriberat² vid Språkarkivet och de standardiserade transkriptionerna har gjorts halvautomatiskt med hjälp av en dialekttranslitterator utvecklad vid Tekstlaboratoriet vid Universitetet i Oslo. Den ljudenliga och den standardiserade transkriptionen är genom tidsstämpling länkade till själva ljudfilen.

Dessa transkriberade klipp ur intervjuerna utgör materialet för samtliga delstudier i denna bok. Den grova ljudskrift som används här är också densamma som används i Talko (för en närmare beskrivning se avsnitt 1.4.3 i del I). Att materialet föreligger i transkriberad form har underlättat arbetet väsentligt. För att hitta de språkdrag man vill analysera går det att göra sökningar i transkriptionerna, och genom tidsstämplingen är det enkelt att gå in och lyssna på uttalet av ett specifikt ord i ljudfilerna eller göra ytterligare annoteringar som länkas till ljudfilen.

Informanterna från Åboland beskrivs i del I, men eftersom informanterna i artiklarna II och III är desamma beskrivs de här. I materialet från de fyra städerna finns sammanlagt 47 informanter fördelade på två åldersgrupper (se tabell 1). Antalet informanter är beroende av antalet invånare i staden i fråga, vilket ger flest informanter från Helsingfors och minst från Mariehamn.

2. I programmet Transcriber (Barras et al. 2000).

Tabell 1. Antal informanter per stad, åldersgrupp och kön.

STAD	ÄLDRE		YNGRE		TOTALT
	KVINNOR	MÄN	KVINNOR	MÄN	
Helsingfors	4	4	4	5	17
Mariehamn	1	1	2	2	6
Vasa	3	3	2	2	10
Åbo	3	3	4	4	14
Totalt	11	11	12	13	47
	22		25		

Tabell 2. Åldersfördelningen bland informanterna. I tabellen anges det lägsta och det högsta födelseåret i varje informantgrupp, samt medianvärdet.

STAD	ÄLDRE			YNGRE		
	MIN.	MEDIAN	MAX.	MIN.	MEDIAN	MAX.
Helsingfors	1921	1945	1950	1974	1980	1985
Mariehamn	1928	1933	1937	1980	1985	1986
Vasa	1940	1944	1953	1980	1982	1985
Åbo	1938	1944	1952	1984	1986	1990

Informanternas genomsnittliga ålder redovisas i tabell 2. De flesta av de äldre informanterna är födda på 1940-talet medan största delen av de yngre informanterna är födda på 1980-talet. De äldre informanterna från Mariehamn är något äldre än medeltalet och av Helsingforsinformanterna är tre stycken födda på 1920-talet. Bland de yngre informanterna är de från Åbo något yngre än genomsnittet och Helsingforsinformanterna något äldre.

För identifiering av informanterna i texten används signum. Dessa är samma signum som används i korpusen Talko, vilket gör att läsaren kan använda korpusen för att lyssna på hur informanterna låter.³ Bara

3. I skrivande stund har endast pilotversionen Talko 0.1 publicerats på nätet (www.sls.fi/talko), och alla de informanter som ingår i analyserna i denna volym finns ännu inte med i nätversionen av Talko.

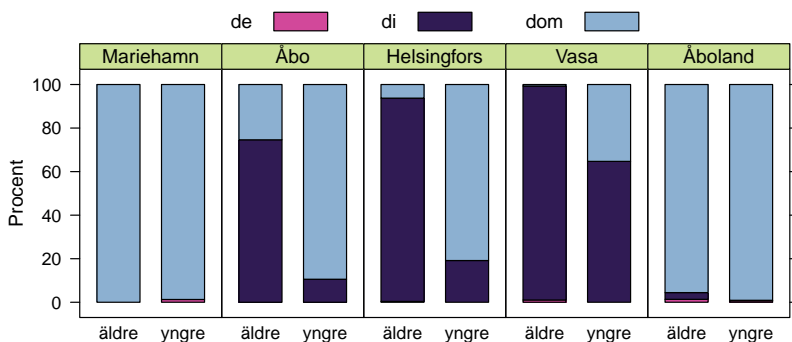
ortnamnsdelen av signumerna har förkortats och består här av endast två bokstäver i stället för hela ortnamnet. I övrigt består signumerna av två bokstäver som anger åldersgrupp och kön samt ett nummer. I signumerna anger *o* äldre och *y* yngre, medan *w* står för kvinna och *m* för man. De två yngre kvinnorna från Vasa har t.ex. signumerna Va.yw10 och Va.yw11.⁴

Jämförelse av språkdrag i delmaterialen

Eftersom hela *Spara talet*-materialet är insamlat på samma sätt och vid ungefär samma tidpunkt är resultaten från analyser av de olika delmaterialen direkt jämförbara med varandra. Nedan följer några exempel.

di ~ dām

I talspråket i Sverige används nästan uteslutande formen *dom* för skriftens *de* och *dem*. Skriftens skillnad mellan *de* som subjektform och *dem* som objektform för tredje personens personliga pronomen i plural saknas alltså i talspråket. I standardfinlandssvenskan har man såsom i skriften gjort skillnad mellan de båda formerna: *de* uttalas *di* och *dem* har även i talspråket formen *dem*.



Figur 2. Stapeldiagram över *di ~ dom*.

- Eftersom materialsamlingen som helhet har fler informanter från varje ort än de som ingår i Talko, förekommer det att informantnumren är högre än antalet informanter i gruppen.

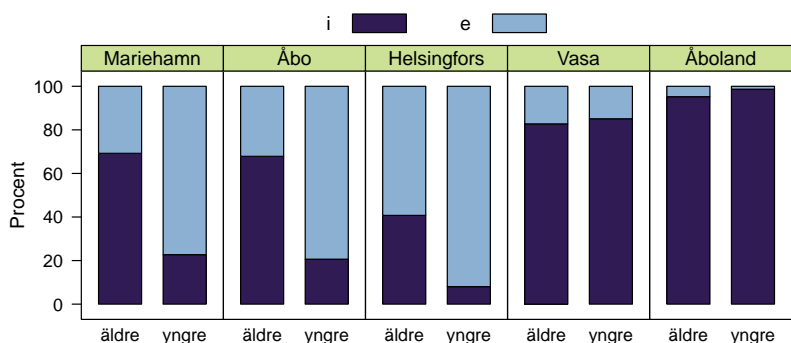
I *Spara talet*-materialet uppvisar variabeln *di* ~ *dâm* stor variation både mellan och inom städerna Helsingfors, Åbo, Mariehamn och Vasa, vilket beskrivs i del II. I Mariehamn är *dâm* den så gott som allenaformande formen. I de övriga tre städerna använder unga personer *dâm* i större utsträckning än äldre personer. I Vasa är *di* vanligare än i södra Finland, vilket kan bero på att staden Vasa omges av landsbygd där dialekterna traditionellt upprätthållit skillnaden mellan subjektform och objektform.

Variabeln *di* ~ *dâm* behandlas inte specifikt för de åboländska dialekternas del i del I, eftersom informanterna inte uppvisar någon större variation. Ser man på variabeln i det åboländska materialet kan man konstatera att alla informanter till övervägande del använder formen *dâm* (eller i västligaste Åboland ibland varianten *tâm*). Detta är också den form som traditionellt finns belagd i dialekterna (se kartorna i figur II.1 på s. 149). Formen *di* förekommer bara i enstaka fall i Åboland (figur 2), och då oftast hos någon av de äldre talarna. Dessa belägg på *di* kan ses som inflytande från det finlandssvenska standardspråket. Men de ytterst få beläggen på *di* och det faktum att formen håller på att minska i stadsspråken tyder på att inflytandet i fråga om denna variabel inte är av någon bestående natur, utan det är formen *dâm* som är den livskraftiga.

i-uttal i *se* och *tre*

En annan variabel, som inte behandlas närmare i någon av de längre artiklarna i denna volym, men där man kan jämföra i de olika materialen, är uttalet av orden *se* och *tre*. Uttal med *i*-vokal, dvs. *si:* och *tri:*, har varit det normala uttalet både i finlandssvenska dialekter och i det finlandssvenska standardspråket (Ahlbäck 1956, 37). Ännu år 2003 konstaterade Ivars (2003, 77) att *i*-formerna av *se* och *tre* hörde till de livskraftiga dragen i svenskan i Finland. Tidigare var detta uttal vanligt också i Sverige både i dialekter och i stadsspråk.

I det åboländska materialet är *i*-uttalet det klart dominerande. I stadsmaterialen kan man dock iaktta en förändring. Figur 3 visar att *i*-uttal används i alla fyra städer och i bägge åldersgrupperna, men *i*-uttalet är ovanligare bland yngre än bland äldre personer i alla städer



Figur 3. Stapeldiagram över andelen *i*-uttal i orden *se* och *tre*.

utom i Vasa.⁵ I Vasa är andelen *i*-uttal överlag högst av städerna, men inte lika hög som i Åboland. I Vasa råder heller ingen egentlig skillnad mellan äldre och yngre personer om man ser på grupperna som helhet. En av informanterna från Vasa saknar dock helt *i*-former – en äldre man med hög utbildningsnivå. Denna informant saknar även andra dialektala drag i uttalet (*se* nedan). Samvariationen mellan *i*-uttal av *se* och *tre* och andra mer utpräglade dialektala drag kan kanske tolkas som att *i*-uttal har kommit att uppfattas som ett dialektdrag.

Ser man närmare på variationen, finns det en skillnad mellan de båda orden *se* och *tre*. *Ser du* används ofta som diskursmarkör och har då oftast uttalet *sidu*. Men även om man bortser från beläggen på *ser du* som diskursmarkör är *i*-uttalet vanligare i olika former av *se* än i räkneordet *tre*. En möjlig orsak kan vara att räkneordet ofta används i formellare sammanhang, medan verbet *se* är mycket frekvent och kan användas både bokstavligt och bildligt.

Tidigare har *i*-formerna av *se* och *tre* varit ett karakteristiskt drag för svenskan i Finland generellt. I det undersökta materialet tyder variationen dels mellan städer och skärgård i södra Finland, dels mellan olika informanter inom Vasa stad på att *i*-formerna har utvecklats till ett dialektalt drag i stället för ett allmänfinlandssvenskt drag. Om detta bara är ett övergångsskede eller om förändringen så småningom kommer att spridas till alla varieteter av finlandssvenska återstår att se.

5. Varje informant har i medeltal sju belägg på variabeln.

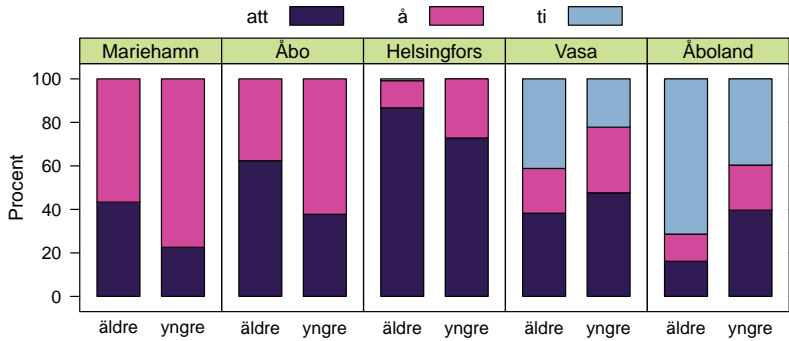
Infinitivmärket

Infinitivmärkets olika former i Åboland beskrivs i avsnitt 3.4.1 i del I. De tre former som förekommer är *ti*, *att* och *å*. Formen *ti* har traditionellt setts som den dialektala varianten, men andelen *ti* ser ut att minska, särskilt i centrala Åboland.

Enligt Nyholm (1984, 62–64) förekom alla tre varianterna av infinitivmärket i Helsingforssvenskan på 1970-talet. Formen *ti* var den typiskt lågspråkliga varianten som främst fanns hos inflyttade helsingforsare. I övrigt användes *att* och *å* sida vid sida, men enligt Nyholm visade variationen mellan dessa två varianter inga tydliga mönster, varken sociala eller kontextbundna. Variabeln var klart polariserad så att informanterna tenderade att favorisera en av varianterna. Så många som hälften av Nyholms informanter använde uteslutande en av de tre varianterna. I *Spara talet*-materialet är variabeln inte lika polariserad. Bara två av de sju informanterna från Helsingfors använder endast en av varianterna. Dessa två använder varianten *att*. Majoriteten av Helsingforsinformanterna använder *att* i större utsträckning än *å*. Bara bland de yngre talarna finns det enstaka personer som verkar föredra *å*. Jämfört med Nyholms material verkar användningen av *att* att ha ökat. Figur 4 visar den relativa användningen av olika varianter efter ort och åldersgrupp.⁶ Andelen *att* är klart högst i Helsingfors. Det är anmärkningsvärt att andelen *att* är lägre bland yngre än bland äldre talare i alla tre sydfinländska städer. I Åbo är variabeln något mer polariserad än i Helsingfors: två av de äldre informanterna använder uteslutande *att* medan en av de yngre uteslutande använder *å*. Eventuellt är variabeln en indikator för graden av formalitet. Man kan t.ex. konstatera att de tre informanter som i kvantitetsanalysen (del III, avsnitt 4.4) visade sig använda det uttal som åtminstone tidigare setts som ett prestigeuttal (Hf.owo₁, Ab.omo₄ och Hf.ymo₁) också uteslutande eller i mycket stor utsträckning använder *att*. Skriftspråksformen *att* verkar alltså även i talspråket ses som standardvarianten.

I det åboländska materialet visade det sig att kontexten verkar ha betydelse för valet av infinitivmärke. En sådan tendens finns också i stadsmaterialet. Analysen försvåras av att det finns tre varianter, men

6. Medelantalet belägg per informant i stadsmaterialet är 14.



Figur 4. Stapeldiagram över infinitivmärket.

om man ser bara på de sydfinländska städerna där *ti* saknas ser man en tendens. När talare som annars föredrar *å* använder *att* är det ofta i en kontext där infinitivmärket föregås av en preposition. Några exempel:

INFORMANT	å	att
Ab.ow01	<i>de va ju bara å velja</i>	<i>de uppmuntra ti att tenkka</i>
Ab.ym12	<i>de kan va li:te jäbbit å le:sa</i>	<i>int ärkkar man rikit fäkuse:ra på att le:sa</i>
Mh.yw08	<i>ja jillar å lä:ra mej</i>	<i>de finns ju en påeng me: att jö:ra de</i>

Kontexten verkar alltså ha betydelse för valet av infinitivmärke. För en mer precis analys av kontextens betydelse och för att hitta flera kontextbundna faktorer skulle man ändå behöva veta exakt vilken typ av konstruktion infinitivfrasen ingår i och inte bara se på ordet innan. Det skulle också behövas ett större material eftersom infinitivmärket inte förekommer särskilt frekvent i löpande tal och antalet olika konstruktionstyper som man behöver kunna beakta är många. En sådan analys kan bli möjlig när *Spara talet*-databasen är fullständigt transkriberad och uppmärkt (jfr del IV).

Vasasvenskan mellan dialekt och standard

I stadsmaterialet förekommer *ti* relativt frekvent i Vasa, men saknas nästan helt i de övriga städerna.⁷ Användningen av *ti* i Vasa kräver en närmare förklaring. I analyser av språket i Vasa har Rabb (2014 och 2015) visat att det finns stora språkliga skillnader inom Vasa och att variationen till en del är geografiskt och åldersmässigt betingad, men att individuella faktorer också spelar roll. De mest utpräglat dialektala Vasaborna använder enligt Rabb i stor utsträckning dialektdrag som apokopering, *o*-vokal i stället för kort *u* och *ti* som infinitivmärke. Även i *Spara talet*-materialet kan man se stora skillnader mellan informanterna. Av de tre äldre männen använder en uteslutande *ti* som infinitivmärke, medan en annan helt saknar belägg på *ti* och i mer än 90 procent av beläggen använder *att*.

Variationen i användningen av infinitivmärket kan kopplas ihop med variation i fråga om andra språkdrag. Jag har även excerperat variabeln *det* ~ *het* (dvs. standarduttalet *de* eller *det* versus det dialektala uttalet *he* eller *e*) i Vasamaterialet (jfr avsnitt 3.3.1 i Åbolandsundersökningen, del I). Ser man på bägge variablerna tillsammans kan man urskilja ett mönster. Bland de äldre informanterna saknar de två informanterna med den högsta utbildningsnivån (en kvinna och en man) helt belägg på både *ti* och *he*. Dessa båda informanter föredrar också *e*-uttal av *se* och *tre*. De övriga äldre informanterna använder de dialektala varianterna i olika utsträckning, men den mest dialektala talaren saknar nästan helt och hållet standardvarianter av dessa variabler. Bland de yngre informanterna saknar de båda kvinnorna nästan helt både *ti* och *het*, medan de båda männen använder dem i olika utsträckning.

Bredden i fråga om dialektala drag är alltså klart större i Vasa än i de tre sydfinländska städerna. De olika dialektdragen samvarierar och informanterna befinner sig på ett kontinuum mellan dialekt och standardspråk i fråga om flera olika dialektdrag samtidigt. Enligt Rabb (2015) är utbildningsnivå och bostadsort inom Vasa (som i sig samvarierar) de mest avgörande faktorerna för den språkliga variationen i den äldsta generationen av Vasabor, medan andra faktorer är mer avgörande bland yngre talare. Skillnaderna i de yngre generationerna

7. Från Helsingfors finns ett belägg hos en äldre man.

är mindre polariserade än i den äldsta generationen. Förutom att vissa utpräglat dialektala drag minskar verkar samtidigt en destandardisering ske, dvs. vissa talspråkliga drag används i större utsträckning av yngre talare än av äldre. Samma tendens kan skönjas i *Spara talet*-materialet även om både antalet informanter och antalet språkdrag som analyserats är relativt litet. Det är också tydligt att Vasa som språkmiljö klart skiljer sig från Helsingfors, Åbo och Mariehamn där utpräglat dialektala drag så gott som helt lyser med sin frånvaro.

Del I

ÅBOLÄNDSKA PÅ 2000-TALET

1

Introduktion

I denna studie presenteras resultat av forskning inom projektet *Svenskan i Finland – i dag och i går* vid Svenska litteratursällskapet i Finland. Forskningen har genomförts som en ettårig punktinsats inom delprojektet *Svenskan i Finland i tal*. Målet med uppsatsen är att öka kunskapen om de åboländska dialekterna så som de talas i början av 2000-talet.

Tack vare främst Kurt Zilliacus forskningsinsatser (se avsnitt 1.2) är de åboländska dialekternas särdrag fram till mitten av 1900-talet väl dokumenterade. Däremot existerar bara några få, geografiskt begränsade undersökningar som beskriver åboländska dialekter under slutet av 1900-talet, och dialektutvecklingen fram till i dag är således till stora delar okänd.

Under senare delen av 1900-talet har de svenska dialekterna drabbats av långtgående utjämning och regionalisering. I Finland har utjämningsstendensen varit starkare i södra Finland än i Österbotten (Ivars 2003). Homogeniseringen av dialekterna leder dels till att dialekterna närmar sig standardspråket, dels till att nya regionala varieteter uppstår när de mest lokala dialekt dragen försvinner och dialekt drag som är gemensamma för större geografiska områden bevaras. Ett mål för denna studie är att beskriva i vilken utsträckning de åboländska dialekterna har utjämnats under senare delen 1900-talet. Vilka språkdrag bevaras och vilka går förlorade? Av intresse är också frågan om de lokala dialekterna har ersatts eller håller på att ersättas av ett åboländskt regionalspråk.

För att besvara dessa frågor har material insamlat i projektet *Spara det finlandssvenska talet* analyserats. Eftersom delar av materialet föreligger i färdigt transkriberad form har det varit möjligt att relativt snabbt

excerpera många språkdrag ur inspelningarna och på så vis inom ett relativt kort projekt ge en övergripande bild av dialektsituationen i Åboland i början av 2000-talet. Det bör dock påpekas att denna undersökning är en ren ”skrivbordsprodukt” eftersom det inte har varit möjligt att, på samma vis som t.ex. Janne Thurman och Kurt Zilliacus (se avsnitt 1.2), vistas längre tider på ort och ställe och göra djupgående iakttagelser av det naturliga språkbruket.

Analysen omfattar dialektdrag främst inom fonologin och morfologin, och resultatredovisningen är på klassiskt vis indelad i ljudlära (kapitel 2) och formlära (kapitel 3). Detta innebär inte att en fullständig beskrivning ges av dialekternas fonologi och morfologi, utan analysen omfattar ett axplock av drag inom ljudläran och formläran. Avgörande vid valet av dialektdrag har varit dragens variabilitet och frekvens i språket. Eftersom materialet omfattar endast tjugo minuter inspelat tal per informant har endast drag som förekommer så frekvent i språket att ett tillräckligt antal belägg produceras under tjugo minuter kunnat inkluderas. I kapitlet om formläran tas förutom rena morfologiska drag även vissa formord (som infinitivmärke och negation) upp.

I avsnitt 1.1 nedan ges en kort beskrivning av Åboland och dess befolkning och i 1.2 ges en översikt över tidigare forskning om åboländska dialekter. Syftet och frågeställningarna för undersökningen preciseras i avsnitt 1.3, och materialet och metoderna beskrivs närmare i avsnitt 1.4. I avsnitt 1.5 ges en allmän bild av dialekternas ställning i området.

Efter redovisningen av de enskilda språkdragen i kapitel 2 och 3 förenas alla kvantifierbara variabler i en sammanlagd kvantitativ analys i kapitel 4. I kapitel 5 diskuteras och sammanfattas resultaten av undersökningen.

1.1 Åboland

De svenskspråkiga skärgårdsområdena i landskapet Egentliga Finland brukar kallas Åboland – ett namn som togs i bruk så sent som på 1910-talet (Zilliacus 1990, 11). Fram till dess kallades området ”Finska skären”, som härrör från den tid då namnet Finland betecknade området för dagens Egentliga Finland. Namnet Åboland slog snabbt igenom och i dag uppfattas Åboland ”som ett särskilt (inofficiellt) landskap” (UVF).



Figur I.1. Karta över Åboland. Åbo är ekonomisk och kulturell centralort för området.

Officiell status som administrativ enhet fick Åboland först 1994 då ekonomiska regioner infördes i Finland. De två åboländska kommunerna Pargas och Kimitoön bildar i dag tillsammans den ekonomiska regionen Åboland. Då indelningen i ekonomiska regioner infördes beaktades kriterier som interkommunalt samarbete, trafikförbindelser, pendling och kundströmmar, vilket visar att Åboland var en funktionell helhet redan innan det fick status som administrativt område 1994 (UVF).

Inom dialektforskningen brukar man i stället för aktuella kommungränser följa den gamla sockenindelningen. I Åboland har socknarna, som finns utprickade på kartan i figur I.1, varit tio till antalet: Iniö, Houtskär, Korpo, Nagu, Pargas, Kimito, Dragsfjärd, Västansfjärd, Hitis och Finby. Hitis sammanslogs med Dragsfjärds kommun redan

1969. Vid den senaste kommunsammanslagningen 2009 bildade Iniö, Houtskär, Korpo, Nagu och Pargas staden Väståboland, som 2012 bytte namn till Pargas. Kommunen Kimitoön bildades av Kimito, Dragsfjärd och Västanfjärd.

Finby slogs vid kommunsammanslagningen 2009 ihop med staden Salo och hör därmed inte till den ekonomiska regionen Åboland. Före kommunsammanslagningen hade Finby 741 invånare varav 10,5 procent svenskspråkiga (31/12 2005, *Statistisk årsbok för Finland 2007*). Eftersom inget material från Finby ingår i den talspråkskorpus som ligger till grund för föreliggande arbete (se avsnitt 1.4), behandlas svenskan i Finby inte i denna undersökning.

Antalet invånare i de olika kommundelarna före kommunsammanslagningen 2009 samt näringsstrukturen framgår av tabell I.1. Servicenäringarna dominerar i hela Åboland. Jord- och skogsbruk sysselsätter endast en liten andel av befolkningen, med de högsta andelarna i Houtskär och Nagu (drygt 16 procent). Industrins roll är relativt sett störst i Dragsfjärd och Pargas. Som jämförelse kan nämnas att jord- och skogsbruket i början av 1960-talet var den dominerande näringen i hela Åboland utom i Dragsfjärd och Pargas där industrin sysselsatte den största andelen av befolkningen. Kolumnen pendling i tabell I.1 anger andelen av den sysselsatta befolkningen som arbetar utanför hemkommunen (2005 års kommungränser). Siffrorna visar att den dagliga rörligheten är stor.

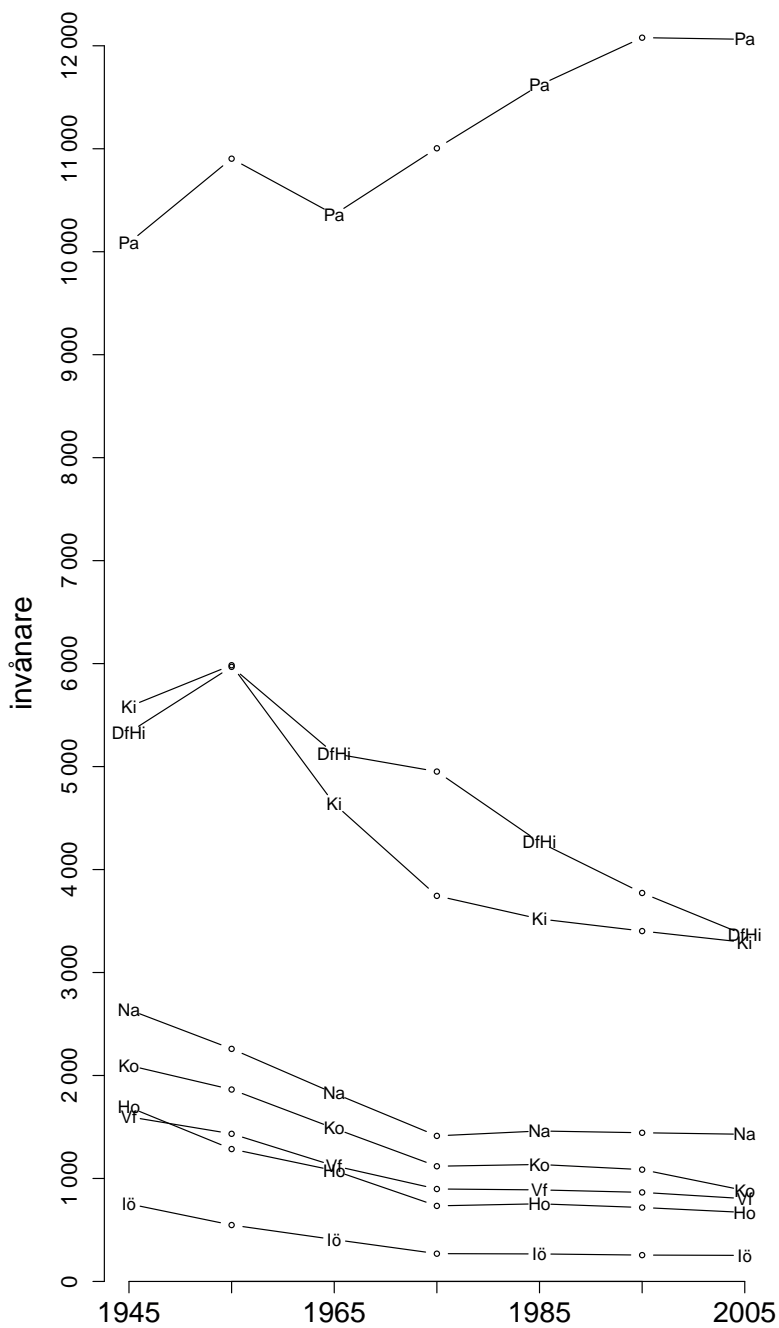
Befolkningsutvecklingen under senare delen av 1900-talet framgår av figur I.2. Antalet invånare har sjunkit drastiskt överallt utom i Pargas, som har haft en positiv befolkningsutveckling. Pargas är den överlägset största av de tidigare åboländska kommunerna och den enda som var stad före 2009. Tätortsgraden, som avser andelen personer som är bosatta i tätorter, är högst i Pargas, medan Iniö, Houtskär och Västanfjärd utgör ren glesbygd. Den svenskspråkiga delen av befolkningen är i knapp majoritet i Pargas, medan andelen svenskspråkiga är betydligt större på övriga håll i Åboland. Andelen svenskspråkiga har under senare hälften av 1900-talet sjunkit överallt i Åboland.

Befolkningstätheten är låg i Åboland, men ökar betydligt under sommarmånaderna. År 2010 hade Väståboland flest fritidshus av alla kommuner i Finland, medan Kimitoön kom på tionde plats (FOS: *Byggnader och fritidshus 2010*). Det totala antalet sommarstugor i Åbo-

Tabell 1.1. Befolkning och näringsstruktur i Åboland år 2005.

ORT	BEFOLKNING		TÄRTORTSGRAD (%)	NÄRINGSSTRUKTUR			PENDLING (%)
	ANTAL	ANDEL (%) SVENSKSPRÅKIGA		JORD- OCH SKOGBRUK, FISKE OCH JAKT (%)	INDUSTRI (%)	TJÄNSTER (%)	
Iniö	253	71,5	-	12,3	10,5	74,6	40,4
Houtskär	670	87,5	-	16,5	7,4	72,8	30,0
Korpo	883	74,1	37,2	8,1	7,6	82,0	44,1
Nagu	1 429	71,4	35,3	16,3	15,7	64,7	34,8
Pargas	12 063	54,2	78,6	1,8	29,5	66,8	38,3
Kimito	3 293	64,5	41,9	10,2	20,9	65,9	36,2
Västanfjärd	804	87,9	-	10,1	23,7	62,7	48,0
Dragsfjärd med Hitis	3 365	75,9	65,2	4,7	33,9	58,7	24,3

(Källa: Statistisk årsbok för Finland 2006 & 2007)



Figur I.2. Antal invånare i de åboländska kommunerna åren 1945–2005.

land var närmare 13 000 år 2010, medan antalet år 1970 var bara knappa 5 000 (StatFin: *Asuminen* > *Kesämökki*). Under 2000-talet byggdes cirka 60 nya fritidshus per år i Väståboland. Trafiken mellan öarna i området sköts av färjor och förbindelsefartyg.

1.2 Tidigare forskning

En av de tidigaste beskrivningarna som behandlar åboländska dialekter är *Anteckningar om Korpo och Houtskärs socknar* av L. W. Fagerlund (1878) som innehåller ett avsnitt om ”folkspråk”. Ett tidigt insamlingsarbete i Åboland utfördes av Janne Thurman som i slutet av 1800-talet samlade in material i Korpo, Nagu och Pargas och som gav ut monografierna *Pargasmålet* (1898) och *Nagumålet* (1929). En släkting till en av Thurmans informanter deltog i Folkkultursarkivets pristävling *Vad talar du? Dialekt – när, var, hur?* år 1992. Hon skriver om Thurman: ”Han tillbringade flera somrar på Åntala (Ontala) Östergård, där han intervjuade min mormors farfar [...] Han frågade alla på gården och var ivrig att få med de allra äldsta benämningarna” (SLS 1746: 492).

En mycket stor insats för insamlandet och beskrivningen av åboländska dialekter har utförts av Kurt Zilliacus. Största delen av hans insamlingsarbete skedde under 1950-talet med kulmen i den stor-samling från Houtskär (SLS 701–702) som han samlade in under en nästan oavbruten vistelse i Houtskär från sommaren 1957 till hösten 1958. Genom att hans uppteckningar bygger på långvariga vistelser i Åboland, där han umgicks med åbolänningar och på så vis ingående lärde känna deras naturliga språkbruk, är forskningsinsatsen ovärderlig. Zilliacus iakttagelser av de åboländska dialekterna finns samlade i boken *Åboländska* (1992). I boken ger Zilliacus en mycket knapphändig beskrivning (s. 8–9) av det material som ligger till grund för hans dialektbeskrivningar. Eftersom dialektsituationen i Åboland på 2000-talet här kommer att jämföras med den situation Zilliacus beskriver, är det viktigt att känna till bakgrunden för Zilliacus arbete. Under det mer än halva sekel som förflutit sedan Zilliacus samlade in material i Åboland har både dialektsynen och insamlings- och analysmetoderna ändrats. För att materialen ska gå att jämföra ges här en lite närmare beskrivning av Zilliacus arbete utifrån hans anteckningar i samlingarna i SLS arkiv.

Målet för Zilliacus insamlingsarbete var att teckna upp den ”ge-
nuina” dialekten. Därmed avsåg han

den dialekt som förefaller att ha talats allmänt och utan nämnvärda variationer i flertalet av de anlitade sagesmännens ungdom i början av 1900-talet, och som kvarlever i rätt oförvanskad form i de dialektkunnigaste ortsbornas vardagstal. Denna dialekt har jag uppfattat som en genom sekler i huvudsak naturligt och under normal yttre påverkan utvecklad språkart. (Ur ”Inledning och kommentar till SLS 701–702”, s. 2.)

Vad som av Zilliacus räknas som ”normal” yttre påverkan och vad som är ”onormalt” är inte helt klart, men kan anas i vissa kommentarer i de ”Rapporter till arkivet” (dvs. Folkkultursarkivet i Helsingfors) som Zilliacus skrev under sin vistelse i Houtskär: ”Ofta resultatlösa utfärder – amerikavistelser och sjömansliv har fördärvat en hel del av karlfolket.” Han menar att det är kvinnorna som är ”avgjort kunnigare i målet”. Detta är en intressant iakttagelse, eftersom män snarare än kvinnor inom dialektologin har ansetts vara uppbärare av den traditionella dialekten. Den klassiska dialektinformanten är en s.k. NORM (eng. *nonmobile, older, rural male*; Chambers & Trudgill 1998, 29). Men i Houtskär var männen alltså geografiskt mer rörliga, medan kvinnorna oftare bodde kvar i Houtskär och på så vis bevarade dialekten.

Det framgår att Zilliacus var medveten om den stora variation som fanns bland dialekttalarna, men att hans mål var att samla in och beskriva en så ren dialekt som möjligt, dvs. själva språkssystemet utan att beakta eventuell variation inom systemet. Zilliacus skriver: ”Den i och för sig givetvis intressanta undersökning, som kunde göras av dialektens upplösning och de eventuella lagarna för därvid skeende ljudövergångar och förändringar i böjningssystemen m.m., synes mig helst böra göras separat.” Zilliacus gör alltså en klar avgränsning i sitt arbete och inser att en annan typ av metoder behövs för att beskriva pågående förändringar i dialekterna. Dessa tankar är nedtecknade på 1950-talet och man kan säga att Zilliacus förutspådde den riktning dialektforskningen skulle komma att ta. Först på 1960-talet introducerades sociolingvistikens med dess variationslingvistiska metoder som möjliggör en systematisk analys av den typen av språkförändringar som Zilliacus beskriver.

Den dialekt som Zilliacus beskriver är alltså delvis ett äldre och idealiserat stadium av dialekten. I inledningen till samling SLS 673 från Iniö skriver han: "Anteckningarna visar huvudsakligen äldre dialekt, som avviker från det dagliga språket av idag både i fråga om ordförråd och uttal och håller på att falla i glömska." Och om samling SLS 663 från Hitis skriver han att den "utgöres huvudsakligen av äldre personers hågkomster av hur man talade i deras barn- och ungdom. Såväl i fråga om ordförrådet som i synnerhet när det gäller det genuina uttalet är *skäriskan* i Hitis ett halvdött språk."

Dialektforskningen i allmänhet har karakteriserats just precis av att de äldsta och minst mobila talarna oftast har valts ut som informanter, och dialektforskningen har därigenom ibland fått stämpeln som en slags "lingvistisk arkeologi" (Chambers & Trudgill 1998, 29–30). Det är talande att Zilliacus på 1950-talet försökte rekonstruera den dialekt som talats i början av 1900-talet, medan Thurman redan på 1890-talet var mån om att anteckna de äldsta benämningarna. Först med sociolingvistikens genombrott på 1960-talet började forskare intressera sig för variationen i språkssystem och inom talgemenskaper. Dagens forskningssyn går ut på att ett stabilt språkssystem knappast existerar, utan språkkontakt och dialektkontakt har funnits i större eller mindre utsträckning i alla tider och alla levande språk innehåller ett visst mått av variation. Zilliacus beskrivningar av de åboländska dialekterna bygger på en annan dialektsyn och forskningsinriktning än dialektforskningen har i dag, vilket man bör hålla i minnet när jämförelser görs med dialektmaterial insamlat och analyserat enligt dagens metoder.

Efter Zilliacus (1992) har ingen samlad bild getts av dialekterna i Åboland, men några punktundersökningar finns. Bergström (1984) har beskrivit fonologin och morfologin i Mattkärrdialekten (Kimito) i början av 1980-talet i en strukturalistisk avhandling. Karlsson (2006) har gjort en fallstudie om dialektutjämning i Nagu. Karlssons studie, som i likhet med Bergströms behandlar dialektens fonologi och morfologi, bygger på tre informanter ur tre olika generationer i en familj, födda under 1900-talet. En av hennes slutsatser är att variationen mellan dialektvarianter och standardvarianter är störst hos informanten född 1936. Den äldsta informanten (f. 1903) använder nästan genomgående dialektvarianter av de undersökta dragen, medan den yngsta informanten (f. 1977) bara sporadiskt använder dialektvarianter av de mest

markerade dialekt dragen och alltså i stor utsträckning har gått över till standardsvenska.

Flera forskare har intresserat sig för dialekternas fonologi och morfologi, men syntax och prosodi har blivit mindre ingående behandlade. För de åboländska dialekternas syntax hänvisas till Ivars (2015).

Den åboländska dialekt vars prosodi har blivit mest studerad är Kimitodialekten, vars tonala drag har analyserats akustiskt av Selenius (1978), Kim (2006) och Aho (2010). Vidare ingår Pargas i Kims och Ahos undersökningar, och i Ahos undersökning ingår även informanter från Västanfjärd och Nagu. Enligt dessa forskare skiljer sig intonationen i de åboländska dialekterna från Helsingforssvenska och mellannyländska och östnyländska dialekter genom ett stigande tonförlopp i den betonade stavelsen i fokuserade ord. I mellersta och östra Nyland är tonen fallande, vilket är karakteristiskt även för den finska prosodin. Selenius (1978) och Aho (2010) menar att det finns rester av ett tvåaccentsystem i Kimitodialekten, men att accentskillnaden inte längre är fonematisk (såsom den är t.ex. i västra Nyland). Enligt Aho (2010) gäller den slumpmässiga fördelningen av ordaccenter förutom för Kimitodialekten även dialekterna i Pargas och Västanfjärd. I Nagu är fördelningen mellan accent I och accent II enligt Aho mer regelbunden än i de öståboländska dialekterna, men huruvida det är fråga om en betydelseskiljande ordaccent är osäkert. Akustiska undersökningar över prosodin i de västligaste och sydligaste dialekterna i Åboland (Iniö, Houtskär, Korpo, Hitis) saknas, men kunde bidra med intressanta resultat. Selenius (1978) har påpekat att Hitisdialekten kan ha bevarat mer spår av ett tvåaccentsystem än Kimitodialekten, och Ahos (2010) resultat tyder på att accentsystemet är bättre bevarat i västra än i östra Åboland. Zilliacus (1992, 24) nämner att det finns spår av accent II i Hitis och i utskärsbyarna i Korpo och Nagu. I nordvästra Åboland har Zilliacus dessutom iakttagit spår av den s.k. jämviktsaccenten som har varit kännetecknande för kortstaviga ord.

Dialektindelning

Den förste som försökte systematisera förhållandena mellan de finlandsvenska dialekterna, och därmed också de åboländska, var Hultman (1894) som delade in de finlandssvenska dialekterna i en nordfinländsk

och en sydfinländsk grupp. Gränsen mellan de två dialektområdena gick enligt Hultman (1894, 168) rakt igenom Åboland mellan Houtskär och Korpo. Hans indelning har senare kritiserats av flera forskare. Ahlbäck (1956, 15–16) menar att de karakteristiska säregenheterna för de två grupperna är få och arkaistiska, och även Harling-Kranck (1981) anser att indelningen bygger på mycket få språkdrag och dessutom på drag som sällan gäller alla dialekter i de båda grupperna. Naert (1995, 17) menar att Hultman hade för lite material från de åboländska dialekterna för att kunna göra en tillförlitlig indelning, och att den gräns som han drog genom Åboland visserligen är en intern dialektgräns i landskapet men inte kan användas för att på lingvistisk grund dela in hela det svenska språkområdet i Finland i en nordlig och en sydlig dialektgrupp.

Trots att indelningen i ett nordfinländskt och ett sydfinländskt dialektområde har kritiserats, är forskare överens om att flera dialektskillnader möts just i Åboland som är ett övergångsområde där österbottniska, åländska och nyländska dialekter möts. Zilliacus (1992, 8) har påpekat att dialektgränsen genom Åboland sammanfaller med en medeltida administrativ gräns mellan Norrfinne och Söderfinne lagsagor och att gränsen inte enbart är en dialektgräns utan även utgör en kulturell utbredningsgräns.

Indelningen av de åboländska dialekterna i undergrupper har varierat mellan olika forskare. En vanlig indelning är den i *västra* och *östra* Åboland med gräns antingen mellan Nagu och Pargas eller mellan Pargas och Kimitoön. Denna indelning återfinns bl.a. i FO och hos Harling-Kranck (1998). Enligt Zilliacus (1992, 121–123) kan man däremot särskilja tre dialektområden inom Åboland: *västra* Åboland med Iniö, Houtskär och delar av Korpo, *östra* Åboland som omfattar norra och mellersta Nagu, Pargas, hela Kimitoön och Finby, samt *södra Åboland* som omfattar de sydligare delarna av Korpo och Nagu samt hela Hitis socken. De väståboländska dialekterna har stått nära de öståländska och sydösterbottniska dialekterna, medan de öståboländska dialekterna har gemensamma drag med de västnyländska. Till det sydåboländska dialektområdet har novationer kommit in via södra Åland.

Naert (1995) går ännu ett steg längre och menar att det finns fyra tydliga dialektområden i Åboland. Enligt henne skiljer sig de arkaiska dialekterna i Nagu och Pargas till den grad från dialekterna i västra, östra och södra Åboland att de bildar en separat fjärde grupp.

En dialektometrisk (dvs. kvantitativ dialektgeografisk) analys (Leinonen 2012) visar att öst–väst-dimensionen förklarar största delen av dialektskillnaderna i Åboland, men att det också finns en tydlig nord–syd-dimension och en dimension där Nagu och Pargas skiljer sig från det övriga Åboland. De motstridiga spridningsmönstren för olika dialektdrag inom ett relativt litet geografiskt område gör det svårt att göra en entydig dialektindelning. De språkliga avstånden mellan dialekterna är också större i Åboland än i de tre övriga svenska landskapen i Finland, vilket innebär att dialektsplittringen har varit särskilt stor just i Åboland (Leinonen 2012, 171).

Dialektkartor som visar hur de åboländska dialekterna förhåller sig till övriga finlandssvenska och nordiska dialekter finns bl.a. hos Zilliacus (1992), Ivars (2015) och Leinonen (2015).

1.3 Syfte

Huvudsyftet med undersökningen är att beskriva dialektutvecklingen i Åboland under senare delen av 1900-talet. Dialekterna i Europa har under 1900-talet utmärkts av två starka tendenser, nämligen utjämning och regionalisering (Auer 2011). Med utjämning avses att (markerade) dialektdrag försvinner. Hypotetiskt kan man räkna med dialektutjämning dels på en vertikal skala mellan dialekt och standardspråk, dels på en horisontell skala mellan olika dialekter inom ett geografiskt område (Auer & Hinskens 1996, 5–15). Både vertikal dialektutjämning och horisontell dialektutjämning kan leda till regionalisering.

Termerna regionalisering och regiolekt har definierats delvis olika av olika forskare. Ursprungligen har regiolekt (eller regional dialekt) använts om ett kontinuum av språknivåer mellan traditionella dialekter och standardspråk (Hoppenbrouwers 1983; se även Auer 2011, 491–492). Numera används termen regionalisering även om horisontell dialektutjämning (även kallad *koineisering*; Hinskens 2011, 101), dvs. att dialekterna inom en region blir språkligt mer lika varandra utan att detta nödvändigtvis sker under inverkan av ett standardspråk. Ofta är horisontell dialektutjämning dock en biprodukt av utjämning i riktning mot standardspråket. Med termen regionalisering betonas ofta minskad differentiering på den lokala nivån till förmån för differentiering på den regionala nivån (Hinskens et al. 2005, 12; Kerswill & Trudgill 2005,

202). Vid regionalisering bevaras ofta dialektdrag med större geografisk spridning, medan drag med en begränsad, lokal utbredning försvinner (se t.ex. Thelander 1980 & 1982; Trudgill 1986, 98; Auer & Hinskens 1996, 14; Auer 2005, 27; Schwarz et al. 2011).

Akselberg (2003) skiljer mellan urban och rural regionalisering. Vid urban regionalisering utgör talspråket i en större stad utgångspunkt och drag ur stadsmålet sprider sig till omkringliggande dialekter. Vid rural regionalisering finns det däremot inget stadsspråk med stort inflytande, utan landsbygdsdialekterna inom ett område påverkar varandra ömsesidigt på så vis att de mest lokala och markerade dialektdragen utjämnas och dialektdrag gemensamma för alla eller de flesta dialekterna i området bevaras.

Dialektutjämning kan leda antingen till stilistisk variation, dvs. att dialektdrag markerar en viss stilnivå i talspråket, eller till att de mest lokala språkformerna helt och hållet dör ut och dialekten ersätts av en regional varietet.

För Åbolands del skulle en rural regionalisering innebära att vissa drag som varit gemensamma för alla åboländska dialekter bevaras medan drag som funnits bara i enskilda sockenmål går förlorade. Den närmaste större staden är Åbo, och vid urban regionalisering skulle talspråket i Åbo följaktligen kunna utgöra en utgångspunkt. Eftersom det finlandssvenska standardspråket vuxit fram just i Åbo (Ahlbäck 1956), går det i detta fall inte att skilja mellan urban regionalisering med utgångspunkt i Åbo och vertikal dialektutjämning i riktning mot överregionalt finlandssvenskt standardspråk. Vidare har stadsspråket i Åbo under senare tid inte varit föremål för detaljerade vetenskapliga studier. Därmed går det inte att avgöra om det är Åbo eller t.ex. Helsingfors som utövar starkast inflytande vid utjämning mot en standardnära varietet. Ytterligare källor till språkligt inflytande utgörs dels av skriftspråket, dels av svenskan i Sverige.

Enligt Ivars (2003) har dialektutjämnningen gått längre i södra Finland än i Österbotten. I söder går utvecklingen mot regionala dialekter, men med olika styrka i olika regioner. Ivars (2003, 61–62) nämner begreppet *åboländska*, men huruvida det står för en stilistisk varietet som används beroende av talsituation eller för en nyutvecklad regional dialekt, det är inte närmare utrett. Åboländska dialektdrag som visar stor motståndskraft mot utjämning är, att döma av Bergströms studie

(1984), sådana som är gemensamma för åboländska och nyländska dialekter. Likheterna mellan östra Åboland och västra Nyland består i stort sett när dialekterna jämnas ut och går över i nya varieteter med större geografisk eller stilistisk räckvidd (Ivars 2003, 75). Vissa markerade drag försvinner dock i processen trots att de haft stor geografisk spridning (Ivars 2003, 68).

Den översikt som Ivars ger över tendenserna i språkutvecklingen i svenskan i Finland bygger för Åbolands del på Bergströms beskrivning av Mattkärrdialekten i Kimito. Slutsatserna kan alltså närmast sägas gälla östra Åboland. Eftersom det framför allt är de öståboländska dialekterna som har uppvisat stora likheter med västnyländska (Zilliacus 1992, 122; Leinonen 2012), är det möjligt att dialektutjämningen i västra Åboland följer andra mönster.

Frågan huruvida ett åboländskt regionalspråk har uppstått, eller håller på att uppstå, har hittills alltså inte utretts empiriskt. De språkliga avstånden mellan sockendialekterna har varit större i Åboland än i Nyland, Åland och Österbotten, vilket gör en undersökning av regionaliseringstendenser särskilt intressant i Åboland. Administrativt och kulturellt bildar Åboland en enhet, men eftersom dialektavstånden har varit stora, kan det antas att det inte finns särskilt goda lingvistiska grunder för att en gemensam åboländsk regionalvarietet ska uppstå (Leinonen 2012, 172).

För analysen av dialektutvecklingen i Åboland har i denna studie material inspelat i början av 2000-talet använts (se inledningen till boken och avsnitt 1.4 nedan). För studiet av den kronologiska utvecklingen jämförs 2000-talsmaterialet främst med Zilliacus (1992) beskrivning av de åboländska dialekterna. Jämförelser görs även med Bergström (1984) eftersom hennes dialektmaterial kronologiskt infaller mellan det språk Zilliacus beskriver och 2000-talsmaterialet. En analys av pågående utvecklingstendenser möjliggörs också av att informanter ur en äldre och en yngre åldersgrupp ingår i det inspelade materialet.

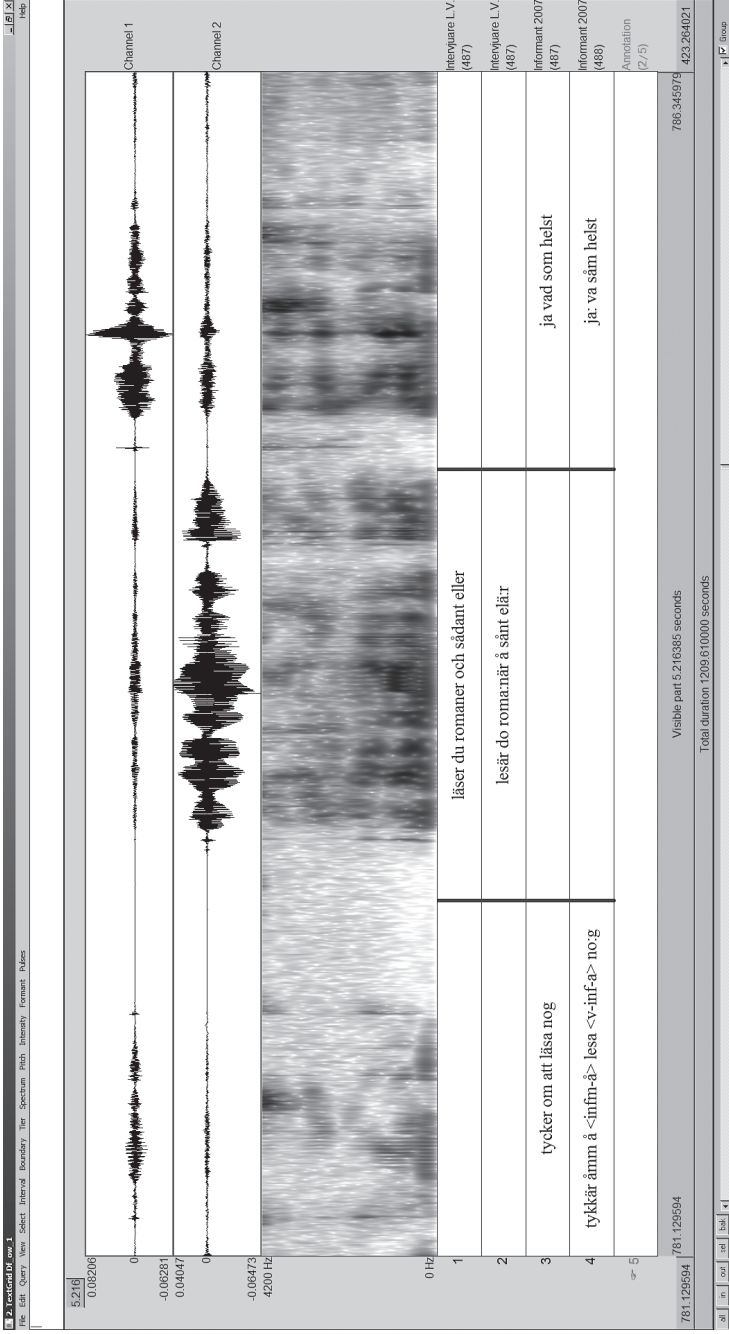
Eftersom både Zilliacus och Bergströms arbeten beskriver dialekternas fonologi och morfologi har dessa språknivåer varit i fokus även för föreliggande arbete. En systematisk genomgång av dialekternas fonetik, syntax och prosodi faller utanför ramen för undersökningen, även om regionala drag i större utsträckning kan tänkas bevaras just på dessa språkliga nivåer.

De konkreta frågeställningarna för undersökningen är:

- Hur långt har dialektutjämningen gått i Åboland under senare delen av 1900-talet?
- Vilka dialektdrag bevaras vid utjämning av dialekterna?
- Vilka dialektdrag försvinner?
- Finns det tecken på uppkomst av en regional åboländsk varietet?

1.4 Material och metoder

Materialet för undersökningen är huvudsakligen insamlat inom projektet *Spara det finlandssvenska talet*. De transkriberade klippena på cirka 20 minuter per informant (se inledningen till denna bok) utgör huvudmaterialet för analysen. I *Spara talet*-samlingen ingår alltså mer material från Åboland än det som beskrivs här. Analys av ett helt obearbetat material skulle ha tagit mycket mer tid i anspråk och hade inte varit genomförbart inom ramen för detta projekt, där avsikten har varit att skapa en överblick över dialektsituationen i Åboland i början av 2000-talet. Tack vare transkriptionerna har det varit möjligt att relativt snabbt ge en sådan övergripande bild av situationen. För djupare studier av enskilda språkdrag eller dialekter skulle det naturligtvis krävas mera material, vilket finns att hämta i de obearbetade ljudfilerna i *Spara talet*-samlingen. Inför analysen importerades materialet i programmet Praat (Boersma & Weenink 2013). Öppnar man filerna för en informant i Praat ligger den ljudenliga och den standardiserade transkriptionen av intervjuarens och informantens repliker i parallella lager. Dessutom visas ett spektrogram och två vågformer (eftersom ljudfilerna är i stereo). Figur 1.3 visar en bildskärmsdump ur Praat. Räknet uppifrån och ner visar de två översta raderna vågformerna och den tredje raden spektrogrammet. Sedan följer först den standardiserade och den ljudenliga transkriptionen av intervjuarens repliker och därefter detsamma för informanten. Den sista raden är reserverad för eventuella kommentarer. Genom att ljudfilen ligger parallellt med transkriptionen är det enkelt att snabbt gå in och lyssna på ett specifikt ställe i inspelningen eller söka efter ett visst ord i den standardiserade transkriptionen och lyssna efter hur informanten uttalar just det ordet.



Figur 1.3. Skärmdump ur Praat. Lagren visar två vågformer, spektrogram, intervjuarens repliker i standardiserad och i ljudenlig transkription samt informantens repliker i standardiserad och i ljudenlig transkription. Två taggar har lagts in manuellt i informantens repliker.

1.4.1 Informanter

Sammanlagt 41 informanter från den åboländska skärgården ingår i SLS talspråkskorpus. Dessa informanter är fördelade på nio socknar och representerar två olika åldersgrupper. De yngre informanterna var cirka 20–30 år vid inspelningsögonblicket och de äldre informanterna var cirka 55–75. Antalet informanter från de nio socknarna är inte jämnt fördelat utan speglar antalet invånare på orten. På så vis finns det flest informanter från Pargas och minst från Iniö, Västanfjärd och Hitis. Förutom de 41 informanterna i *Spara talet*-samlingen togs två informanter ur projektet *SweDia* (Eriksson 2004), där liknande dialektinspelningar gjordes, med i analysen. Dessa är två äldre personer från Houtskär. Skälet till att de inkluderades beskrivs närmare i avsnitt 1.5 nedan.

Antalet informanter per ort, åldersgrupp och kön framgår av tabell I.2. Av de totalt 43 informanterna hör 22 till den äldre åldersgruppen och 21 till den yngre.

För identifiering av informanterna i texten används signum. Dessa är samma koder som används i de båda databaserna *Spara talet* och *SweDia*, bara ortnamnsdelen av signumerna har förkortats och består här av endast två bokstäver. Förkortningarna ses i tabell I.2. Sist i sig-

Tabell I.2. Antal informanter per ort, åldersgrupp och kön.

ORT	ÄLDRE		YNGRE		TOTALT
	KVINNOR	MÄN	KVINNOR	MÄN	
Iniö (Io)	1	1	1	0	3
Houtskär (Ho)	2	2	1	1	6
Korpo (Ko)	2	1	0	1	4
Nagu (Na)	1	2	2	2	7
Pargas (Pa)	2	2	2	2	8
Kimito (Ki)	1	1	1	1	4
Dragsfjärd (Df)	1	1	2	2	6
Västanfjärd (Vf)	1	0	1	0	2
Hitis (Hi)	1	0	1	1	3
Totalt	12	10	11	10	43
	22		21		

numen står antingen *sp* (= *Spara talet*) eller *sw* (= *SweDia*). De två yngre kvinnorna från Dragsfjärd har alltså signumena Df.ywo1.sp och Df.ywo2.sp etc.

De flesta informanter är både uppvuxna och bosatta på den ort de representerar. Undantag från denna regel utgör Hi.owo1.sp som vid tiden för intervjun bodde i Dragsfjärd dit hon flyttat i vuxen ålder. Ko.ymo2.sp har inte heller varit bofast men har största delen av livet bott någonstans i den yttre skärgården varifrån även bägge föräldrarna kommer. Ko.omo2.sp och Ko.owo3.sp (som intervjuades tillsammans) har levt största delen av sina liv i utskären i Korpo men flyttat till kyrkbyn som pensionärer. Från Korpo finns bara en enda informant i korpusen som kan sägas representera Kyrklandet (Ko.owo1.sp).

Bland de yngre informanterna finns fyra stycken som har angett dubbla bostadsorter: dels den ort där de vuxit upp och som de representerar i korpusen, dels en annan (tillfällig?) studie- eller bostadsort (i tre fall Åbo, i ett fall Pargas).

I *Spara talet*-projektet bemödade man sig om att rekrytera intervjuare som själva hade anknytning till orten och talade det lokala språket. För Åbolands del kan man säga att detta lyckades något bättre för de östra delarna än för de västra. Alla intervjuare har en personlig anknytning till orten, men de som utförde intervjuerna i Iniö, Houtskär, Korpo och Nagu talar inte själva det lokala språket. Exempel på hur detta har påverkat informanternas språkbruk ges i avsnitt 1.5.

1.4.2 Val av språkliga variabler och excerpering

För att en diakronisk jämförelse skulle vara möjlig har undersökningen utgått från de språkdrag som beskrivs i *Åboländska* (1992) av Kurt Zilliacus. Där beskrivs variation i uttal, böjningsformer, ordbildning och ordförråd. Beskrivningen av ordförrådet behandlar dock främst termer ur bondesamhället som det inte finns belägg på i 2000-talsinspelningarna. Föreliggande undersökning är därför främst inriktad på variation på den fonologiska och morfologiska nivån.

Flera av de drag som Zilliacus (1992) beskriver är för infrekventa i språket för att man ska kunna hitta tillräckligt många belägg för kvantifiering i en spontantalskorpus som omfattar bara tjugo minuter per informant. Sådana drag har fått utgå helt eller nämns i vissa fall bara i

form av enskilda exempel i texten. Analysen har dessutom i första hand koncentrerats till drag som uppvisar faktisk variation i materialet. Ett drag kan vara invariant antingen för att alla (eller nästan alla) informanter använder bara standardvarianten av ett visst dialektdrag, eller för att informanterna uteslutande använder dialektvarianten. Det förstnämnda fallet gäller för flera av de drag som har beskrivits av Zilliacus (t.ex. *kv-* av gammalt *hv-* i ord som *vit* och *vete*). Vad gäller dessa drag har dialektutjämnningen gått så långt att dialektvarianten verkar vara på väg att försvinna helt. I sådana fall kanske det bara finns några ströbelägg på dialektvarianten i materialet, medan majoriteten av informanterna uteslutande använder den standardsvenska formen. Sådana språkdrag har inte kvantifierats, men för vissa drag ges exempel på förekomst i materialet (se t.ex. avsnitt 2.2.3 om inskottskonsonanter, 2.2.5 om initialt *g-*, *k-* och *sk-* och 3.3.3 om pronomenet *token/tolin*).

Språkdrag där samtliga informanter använder en icke-standard-svensk variant förekommer också i materialet. Men inget av dessa drag är ett specifikt dialektdrag, utan det är närmast frågan om allmänfinlandssvenska talspråksdrag, som bortfall av *-t* hos neutrer i bestämd form (*ta:ke* 'taket'), bortfall av *-de* i preteritum av första konjugationens verb (*ba:da* 'badade') och uttal som *si:* 'se' och *tri:* 'tre' (jfr Ivars 2003, 77–78).

Tack vare att materialet är ljudenligt transkriberat och länkat till standardiserad ortografi har det varit möjligt att exciperera flera av variablerna automatiskt eller halvautomatiskt. För att ta reda på hur informanterna uttalar t.ex. ord på *dj-* (avsnitt 2.2.4) har det gått att automatiskt ta ut en lista över alla ord som börjar på *dj-* i den standardiserade transkriptionen och räkna hur många av beläggen som motsvaras av *dj-* respektive *j-* i den ljudenliga transkriptionen.

Excerpteringen har baserats direkt på ljudskriften i sådana fall där variabeln är så tydligt framträdande att man kan räkna med att transkriptionerna är pålitliga trots att de utförts av flera olika transkriberare. Detta gäller t.ex. variationen *inga* ~ *inte* (avsnitt 3.5). I andra fall, där man inte självklart kan utgå från att en transkriberare lägger märke till variationen (t.ex. variationen *ng* ~ *ngg*, avsnitt 2.2.1), har jag lyssnat igenom beläggen och vid behov gjort ändringar i ljudskriften.

Eftersom materialet inte var ordklasstaggat när analysen av det åboländska materialet gjordes, var en automatisk excerptering inte möjlig

för alla variabler. Exempel på variabler som excerperades manuellt är infinitivmärket (3.4.1) och infinitiv (3.4.2). För sådana drag har jag gått igenom materialet manuellt och lagt in taggar i transkriptionen. I figur I.3 (s. 40) finns t.ex. en tagg för varianten *å* av infinitivmärket och en tagg för verb i infinitiv med ändelsen *-a*. Efter den manuella taggningen kunde antalet taggar av varje sort räknas automatiskt för varje informant. Tack vare excerpering i form av taggning var det också möjligt att i efterhand gå in i transkriptionen och göra ytterligare noteringar eller begränsa kontexten för en viss variabel.

Egennamn ingår inte i de kvantifierade variablerna.

1.4.3 Materialredovisning

Ljudskrift

Vid genomgången av de olika språkdragen ges ett antal exempel i texten. I exemplen återges replikerna med en grov ljudskrift med det svenska alfabetets tecken som i stort sett motsvarar den grova ljudbeteckning som används i FO (bd I, s. XIV–XV, XVII). Långa vokaler markeras med kolon och långa konsonanter dubbeltecknas. Tecknet *o* står för *o*-ljudet i *sol* och *rott* och *å* för vokalen i *håll* och *rått*.¹ *Sje*-ljud anges med *sj* och *tje*-ljud med hörbart *t*-förslag anges med *tj*. *Ng*-ljudet markeras med *ng* eller om *g* hörs med *ngg*. Supradentalt uttal av *rs* markeras med *ssj*, övriga supradentaler markeras inte utan skrivs ut som konsonantkombinationer (*rt*, *rn*). Pauser markeras med (*.*), (*..*) eller (*...*) (ju fler punkter desto längre paus), och avbrott markeras med [...]. Efter ljudskriften följer i exemplen den standardiserade transkriptionen.

1. I vissa dialekter i Åboland finns en vokal som ligger mitt emellan dessa två ljud – landsmålsalfabetets *o* med en boll inuti nere (*o*) – varvid transkriberaren har varit tvungen att välja något av de två tecknen för *o*/*å*-ljud i den grova ljudskriften.

Lådagram

Resultaten för de kvantifierade variablerna visualiseras med hjälp av s.k. lådagram med morrhår. Se t.ex. figur I.15 på s. 121, som anger andelen apokoperade preteritumformer. Lådagrammen visar distributionen av ett visst språkdrag per ort och åldersgrupp. De anger andelen dialektvarianter av det totala antalet belägg för en viss variabel på en skala mellan 0 (= bara standardvarianter) och 1 (= bara dialektvarianter). Andelen är beräknad för varje informant. Bollen inuti lådan anger medianvärdet för informanterna i varje grupp och lådan rymmer värdena för den mittersta hälften av informanterna. Morrhåren (de lodräta linjerna) sträcker sig till det högsta och det lägsta värdet i gruppen. Ifall cirklar ligger utanför lådan, anger dessa extremvärden. Diagrammen är gjorda med paketet *lattice* (Sarkar 2008) i programmet R (R Core Team 2013).

Lådagrammen gör det möjligt att jämföra grupper av informanter med varandra och samtidigt se hur mycket variation det finns mellan informanterna inom varje grupp. Vad gäller t.ex. informanterna från Pargas i figur I.15 (s. 121) kan man se att de yngre informanterna apokoperar preteritumformerna i genomsnitt betydligt oftare än de äldre informanterna. De yngre informanterna har apokopering av preteritum i genomsnitt i cirka 85 procent av fallen, medan de äldre apokoperar i genomsnitt bara i knappa 60 procent av fallen (bollen inuti lådan). Men man kan också se att variationen bland de äldre informanterna från Pargas är betydligt större än bland de yngre (större låda och längre morrhår för de äldre informanterna), och trots att de äldre i genomsnitt apokoperar mindre än de yngre så finns det också enskilda äldre informanter som apokoperar i lika stor utsträckning som de flesta yngre.

Ett streck i stället för en låda innebär att det inte finns någon variation alls. Detta kan bero på att det bara finns en enda informant i en ”grupp”, det finns t.ex. bara en yngre informant från Iniö (se tabell I.2 ovan). Även när alla informanter i en grupp betar sig precis lika uppkommer ett streck i figuren. De bägge yngre informanterna från Houtskär apokoperar preteritumformerna i 100 procent av fallen, vilket leder till ett streck vid Houtskär yngre i figur I.15.

Trots att det finns få informanter i vissa grupper har jag valt att visualisera resultaten på detta sätt enligt ort och åldersgrupp. Diagrammen ger möjlighet att snabbt bilda sig en uppfattning om olika drags

relativa frekvens i materialet, och eftersom dialekterna i Åboland har varit så olika sinsemellan har en redovisning per ort känts som det enda rimliga. Vidare gör indelningen i åldersgrupper det möjligt att se om en tendens till språkförändring verkar föreligga. De två åldersgrupperna får så att säga representera olika tider (undersökning av språkförändring i s.k. skenbar tid). På grund av det ringa antalet informanter måste vi hålla i minnet att konklusioner om den språkliga variationen inte kan dras enbart utifrån diagrammen.

1.4.4 Kvantitativa analysmetoder

Variabler som uppvisar variation mellan informanterna och som de flesta informanter har minst fem belägg på har kvantifierats. I kapitel 4 redovisas en sammanlagd statistisk analys med hjälp av principalkomponentanalys av de variabler som har ansetts tillräckligt frekventa för kvantifiering. I genomgången av de enskilda variablerna i kapitel 2 och 3 analyseras variationen med statistiska medel bara i de fall då det finns inomspråkliga faktorer (som t.ex. betoning, syntaktisk ställning eller ord) som hypotetiskt kan förklara en del av variationen. Detta gäller för variablerna *ei*-diftongen (avsnitt 2.1.1), kort *u* (2.1.5), *ng*-ljudet (2.2.1), infinitivmärket (3.4.1), infinitiv (3.4.2) och preteritum (3.4.3). Den metod som har använts i dessa fall är logistisk regression med blandade effekter.

Logistisk regressionsanalys tillämpas inom sociolingvistik ofta för att klarlägga sambandet mellan olika inom- eller utomspråkliga faktorer och en språklig variabel med två varianter (vanligen en dialektvariant och en standardspråklig variant). Sådan analys brukar ofta kallas Varbrul-analys (av eng. *variable rule*; Sankoff 1988; Paolillo 2001). Under de senaste åren har Varbrul-analysen kritiserats för att inte beakta variationen mellan individer (K. Johnson 2008, 264; D. E. Johnson 2009). Med hjälp av Varbrul-analys har man ofta undersökt hur sociala faktorer som t.ex. ålder och kön påverkar variationen, men analysen beaktar inte faktorer som variation mellan individer, hur många informanter som ingår i varje grupp eller hur många belägg som finns för varje informant. Genom att inte beakta variationen på individnivå överskattar man ofta signifikansen av de sociala faktorerna.

En lösning på problemet är att använda logistisk regression med

blandade effekter. Tillämpningen av regressionsanalys med blandade effekter inom språkvetenskapen har behandlats av bl.a. Baayen (2008, 241–302), K. Johnson (2008, 259–264), D. E. Johnson (2009), Tagliamonte & Baayen (2012) och Drager & Hay (2012). I en klassisk Varbrul-analys ingår endast fixa effekter. I en modell med blandade effekter skiljer man däremot mellan fixa effekter och stokastiska effekter. Fixa effekter är faktorer med ett begränsat, relativt litet antal faktornivåer som kan upprepas från undersökning till undersökning. Exempel på sådana faktorer är kön (man/kvinna) och betoning (betonad/obetonad). I motsats till de fixa effekterna har de stokastiska effekterna ett obegränsat antal nivåer som dras slumpmässigt från en population. Medan fixa effekter har nivåer som kan hållas konstanta genom olika undersökningar, kommer faktornivåerna för de stokastiska effekterna i regel att variera från undersökning till undersökning.

En typisk stokastisk effekt är informant. Informanterna dras i bästa fall slumpmässigt från en större population och det är osannolikt att precis samma informanter ingår i två olika undersökningar. Informanterna kommer därmed också att uppvisa slumpmässig variation. Det kan t.ex. finnas informanter som är mer benägna att använda dialektvarianten av en viss variabel än vad man skulle förvänta sig utifrån informantens ålder, kön och sociala tillhörighet. En annan typisk stokastisk effekt i språkvetenskapliga undersökningar utgörs av faktorn ord. Om man t.ex. undersöker apokopering av preteritumformer kan det hända att vissa verb apokoperas oftare än andra. Om undersökningsmaterialet består av spontant tal kommer antalet belägg per verb att variera och sannolikheten att precis samma verb ingår i en ny undersökning är liten.

Modellen med blandade effekter minskar risken för felaktiga slutsatser om de fixa effekterna. När man lägger till informant som stokastisk faktor i en modell kommer en social faktor som kön att bli signifikant bara om variationen mellan män och kvinnor är större än variationen mellan enskilda informanter. I en klassisk Varbrul-analys som inte beaktar individuell variation finns det däremot ingen möjlighet att veta om variationen mellan grupper verkligen överskrider variationen mellan individer.

Informanterna i undersökningen kan delas in i grupper utifrån faktorerna *ort*, *ålder* och *kön*. Dessa tre variabler har utgjort fixa effekter vid regressionsmodelleringen. Beroende på vilket språkdrag som un-

dersöks ingår olika lingvistiska förklarande variabler som fixa effekter. Informant används som stokastisk faktor, som ger en slumpmässig variation i interceptet (dvs. regressionslinjens skärningspunkt med y-axeln), i alla analyser. Beroende på vilket drag som undersöks kan även andra stokastiska effekter ingå i modellen.

Likelihood-kvottest har använts för identifiering av den bästa modellen (Baayen 2008, 253), och modellens anpassningsgrad anges med *Akaike Information Criterion* (AIC; Akaike 1974) och *Somers D_{xy}* . För AIC gäller principen ju lägre värde desto bättre anpassningsgrad. *Somers D_{xy}* anger korrelationen mellan förväntade sannolikheter och observerat utfall; 0 innebär slumpmässigt samband och 1 perfekt prognos (Baayen 2008, 204). Alla analyser är gjorda i programmet R (R Core Team 2013), och det paket som har använts är lme4 (Bates et al. 2011).

1.5 Dialekternas ställning i Åboland

Beskrivningen i detta avsnitt bygger dels på informanternas kommentarer om dialekten och dess ställning i de inspelningar som utgör materialet för den språkliga analysen, dels på svar på två frågelistor från Folkkultursarkivet och Språkarkivet under 1990-talet. Frågelistorna är Folkkultursarkivets pristävling nr 32 *Vad talar du? Dialekt – när, var, hur?* (SLS 1746) och Språkarkivets specialfrågebrev nr 47 *Dialekt och talspråk idag* (SLS 1954). Av dem som svarat på frågelistorna är de flesta från Pargas (6 stycken) medan det från övriga orter i den åboländska skärgården kom in bara enstaka svar på frågelistorna.

1.5.1 Dialekt och standardspråk

En tendens som de flesta meddelare är eniga om är att dialektens ställning har försvagats under 1900-talet. Förutom den allmänna åsikten att det bara är äldre personer som kan "genuin" eller "ren" dialekt, finns det också mer konkreta exempel på hur dialektens ställning har försvagats. Det finns meddelare som berättar att de under uppväxten på 1930- och 1940-talen sällan hörde annat språk än den lokala dialekten. På grund av skola, tv och (finsk) inflyttning har språkförhållandena blivit helt andra. Finska hördes tidigare nästan inte alls i skärgården. Många meddelare vittnar också om att det är skolan som har infört ett mer standardnära

språk – som meddelarna ofta kallar skolspråk, skriftspråk eller bokspråk. De yngres språk beskriver många som en blandning mellan dialekt och standardspråk. En meddelare från Pargas (f. 1927) skriver att hans systrar som gick i samskola tidigt lade bort dialekten (SLS 1746: 487). Andra meddelare berättar om hur de själva har övergått till att tala mer standardnära eller valt att tala standardspråk med sina barn trots att de själva är uppvuxna med dialekt som modersmål (SLS 1746: 467, 500, 508). Barnen har sedan ändå lärt sig en sorts dialekt från omgivningen – ”färglösare och utslätad” jämfört med den som talades av tidigare generationer (SLS 1746: 501). Flera meddelare berättar att man tidigt kom i kontakt med standardspråket via sommargäster och att man också själv försökte ”bättra på språket” när man talade med sommargästerna (SLS 1746: 487, 527, 542).

Som orsaker till att välja att (i vissa situationer) tala dialekt i stället för att övergå till standardspråk nämns av dem som svarat på frågelis-torna att man känner sig som traditionsbärare och att dialekten ger samhörighetskänsla. Det känns också tryggt och bekvämt att tala så som man har lärt sig av de äldre generationerna.

Att det funnits skillnader i dialekten också inom varje socken är tydligt. Språket har ofta varit mer standardnära i kyrkbyarna än i mer perifera byar. I Pargas har bruken och herrgårdarna spelat en stor roll eftersom de dragit till sig inflyttare från andra orter. Skillnaderna mellan byarna kvarstår ofta trots att dialekterna utjämnas. Meddelarna beskriver det ofta som att det är ”betoningen” eller ”tonfallet” som är olika. Det handlar då antagligen om skillnader i prosodin eller om fonetiska skillnader i uttalet av t.ex. enskilda vokaler. Meddelarna berättar också att man trots att man övergår till att tala mer standardnära håller kvar vissa drag av dialekten för att visa var man har sina rötter. Med ordet *dialekt* avser man då ofta den dialekt som talas av äldre personer, vilket kan exemplifieras med följande utdrag:

- (1.1) *ja tro: int ens ja kan den hä dialektn* (.) *övärhovota:ge* (Hi.ymo1.sp)
jag tror inte ens jag kan den här dialekten (.) överhuvudtaget

Informanten anser sig inte kunna den lokala dialekten, men använder ändå vissa typiska dialektdrag som t.ex. synkopering (*dialektn*) och o-uttal av kort *u*-ljud (*övärhovota:ge*).

Graden av bidialektalitet bland informanterna är växlande. Från Kimitoön finns vittnesmål om att dialekten är gångbar på hela Kimitoön, men att man försöker ändra sitt språk när man "åker över bron" (SLS 1954: 264; Bergström 1984, 4). Tydligast blir bidialektaliteten i några av *Spara talet*-intervjuerna från Iniö och Korpo. Den som utför intervjuerna talar inte själv det lokala språket. De två äldre informanterna från Iniö (Io.omo1.sp och Io.owo1.sp) menar att om de pratar riktig dialekt så förstår utomstående inte vad de säger. De anpassar sig automatiskt till intervjuaren i intervjusituationen och säger att det skulle kännas onaturligt för dem att tala dialekt i en sådan situation. I en sådan talsituation avstår informanterna från de mest markerade dialektdragen, medan mindre markerade drag finns kvar.

De två äldre informanterna från Korpo som har vuxit upp och levt största delen av sina liv i utskären (Ko.omo2.sp och Ko.owo3.sp) berättar om sin bidialektalitet. Informanterna har flyttat från utskärgården till Kyrklandet, och i utskären vistas de numera bara sommartid. De förklarar att de talar annorlunda när de är på tumanhand ute på ön än vad de gör när de befinner sig på Kyrklandet. För dem är det alltså inte bara samtalspartnern utan också orten i sig som betingar valet av språkvarietet. Intervjun är inspelad på Kyrklandet.

För de yngre informanterna i *Spara talet*-materialet verkar det kännas naturligare att tala dialektalt i intervjusituationen än vad det gör för flera av de äldre informanterna. Denna tendens har ofta noterats i sociolingvistiska undersökningar (se t.ex. Gullmets-Wik 2004, 213). Det finns också informanter som tydligt bemödar sig om att tala dialekt även om dialekten kanske inte är det helt naturliga uttrycksmedlet i en sådan situation. Detta märker man när informanterna rättar sig själva efter att först ha använt ett standardsvenskt ord eller uttryck:

- (1.2) *ingen såm (.) såm stå:r i ve:gen (.) sta:r i ve:j in ska ja säjj*
(Ho.ywo1.sp)

ingen som (.) som står i vägen (.) står i vägen ska jag säga

- (1.3) *de va nån vekka före ju:len nån (.) ja: he va kansk i en viko före ju:len* (Na.omo2.sp)

det var någon vecka före julen någon (.) ja det var kanske en vecka före julen

I en del av *Spara talet*-intervjuerna från Åboland kan man också märka att samtalsämnet har betydelse för valet av språkform. När det är tal om studier och arbetsliv utanför hemorten blir språket mer standardnära än när man talar om det lokala dagliga livet, barndomsminnen och traditioner.

1.5.2 Stora skillnader inom Houtskär

I undersökningen ingår, som redan nämnts, två informanter som inte är inspelade i *Spara talet*-projektet utan i *SweDia*-projektet. Skälet till att dessa två informanter inkluderades var att det i ett tidigt skede av analysen blev klart att dialektskillnaderna är särskilt stora inom Houtskär.

De två äldre Houtskärsinformanterna i *Spara talet*-materialet kommer från Hönsnäs by nära kyrkbyn, medan de två yngre informanterna kommer från Björkö och Mossala, de s.k. Norrbyarna. Det visade sig att dialekten i dessa Norrbyar avviker från den dialekt som talas i de sydligare delarna av Houtskär. Eftersom de yngre *Spara talet*-informanterna är från Norrbyarna och de äldre från Hönsnäs, var det inte möjligt att veta om skillnaderna i språket var geografiskt eller åldersmässigt betingade. Samtliga inspelningar från Houtskär i *SweDia*-projektet gjordes i Norrbyarna, och därför kunde två äldre informanter plockas ur *SweDia*-materialet.

Skillnaden mellan Norrbyarna och det övriga Houtskär noterades redan av Kurt Zilliacus, vars beskrivning av Houtskärsdialekten till största delen bygger på material insamlat i Norrbyarna. Befolkningen i Norrbyarna var enligt Zilliacus känd ”för sin dialektkunnighet” (SLS 701–702, s. 1).

Även 2000-talsinformanterna vittnar om dialektskillnaderna inom Houtskär. Den äldre kvinnan från Hönsnäs berättar om släktingar i Mossala:

- (1.4) *däm pra:ta li:te ry:slit gammal (.) ho:tsjä:rsko däm (..) så sån här ho:tsjä:rdialekt såm ingga ja:g (.) kuna pass fösto: hellär ibland (Ho.ow01.sp)*

de pratade lite rysligt gammal (.) houtskärska de (..) så sådan här houtskärdialekt som inte jag (.) kunde pass förstod heller ibland

Även informant Ho.omo1.sp håller med om att Houtskärsdialekten har bevarats bäst i Björkö och Mossala, där man ganska länge talade ”*re:n ho:tsjä:rsko*”. En *SweDia*-informant från Björkö berättar å sin sida att det nästan inte talas någon dialekt i kyrkbyn Näsby numera, medan Björkö och Mossala har mera dialekt.

Att de geografiskt betingade skillnaderna i materialet från Houtskär är större än skillnaden mellan äldre och yngre informanter bekräftas av den kvantitativa analysen som redovisas i avsnitt 4.3.

2

Ljudlära

Nedan beskrivs drag inom ljudläran där de åboländska dialekterna visar avvikelser från standardspråket. Först beskrivs drag inom vokalismen (2.1) och sedan drag som har med konsonantuttalet (2.2) att göra. Sist i kapitlet (2.3) behandlas det fonologiska kvantitetssystemet, dvs. förekomsten av korta och överlånga stavelser.

2.1 Vokaler

Vokaluttal inrymmer i allmänhet ett stort mått av dialektal variation (Bruce 2010, 34–35). I detta avsnitt behandlas främst fonologisk variation i vokaluttal, dvs. variation som överskrider fonemgränserna. För analys av fonetisk variation, dvs. variation i det exakta uttalet av enskilda vokalfonem, lämpar sig instrumentella analyser bättre än transkriptionsbaserade.

Två av de åboländska orterna, Houtskär och Dragsfjärd, ingår i en instrumentell analys av vokaluttal baserad på material från *SweDia*-projektet insamlat kring år 2000 (Leinonen 2010). Analysen visar att de åboländska orterna i fråga om vokaluttal hör till ett kontinuerligt sydfinländskt dialektområde (Leinonen 2010, 132–136). Houtskärsdialekten skiljer sig från de nyländska dialekterna bl.a. i fråga om det långa *u*-ljudet, som ligger närmare det sverigesvenska *u*-ljudet (Leinonen 2010, 220), och i fråga om *ä*-ljudet som uttalas slutet även före *r* (dvs. orden *lär* och *nät* uttalas med ungefär samma vokal; Leinonen 2010, 95, 162). Vidare är diftongerade uttal vanligare i Åboland än i de nyländska dialekterna (Leinonen 2010, 91, 93, 99, 119).

Av de dialektala dragen i vokaluttal som behandlas nedan har tre variabler kvantifierats och ingår i den sammanlagda kvantitativa analysen i kapitel 4. Dessa är *ei*-diftongen (avsnitt 2.1.1), bevarat *a* (2.1.4) och kort *u* (2.1.5).

2.1.1 Primära diftonger

De primära diftongerna *ei*, *öy* och *öu* härstammar från urnordisk tid. Av dessa diftonger är *ei*-diftongen den vanligaste (Zilliacus 1992, 47). Enligt Zilliacus (1992, 48) hade *öu*-diftongen redan i hans äldre åboländska material gått över till långt *ö* i ord som *död*, *gröt*, *hög* och *höst*. Diftongen har enligt Zilliacus bevarats främst i typiska dialektord som *böula* 'böla', *röuta* 'råma', *söud* 'får; ungnöt', *göut* 'halvvuxet barn' och *möura* 'myra', för vilka belägg saknas i 2000-talsmaterialet.

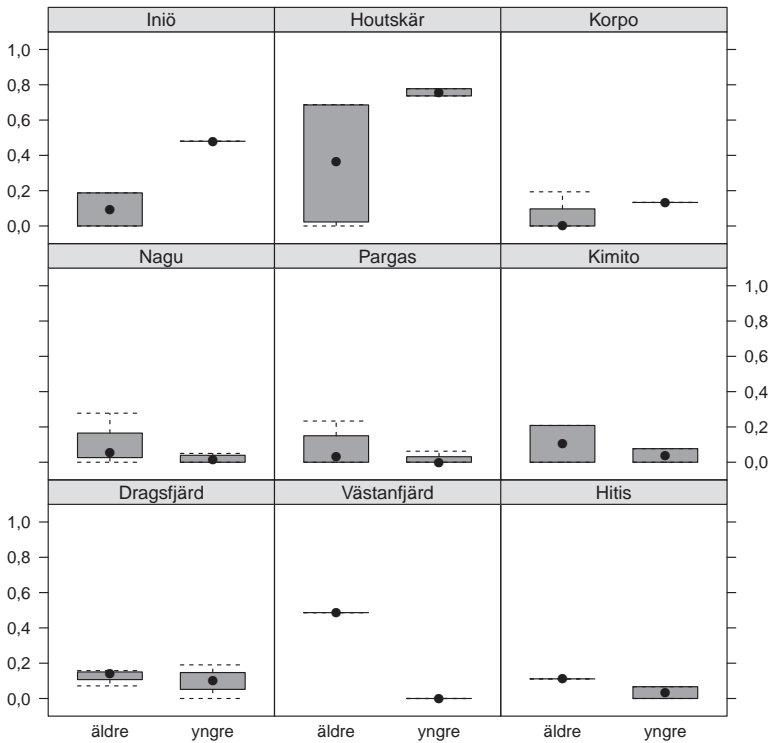
ei

ei-diftongen har enligt Zilliacus (1992, 47) varit utbredd i hela Åboland. Eftersom den förekommer i en rad mycket frekventa ord i materialet har det varit möjligt att kvantifiera variabeln. Variabeln har excerperats ur följande ord för den kvantitativa analysen:

- preteritum av verbet *bliva*, dvs. *blev* (belägg hos 41 informanter)
- adjektivet *hel* inklusive böjningsformerna *helt* och *hela* (belägg hos 41 informanter)
- adjektivet och adverbet *mer(a)* (belägg hos 42 informanter)
- presensformen av verbet *veta*, dvs. *vet* (belägg hos 39 informanter).

Variationen hos denna variabel sker mellan den dialektala diftongen *ei/äi* och standardspråkets monoftong *e*. Antalet belägg med diftong respektive monoftong av ovannämnda ord har räknats för alla informanter. Det totala antalet belägg på variabeln är 962. Antalet belägg per informant varierar mellan 7 och 48 med medianvärdet 21. Andelen diftonger ges som den procentuella andelen av det totala antalet belägg på variabeln hos varje informant.

Endast få informanter har en hög andel diftonger. Bara fyra av informanterna har mer än 50 procent diftonger. Dessa är de fyra infor-



Figur I.4. Lådagram över andelen dialektvarianter av variabeln 'ei-diftong' efter ort och ålder.

manterna från Norrbyarna i Houtskär (dvs. Ho.omo1.sw, Ho.owo1.sw, Ho.ymo1.sp och Ho.ywo1.sp). Fem informanter har mellan 20 och 50 procent diftonger: Io.ywo2.sp, Na.omo2.sp, Pa.omo3.sp, Ki.owo1.sp, Vf.owo1.sp. Åtta av informanterna har mellan 10 och 20 procent diftonger och nio informanter har belegg på ei-diftongen men en diftongandel under 10 procent. Hela 17 informanter saknar helt diftonger.

Lådagrammen i figur I.4 visar andelen diftonger bland informanterna enligt ort och åldersgrupp. Andelen diftonger är högst i Iniö och Houtskär. De yngre informanterna på dessa orter samt i Korpo har genomsnittligt högre andelar diftonger än de äldre. På alla de övriga orterna är andelen diftonger lägre bland de yngre informanterna än bland de äldre.

Resultaten ovan visar på samma tendens som en akustisk analys av vokaluttal baserad på *SweDia*-material, enligt vilken den primära

diftongen *ei* används i ordet *leta* i Houtskär men inte i Dragsfjärd (Leinonen 2010, 91). En manuell genomgång av materialet för den akustiska undersökningen visar att *ei*-diftongen finns hos enskilda informanter från Dragsfjärd i ordet *leta*, men att dessa inte är tillräckligt många för att diftongen ska komma fram tydligt i de gruppmedeltal (11–12 informanter per ort) som redovisas i den akustiska undersökningen.

En logistisk regressionsanalys av det föreliggande materialet med ord som stokastisk effekt (se bilaga A, 1.1) visar att tendensen till bruk av diftong varierar något mellan de fyra ord (*blev*, *hel*, *mer(a)*, *vet*) som excerperades för variabeln. Tendensen till diftonguttal är högst för *blev* och lägst för *hel*. Modellen visar också att interaktionen mellan ort och ålder är signifikant (dvs. det finns fler diftonger bland yngre än bland äldre informanter i väster och motsatt förhållande i öster).

Nedan följer några exempel på *ei*-diftongen ur materialet. Även ord som inte ingår i den kvantitativa analysen (dvs. andra ord än de fyra angivna) är markerade med fet stil.

- (2.1) *föru:t så va e s- (.) syrran min å å de här (.) mamma tå såm at mamma tjö:rd hesti å å å syrran sa:t på höirefson å å trampa u:r (.) mi (.) så va e va e in da:g tå såm e (.) så:g u:t ti kåm oska (..) så he **blei** bråttåm tå ti ti set ho:p höjje å (...) å så sku e inga rikit va: ti:d tå ti (.) ti va: **fleir** å å tjö:r e så så **blei** ja no ti sess på: tå å böja tjö:r sjölvår int va e nu vel så lyckat i början (..) he å:rdna opp se sen å eftår he:d så **blei** e så så tjö:rd ja (Ho.omo1.sw)*

förut så var det s- (.) syrran min och och det här (.) mamma då som att mamma körde hästen och och och syrran satt på höråfsan och och trampade ur (.) men (.) så var det var det en dag då som det (.) såg ut att komma åska (..) så det blev bråttom då att att sätta ihop höet och (...) och så skulle det inte riktigt vara tid då att (.) att vara flera och och köra det så så blev jag nu att sitta på då och började köra själv inte var det nu väl så lyckat i början (..) det ordnade upp sig sedan och efter det så blev det så så körde jag

- (2.2) [Int.:] *si:r ho:tsjä:rs ju:lgåbben annårlunda u:t en (.) en han i kårvatontori [Inf.:] nn- nä:ä kansje at in int int e så **feitår** (.) han e lite **meir** sj- inga sjinn å **båin** (Ho.ywo1.sp)*

[Int.:] ser Houtskärs julgubben annorlunda ut än (.) än han i Korvatunturi [Inf.:] nn- nä kanske att han inte inte är så fet (.) han är lite mer sj- inte skinn och ben

- (2.3) *de käämmär ja ihå:g at he va ka:lt ry:slit (.) man frö:s så sa:keli (.) å ju eldre man **blei** destu je:vlägare **blei** de man vild inga bo: dä:r sen **meira** (Pa.omo3.sp)*

det kommer jag ihåg att det var kallt rysligt (.) man frös så sakeli (= satan) (.) och ju äldre man blev desto djävligare blev det man ville inte bo där sedan mera

- (2.4) *nu **veit** ja jo at ja får å:ka me me andra men int e de sama sa:k (.) he he e **heilt** annårlonda (Vf.owo1.sp)*

nu vet jag ju att jag får åka med med andra men inte är det samma sak (.) det det är helt annorlunda

öy

Liksom *ei*-diftongen har *öy*-diftongen funnits i hela Åboland (Zilliacus 1992, 48). Variabeln är dock så ovanlig att det i ett så litet material som 20 minuter per informant är svårt att hitta belägg på den hos alla informanter. Av ord med potentiell *öy*-diftong är verbet *sköta* med avledningen *skötare* i olika sammansättningar vanligast (11 informanter). Substantiven *ö* och *hö* finns hos några informanter och verbformen *dör* hos enstaka. I materialet finns bara några få belägg på diftonguttal. Se *hö* i exempel 2.1 ovan samt exemplen nedan.

- (2.5) *vi hadd ein sån dä: veivtjä:rno sām va li:te möjligan vi vild inga stå å **stöit** (Ho.owo1.sw)*

vi hade en sådan där vevkärna som var lite möjligare vi ville inte stå och stöta

- (2.6) *så hö:rd vi at de kām bombpla:n å vi hadd jo eld de va på vintärn å vi hadd laga bra:sa (.) för vi sko värma åt åss (.) vi hadd mjölk me: at vi sko få värmd den at de he va så kalt å (.) vi fikk no:g lå:v å slekk bra:san att int dåm sko si: att vi hadd (.) de rö:k jo åpp **røyken** så: dåm (.) när vi va i sko:gen å (..) men sen nä:r dåm fo:r vi:dare sen så (.) så laga vi bra:sa på nytt å så värmd vi vå:r mjölk å så (.) hadd vi en li:tn paus å så börja vi hogga ijen (Ki.owo1.sp)*

så hörde vi att det kom bombplan och vi hade ju eld det var på vintern och vi hade lagat brasa (.) för vi skulle värma åt oss (.) vi hade mjölk med att vi skulle få värmd den att det var så kallt och (.) vi fick nog lov att släcka brasan att inte de skulle se att vi hade (.) det rök ju opp röken såg de (.) när vi var i skogen och (..) men sedan när de for vidare sedan så (.) så lagade vi brasa på nytt och så värmd vi vår mjölk och så (.) hade vi en liten paus och så började vi hugga igen

(2.7) *nå ja bru:kar vara u:t i engesöyn nestan nu på såmmarn då de ha vari fi:nt* (Df.omo2.sp)

nå jag brukar vara ut i Ängesön nästan nu på sommaren då det har varit fint

Den ovannämnda akustiska analysen av material insamlat inom *SweDia*-projektet gav tydligt utslag för bevarad öy-diftong i ordet *dör* i Houtskär men däremot inte i Dragsfjärd (Leinonen 2010, 99). I Dragsfjärd eliciterades även ordet *rök* vid *SweDia*-inspelningarna. Detta ord ingick inte i den akustiska analysen (eftersom det inte eliciterats på alla orter), men en manuell genomgång visar att flera av Dragsfjärdinformanterna uttalar *rök* med diftong trots att *dör*, som eliciterades direkt efter *rök*, uttalas med monoftong av de flesta informanterna. Bland de yngre informanterna är diftongen ovanligare än bland de äldre, och några av de yngre informanterna säger under inspelningen att de känner till uttalet med diftong men att de själva inte använder det.

2.1.2 Diftongering av långa vokaler

Ursprungliga långa vokaler som uttalas diftongerat kallas sekundära diftonger. De vanligaste sekundära diftongerna i Åboland är enligt Zilliacus (1992, 52–55) *oå/oa* (< långt *å*), *åo* (< långt *o*) och *yö* (< långt *ö*). Sekundära diftonger har funnits framför allt på Kimitoön och i Hitis samt i utskären i Nagu och Korpo. Dessutom diftongeras långt *o* > *åo* i Pargas. De sekundära diftongerna har alltså haft en östlig och sydlig utbredning i Åboland. På övriga håll har sekundära diftonger varit ovanliga. Förutom de nämnda diftongerna har diftongering av långt *e* > *ie* förekommit i Kimitosocknarna och diftongering av långt *u* > *yu/öu* i sydligaste Åboland.

Den enda av de sekundära diftongerna som det finns mer än några ströbelägg på i 2000-talsmaterialet är *åo*-diftongen, som behandlas nedan.

Vad gäller övriga sekundära diftonger finns bara några enstaka belägg:

- (2.8) *så de **bliev** ju slu:t me den där (..) vassbuksfisken at (.) de va inggen rikit såm tjöpt någå såna dä:r (.) sto:ra mengdär me:ra*
(Ko.omo2.sp)

så det blev ju slut med den där (..) vassbuksfisken att (.) det var ingen riktigt som köpte några sådana där (.) stora mängder mera

- (2.9) *dä:r såm du **viek** åpp hi:t ti kå- ti kå:rkollave:gen så dä:r va et mjölkbo:rd såm alla fö:rd nä:r de sen blei bi:ltranspårt ijen*
(Ki.omo1.sp)

där som du vek upp hit till kå- till Kårkullavägen så där var ett mjölkbord som alla förde när det sedan blev biltransport igen

Dessutom finns följande kommentar av den äldre informanten från Västansfjärd:

- (2.10) *tili:ka så tykkär ja no: att att de ha (.) att de ha blivi sen sen ja va li:tn så no (.) no ha e ju kansje endra li:te (.) ett eksempel åm ja no sko sejj så (.) så kond man no sejj til eksempel at hon den mensjan hon ventar **boån** hon (.) kond man sejj då: (..) men no sko man no sejj no att de:- att den mensjan (.) ventar **ba:rn** (.) sko man sejj nu:*
(Vf.owo1.sp)

tillika så tycker jag nog att att det har (.) att det har blivit sedan sedan jag var liten så nog (.) nog har det ju kanske ändrat lite (.) ett exempel om jag nu skulle säga så (.) så kunde man nog säga till exempel att hon den människan hon väntar barn hon (.) kunde man säga då (..) men nog skulle man nog säga nu att de- att den människan (.) väntar barn (.) skulle man säga nu

Enligt informanten har alltså diftongen som användes när hon var liten ersatts av lång vokal (dessutom har fonemet *å*: i detta ord bytts ut mot det standardspråkliga *a*:). Sekundära diftonger lyser så gott som helt

och hållet med sin frånvaro bland yngre informanter på hela Kimitoön i 2000-talsmaterialet. Detta dialektdrag verkar alltså ha försvunnit och ersatts av standardspråkets långa vokal.

Långt o

Den bäst bevarade av de sekundära diftongerna är som sagt *åo*-diftongen, som är kännetecknande för Pargasdialekten. Vanligast är denna diftong i småorden *jo* och *nog*, men dessa ord har så svag betoning att det ofta är svårt att avgöra om det är fråga om en diftong eller en lång vokal. Därför har analysen begränsats till mer innehållstunga ord.

Långt *o* är ett så vanligt fonem i språket att alla informanter har många belägg på variabeln (i orden *stor*, *for*, *gjorde/gjort* etc.). Diftonger är dock relativt sällsynta. Ingen av informanterna har konsekvent diftongering av långt *o*, utan även de informanter som har belägg på diftongen använder betydligt oftare ett monofongiskt än ett diftongerat uttal. Det är framför allt tre informanter som har lite flera belägg på *åo*-diftongen: Pa.omo3.sp (3 belägg), Pa.ywo1.sp (6 belägg) och Vf.owo1.sp (7 belägg). Övriga informanter, av vilka de flesta är från Pargas, har bara enstaka belägg. Av de totalt åtta informanterna från Pargas har fem stycken åtminstone ett belägg på *åo*-diftongen. Diftongen verkar alltså finnas i dialekten, men åtminstone i föreliggande intervjumaterial intar den en mycket liten plats.

- (2.11) *lä:rarn så han va myki intrese:ra av jumnasti:k å sånt så vi hadd åfta sån hä:na pyrami:där (.) vi stelde åpp i o:lika sån hä: fä:masjäonär å sto: på varandra å sån hä:nt (Pa.omo4.sp)*

läraren så han var mycket intresserad av gymnastik och sådant så vi hade ofta sådana här pyramider (.) vi ställde upp i olika sådana här formationer och stod på varandra och sådant här

- (2.12) *sko vell va: förståss å (.) ha en (.) ståorär jenkare å bara kroisa åmkring här i (.) pargas (Pa.ymo5.sp)*

skulle väl vara förstås att (.) ha en (.) stor jänkare och bara cruisa omrking här i (.) Pargas

- (2.13) *så va dām no å dansa me va:r å varannan dā:r å (.) sen så ble: ju alla tilreklit slu:t at dām liksåm årka inga dām vila ta: de li:te longt imellan å (.) så va dām hä: svå:grana å to: va:rsin ståol (.) å så sto: dām dā:r mitt på gålve å dansa me dām (Pa.ywo1.sp)*

så var de nu och dansade med var och varannan där och (.) sedan så blev ju alla tillräckligt slut att de liksom orkade inte de ville ta det lite lugnt emellan och (.) så var de här svågrarna och tog varsin stol (.) och så stod de där mitt på golvet och dansade med dem

- (2.14) *man måst jo få ha pla:när fast man int få:r e jäort hellär (Vf.owo1.sp)*

man måste ju få ha planer fast man inte får det gjort heller

Akustisk analys av den långa *o*-vokalen i ordet *sot* visar att ett diffongerat uttal förekommer bland äldre informanter från Dragsfjärd men saknas i den yngre generationen (Leinonen 2010, kartor på s. 119 och 232).

2.1.3 *a* och *ä* före *ld*, *mb*, *nd* och *ng*

Före konsonantkombinationerna *ld*, *mb*, *nd* och *ng* har vokalerna *a* och *ä* förlängts i dialekterna, vilket har lett till uttal som *la:nd* 'land' och *ma:ngg* 'många'. Zilliacus (1992, 105–107) menar att denna förlängning skett över hela det finlandssvenska språkområdet och att kortvokaliska former senare (p.g.a. en allmän tendens till vokalförkortning före konsonantkombinationer) har återställts på många håll. I Åboland menar Zilliacus att långa vokaler har bevarats i väster, söder och öster, medan mellersta Åboland har kortvokaliska former.

I det svenska standardspråket har *a* före *ng* vidare gått över till *å* (*mangga* > *ma:ngga* > *må:ngga* > *mångga*). I vissa fall har det också skett före *ld* (*hålla* < *halda*, men *kall* < *kald*). Enligt Zilliacus har vokalkvaliteten bevarats som *a* i hela Åboland.

Belägg på bevarade långa *a* och *ä* finns i riklig mängd hos de yngre informanterna från Iniö och Houtskär samt hos de två äldre informanterna från Björkö (Houtskär). Lång vokal finns hos dessa talare t.ex. i *kve:ld* 'kväll', *ba:nd* 'band', *ha:nd* 'hand', *le:ng* 'länge', *ma:ng* 'många', *spje:lde* 'spjället'. Från övriga orter finns bara några få belägg: *e:ld* 'eld'

och *kve:ldn* 'kvällen' (Na.omo2.sp), *ila:nd* 'iland' (Pa.omo4.sp) samt *einfa:ldåga* 'enfaldiga' (Pa.ywo1.sp). I dessa få belägg från övriga orter handlar det om konsonantkombinationerna *ld* och *nd*. Zilliacus (1992, 18–19) har noterat att förkortningstendensen har varit starkare före *ng* än före *ld* och *nd*.

2.1.4 Bevarat *a*

Gammalt långt *a* har i standardsvenska gått över till *å* (t.ex. *ga*: > *gå*:). I Åboland har *a* bevarats i vissa ställningar, bl.a. före *ng* där vokalförlängningen enligt Zilliacus (1992, 41) har skett efter att övergången från *a* till *å* har upphört (jfr ovan). Eftersom potentiellt bevarat *a* förekommer i några frekventa ord har variabeln kvantifierats.

Belägg på åtminstone något av de frekventa orden *gång*, *lång* och *många* med böjningsformer finns hos alla informanter. I regel följs vokalkvalitet och längd åt i dessa ord, så att vokalen antingen är ett långt *a* eller ett kort *å*, men belägg med kort *a* förekommer också i viss utsträckning. Endast de tre yngre informanterna från Iniö och Houtskär och de två äldre informanterna från Björkö (Houtskär) har belägg på *a* i dessa ord. I gengäld är andelen *a* mycket hög hos dessa informanter (100 procent av fallen hos alla utom Io.ywo2.sp som har *a* i 91 procent av fallen).

- (2.15) *he va int så ma:ng gangor i vekkan såm du fo:r å handel utan tu ringd istelle te handli å så kām e en lå:do me me bå:t di: å så fo:r du ne:r eftär den ti ti bro- bro:en* (Io.ywo2.sp)

det var inte så många gånger i veckan som du for och handlade utan du ringde i stället till handeln och så kom det en låda med med båt dit och så for du ner efter den till till bro- bron (= bryggan)

- (2.16) *he va en ga:nggo (.) ja vi hold på: me höibärngningen* (Ho.omo1.sw)
det var en gång (.) ja vi höll på med höibärngningen

- (2.17) *he ha gatt (.) så **la:ngt** elär va vi ska sejj att (.) bö:r ju sjakt bårt snö:in ren från å:kran liksäm då so:len bö:r värm så fa:r däm me traktårin å sjaktar bårt snö:in* (Ho.omo1.sw)

det har gått (.) så långt eller vad vi ska säga att (.) börjar ju schakta bort snön redan från åkrarna liksom då solen börjar värma så far de med traktorn och schaktar bort snön

För analysen av variationen mellan *a* och *å* har dessutom pronomenet *någon* excerperats. Alla informanter har minst fem belägg på *någon* om man räknar med alla böjningsformer. Pronomenet har enligt Zilliacus (1992, 82) haft följande former i de åboländska dialekterna:

na(in) karl, naen kvinna, na barn, plur. nagär barn

I detta ord har nio av de 43 informanterna belägg på *a*-former. Av dessa har fyra (Io.omo1.sp, Io.owo1.sp, Ko.ymo2.sp och Na.omo2.sp) en mycket låg andel *a*-former och använder oftast *å*-former. Fem av informanterna har däremot en betydligt högre andel *a*-former än *å*-former. Dessa är de yngre informanterna från Iniö och Houtskär och de två äldre informanterna från Björkö (Houtskär) (dvs. samma informanter som har bevarat *a* i orden *gång, lång, många*). Alla fem använder *a*-former i över 75 procent av fallen.

- (2.18) *ja mins ju att man va bårt från sko:lan då unjejä:r **nain** da:g åm man va te å:bo me mamma å pappa* (Ho.ywo1.sp)

jag minns ju att man var borta från skolan då ungefär någon dag om man var till Åbo med mamma och pappa

- (2.19) *ja va no:g ganska gamol innan ja hö:rd **na** finsko als överhuvuta:je* (Ho.owo1.sw)

jag var nog ganska gammal innan jag hörde någon finska alls överhuvudtaget

- (2.20) *ja hadd två: så:ar fulla me simpår dä:r va vel se:kärt en åtta ni:e hundra (..) på **nagär** ne:t* (Ko.ymo2.sp)

jag hade två såar fulla med simpår där var väl säkert en åtta nio hundra (..) på några nät

Vid kvantifieringen av variabeln 'bevarat *a*', som ingår i den sammanlagda kvantitativa analysen (kapitel 4), har beläggen för de fyra orden *gång*, *lång*, *många* och *någon* räknats ihop.

2.1.5 Kort *u*

Dialekter där kort *u* uttalas som *o* brukar kallas *o*-mål. Alla åboländska socknar utom Korpo och Nagu har hört till *o*-målsområdet (Wessman 1936, 6–7). Det korta *u*-ljudets historiska utveckling är dock rätt komplicerad (utvecklingen i de finlandssvenska dialekterna har utretts i Wessman 1936 och 1938). I Åboland har, enligt Zilliacus (1992, 28–29), *u* öppnats till *å* framför långa konsonanter och konsonantkluster utom då den efterföljande konsonanten är *g*, *l*, *n* eller *ng*. Exempel på detta är uttal som *gåbb(e)* 'gubbe' och *båska* 'buske'. För variabeln 'kort *u*' har följaktligen belägg excerperats i betonad stavelse framför ursprungligt långt eller av annan konsonant följt *g*, *l*, *n* eller *ng*.

Enligt Zilliacus (1992, 28) ligger uttalet av det korta *u*-ljudet i Korpo och Nagu på gränsen mellan *u* och *o*, vilket bekräftas då man lyssnar på 2000-talsinspelningarna. Eftersom vokalen ofta är mycket kort kan det vara svårt för den som transkriberar att avgöra den exakta vokal-kvaliteten. Eftersom transkriptionerna dessutom är gjorda med en grov ljudskrift har det inte varit möjligt att välja ett vokaltecken mitt emellan *u* och *o*, utan transkriberaren har tvingats klassificera beläggen som antingen *u* eller *o*. För en mer exakt analys av variabeln skulle en akustisk analys antagligen lämpa sig bättre. Resultaten nedan måste därför ses som tentativa.

Det totala antalet belägg på variabeln är 637 med ett medianvärde på 14 belägg per informant. Beläggen är excerperade ur sammanlagt 176 olika ord, varav 20 förekommer minst 5 gånger i materialet. Det mest frekventa ordet är preteritumformen *kunde* som står för 103 av beläggen. Utdragen nedan ger exempel på *o*-former i kontext:

(2.21) *man kan vel sejj at de fans en tri: fy:ra hondra ko:r* (Io.omo1.sp)

man kan väl säga att det fanns en tre fyra hundra kor

(2.22) *mång gångär va de no: så: at vi glömd bårt ti he dä ve:dhoggande*
då: å å fikk lite gre:lår åkså (Pa.omo4.sp)

många gånger var det nog så att vi glömde bort att det där vedhuggandet då och och fick lite grälör också

- (2.23) *man kond fa: li:te u:t å å:ka polka å sji:da* (Pa.owo3.sp)

man kunde fara lite ut och åka pulka och skida

- (2.24) *å: så: pressadis de å tårkadis å så sattes (.) å brendes dām i en i en ongn såm byggdes då: de ble:v två: ongnar sennan i: (.) i en gammal ladugå:rd (.) i lammkolla ladugå:rden så: ble:v de två: ongnar såm (.) såm såm brende te:gel elår rö:- rör: i jentligen sen fö tekkdi:kningen (.) men de fo:r bårt sennan fö de ble:v jo plast sennan* (Ki.omo1.sp)

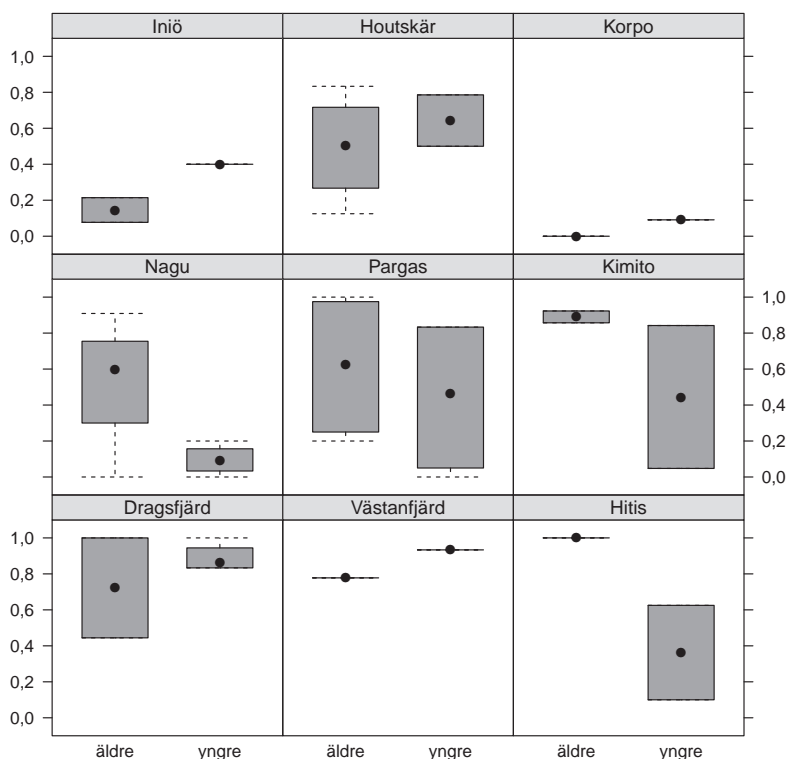
och så pressades det och torkades och så sattes (.) och brändes de i en i en ugn som byggdes då det blev två ugnar sedan i (.) i en gammal ladugård (.) i Lamkulla ladugården så blev det två ugnar som (.) som som brände tegel eller rö- rör egentligen sedan för täckdikningen (.) men det for bort sedan för det blev ju plast sedan

- (2.25) *int va de ju frå:ga åm at de aldri sko ha vari bara e:n elår två: pārso:når på nå:n klass int u:tan at no va de jo alla klassår sån dāna folla klassår* (Df.ywo1.sp)

inte var det ju fråga om att det aldrig skulle ha varit bara en eller två personer på någon klass inte utan att nog var det ju alla klasser sådana där fulla klasser

Lådagrammen i figur I.5 visar andelen *o*-uttal efter ort och åldersgrupp. Andelarna är överlag relativt höga, vilket visar att detta dialektdrag är väl bevarat. Variationen är påfallande stor i Pargas där andelen *o*-uttal varierar mellan 0 och 100 procent bland informanterna. I Korpo, där dialekten har hört till *u*-målen, saknas belägg på *o*-uttal nästan helt. I Nagu som också haft *u*-mål är *o*-uttal däremot ganska vanligt, framför allt bland de äldre informanterna. Detta kan tyda på att *o*-uttalet är ett dialektdrag som har spridit sig till Nagu från de omkringliggande dialekterna, men det kan också bero på svårigheten att vid transkriberingen klassificera vokaler tillförlitligt.

För att testa huruvida variationen mellan *u* och *o* är lexikalt betingad utfördes en logistisk regressionsanalys med blandade effekter



Figur 1.5. Lådagram över andelen dialektvarianter (o) av variabeln 'kort u' efter ort och ålder.

(se bilaga A, 1.2). Av de utomspråkliga variablerna uppvisar ort och kön signifikanta skillnader. Mest utbrett är *o*-uttalet i Dragsfjärd och Västanfjärd och minst förekommer det i Korpo och Nagu. Bland män är *o*-uttal vanligare än bland kvinnor. Mellan de två åldersgrupperna finns däremot ingen signifikant skillnad. Modellen förbättras inte signifikant av att man lägger till ord som stokastisk effekt (likelihood-kvottest, $p = 0,201$), vilket tyder på att benägenheten till *o*-uttal inte varierar mellan olika ord utan det handlar om en rent fonologisk variabel.

2.2 Konsonanter

I avsnitten nedan behandlas olika typer av konsonantkombinationer samt förmjukning av *g* och *k*. Två av de variabler som har med konso-

nantuttal att göra har kvantifierats, nämligen *ng*-ljudet (avsnitt 2.2.1) och norrländsk förmjukning (2.2.6).

Ytterligare särdrag i konsonantuttalet som nämns av Zilliacus (1992) men som inte behandlas mer ingående nedan är bl.a. tonlösa sonoranter och palatalisering av *l* och *n*. De tonlösa sonoranterna förekommer framför *p*, *t* och *k* och har varit särskilt vanliga i sydligaste Åboland (Zilliacus 1992, 66). I det föreliggande materialet har tonlösa sonoranter noterats åtminstone hos informanterna Ho.omo1.sp, Ko.omo2.sp och Ko.owo3.sp i ord som *allt*, *klant*, *sumpa*, men draget har inte excerperats systematiskt.

Palatalisering av *l* och *n* har förekommit i Iniö och Houtskär (Zilliacus 1992, 67). Palataliseringen har gällt långt *l* och *n* och har varit vanlig särskilt i den maskulina slutartikeln, t.ex. *bå:tin* 'båten'. Entydiga exempel på palataliserat uttal är svåra att hitta i det aktuella materialet. Den maskulina slutartikeln har i många fall reducerats till ett *-i* (se 3.1), och i de fall då *n*:et står kvar är det i regel fråga om ett vanligt alveolart *n*.

2.2.1 *ng*-ljudet

Dialekterna i Åboland har, liksom de övriga svenska dialekterna i Finland, kännetecknats av ett *ng*-ljud med hörbart *g* (Wessén 1969, 46). Bergström (1984, 43, 48) har konstaterat att *ng*-ljud med hörbart *g* i Mattkärrdialekten (Kimito) på 1980-talet endast förekom hos äldre personer, vilket tyder på att draget är på väg att försvinna.

Ivars (1986, 122 och 1988, 118) har för Närpesdialektens del konstaterat att *g*:et är hörbart bara då *ng*-ljudet följs av en vokal. Vidare har hon (1986, 122) poängterat att betoning och fonotaktisk ställning (inljud/slutljud) är faktorer som kan påverka om ljudet realiserar med eller utan hörbart *g*. Beläggen på *ng*-ljud i materialet har därför delats in i inljud och slutljud, respektive om de står i en betonad eller obetonad stavelse. Belägg före konsonant i inljud har inte inkluderats (t.ex. *långt*, *ringde*). Vidare har belägg på *ng* i slutljud uteslutits i fall där följande ord inleds med *g* eller *k*, eftersom det i sådana fall är omöjligt att avgöra om ett hörbart *g* är en del av *ng*-ljudet eller om det hör till följande ord. Ord som har (eller kan ha) ett uttal med hörbart *g* även i standardspråket har inte tagits med i analysen (*bingo*, *engagera*, *engagemang*, *fungera*, *tango* och dessutom *evangelium* och *ungefär* som har uttalet *-nj-*; Garlén 2003).

Exempel 2.26 visar belägg på *ng*-ljud med hörbart *g* i betonad stavelse både i slutljud (*ung*) och i inljud (*unga*):

(2.26) *ja no va de bra: ti va ongg (.) då: nä ja va (.) ellär nä vi va ongga bru:kar ja sejj* (Hi.owo1.sp)

ja nog var det bra att vara ung (.) då när jag var (.) eller när vi var unga brukar jag säga

Belägg i obetonad stavelse finns i exempel 2.27 (slutljud) och exempel 2.28 (inljud). I exempel 2.28 finns dessutom ett belägg i negationen *inga*.

(2.27) *ja: de va: de va: jedda (.) såm man fiska å (.) å så va de vell li:te strömming å: vassbu:k* (Ko.omo2.sp)

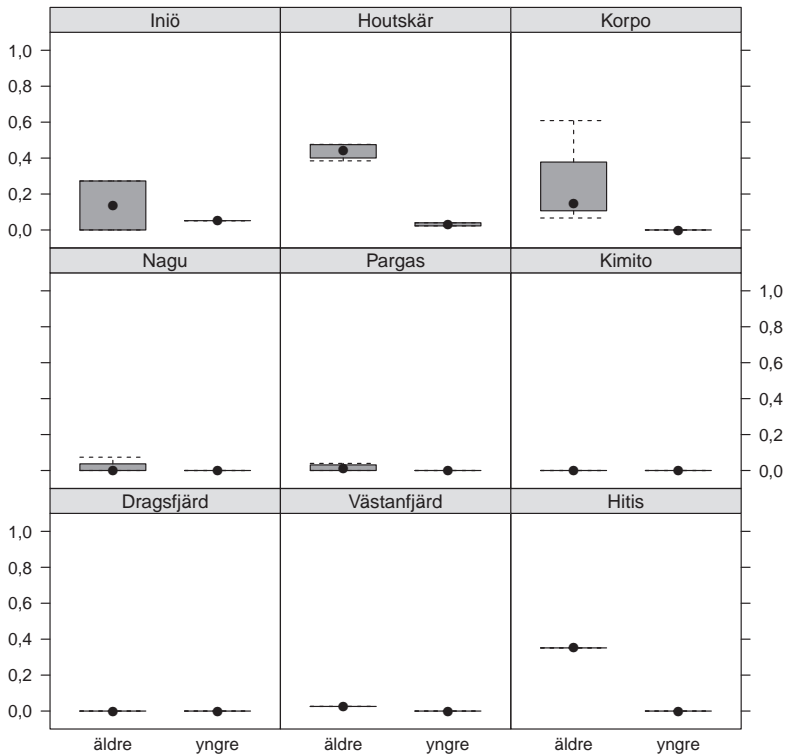
ja det var det var gädda (.) som man fiskade och (.) och så var det väl lite strömming och vassbuk

(2.28) *de ha vari nån gång såm dām ha hitta e:n he:l massa me: me: gu:dīnggar såm såm hadd vari dö:da på såmmarn så (.) men va de ha va dām hadd dött a:v vist man ingga elär så: int* (Ho.omo1.sp)

det har varit någon gång som de har hittat en hel massa med med gudingar som som hade varit döda på sommaren så (.) men vad det har vad de hade dött av visste man inte eller så inte

Det totala antalet belägg på variabeln i materialet är 1 992, varav endast 174 med hörbart *g* (en kontrollsyssning har utförts av samtliga belägg eftersom detta drag inte har transkriberats konsekvent i de ursprungliga transkriptionerna). Antalet belägg per informant varierar mellan 23 och 101, och medianvärdet är 42. Endast 16 av de 43 informanterna har belägg på uttal med hörbart *g*. Figur I.6 visar andelen dialektvarianter (*ngg*) efter ort och ålder. Som synes är andelen *ngg*-uttal överlag liten och hos de yngre informanterna rentav försvinnande liten.

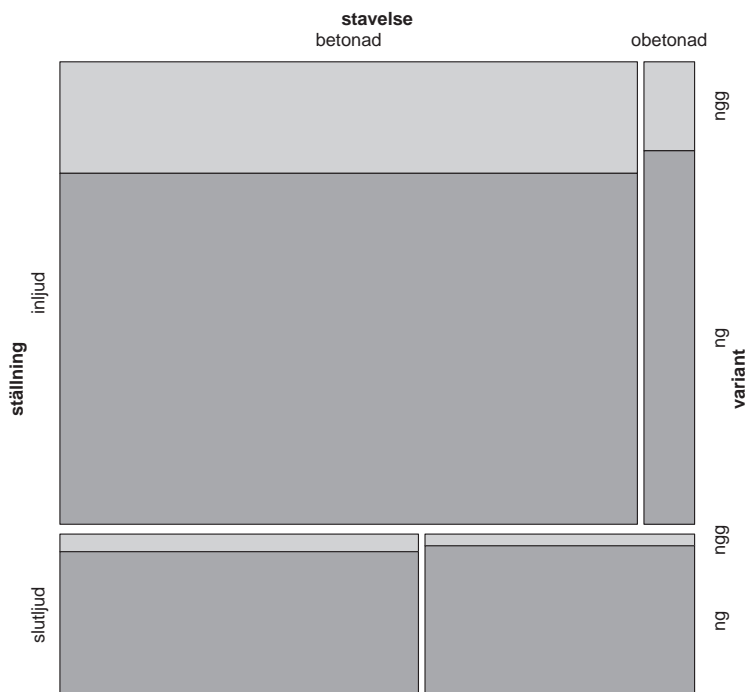
Figur I.7 visar fördelningen mellan uttalen *ng* och *ngg* efter de inomspråkliga faktorerna. Bilden bygger på material endast från de 16 informanter som har belägg på *ngg*-uttal. Lådornas storlek står i proportion till antalet belägg i de olika kategorierna och de gråa fälten anger andelen *ng* respektive *ngg*. Bilden visar att uttalet *ng* är betydligt



Figur 1.6. Lådagram över andelen dialektvarianter (*ngg*) av *ng*-ljudet efter ort och åldersgrupp.

vanligare än *ngg* även hos dessa 16 informanter. Vidare ser man att *ngg*-uttal är vanligare i inljud än i slutljud (den ljusa nyansen fyller en större andel av lådan vid inljud än vid slutljud). Betoningen verkar också spela en viss roll både i inljud och i slutljud – andelen *ngg* är något större i betonad stavelse än i obetonad, men skillnaderna i fråga om betoning är små. Vanligast ser *ngg* ut att vara i betonad stavelse i inljud, ovanligast i obetonat slutljud.

En logistisk regressionsanalys (bilaga A, 1.3) med informant som stokastisk effekt utfördes för att testa de olika faktorernas betydelse för variationen mellan *ng* och *ngg*. Alla 43 informanter ingår i analysen. Modellen visar att faktorerna ålder, ställning i ordet och betoning är signifikanta. Uttal med hörbart *g* är vanligare bland äldre talare än bland yngre, vilket antyder att draget håller på att försvinna. Faktorerna ort

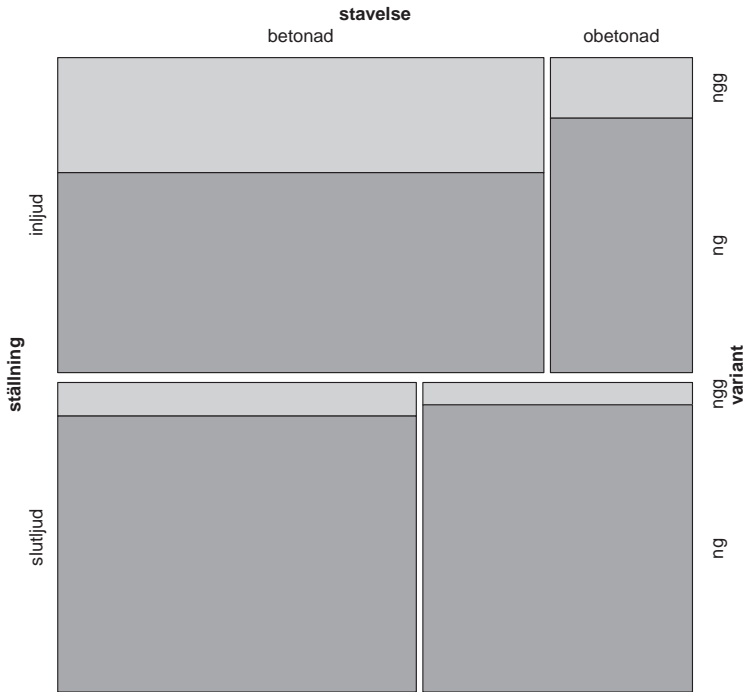


Figur I.7. Mosaikdiagram över variationen *ngg* ~ *ng* i inljud och slutljud samt i betonad och obetonad stavelse. Material bara från de 16 informanter som har belägg på dialektvarianten. Lådornas storlek står i proportion till antalet belägg i de olika kategorierna.

och kön är däremot inte signifikanta för variationen mellan *ng* och *ngg*.

Variabeln *ng(g)* förekommer särskilt ofta i ordet *inga*. Eftersom frekventa ord kan ha en annan fördelning än den genomsnittliga lades en stokastisk effekt för faktorn ord (med de två varianterna *inga* eller vilket som helst annat ord) in i modellen. Modellen förbättrades signifikant (likelihood-kvottest, $p < 0,001$) och visar att andelen *ngg*-uttal är lägre i *inga* än i övriga ord. Ordet *inga* går alltså stick i stäv mot regeln att *ngg* är vanligast i betonad stavelse i inljud. Om belägg på *inga* utesluts ur analysen blir skillnaderna i fråga om ställning i ordet och betoning ännu tydligare. Detta åskådliggörs av figur I.8 som upprepar figur I.7 men utan belägg på *inga*.

Variationen mellan uttalen *ng* och *ngg* påverkas alltså av en rad inomspråkliga faktorer. I den kvantitativa analysen som redovisas i kapi-



Figur 1.8. Mosaikdiagram över variationen *ngg* ~ *ng* i inljud och slutljud samt i betonad och obetonad stavelse. Material bara från de 16 informanter som har belägg på dialektvarianten och utan belägg på *inga*. I jämförelse med figur 1.7, där *inga* ingår, är skillnaderna i fråga om ställning i ordet och betoning tydligare i denna figur.

tel 4 har trots allt samtliga belägg på variabeln slagits ihop, eftersom antalet belägg per informant annars skulle sjunka drastiskt. Variabeln ger på så vis en något grov bild av variationen bland informanterna.

2.2.2 Bevarade konsonantkombinationer: *ld*, *mb*

Konsonantkombinationerna *ld* och *mb*, som fanns i fornspråket, har assimilerats till *ll* och *mm* i det svenska standardspråket.² I finlands-svenska dialekter har dessa konsonantkombinationer i stor utsträck-

2. Detta är i grunden samma typ av assimilation som leder till att det hörbara *g*:et i *ng*-ljudet försvinner.

ning bevarats (Harling-Kranck 1998, 139). Som exempel kan nämnas *hald(a)* 'hålla', *kväld* 'kväll', *muld* 'mull', *damb* 'damm' och *lamb* 'lamm'. Detsamma gäller *ng*-ljudet som i dialekterna har bevarats med ett slutljudande *g* (se föregående avsnitt).

Ord med potentiellt bevarat *mb* saknas så gott som helt i materialet. Däremot använder de flesta av informanterna under de 20 minuter långa klippen någon form av verbet *hålla* eller substantivet *kväll* som båda ursprungligen har haft *ld* i stället för det assimilerade *ll*. Belägg på uttal med *ld* av ordet *kväll* finns hos de yngre informanterna från Iniö och Houtskär och den äldre kvinnan från Iniö. Från övriga orter finns bara ett belägg med bevarat *ld* i ordet *kväll* hos en av de äldre Naguinformanterna. Denna använder först formen *kveln* och korrigerar sig sedan och använder den mer dialektala varianten *kve:ldn* (se 2.31 nedan).

- (2.29) *å sen på lö:rdakve:ldin bli:r e no n- no: nestan ti fa:r ti pubben*
(Ho.ywo1.sp)

och sedan på lördagkvällen blir det nog n- nog nästan att fara till puben

- (2.30) *vi ha:r va ha:r vi (.) tri: fy:ra dansär pär å:r (.) såm vi å:rnar å så ha:r vi försö:kt å:rn (.) såm spe:lkve:ldar* (Ho.ywo1.sp)

vi har vad har vi (.) tre fyra danser per år (.) som vi ordnar och så har vi försökt ordna (.) som spelkvällar

- (2.31) *no va de no en ti:e kilåme:tär ti skoff hem den dä mope:dn så han kåm hem (.) halv seks på kveln sen å s- (.) på kve:ldn å å sko fara ti (.) å då hadd mamma fari ti föyse* (Na.omo2.sp)

nog var det nog en tio kilometer att skuffa hem den där mopeden så han kom hem (.) halv sex på kvällen sedan och s- (.) på kvällen och och och skulle fara till (.) och då hade mamma farit till fähuset

Alla infinitivformer av verbet *hålla* i materialet uttalas med assimilerat *ll* (sammanlagt 26 belägg) i stället för *ld*. I presens finns belägg med *ld* hos två informanter (Io.ywo2.sp, Ho.owo1.sw). I preteritum av *hålla* finns flera former med *ld* (*hold/håld*). Men för preteritumformerna är det svårt att avgöra om det är frågan om en äldre stark böjning med bevarat *ld* (se t.ex. FO *hålla* v.1 och Bergström 1984, 95) eller om en

nyare böjning enligt svag böjningsklass, där *d*:et är ändelsemorfem i preteritum. Bergström (1984) nämner att *hålla* kan ha både stark och svag böjning i Mattkärrdialekten. I materialet finns fem belägg på supinumformen *hålt* (4 från Kimitoön och 1 från Pargas), vilket indikerar en svag böjning av verbet.

- (2.32) *he va vel na:n ga:ngg såm (.) han ti:dins kāmuna:ldirektö:r hold för åss (.) jippokve:ldnar (Io.ywo2.sp)*

det var väl någon gång som (.) den tidens kommunaldirektör höll för oss (.) jippokvällar

- (2.33) *ja sku sejj at e no:g fins såna se:kärt såna hu:shåll såm tjö:pär brö:d (.) å sen ä e ju bra: at he fins i såna dä:r li:te brå:da ti:där såm i pä:ronti:din så (.) fålk haldär på: såm värst (.) kan he no: va: (.) va: bra: ti kon ga: buti:k- ti buti:ken ett kaffebrö:dä tå (Ho.owo1.sw)*

jag skulle säga att det nog finns sådana säkert sådana hushåll som köper bröd (.) och sedan är det ju bra att det finns i sådana där lite bråda tider som i potatistiden så (.) folk håller på som värst (.) kan det nog vara (.) vara bra att kunna gå butik- till butiken efter kaffebrödet då

För övriga ord än *kväll* och *hålla* finns bara några enstaka belägg på bevarat *ld*.

- (2.34) *sen så måst de ju ha glö:d innan i sju:tär i spje:lde (Io.ywo2.sp)*

sedan så måste det ju ha glöd innan en skjuter i spjället

- (2.35) *fålk hadd ju tjällarar på ana vi:s i å:bo tå: såm dām ha:r nu: nu: ha:r dām ju inga na: ka:lida u:try:men me:ra i tām den hö:ghu:sen (Ho.owo1.sw)*

folk hade ju källare på annat vis i Åbo då än de har nu nu har de ju inte några kalla utrymmen mera i de där höghusen

I den mån konsonantkombinationen *ld* finns bevarad i dialekterna i Åboland, verkar det vara bara i de västligaste delarna.

2.2.3 Inskottskonsonanter: *ld, mb, nd*

Bevarade runsvenska inskottskonsonanter har varit kännetecknande för dialekterna i hela det finlandssvenska språkområdet (Wessén 1969, 47; Ivars 1988, 87). I runsvensk tid fogades *d* in mellan *l* och *r* och mellan *n* och *r*, och *b* infogades mellan *m* och *l* och mellan *m* och *r*. Inskottskonsonanter ses i presensformer som *tjendär* 'känner' och *kåmbär* 'kommer'. I dessa former har ändelsevokalen lagts till senare, och utvecklingen har alltså varit *komr* > *kombr* > *komber*. Inskottskonsonanter ses också i uttal som *gambel* 'gammal' (av *gaml*) och *sembär* 'sämre'.

Ord i presensform med potentiell inskottsvokal finns hos alla informanter i materialet (t.ex. ordet *kommer* förekommer flera gånger hos alla informanter), men inskottskonsonanter är ytterst sällsynta. Belägg på inskottskonsonant i presensformer finns bara hos följande sex informanter:

- Io.owo1.sp: *kåmbär* 'kommer'
- Io.ywo2.sp: *brindär* 'brinner', *fylldär* 'fyller', *kåmbär* 'kommer'
- Ho.omo1.sw: *kåmbär* 'kommer'
- Ho.owo1.sw: *brendär* 'bränner'
- Ho.ywo1.sp: *kåmbär* 'kommer'
- Na.omo2.sp: *stembär* 'stämmer'

Ingen av dessa sex informanter har konsekvent konsonantinskott i denna ordtyp, utan också uttal utan inskottskonsonant förekommer.

Också i övriga ordtyper saknas inskottskonsonanter nästan helt. De enda informanterna som har belägg på inskottskonsonanter i andra typer av ord än i presensformer är de två äldre informanterna från Björkö (Houtskär).

(2.36) *mi he värkar no såm e sku bli: barå se:mbär å se:mbär me me hela fålkhö:gsko:lan* (Ho.omo1.sw)

men det verkar nog som det skulle bli bara sämre och sämre med med hela folkhögskolan

(2.37) *alla tåm dä:r sto:r ga:mbel trebå:tan så fösvann ju* (Ho.owo1.sw)

alla de där stora gamla träbåtarna så försvann ju

Vid konsonantkombinationen *lr* är regeln om inskottskonsonanter fortfarande aktiv och leder till uttal som *aldra* 'allra', *heldre* 'hellre', *skvaldra* 'skvallra', *taldrik* 'tallrik' etc. I detta fall är inskottskonsonanten fonetiskt betingad och innebär uttalslättnad.

2.2.4 *dj*-

I svenskan i Finland, särskilt södra Finland, har konsonantkombinationen *dj*- bevarats (Ahlbäck 1956, 37–38). I många finlandssvenska dialekter liksom i sverigesvenskan har *d*:et i denna konsonantkombination dock fallit bort. Av de åboländska dialekterna har endast dialekterna i Nagu och Pargas bevarat *dj*- (Ahlbäck 1956, 12; Zilliacus 1992, 69). Eftersom uttal med hörbart *d* är det vanliga bl.a. i Helsingforssvenskan verkar *dj*- återinföras i vissa dialekter. Enligt Bergström (1984, 31) saknades *dj*- hos äldre personer i Kimito på 1980-talet, men var allmänt bland yngre talare.

Denna konsonantkombination är sällsynt i materialet, eftersom bara ett fåtal arvord i svenskan har *dj*-. Dessa ord är *djup*, *djur*, *djäkne*, *djärv* och *djävul/djävul* samt avledningar av och sammansättningar med dessa ord (SO).

I materialet förekommer *djup* (i formerna *djup*, *djupare*, *djupet*) hos två informanter och *djur* hos nio informanter. Ordet *djävul* med alla slags avledningar, inklusive former med *k* (se FO *djängel*, *djäkla*, *djäklas*, *djäklig*), finns hos nio informanter. Sammanlagt har endast 14 av de 43 informanterna belägg på något av *dj*-orden. Dessa informanter kommer från orterna Houtskär, Korpo, Nagu, Pargas och Kimito.

Ordet *djävul* med dess avledningar uttalas av samtliga informanter utan *d* (jfr Ahlbäck 1956, 38). Variation mellan *dj*- och *j*- finns således bara vid orden *djup* och *djur*. Informanterna från Houtskär använder konsekvent *j*-uttal av dessa två ord, medan det i Korpo finns uttal både med *dj*- och *j*-. Informanterna från Nagu, Pargas och Kimito använder konsekvent ett *dj*-uttal.

Materialet är för litet för att det ska gå att säga något entydigt om variabeln *dj*- i Åboland. Man kan försiktigt konstatera att *j*-former verkar finnas kvar längst i väster (dvs. i Houtskär och Korpo, från Iniö saknas belägg). Från Nagu och Pargas, som sedan gammalt har hört till *dj*-området, finns bara belägg på *dj*-uttal. Kimitoön har tidigare

hört till *j*-området, men de två Kimitoinformanterna som har belägg på variabeln har gått över till att använda det allmänfinlandssvenska *dj*-uttalet, vilket motsvarar Bergströms (1984) observation.

(2.38) *ennu åkså: få:r man (.) kan man ju få: en åm man ha:r u:t en s- fem seks ne:t så kan man ju få: en tju:gu tretti simpår [...] att dåm ha inga gatt u:t på ju:pe ennu* (Ko.ymo2.sp)

ännu också får man (.) kan man ju få en om man har ut en s- fem sex nät så kan man ju få en tju:go trettio simpør [...] att de har inte gått ut på djupet ännu

(2.39) *de fins ju int betre en att ba:rn ska få: veksa åpp me dju:r*
(Ki.ywo2.sp)

det finns ju inte bättre än att barn ska få växa opp med djur

2.2.5 Initialt *g*-, *k*- och *sk*-

Att *g*-, *k*- och *sk*- står oförmjukade i början av ord före främre vokal (*i*, *e*, *y*, *ä*, *ö*) har tidigare varit kännetecknande för dialekterna i Nagu och Pargas (Zilliacus 1992, 64). Övriga åboländska dialekter har liksom standardspråket haft förmjukat uttal. Bland ord där detta drag kan förekomma finns en rad vanliga ord som finns belagda under intervjuerna med de flesta av informanterna (t.ex. *ge*, *genom*, *göra*, *känna*, *köpa*, *köra* etc.). Trots att Nagu och Pargas är de två orter i materialet som representeras av flest informanter (7 respektive 8) är det få informanter som har belägg på oförmjukat uttal.³

Belägg på oförmjukat uttal finns bara hos följande fyra informanter: Na.om02.sp, Pa.om03.sp, Pa.ymo5.sp och Pa.ywo1.sp. Alla nämnda informanter har dessutom betydligt fler belägg på förmjukade varianter än på oförmjukade. Oförmjukade former av *g*-, *k*- och *sk*- förekommer alltså fortfarande i Nagu och Pargas, men detta dialektdrag har åtminstone i det föreliggande intervjumaterialet till största delen ersatts av standardspråkets förmjukade former.

3. Lånord från finska (*kiva*, *kännykkä*, *keikka*) eller andra språk (*geisha*, *kick*, *keyboard*, *kö*) har givetvis inte beaktats, liksom inte heller *kilometer* och *kiosk* som allmänt uttalas med hårt *k* i svenskan i Finland.

- (2.40) *en fastlasti:sda såm vi va: på på vi:ken å (.) å:ka slengkelka*
(Na.omo2.sp)
en fastlagstisdag som vi var på på viken och (.) åkte slängkälke
- (2.41) *ja int brukär vi jöra så myky vi brukär fa: u:t å (.) kö:ra me bi:ln å*
(.) å de hä:r å (.) hö:ra på musi:k å pra:ta ski:t (Pa.ymo5.sp)
ja inte brukar vi göra så mycket vi brukar fara ut och (.) köra med
bilen och (.) och det här och (.) höra på musik och prata skit
- (2.42) *å sen så ha:r man ju hela resten av såmarn såm man inga gä:r någä*
annat (Pa.ywo1.sp)
och sedan så har man ju hela resten av sommaren som man inte
gör något annat

Av de meddelare från Nagu och Pargas som har svarat på SLS frågelistor (SLS 1746 och SLS 1954) är det flera som nämner ”hårt k” som ett drag som är typiskt för dialekten. Det är alltså fråga om ett drag som dialektalarna själva är i hög grad medvetna om. Draget kan klassificeras som markerat (Trudgill 1986, 98) i och med att dess geografiska spridning varit begränsad till endast två av de åboländska sockendialekterna. Samtidigt är draget framträdande genom att det fonetiska avståndet mellan dialektvarianten och standardvarianten är stort (Trudgill 1986, 11; Kerswill & Williams 2002, 105; Schwarz et al. 2011, 19). Bägge faktorerna, begränsad geografisk spridning och faktumet att draget är perceptuellt tydligt framträdande, är sådana som bidrar till att ett dialektdrag lätt kan gå förlorat vid utjämning.

2.2.6 Norrländsk förmjukning

Norrländsk förmjukning innebär att *g* och *k* förmjukas i medial ställning framför främre vokal. Förmjukningen är särskilt vanlig i böjda ordformer:

<i>da:g</i>	>	<i>da:jín</i>	'dag, dagen'
<i>hogg(a)</i>	>	<i>hoddji</i>	'hugga, huggit'
<i>ta:k</i>	>	<i>ta:tji</i>	'tak, taket'
<i>flesk</i>	>	<i>flestji</i>	'fläsk, fläsket'

Liksom initialförmjukning (se föreg. avsnitt) har norrländsk förmjukning förekommit i hela Åboland utom Nagu och Pargas, men eftersom norrländsk förmjukning saknas i standardspråket har hårda former senare börjat tränga undan de förmjukade formerna (Zilliacus 1992, 64). Bergström (1984, 31) har konstaterat att norrländsk förmjukning var mycket vanlig i Kimito ännu på 1980-talet, men enligt Harling-Kranck (1998, 152) växlar hårda och förmjukade former i ord som *mycket* (*myky* ~ *mytjy*).

Förmjukade former är ytterst sällsynta i 2000-talsmaterialet. Att den norrländska förmjukningen verkar vara på väg ut ur dialekterna exemplifieras av informant Vf.owo1.sp. I informantens spontana tal finns inte ett enda belegg på norrländsk förmjukning. Belegg på ord med potentiell förmjukning är t.ex. *mycket* (uttal *myki*, *myky*, sammanlagt 7 belägg), *duken* (*du:ken*, 2 belägg) och *sängen* (*sengen*, 1 belägg). Intervjuaren frågar informanten explicit hur hon uttalar ordet *dike* och undrar om hon någon gång har sagt *di:tji*. Informanten svarar då:

(2.43) **di:tji** ja: (.) jo: att att man man eh ramlade i **di:tji** hellär elär den å den ramlade i **di:tji** (.) ja: i stället för **di:ke** (Vf.owo1.sp)

diket ja (.) ja att att man man eh ramlade i diket eller eller den och den ramlade i diket (.) ja i stället för diket

Informanten nämner att barnbarnen (som inte bor i Västanfjärd) brukar skratta åt detta uttal, och hon fortsätter direkt med ett annat exempel:

(2.44) *dåm bru:ka ta: e ti eksempel att (.) att att att mommo hon fö:r tettji ti ledon (..) ja: (.) att sånt så: så: sejjär man no int riktigt meira kansji int (.) men nu pra:tar no no no veit ja at ja sejjär många gångär sånt då:rnt (.) såna u:trykk åkså nåo* (Vf.owo1.sp)

de brukar ta det till exempel att (.) att att att mormor hon för täcket till ladan (..) ja (.) att sådant så så säger man nu inte riktigt mera kanske inte (.) men nog pratar nog nog nog vet jag att jag säger många gånger sådant där (.) sådana uttryck också nog

Utsagan kan kanske tolkas som att informanten själv ibland använder detta uttal, men att det överlag inte är särskilt vanligt längre. Talande är också ett belegg *røyken* 'röken' hos Ki.owo1.sp (se exempel 2.6). Av

de dialektala dragen i uttalet av detta ord är diftongen bevarad men inte förmjukningen.

Ett ord med potentiell norrländsk förmjukning som är mycket frekvent i materialet är ordet *mycket*. Alla informanter har belägg på detta ord. Antalet belägg per informant varierar mellan 3 och 28, med ett medianvärde på 8. Det totala antalet belägg i materialet är 392. Bara fem av informanterna har belägg på förmjukade former (*mytji*, *mytjy*). Dessa är Io.owo1.sp (som förmjukar 10 av de totalt 17 beläggen), Io.ywo2.sp (2/5), Ho.omo1.sw (5/6), Ho.owo1.sw (6/7) och Ho.ywo1.sp (1/11). Det är alltså tydligt att i den mån draget finns bevarat så är det främst på de två västligaste orterna.

När det gäller andra ord än *mycket* finns bara ströbelägg på norrländsk förmjukning i materialet.

(2.45) *nu få:r trafi:tjin kãm ijenãm för mejj dä:r no* (Io.owo1.sp)

nu får trafiken komma igenom för mig där nog

(2.46) *ja tro:r no:g (.) att att sku man läså li:te meir o:rd så (.) å va: tvonji ti pra:t engelska (.) så tro:r ja no:g på et å:r så så så kla:rar man vel no:g å:v e* (Ho.omo1.sw)

jag tror nog (.) att att skulle man läsa lite mera ord så (.) och vara tvungen att prata engelska (.) så tror jag nog på ett år så så så klarar man väl nog av det

(2.47) *i: å me: att ve:jin kãm så (.) alla tãm dä:r sto:r ga:mbel trebå:tan så fösvann ju* (Ho.owo1.sw)

i och med att vägen kom så (.) alla de där stora gamla träbåtarna så försvann ju

2.3 Kvantitet

I standardsvenska gäller regeln att varje betonad stavelse måste ha antingen en lång vokal eller en lång konsonant. De två stavelsetyper som förekommer är följaktligen V:C (t.ex. *da:g*) och VC: (t.ex. *dagg*). I de åboländska dialekterna har ett äldre kvantitetssystem bevarats, där förutom de två nämnda stavelsetyperna även korta (VC) och överlånga (V:C:) stavelser förekommer i betonad ställning. I östra Åboland (från

Nagu österut) har kortstavigheten bevarats både i en- och i flerstaviga ord. De väståboländska dialekterna (Iniö, Houtskär, Korpo) har där-
emot haft bevarad kortstavighet bara i två- och flerstaviga ord, medan
enstaviga ord har förlängts (Zilliacus 1992, 11). I utskärtdialekterna
i södra Korpo, södra Nagu och Hitis har stavelseförlängningen varit
genomförd så att kortstaviga ord saknas. I östra Åboland från Korpo
österut har även överlånga stavelser bevarats (t.ex. preteritum *ro:dd* av
verbet *ro*), medan de i västra Åboland ersatts av långa stavelser (*rodd*).

2.3.1 Kortstavighet

Kortstavigheten har undersökts med hjälp av ett index där ett antal
potentiellt kortstaviga ord ingår. Härvid har enstaviga ord inte beaktats
utan analysen av kortstavighet begränsas till ord med minst två stavel-
ser. Till att börja med extraherades en lista över ord som förekommer
med kortstavigt uttal i materialet. Eftersom materialet är begränsat till
20 minuter per talare är det få ord av de potentiellt kortstaviga orden
som finns belagda hos varje informant. Vissa av de potentiellt kort-
staviga orden är dessutom ointressanta för en vidare analys eftersom
variationen är ringa, antingen för att nästan alla informanter har ett
kortstavigt uttal eller för att nästan alla informanter har ett långstavigt
uttal.

Exempel på frekventa ord i materialet som nästan genomgående
uttalas kortstavigt är pronomenen *mina*, *dina*, *sina* och *såna* (av *sådan*),
infinitiverna *fara*, *göra*, *laga* och *vara* (som dessutom ibland är enstaviga
p.g.a. apokopering, se 3.4.2), verbformen *blivit* samt prepositionerna *före*
och *över*, adverbet *bara* och substantivet *morgon*. Dessa ord (utom möj-
ligen *laga*) hör till en grupp ord som även i Helsingforssvenska i ledigt
talspråk ofta uttalas kortstavigt. Medan kortstavigheten i dialekterna
i stort sett följer distributionen i fornsvenskan styrs kortstavigheten i
Helsingforssvenskan av andra regler. Kortstavigheten i nämnda ord kan
enligt Reuter (1980) förklaras som sekundär betoning av syntaktiskt
normalt obetonade ord (pronomina, prepositioner, adverb, hjälpverb
etc.) eller som bortfall av konsonant i tvåkonsonantkombinationer
(*mårgån* > *mårån*). Även i mindre städer i södra Finland följer det
högre språket i stort sett samma kvantitetsregler som gäller i Helsing-
forssvenskan (se Ivars 1996, s. 131 Ekenäs och s. 158 Lovisa). Eftersom

ovannämnda ord av alla informanter nästan genomgående uttalas kortstavigt har de inte inkluderats i kortstavighetsindexet. Därigenom kan kortstavighetsindexet anses mäta en ren dialektal typ av kortstavighet och inte allmänfinlandssvensk kortstavighet.

Följande kriterier ställdes upp för inkluderandet av ett ord i kortstavighetsindexet: 1) minst 15 av de 43 informanterna har belägg på ordet, 2) minst 4 informanter har belägg på kortstavigt uttal och minst 4 informanter har belägg på långstavigt uttal samt standardavvikelsen av andelen kortstavigt uttal hos de informanter som har belägg på ordet är större än 0,1. Det första kriteriet försäkrar att ordet är någorlunda frekvent och det andra att det finns en verklig variation mellan informanterna. För de ord som återstod efter denna uteslutningsmetod gjordes en korrelationsanalys (produktmomentkorrelation), och på basis av denna analys valdes ett antal ord ut för att representera kortstavighet. Ord med relativt hög korrelation med varandra valdes ut, eftersom dessa kan anses mäta ungefär samma sak.⁴ Följande ord ingår i kortstavighetsindexet:

- substantivet *dag* i plural bestämd och obestämd form: *dagar(na)*
- adjektivet *gammal* i stark böjning singular utrum och neutrum: *gammal(t)*
- verbet *komma* i infinitiv och supinum: *komma, kommit*
- verbet *läsa* i infinitiv och presens: *läsa, läser*
- substantivet *sommar* i singular bestämd och obestämd form: *sommar(en)*
- alla former av verbet *spela*: *spela, spelade, spelar, spelat.*

Belägg ur tvåledade sammansättningar där ett av de angivna orden utgör en led i sammansättningen har också excerperats. Indexet, som räknats ut för varje informant, anger andelen kortstavigt uttal av alla belägg på dessa ord. Indexvärdet 0 innebär uteslutande långstavigt uttal och indexvärdet 1 innebär uteslutande kortstavigt uttal av angivna ord. Utdragen nedan ger exempel på variabeln i kontext:

4. Inget av orden finns belagt hos alla informanter, vilket betyder att korrelationerna endast kan ses som approximativa, eftersom varje korrelation av två ord bygger på olika samplar av informanter.

- (2.48) *å sen ha de int vari bara såna kala:sär me me me både fö:delse-dagar å (.) å å å: (.) de ha vari studentkala:sär å de ha vari alla möjliga (.) tillstelningar dä:r på kāmuna:lgå:rdn (Ko.owo1.sp)*

och sedan har det inte varit bara sådana kalas med med med både födelsedagar och (.) och och och (.) det har varit studentkalas och det har varit alla möjliga (.) tillställningar där på kommunalgården

- (2.49) *ām man sti:gär åpp fem börjar lesa då så (.) hinnär man jo bra: fören man gå:r på jäbb (Na.omo1.sp)*

om man stiger opp fem börjar läsa då så (.) hinner man ju bra förrän man går på jobb

- (2.50) *sen nä man fick måpån sen hadd man ju me:ra liksom (.) du kund ju fara me:ra å kāmā me:ra hu du va du tyckte sen liksom å passa så dä:r du behöva int alti funde:ra att pappa elär mamma sku kāmā å sö:ka e:n (Na.ywo3.sp)*

sedan när man fick mopen (= mopeden) sedan hade man ju mera liksom (.) du kunde ju fara mera och komma mera hur du vad du tyckte sedan liksom och passa så där du behövde inte alltid fundera att pappa eller mamma skulle komma och söka (= hämta) en

- (2.51) *så spe:la vi sjelv åkså vi va några flickår såm hade (.) en årkestär såm hette harposkramli ja spela mandoli:nen såm e dä:r på veggen (.) å: å å vi: upptredde mykke me de å spela ti dans no:g å alt (Pa.owo2.sp)*

så spelade vi själva också vi var några flickor som hade (.) en orkester som hette Harposkramlet jag spelade mandolinen som är där på väggen (.) och och och vi uppträdde mycket med det och spelade till dans nog och allt

- (2.52) *på såmarn va ja ju jäbba ja ju på (.) sentra:ln hä:r (.) men de va bara såmarjäbb (Df.owo1.sp)*

på sommaren var jag ju jobbade jag ju på (.) centralen här (.) men det var bara sommarjobb

- (2.53) *dåm hadd lite då:li bra:sa i å:r dåm bru:kar ha u:te på en (.) på et grond di:t (.) så där närma:lt så sle:par dåm u:t en massa sji:t å tendär eld på (.) en massa **gamalt** brå:te såm ha blivi övär*
(Hi.ymo1.sp)

de hade lite dålig brasa i år de brukar ha ute på en (.) på ett grund där (.) så där normalt så släpar de ut en massa skit och tänder eld på (.) en massa gammalt bråte som har blivit över

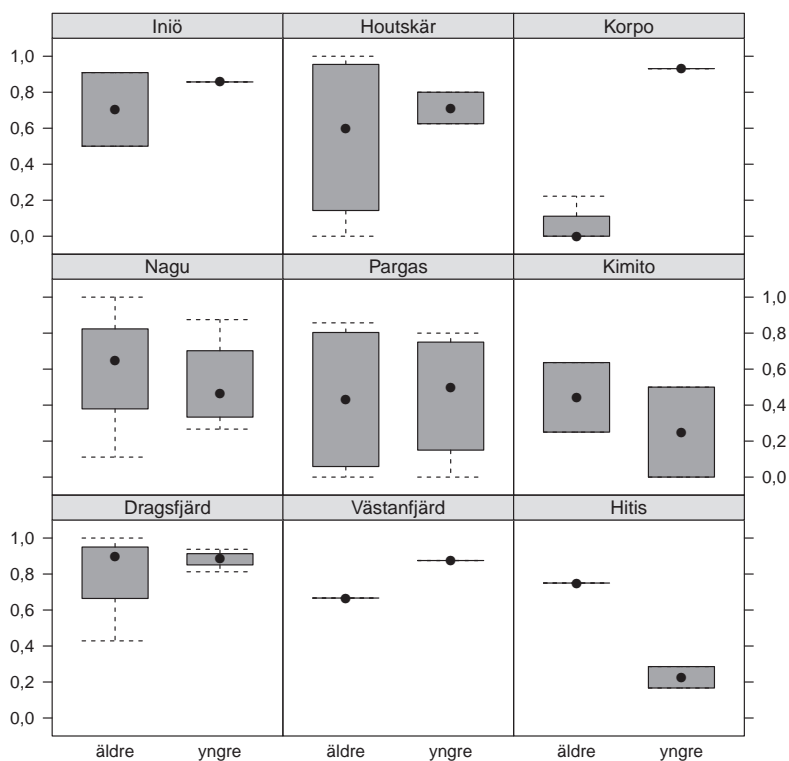
Vissa ord har som sagt inte inkluderats i kortstavighetsindexet trots att de är relativt frekventa och uppvisar variation i fråga om kvantitet bland informanterna. Orden har en låg korrelation med de övriga kortstaviga orden vilket tyder på att de följer andra distributionsmönster. Detta gäller t.ex. orden *hela* och *mycket* där det avvikande variationsmönstret kan ha att göra med att det förekommer tre varianter av orden. Dessutom är dessa ord ofta obetonade. Av *hela* förekommer förutom det kortstaviga uttalet *hela* och det långstaviga *he:la* även ett uttal med diftong *heila*, och av *mycket* förekommer förutom det kortstaviga uttalet *myki* och det långstaviga *mykke* även en variant med norrländsk förmjukning *mytji* (de diftongerade respektive förmjukade beläggen har inte räknats med i någon av kategorierna utan helt lämnats utanför vid excerperingen). Ytterligare ett exempel på ord med avvikande distribution är *riktigt* (som enligt Reuters, 1980, regler för kortstavighet i Helsingforssvenskan kan klassas som ord med ”bortfall av konsonant i tvåkonsonantkombination”).

En enda informant har inte ett enda belägg på något av de ord som ingår i kortstavighetsindexet, nämligen Ko.omo2.sp. Denna informant tilldelades dummyindexet o eftersom en rad andra potentiellt kortstaviga ord uttalas konsekvent långstavigt av informanten (*hela*, *mycket*, *samma*, *äta*). Informanten ger allmänt sett ett mycket långstavigt intryck och Ko.owo3.sp som deltar i samma intervju har också indexet o (dvs. bara långstaviga uttal av orden i fråga). Avsaknaden av kortstavighet hos dessa informanter förklaras av att de kommer från den yttre skärgården i Korpo där dialekterna har genomgått stavelseförlängning, vilket gör att kortstaviga former inte heller är att vänta.

För de övriga 42 informanterna varierar antalet belägg mellan 3 och 30. Medianvärdet för antalet belägg per informant är 10 och det totala antalet belägg är 474. Det lägsta indexvärdet är o (dvs. bara långstaviga

belägg, 6 informanter) och det högsta indexvärdet är 1 (dvs. bara kortstaviga belägg, 3 informanter). Medelvärde är 0,55 och medianvärdet 0,64. Distributionen efter ort och åldersgrupp ses i figur I.9. Andelen kortstaviga uttal är överlag hög, vilket visar att det är fråga om ett dialektdrag som är relativt väl bevarat. Variationen är stor framför allt i mellersta Åboland samt bland de äldre informanterna från Houtskär (där Björköinformanterna har mycket höga indexvärden medan Näsbyinformanterna har låga värden).

Andelen kortstaviga uttal är anmärkningsvärt hög hos den yngre informanten från Korpo (Ko.ymo2.sp) med tanke på att han kommer från utskären i södra Åboland där stavelseförlängningen har varit genomförd. Även den äldre informanten som representerar Hitis (Hi.owo1.sp) har en förvånansvärt hög andel kortstaviga former. Hon är



Figur I.9. Lådagram över kortstavighetsindex efter ort och åldersgrupp. Indexet anger andelen kortstaviga uttal av ett urval ord.

uppvuxen i den yttre skärgården i Hitis, men hennes mor härstammar från västra Nyland och själv har hon flyttat till Dalsbruk (Dragsfjärd) i vuxen ålder, vilket delvis kan förklara de kortstaviga formerna. Även hos de yngre informanterna från Hitis finns det belägg på kortstavigt uttal. Kortstavigheten verkar alltså vara ett dialektdrag som har potential att sprida sig.

2.3.2 Överlång stavelse

I motsats till kortstavighet verkar överlånga stavelser inte längre vara särskilt vanliga i dialekterna (Bergström 1984; Ivars 2003, 68). Enligt Zilliacus (1992, 20) har överlånga stavelser varit särskilt vanliga i böjningsformer av verb, som *fö:dd* 'födde' och *ro:dd* 'rodde'. Preteritumformer av denna verbklass är vanliga i materialet, t.ex. preteritum av *bo*, uttalat *bodd(e)*, finns belagt hos 27 informanter, varav 14 från östra Åboland där överlånga stavelser tidigare varit vanliga. Men i materialet finns bara ett enda belägg på en preteritumform med överlång stavelse:

(2.54) *no va hä:r jo såna såm fikk sej utav råttingen å när dâm int ly:dd*
(Ki.owo1.sp)

nog var här ju sådana som fick sig utav rottingen också när de inte lydde

3

Formlära och formord

I detta kapitel behandlas dialektdrag i formläran. Förutom ren morfologi behandlas även några formord. Först behandlas substantiv (3.1), sedan adjektiv (3.2) och pronomen (3.3) följt av drag som har med verb och verbböjning att göra (3.4). I samband med verben behandlas infinitivmärket som är en del av verbfrasen. Till sist beskrivs användningen av negationsordet *inga* (3.5).

3.1 Substantiv

Substantiven har i de åboländska dialekterna liksom i fornsvenskan varit indelade i tre genus: maskulinum, femininum och neutrum. Därtill har det funnits flera olika böjningsklasser och böjningen av substantiv har varierat efter genus och böjningsklass. Dessutom har det funnits variation i substantivböjningen mellan de olika dialekterna inom Åboland.

I de 20 minuters klipp som finns transkriberade för varje informant är antalet belägg per substantivklass i de flesta fall så litet att en systematisk analys inte har varit möjlig. I detta avsnitt beskrivs därför endast böjningen av s.k. starka maskuliner och femininer i bestämd form singular närmare.

Vid dialektutjämning sker ofta en övergång från ett tidigare tre-genussystem till standardsvenskans system med bara två grammatiska genus: utrum och neutrum. Historiskt sett har utrum uppkommit genom sammanfall av kategorierna maskulinum och femininum (SAG 2, 61). Den bestämda slutartikeln för de starkt böjda maskulina och

feminina substantiven är av särskilt intresse när man studerar dialektutjämning. Medan femininerna har haft en bestämd slutartikel som i stort sett motsvarar den standardsvenska ändelsen *-en* (t.ex. *granen, handen*), har maskulinerna haft en avvikande böjning. När tregenussystemet börjar svikta i dialektutjämningsprocessen sker ofta en övergeneralisering av maskulinum, så att ord som traditionellt varit femininer får den maskulina bestämda slutartikeln (Rabb 2007, 191; Sandström 2010, 165). Vid en fullbordad övergång från ett tregenussystem till ett tvågenussystem tar de standardsvenska böjningsmönstren överhand.

Den ursprungliga ändelsen för maskuliner i bestämd form är *-in* (Zilliacus 1992, 73). Denna ändelse har bevarats som sådan i dialekterna i Iniö och Houtskär. I de övriga åboländska dialekterna har ändelsen däremot genomgått synkopering, vilket innebär att ändelsevokalen faller bort. Synkoperingen har skett då substantivet slutar på *d, l, n, r, s* eller *t* (dvs. efter koronala konsonanter). På orterna från Pargas österut (Pargas, Kimito, Dragsfjärd, Västanfjärd, Hitis) har synkoperingen dessutom också skett efter labiala konsonanter (*b, m, p, v*).

Böjningen av maskulinerna *båt, kniv* och *dag* i bestämd form har alltså varit följande:

Iö, Ho	<i>bå:tin</i>	<i>kni:vin</i>	<i>da:jin</i>
Ko, Na	<i>bå:tn</i>	<i>kni:vin</i>	<i>da:jin</i>
Pa, Ki, Df, Vf, Hi	<i>bå:tn</i>	<i>kni:vn</i>	<i>da:jin</i>

Analysen av den bestämda slutartikeln har begränsats till enstaviga substantiv som slutar på *d, l, r, s* eller *t*. Dessa har excerperats manuellt i materialet. Variabeln har tre varianter: *-in* (traditionell maskulin slutartikel i Iniö och Houtskär), *-n* (traditionell maskulin slutartikel på alla övriga orter) och *-en* (slutartikel för utrum i standardsvenska och för femininer i dialekterna). Ursprungligt genus för varje ord i de traditionella sockendialekterna har kontrollerats i FO och FOreg. Trots att bara enstaviga substantiv har inkluderats, har även enstaviga substantiv som slutled i sammansättningar räknats med, så att t.ex. *granngården* räknas som belägg på *gården*.

Förutom de tre nämnda varianterna av slutartikeln förekommer i materialet från Iniö och Houtskär också en form utan slutljudande konsonant, dvs. *-i*. Zilliacus nämner ingen sådan form i *Åboländska* (1992),

men i ”Inledning och kommentar till SLS 701–702” (s. 11) nämner han att ändelsen i maskulinum bestämd form singular ger ”intryck av att vara ett nasalerat *i*”, vilket också gäller för en del av beläggen i det aktuella materialet. Utvecklingen har alltså varit *-in* > *-ī* > *-i*. En meddelare från Houtskär som svarat på en frågelista från SLS år 1992 ger ett kort prov på ljudenligt transkriberad dialekt (SLS 1746: 458–460) där han konsekvent har utelämnat *n*:et i maskulinum bestämd form singular: *såmári* ’sommaren’, *himli* ’himlen’ etc. Bortfallet av *n* verkar alltså vara genomfört i så hög grad att lekmän uppfattar ändelsen som enbart *-i*. Ahlbäck (1946, 99) nämner att former med *n*-bortfall förekommer i Brändö och Kumlinge, men han nämner inga sådana former från Iniö eller Houtskär.

I utdrag 3.1 nedan finns flera exempel på den reducerade maskulina slutartikeln. Av dessa fyller bara ordet *kväll* kriterierna ”enstavigt substantiv som slutar på koronal konsonant”, och de övriga orden ingår följaktligen inte i den kvantitativa analysen. Men exemplet visar en mycket konsekvent reduktion av slutartikeln hos denna informant. Intervjuaren har här frågat om trafikförbindelserna var lika bra under informantens skoltid som de är nu.

- (3.1) *he ha mytjy förendras (..) tå var e jurmå såm jikk (..) två: ganggor åm da:i (..) ein gang på måråni å ein gang på kveldi (..) å nu er e vel (..) e de seks tu:rår åm dai e ga:r (..) elår er e meir* (Io.ywo2.sp)

det har mycket förändrats (..) då var det Jurmo som gick (..) två gånger om dagen (..) en gång på morgonen och en gång på kvällen (..) och nu är det väl (..) är det sex turer om dagen det går (..) eller är det mer

Formen *-i* har räknats som en variant av *-in* vid excerperingen av substantiv.

Antalet maskulina substantiv är större än antalet feminina substantiv, vilket har gjort en kvantifiering av de starka maskulinerna möjlig. För femininernas del har en systematisk analys däremot inte varit möjlig. De flesta informanter (alla utom sju stycken) har minst fem belägg på maskulina substantiv i bestämd form, och ingen informant saknar helt belägg. Medianvärdet är åtta belägg per informant och det totala antalet belägg i materialet är 407. För femininernas del saknar

däremot tolv informanter helt belägg, och tio informanter har bara ett enda belägg. Bara fyra informanter har fem belägg eller flera på starka femininer i bestämd form.

Exempel på maskulina substantiv i materialet är *båt, del, för, gård, höst, is, kväll, mat, präst, spis, tid* och *ved*. Exempel på femininer är *hand, hälft, jul, natt, sol, strand* och *vår*.

Fördelningen av varianterna *-i(n)* och *-n* av den maskulina slutartikeln är i stort sett den väntade. Beläggen på *-i(n)* kommer från Iniö och Houtskär medan beläggen på *-n* kommer från de övriga orterna. Men det finns två informanter som bryter mot detta mönster. De två äldre informanterna från Hönsnäs (Houtskär) använder båda i huvudsak den standardsvenska formen *-en* men har även belägg på synkope som enligt Zilliacus (1992) inte borde förekomma i Houtskär.⁵ Det handlar framför allt om ordet *kvällen* i formen *kveln* som i exemplet nedan, där den äldre mannen beskriver hur man klär midsommarstången:

(3.2) *på eftermiddag:n börjar vi å så kle:der vi ti (.) ti se:nt på **kveln***
(Ho.omo1.sp)

på eftermiddagen börjar vi och så klär vi till (.) till sent på kvällen

Figur I.10 visar i hur stor utsträckning informanterna använder den traditionella bestämda slutartikeln för maskuliner. För informanterna från Iniö och Houtskär har andelen *-i(n)* beräknats och för övriga informanter andelen synkoperade belägg. För alla informanter, utom de två ovannämnda från Hönsnäs (Houtskär), utgörs den resterande delen av belägg på standardvarianten *-en*. Lådornas storlek visar att det finns stor variation mellan informanterna, men man ser också att den traditionella slutartikeln är relativt väl bevarad. De informanter som helt saknar belägg på den traditionella formen och bara har belägg på den standardsvenska formen kommer framför allt från mellersta Åboland.

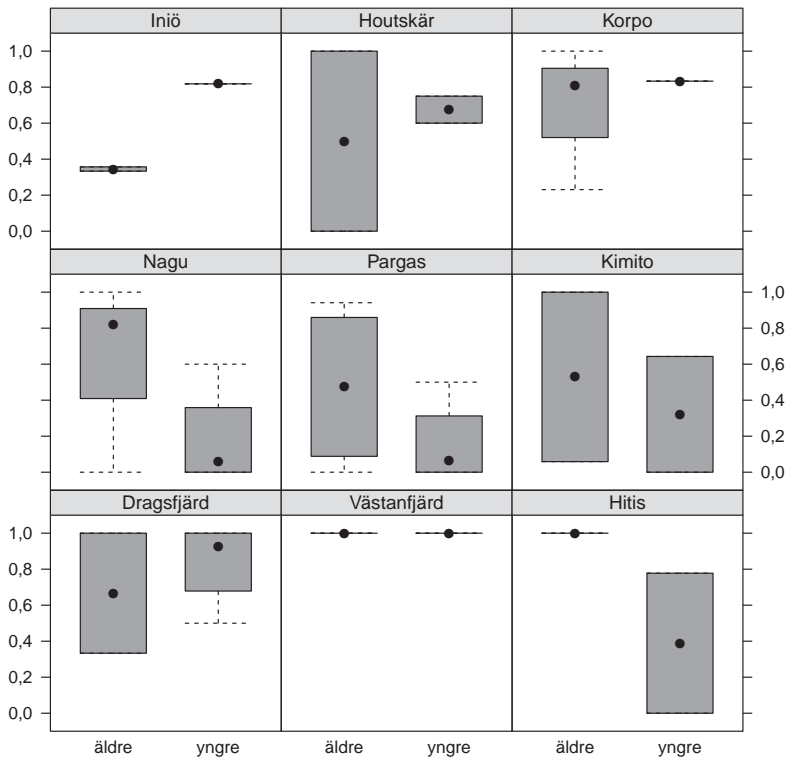
Eftersom antalet belägg på starka femininer är så litet är det svårt att säga i hur stor utsträckning genussystemet upprätthålls bland informanterna. Några få informanter har dock så många belägg att man kan dra vissa slutsatser om deras språkbruk. En av de äldre männen från Nagu

5. Zilliacus beskrivning av dialekten i Houtskär bygger främst på språket i Norrbyarna, så variation inom Houtskär kan ha funnits också tidigare.

upprätthåller tydligt distinktionen mellan maskulinum och femininum. Han har 21 belägg på maskuliner och 11 på femininer. Maskulinerna har i samtliga fall den synkoperade slutartikeln, medan femininerna får ändelsen *-en* i tio av fallen och bara en enda gång en synkoperad slutartikel. I utdraget nedan har *hösten* som är ett maskulinum ändelsen *-n* medan *våren* som är ett femininum har ändelsen *-en*:

- (3.3) *de va: ijentligen två: år såm man jikk i i fo:rtsetningsko:lan (..) de va: vissa vissa ti:där ja tykkär att (.) att jikk man två: vekkär på höstn å två: vekkär på på vå:ren i fo:rtsetningsko:lan*
(Na.omo2.sp)

det var egentligen två år som man gick i i fortsättningsskolan (..) det var vissa vissa tider jag tycker att (.) att gick man två veckor på hösten och två veckor på på våren i fortsättningsskolan



Figur I.10. Lådagram över andelen dialektvarianter (Iniö och Houtskär *-in*, övriga orter *-n*) av den maskulina slutartikeln efter ort och ålder.

En annan informant som verkar upprätthålla genusdistinktionen är den äldre kvinnan från Hitis. Alla hennes belägg följer det traditionella genussystemet. Hon har sju belägg på synkoperade maskuliner och fyra belägg på femininer med *en*-ändelse. Även Ko.omo2.sp och Ho.owo1.sw verkar följa det ursprungliga genussystemet, men antalet belägg hos dessa informanter är rätt litet.

Några av informanterna använder den standardsvenska slutartikeln *-en* i stor utsträckning även för maskuliner, men i de fall där de trots allt använder den maskulina slutartikeln *-i(n)/-n* sker detta bara för ursprungliga maskuliner och aldrig för femininer. Detta kan tyda på att informanterna har kunskap om det ursprungliga genussystemet trots att de i intervjusituationen talar mer standardnära. Informanter som uppvisar detta mönster är: Io.owo1.sp, Ho.omo1.sp, Ko.owo3.sp, Na.ymo1.sp, Pa.ymo5.sp, Pa.ymo4.sp, Ki.omo1.sp och Ki.ymo3.sp. Av dessa informanter nämner åtminstone Io.owo1.sp och Ko.owo3.sp själva under intervjun att de är bidialektala. Intervjuaren i dessa intervjuer talar däremot inte lokal dialekt, och det framkommer under intervjun att det känns onaturligt för informanterna att tala dialekt i en sådan situation. Att vissa informanter använder den maskulina slutartikeln bara i några fall kan också bero på att artikeln finns bevarad lexikalt för vissa ursprungliga maskuliner trots att ett tregenussystem inte längre upprätthålls.

Sammanlagt fem informanter använder den maskulina slutartikeln för ord som är ursprungliga femininer. Tre av dessa kommer från Dragsfjärd och två från Nagu. Det handlar alltså i samtliga fall om synkoperade dialekter. Dessa fem informanter har till största delen synkoperade belägg på maskuliner och verkar tillämpa ett system där den maskulina slutartikeln har generaliserats till att kunna användas för både maskuliner och femininer.

Eftersom belägg på feminina substantiv är få, och eftersom användningen av anaforiska pronomen för maskulina och feminina korrelerat inte har undersökts, går det inte att avgöra hur väl bevarat tregenussystemet är i dialekterna. Maskulinändelsen i sig används dock i relativt hög grad av många av informanterna. Mycket höga andelar av standardvarianten *-en* för ursprungliga maskuliner finns främst bland yngre informanter från Nagu och Pargas, vilket tyder på att genusutjämningsprocessen har nått längst på dessa orter.

3.2 Adjektiv

Adjektivböjningen i Åboland har uppvisat flera avvikelser från standardspråket (Zilliacus 1992, 80–82). Adjektiven har inte excerperats systematiskt för denna undersökning, men exempel på några drag inom adjektivböjningen ges här.

Ahlbäck (1945, 27) nämner att slutvokalen kan saknas hos starkt böjda adjektiv i plural i Korpo och Houtskär, medan ändelsevokalen bevaras i Nagu och Pargas. Predikativ pluralform utan ändelse förekommer i materialet, men används inte konsekvent av informanterna. Intressant är att det finns belägg även från Nagu där slutvokalen enligt Ahlbäck har bevarats.

- (3.4) *vi va **redda** för att åskan sku slå ne:r hä:r [...] för den slo: ne:r i granngå:rden hä:r (.) så va vi **redd** åm vi sku få åskan ne:r hä:r (Ho.ow01.sp)*

vi var rädda för att åskan skulle slå ner här [...] för den slog ner i granngården här (.) så var vi rädda om vi skulle få åskan ner här

- (3.5) *å sen så kåm ju (.) sko:lans skrekk dä:r (.) de härän hon lä:rarinnan såm alla va **redd** fö:r (Na.ymo1.sp)*

och sedan så kom ju (.) skolans skräck där (.) det här hon lärarinnan som alla var rädda för

Apokopering av slutvokalen i komparativer av typen *bättre, längre, äldre* förekommer frekvent bara hos informanterna från Björkö och Mossala (Houtskär):

- (3.6) *ibla:nd ha e vari värre å ibla:nd **betär** (Ho.ow01.sw)*

ibland har det varit värre och ibland bättre

- (3.7) *veit tu att dām e no all (.) meir elär **mindär** li:kadan mä:g att kånse ja ha va:lt rett tå eintå (Ho.ywo1.sp)*

vet du att de är nog alla (.) mer eller mindre likadana mig att kanske jag har valt rätt då ändå

I Iniö och Houtskär har det enligt Zilliacus (1992, 81) funnits komparativformer med ändelsen *-an*. Endast ett belägg på en sådan komparativ har hittats i 2000-talsmaterialet:

- (3.8) *vi hadd ein sån dä: veivtjä:rno såm va li:te möjligan vi vild inga stå å stöit* (Ho.owo1.sw)

vi hade en sådan där vevkärna som var lite möjligare vi ville inte stå och stöta

Ytterligare ett belägg på adjektiv med ändelsen *-an* finns från Korpo. Informanten korrigerar sig själv direkt och använder då en komparativform. Trots det handlar det i detta fall möjligen om en gammal ackusativform. Informanten kommer nämligen från Korpo utskärgård där gamla ackusativformer enligt Zilliacus (1992, 81) har funnits bevarade.

- (3.9) *vi hadd nu e:n (.) sto:ran större hälm såm hette a:lsjär*
(Ko.omo2.sp)

vi hade nu en (.) stor större holme som hette Alskär

För adjektivens del har två drag undersökts närmare, nämligen den gamla nominativändelsen *-er* och adjektivsuffixet *-og*.

3.2.1 *-er*

Adjektiv i stark böjning singular har kongruensböjts enligt tregenussystemet i Åboland (Zilliacus 1992, 80):

storer karl, stor kvinna, stort barn

Redan Ahlbäck (1956, 16–17) menar dock att skillnaden mellan maskulinum och femininum är upphävd och att former med och utan *-er* kan gälla som både maskulina och feminina. Bergström (1984, 124) beskriver att ändelsen *-er* kan användas för alla icke-neutrala adjektiv i Mattkärrdialekten (Kimito) och det råder fri variation mellan former med och utan *-er*.

Det föreliggande materialet är för litet för kvantifiering av adjektivändelsen *-er*. Men då man räknar adjektiv i materialet (i både

attributiv och predikativ funktion) kan man konstatera att 14 av de 43 informanterna har minst ett belägg på ändelsen *-er* under det 20 minuter långa klippet. Dessa informanter fördelar sig över nästan hela undersökningsområdet – bara Kimito och Hitis saknar belägg. Bara sex av informanterna har mer än ett belägg. Ändelsen förekommer både för adjektiv (t.ex. *argär*, *bly:gär*, *enggelskär*, *gla:där*, *sto:rär/ståorär*, *ongär*, *vatogär*) och particip (*föståddär*, *så:ldär*). Flest belägg med *er*-ändelse, fem, finns det för adjektivet *stor*.

Adjektivändelsen *-er* finns alltså bevarad i hela Åboland, men materialet är för litet för att konklusioner om användningsfrekvens ska kunna dras. Materialet är också för litet för en analys av huruvida det finns informanter som upprätthåller distinktionen mellan maskulinum och femininum. Intressant nog står de allra flesta adjektiv och particip med *er*-ändelse i materialet i nominalfraser eller som predikativ till en referent som utgörs av ett maskulinum. Ett enda entydigt belägg finns där en form med *-er* syftar på ett femininum. Det är då en av de kvinnliga informanterna använder adjektivet *då:ligär* om sig själv. Å andra sidan är feminina substantiv mindre frekventa än maskulina, och det finns också många fall av adjektiv utan *er*-ändelse som syftar på maskuliner i materialet.

- (3.10) *he to: e:n å in halvär ti:m på hö:stan nä:r e va rikit då:lit ti kām fram ti kām me traktår åpp ti byss* (Ho.omo1.sw)

det tog en och en halv timme på höstarna när det var riktigt dåligt att komma fram att komma med traktor upp till byn

- (3.11) *ja ha jo aldri vari bly:gär så ja (.) ja pra:ta jo fast ja inga kond* (Pa.ymo4.sp)

jag har ju aldrig varit blyg så jag (.) jag pratade ju fast jag inte kunde

- (3.12) *så leng man e ba:rn så va setär man no vä:rde på: åm man få:r en sån då:n ståorär du:k* (Vf.owo1.sp)

så länge man är barn så vad sätter man nu värde på om man får en sådan där stor duk

3.2.2 -og

Adjektiv som i standardsvenska slutar på *-ig* (t.ex. *avig*) har flera motsvarigheter i Åboland. Dialektformen är *-og* (*a:vogär*), men dessutom har former på *-in* (*a:vin*) och *-ig* (*a:vigär*) använts och även förkortade maskulina former på *-or* (*å:vår*) och växelformer som *å:våtär* (Zilliacus 1992, 80). Adjektiv som i standardspråket har suffixet *-ig* förekommer i riklig mängd i materialet. Uttal med *-og* är dock sällsynta. En systematisk analys av variabeln försvåras av att nyare lånord från standardspråket ofta har fått suffixet *-ig* och det således är svårt att avgöra vilka adjektiv som ursprungligen har kunnat ha suffixet *-og* (jfr Ivars 1988, 163; Gullmets-Wik 2004, 75).

Bara sju av informanterna har belägg på *og*-former. Tabell I.3 återger alla belägg som har hittats i materialet. Flest belägg, fyra, har informanten Na.om02.sp.

Det ringa antalet belägg kunde tänkas bero på att denna ordtyp inte är särskilt frekvent i språket. Men en rad ord som enligt FO tidigare

Tabell I.3. Belägg på suffixet *-og* i materialet.

INFORMANT	STANDARDISERAD FORM	UTTAL
Io.yw02.sp	konstigt	<i>kånståt</i>
	tokigt	<i>to:kåt</i>
Ho.om01.sp	snöiga	<i>snö:å</i>
Ho.yw01.sp	hungrig	<i>hongrå</i>
Ko.om02.sp	kinkigare	<i>tjinkågare</i>
Na.om02.sp	hastigt	<i>haståt</i>
	krokiga	<i>kråkoga</i>
	rostiga	<i>ro:ståga</i>
	vattig	<i>vatogär</i>
Pa.om03.sp	djävligare	<i>je:vlågare</i>
	djävligt	<i>jevlåt</i>
	tokig	<i>to:kå</i>
Pa.yw01.sp	enfaldiga	<i>einfa:ldåga</i>
	tontigt (= dumt)	<i>tåntåt</i>

har haft suffixet *-og* i Åboland förekommer relativt frekvent i materialet, t.ex. *färdig, hastig, konstig*. Dessa ord finns hos flera av informanterna, men hos de allra flesta bara med suffixet *-ig*. Följaktligen kan man dra slutsatsen att det traditionella suffixet *-og* till stora delar har ersatts eller håller på att ersättas av standardspråkets *-ig*.

3.3 Pronomen

3.3.1 *het*

I tregenussystemet används *han, hon* och *het* som personliga pronomen i tredje person singular. I ett tregenussystem syftar alltså pronomenen *han* och *hon* inte bara på levande väsen som i standardsvenskan utan används också vid inanimata maskulina och feminina substantiv.

Sandström (2010) har beskrivit utvecklingen från tregenussystem till tvågenussystem för de östnyländska dialekternas del. Då tregenussystemet börjar vackla drabbas de maskulina och feminina pronomenen för inanimata korrelat först av utjämning och ersätts av det standardsvenska *den* (Sandström 2010, 168–169). I följande skede ersätts *het* av *det*, vilket fullbordar paradigmskiftet från ett tredelat *h*-paradigm (*han, hon, het*) till ett tvådelat *d*-paradigm (*den, det*) för pronomenen (Sandström 2010, 132). Den maskulina slutartikeln är den sista resten av ett tregenussystem och kan finnas kvar hos östnyländska talare som i övrigt har gått över till ett tvågenussystem (Sandström 2010, 165).

Eftersom genussystemet och substantivböjningen i östra Nyland har uppvisat stora likheter med det åboländska systemet kan man tänka sig att utjämningsprocessen också i Åboland följer liknande mönster. Som analysen av substantivböjningen i avsnitt 3.1 ovan visar, är den maskulina slutartikeln väl bevarad hos många informanter. Utöver substantivböjningen kan en analys av det neutrala pronomenet belysa utjämnningen av genussystemet.

Det neutrala pronomenet *het* används både som pronomen för neutrala korrelat och i expletiva funktioner (t.ex. *det fanns inte mycket bilar här*). För de östnyländska dialekternas del hittade Sandström (2010) ingen skillnad i andelen *het* mellan de båda funktionerna. Genom att det neutrala pronomenet också används expletivt är antalet belägg på detta mycket större än för *hon* och *han* (eller *den*), vilket gör

att variabeln lämpar sig väl för kvantifiering. Det neutrala pronomenet har dessutom enkelt kunnat excerperas automatiskt eftersom alla ord som är översatta till *det* i den standardsvenska ortografin har kunnat inkluderas. För de feminina och maskulina substantivens del skulle man däremot vara tvungen att beakta om pronomenet syftar på ett levande väsen eller inte, vilket skulle kräva en manuell genomgång.

Några begränsningar har dock gjorts vid excerperingen. Belägg där ordet före pronomenet slutar på ett *-d* eller *-t* har uteslutits eftersom det i sådana fall inte går att avgöra om pronomenet inleds med ett *d* eller inte. Vidare har alla fall uteslutits där följande ord är antingen *här* eller *där*. Orsaken till detta är att diskurspartiklarna *det här* och *det där* (ofta med uttalen *de häran* och *de därän*) nästan uteslutande har varianten *det* (jfr Wide 2011a). Diskurspartikeln är mycket frekvent hos vissa informanter (Wide 2011a, 264). Efter denna reducering genom kontext är medianvärdet för antal belägg per informant 100 och det totala antalet belägg på variabeln 4 780.

Varianten *det* uttalas så gott som undantagslöst utan slutljudande *t*. I de flesta fall är uttalet kort (*de*), men ibland förekommer uttal med lång vokal (*de:*).

Av dialektvarianten finns det många olika uttalsvarianter. FO ger uttal som *he* och *he:d* i tryckstark ställning och *e*, *ä* och *je* i trycksvag ställning. Varianterna *he* och *e* förekommer frekvent bland 2000-talsinformanterna. Varianterna *he:d* och *je* förekommer också, men inte lika ofta. En indelning i trycksvaga och tryckstarka belägg har inte gjorts vid kvantifieringen, utan alla nämnda uttalsvarianter räknas som varianter av *het*.

Exempel 3.13 nedan visar variationen mellan tryckstarka och trycksvaga former hos en av informanterna. Varianten *he:d* har stark betoning i satsen, medan de båda *e* har svag betoning.

(3.13) *fössjt va e nu vel nä:r man le- le:rd se sy:kel (.) sen eftär he:d va e måpe:d* (Ho.ymo1.sp)

först var det nu väl när man lä- lärde sig cykla (.) sedan efter det var det moped

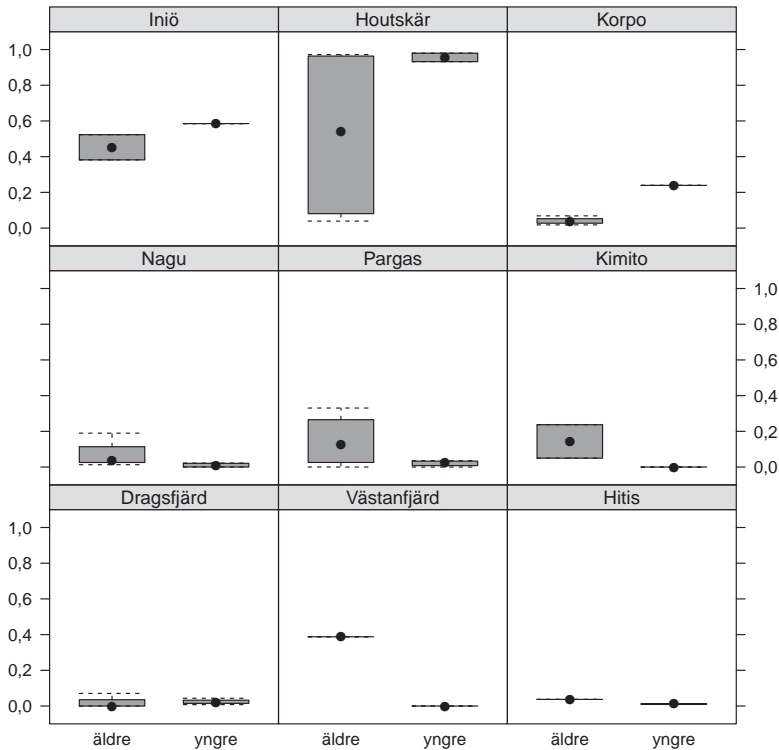
I Iniö verkar det vara särskilt vanligt att ord som *är* och *var* (av *vara*) uttalas med hörbart *r* när de efterföljs av ett trycksvagt *e*, trots att de i

andra kontexter uttalas utan *r* (*e, va*). Detta gäller också andra småord som *hur* och *för*. Från övriga orter än Iniö finns bara ströbelägg på *r*-former.

(3.14) *dåm dansa no:g tiss so:len jikk åpp på brendö (.) på den ti:den men nu: så er e vell andra ti:där dä:r (.) dä:r å förstås* (Io.owo1.sp)

de dansade nog tills solen gick opp på Brändö (.) på den tiden
men nu så är det väl andra tider där (.) där också förstås

Figur I.11 visar andelen dialektvarianter av det totala antalet belägg på det neutrala pronomenet efter ort och åldersgrupp. Lådagrammet visar höga andelar *het* i Iniö och Houtskär och låga andelar på övriga håll. Bland de äldre informanterna från Houtskär är variationen stor på grund av att informanterna från Björkö har mycket höga andelar medan



Figur I.11. Lådagram över andelen dialektvarianter (olika uttalsvarianter av *het*) av neutralt pronomen efter ort och ålder.

informerarna från Hönsnäs har låga andelar. Den äldre informanten från Västanfjärd använder också *het* i ganska stor utsträckning (39 procent), och en äldre man från Pargas använder *het* i 33 procent av fallen. Från Nagu och österut har alla yngre informanter andelar under 10 procent. Sammanfattningsvis kan man säga att det traditionella neutrala pronomenet *het* verkar ha en stark ställning bara i Iniö och Houtskär.

Att andelen dialektvarianter är låg behöver inte betyda att dessa inte har någon funktion. En yngre man från Dragsfjärd använder tydligt *het* som dialektmarkör (jfr Sandström 2010, 308–310). Denna informant har mer än 100 belägg på *det* och bara sex belägg på *het*. De flesta av dessa *het*-belägg kommer när informanten berättar om sina erfarenheter från värnpliktstjänstgöringen. Han delade bostad med flera österbottningar och berättar att han inte hade några problem alls med att förstå de österbottniska dialekterna, så de brukade tala dialekt sinsemellan. För Helsingforsborna var det däremot svårare:

(3.15) [Inf.:] *vi hadd ju några helsingfårsbo:r enno till i vå:r stu:ga i arme:n såm m- bodd dä:r me äss de däran så (.) åm vi pra:t- åm vi vila pra:ta nånting sånt så at int han sko förstå: så då: tala v- pra:ta vi li:ksåm på riktigt årdentlit på dialekt [...]* [Int.:] *men hadd do bara (.) den hä dialektn då do elär slengd do in li:te* [Inf.:] *ja slengd no: in lite östärbåt'niska de då: de däran fö ja: ja ha:r den tendensn at ja nappar åpp e:h snappar åpp ganska hastit o:rd från andra dialektär de däran så (.) så de däran så **he va he va** (.) **he va ro:lit i alla fall de där fö han så: u:t såm et frå: getekken han** (Df.ymo3.sp)*

[Inf.:] *vi hade ju några Helsingforsbor ännu till i vår stuga i armén som m- bodde där med oss det där så (.) om vi prat- om vi ville prata någonting sådant så att inte han skulle förstå så då tala v- prata vi liksom på riktigt ordentligt på dialekt [...]* [Int.:] *men hade du bara (.) den här dialekten då du eller slängde du in lite* [Inf.:] *jag slängde nog in lite österbottniska det då det där för ja jag har den tendensen att jag nappar opp eh snappar opp ganska hastigt ord från andra dialekter det där så (.) så de där så det var det var (.) det var roligt i alla fall det där för han såg ut som ett frågetecken han*

Informanten berättar i utdraget om en incident där han talade mer dialektalt än han annars brukar göra och lägger då in *het* som ett slags

exempel på ett mer dialektalt uttrycks sätt. Trots att denna informant i intervjun i övrigt mest använder *det*, har han alltså kvar också dialektvarianten i sin repertoar och kan använda den medvetet när han vill ge mer dialektal färg åt sitt språk.

3.3.2 *han* och *hon* som objektsformer

I dialekterna i Åboland används *han* och *hon* (med olika uttalsformer som *hon*, *hån*, *hoån*) som objektsformer i stället för *honom* och *henne* (FO; Fagerlund 1878, 159; Thurman 1898, 162; Thurman 1929, 127; Bergström 1984, 134). I vissa socknar, framför allt i de östligare delarna, har formen *henna(r)* också använts vid sidan av *hon* i femininum (FO; Thurman 1898, 162; Thurman 1929, 127). Beläggen på de maskulina och feminina personliga pronomenen som objekt är för få i materialet för en systematisk analys. Femton av informanterna saknar helt och hållet belägg. Vissa informanter har dessutom belägg bara på maskulinum eller femininum, men inte båda.

Trots det ringa antalet belägg kan vissa tendenser urskiljas. För det maskulina pronomenet finns det 28 belägg på *han* och 14 på *honom*. Den icke-standardliga formen är alltså vanligast. För det feminina pronomenet finns det däremot 15 belägg på *hon* och 21 på *henne*.

Från hela Kimitoön inklusive Hitis finns det inte ett enda belägg på *honom*, utan de sex informanter som har belägg använder alla *han* som objektsform. Vidare finns det från Dragsfjärd och Västanfjärd bara belägg på *hon* men inget *henne*, medan Kimitoinformanterna verkar föredra *henne*. Dessutom finns ett belägg *hennar* från Kimito.

Pargasinformanterna, som är flest till antalet, har inga belägg på *hon* som objektsform, utan här används bara *henne*. För det maskulina pronomenet har tre av Pargasinformanterna bara belägg på *han*, medan två informanter uteslutande använder *honom*.

I västra Åboland är variationen för stor och beläggen för få för att några regelbundenheter ska kunna urskiljas.

(3.16) *han jo:rd någå harm åt mej elär va de va så slo: ja han ra:kt på nesan me (.) me kny:tnevan* (Pa.omo3.sp)

han gjorde något harm åt mig eller vad det var så slog jag honom rakt på näsan med (.) med knytnäven

- (3.17) *ja tjend hon me sama fö ja ha träffa hon på såmarn u:te i sjä:ren*
(Hi.owo1.sp)

jag kände henne med samma för jag har träffat henne på sommaren ute i skären

Informanten Na.omo2.sp har många belägg på enklitiska former. Hos övriga informanter finns bara få belägg på sådana former.

- (3.18) *ti slu:t hadd mamma sakt sen åt'n att (.) att va ska do no va arg på påjkana för att däm narra dej aprill* (Na.omo2.sp)

till slut hade mamma sagt sedan åt-honom att (.) att vad ska du nu vara arg på pojkarna för att de narrade dig april

3.3.3 *tocken/tolin*

Pronomenet *tocken* i betydelsen 'sådan' har varit utbrett i hela Åboland utom i Houtskär där den alternativa formen *tolin* har använts (Zilliacus 1992, 82). Enligt Ivars (2003, 63) har formord av detta slag en tendens att försvinna vid dialektutjämning, samtidigt som just denna typ av ord har gett dialekterna deras särprägel och lokala färg. Att *tocken* ersätts av *sådan* åtminstone på Kimitoön framkommer av Bergströms (1984) undersökning.

Av informanterna i föreliggande undersökning är det bara tre stycken som har belägg på *tocken/tolin*: de två äldre informanterna från Björkö (Houtskär) och den yngre informanten från Iniö. Dessa tre informanter använder *tocken/tolin* vid sidan av *sådan* i intervjuerna. Av informanterna från Houtskär har den ena ett belägg på *tocken* och ett på *tolin* medan den andra bara använder *tolin* vid sidan av *sådan*. Hos alla de övriga informanterna i materialet har *tocken/tolin* helt ersatts av det standardspråkliga *sådan*.

- (3.19) *fast ja int e eldär en va ja e: så mins ja ju tåkna i:svintrar såm såm int e nuförti:din meir* (Io.ywo2.sp)

fast jag inte är äldre än vad jag är så minns jag ju sådana isvintrar som som inte är nuförtiden mer

- (3.20) *et sla:g fans e barå tåkå ten små:a min fålk ha börja tjö:p li:te störr
bå:tar me:r nu: (Ho.owo1.sw)*

ett slag fanns det bara sådana där små men folk har börjat köpa
lite större båtar mer nu

3.4 Verb

Zilliacus (1992, 83) konstaterar att verben i Åboland allmänt sett uppvisar "en massa små avvikelser från riksspråket". Detta gäller även i 2000-talsmaterialet. De flesta av dessa drag är dock så infrekventa att kvantifiering inte är möjlig. Därför kommer endast några exempel att ges på drag som har med verb och verbböjning att göra. Bara tre variabler är så frekventa att de har analyserats kvantitativt. Dessa är *infinitivmärket*, *infinitiv* och *preteritum* som behandlas i avsnitt 3.4.1–3.4.3 nedan.

Enligt Zilliacus (1992, 83) förekommer enstaviga presensformer i västra Åboland t.o.m. Nagu och vad gäller enstaka ord också i Pargas. I det aktuella materialet finns det några belägg på presensformer utan ändelse från Iniö och Houtskär, t.ex. i följande utdrag där informanten berättar om hur han sydde segel till segelbåtarna i midsommarstången:

- (3.21) *ja to: uta:v sån dä:r (.) såm dām ha:r eh sement å sån dä:r eh
(.) bru:k i bruk i nu: plastsekkar så (.) så sydd ja se:gel uta:v
de: å då hållär dām i fle:ra å:r (.) ja: (..) så (.) så **slipp** man li:te
lindrigari (Ho.omo1.sp)*

jag tog utav sådan där (.) som de har eh cement och sådan där
eh (.) bruk i bruk i nu plastsäckar så (.) så sydde jag segel utav
det och då håller de i flera år (.) ja (..) så (.) så slipper man lite
lindrigare

Från Pargas finns former utan presensändelse av verbet *orka* hos två informanter:

- (3.22) *no e de ju nån sårts påskmidda nånstans (.) inga **årk** vi sjelv stell
åpp me de inga (Pa.ymo4.sp)*

nog är det ju någon sorts påskmiddag någonstans (.) inte orkar vi
själva ställa opp med det inte

Supinumformer av typen *slaji* 'slagit' ska enligt Bergström (1984) stå sig väl i åboländskan. I Nyland förekommer detta drag även i stadsspråken (Gullmets-Wik 2004, 150–152, 172–173). Bara 13 av de 43 informanterna i materialet har belägg på supinum av något av verben *dra*, *slå* eller *ta*, där denna form skulle kunna förekomma. Därför går det inte att säga något bestämt om i hur stor utsträckning uttal med *-j-* förekommer. Men av de tretton informanterna som har belägg är det bara två som har uttal med hårt *-g-*, vilket antyder att *j-*formerna är utbredda. Uttalet är i allmänhet kortstavigt.

- (3.23) *ja hadd en så hä spesia:lsy:kel såm ja hadd laga bara bara för de
dä:r i prins:ip såm ja hadd må:la å (.) å taji bårt alt ekstra vett du
så att nä man kåm sy:klande (.) sy:klande fult då at man liksåm
kund håppa a:v i fa:rten bara å lå:t sy:keln fara vi:dare
(Na.ymo1.sp)*

jag hade en sådan här specialcykel som jag hade lagat bara bara för det där i princip som jag hade målat och (.) och tagit bort allt extra vet du så att när man kom cyklande (.) cyklande fullt då att man liksom kunde hoppa av i farten bara och låta cykeln fara vidare

- (3.24) *mitt på timman så fikk ja sti:g åpp (.) å gå: å knacka på dören på
den andra klassens dörr dä han dä jassn satt (.) så fikk ja gå: ti hans
plats å be: åm fölä:telse för att ja hadd slaji han (.) så att han fikk
nesablåod (Pa.omo3.sp)*

mitt på timmen så fick jag stiga opp (.) och gå och knacka på dörren på den andra klassens dörr där den där jassen (= killen) satt (.) så fick jag gå till hans plats och be om förlåtelse för att jag hade slagit honom (.) så att han fick näsblod

Enligt Ivars (2006, 28) har dubbelsupinum tidigare använts relativt frekvent i västra Åboland och mer sporadiskt i östra Åboland. I det föreliggande materialet finns belägg på dubbelsupinum, men inte konsekvent och inte hos alla informanter:

- (3.25) *ti: ho:tsjä:r va vi (.) å: (.) hu lenge sku de ha **kuna rekt** mánne de va: en vekka de bara han man faktist lä:ra sej alt sãm sãm stã:r i bi:beln på en li:ten vekka då svã:rt å tro: men de kan henda*
(Na.ymo2.sp)

till Houtskär var vi (.) och (.) hur länge skulle det ha kunnat räckt mánne det var en vecka det bara hann man faktiskt lära sig allt som som står i Bibeln på en liten vecka då svårt att tro men det kan hända

- (3.26) *ha blivi ganska li:te me mitt semestãrfi:rande på dãm hä se:naste å:rena at de e just just endast undãr s- stu:dieti:den sãm (.) när de ha vari nått li:tet glapp så ha man **kuna fari** på nãn re:sa*
(Pa.ywo4.sp)

har blivit ganska lite med mitt semesterfirande på de här senaste åren att det är just just endast under s- studietiden som (.) när det har varit något litet glapp så har man kunnat farit på någon resa

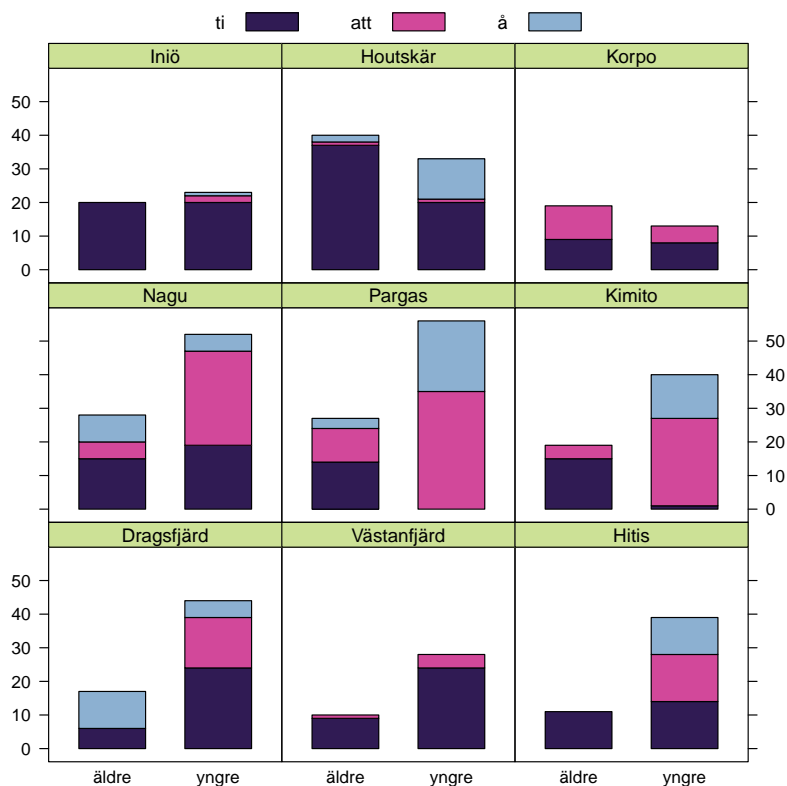
Verbens *s*-former är enligt Zilliacus (1992, 83) ”brokiga”. Någon systematisk genomgång av dessa har inte gjorts i 2000-talsmaterialet, men ett antal belägg på passivformer med *st*-ändelse förekommer bland informanterna från västra Åboland.

- (3.27) *de va no:g mykky lenggi sen de i så: fall (.) sãm e **fo:rest** men (.) men tjö:karbãor från tjö:kar (.) **fo:rest** de ti ti marknan i synnãrhe:ti ti helsingfãssj* (Ko.omo2.sp)

det var nog mycket länge sedan det i så fall (.) som det fors men (.) men Kökarbor från Kökar (.) fors det till till marknaden i synnerhet till Helsingfors

3.4.1 Infinitivmärket

Infinitivmärket förekommer i tre olika varianter i Åboland. Dessa är *ti*, *att* och *å*. Varianten *ti* (med uttalsvarianten *te*) är en dialektvariant som har funnits i alla dialekter i området. Den standardsvenska och skriftspråkliga varianten är *att*, medan *å* är en talspråksform och uttalsvariant av *att* (SAG 2, 745). Enligt Nyholm (1984, 62–63) är *att* den vanligaste varianten i Helsingforssvenskan och variabeln är klart polariserad så



Figur I.12. Stapeldiagram över antalet belägg på varianterna *ti*, *att* och *å* av infinitivmärket efter ort och åldersgrupp.

att talare tenderar att favorisera någon av de tre varianterna. Vad gäller användningen av infinitivmärke i de svenska dialekterna i Sverige har Hagren (2008) visat att varianterna *ti* och *att/å* används sida vid sida över hela landet. Fördelningen mellan de två varianterna är beroende av den syntaktiska eller lexikala konstruktion som infinitivfrasen ingår i. Mellan olika geografiska områden finns gradvisa skillnader i vilken av varianterna som är den vanligaste.

Infinitivmärket har excerperats manuellt i det åboländska materialet. Vid excerperingen har belägg som kan tolkas antingen som infinitivmärke eller som konjunktionen *och*, så som i exempel 3.28 nedan, inte tagits med. Fall där infinitivmärke saknas (där infinitivmärke skulle användas i standardspråket) har inte heller noterats.

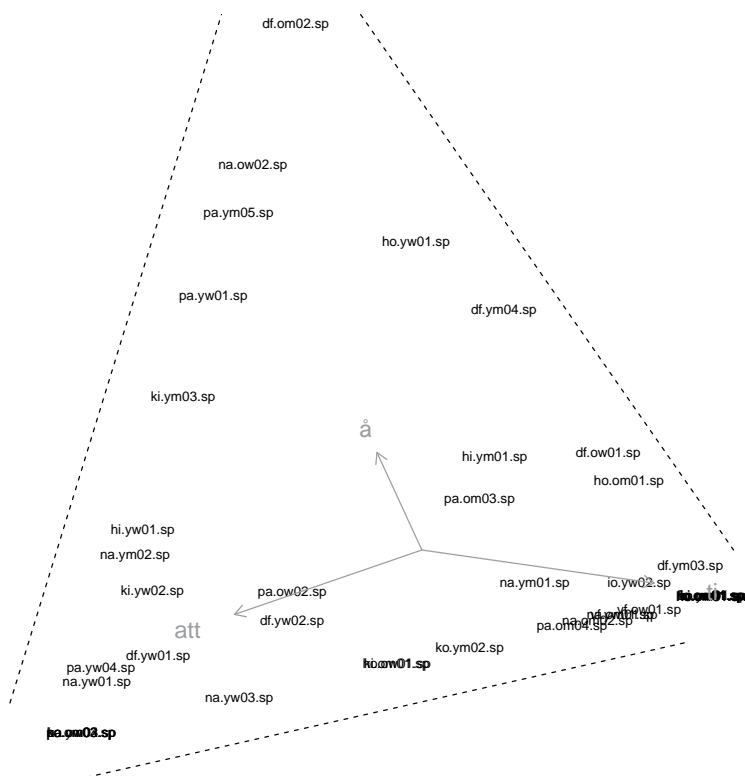
(3.28) *de va NN såm ringd å hadd en ka- en ko: såm höll på: å kalva å inga kla:ra u:t de* (Na.omo2.sp)

det var NN som ringde och hade en ka- en ko som höll på att/
och kalva/kalvade och inte klarade ut det

Det totala antalet belägg på variabeln i materialet är 520. Medianvärdet för antalet belägg per informant är elva, och sex informanter har färre än fem belägg. Varianten *ti* är vanligast i materialet med sammanlagt 266 belägg, följd av *att* med 161 belägg och *å* med 92 belägg. Figur I.12 visar antalet belägg på de tre varianterna efter ort och åldersgrupp. Varianten *ti* är den dominerande bland informanterna från Iniö, Houtskär och Västanfjärd. På övriga orter är variationen stor. Bland de yngre informanterna från Pargas och Kimito saknas dialektvarianten *ti* så gott som helt.

Figur I.13 visar i hur stor utsträckning de enskilda informanterna använder de tre olika varianterna. Signumen anger varje informants placering i figuren. Informanter som placeras i något av de tre hörnen i triangeln använder konsekvent en enda variant. Åtta av informanterna har bara belägg på *ti* (Io.omo1.sp, Io.owo1.sp, Ho.omo1.sw, Ho.owo1.sw, Ho.ymo1.sp, Ko.owo1.sp, Ki.owo1.sp, Hi.owo1.sp) och fyra informanter har bara belägg på *att* (Ko.omo2.sp, Ko.owo3.sp, Pa.owo3.sp, Pa.ymo4.sp). (Signumen för de informanter som konsekvent använder *ti* respektive *att* är oläsliga i figuren eftersom de placeras på varandra.) Endast en informant (Df.omo2.sp) använder bara varianten *å* under inspelningen. De informanter som i figuren ligger längs någon av de streckade linjerna växlar mellan två av varianterna, t.ex. Ki.ymo3.sp som har 9 belägg på *å* och 10 belägg på *att* och Ko.ymo2.sp som har 8 belägg på *ti* och 5 belägg på *att*. Informanter som ligger placerade längre in i triangeln växlar mellan alla tre formerna. Detta gäller nio av informanterna, t.ex. Pa.owo2.sp och Df.ywo2.sp som bägge har flest belägg på *att* men också enstaka belägg på de två övriga varianterna.

För att utforska vilka faktorer som styr valet av infinitivmärke användes logistisk regressionsanalys med blandade effekter (se bilaga A, 2.1). Analysen kompliceras av att variabeln har tre varianter och delades därför upp i två faser. I den första analysen studerades variationen mellan *ti* å ena sidan och *att* med uttalsvarianten *å* å andra sidan. I nästa steg analyserades variationen mellan *att* och *å*.



Figur I.13. Diagram över informanternas användning av de tre formerna av infinitivmärket. Pilarna anger i vilken riktning i figuren som användningen av var och en av de tre varianterna ökar. Ett antal informanter har bara belägg på en enda variant och placerar sig i något av triangelns hörn (vilket leder till informantkoder som överlappar varandra helt i de två nedre hörnen).

I en modell med informant som stokastisk effekt och ort och ålder som fixa effekter (bilaga A, 2.1, $t_i \sim att/\dot{a}$) uppvisar faktorerna ort och ålder signifikanta skillnader i fråga om andelen t_i . Som framkommer av figur I.12 ovan, är andelen t_i störst i Iniö, Houtskär och Västansfjärd och minst i mellersta Åboland. Äldre personer använder t_i i större utsträckning än de yngre. Mellan män och kvinnor finns ingen signifikant skillnad i användningen av t_i .

Eftersom det åboländska materialet inte är syntaktiskt taggat har en fullständig syntaktisk analys av infinitivmärket inte varit möjlig. Däremot har en grov analys av kontextens betydelse för valet av infi-

nitivmärke gjorts genom att föregående och följande ord i kontexten automatiskt tagits ut för varje belägg. Följande ord i kontexten utgörs i samtliga fall av ett verb i infinitiv. I regressionsmodellen lades stokastiska effekter in för verb och för föregående ord i kontexten. Verb som stokastisk effekt förbättrar inte modellen signifikant (likelihood-kvot-test, $p = 0,073$), vilket innebär att verbet som följer efter infinitivmärket inte spelar någon signifikant roll vid valet av variant. Däremot förbättras modellen signifikant av att man lägger till föregående ord i kontexten som stokastisk effekt (likelihood-kvot-test, $p < 0,001$). Modellen visar att en hög andel *ti* förekommer särskilt efter orden *svårt* och *bara*. Ordet *svårt* förekommer särskilt i uttrycken *svårt att säga* och *svårt att få*. Här är ett exempel från Västanfjärd:

- (3.29) *på seinare å:ren så ha: ja svå:rt ti få: li:ksåm (.) sån hä:r vitt
pårslin (.) å så ha he blivi så dy:rt imo:t va he va fössjt
(Vf.owo1.sp)*

på senare åren så har jag svårt att få liksom (.) sådan här vitt
porcelain (.) och så har det blivit så dyrt emot vad det var först

Informant Hi.ymo1.sp i följande exempel varierar under inspelningen mellan alla tre varianterna. Han har sammanlagt 25 belägg på infinitivmärket varav 14 utgörs av *ti*, som hos denna informant förekommer särskilt ofta efter *bara* och efter adjektiv (*bäst ti, lätt ti, roligt ti* etc.). I utdraget använder han alla tre varianter. Exemplet visar att även *att* och *å* kan förekomma efter adjektiv, men om man ser till alla belägg efter adjektiv hos denna informant är *ti* klart dominerande.

- (3.30) *den hä gå:r nu lite åftare å ta:r me:r bi:lar me: u:t så att (..) så att
de e enklare å kåm u:t på de vi:se (..) åkk så ha de endra på de vi:se
at nu: e den gra:tis den hä färjan så nu: e de enklare fö fålk att kåm
u:t de e ba: ti tjö:r åmbo:rd å kåm u:t föru:t så kåsta de (.) fö:r dåm
såm int bodde hä:r då (Hi.ymo1.sp)*

den här går nu lite oftare och tar mer bilar med ut så att (..) så att
det är enklare att komma ut på det viset (..) och så har det ändrat
på det viset att nu är den gratis den här färjan så nu är det enklare
för folk att komma ut det är bara att köra ombord och komma ut
förut så kostade det (.) för dem som inte bodde här då

Varianterna *att* och *å* används enligt regressionsmodellen särskilt efter orden *så* och *om*. De flesta beläggen på *så* + infinitivmärke finns i uttrycket *så att säga*, där samtliga informanter använder varianten *att*. Beläggen på *om* + infinitivmärke kommer ur uttrycket *tycka om att*, där alla tre varianterna av infinitivmärket används, men *att* och *å* dominerar. Informant Hi.ywo1.sp i föregående exempel har t.ex. tre belägg på *tycka om* + infinitivmärke: ett med *att* och två med *å*. Som ytterligare ett exempel kan informant Vf.ywo1.sp (exempel 3.31) nämnas. Hon har sammanlagt 28 belägg på infinitivmärket, varav 24 *ti* och 4 *att*. Två av de fyra beläggen på *att* finns i uttrycket *tycka om att*.

(3.31) *lesär jö:r vi int så åfta fö ja tykkär int åm att lesa* (Vf.ywo1.sp)

läser gör vi inte så ofta för jag tycker inte om att läsa

I följande steg analyserades variationen mellan varianterna *att* och *å* med hjälp av en regressionsmodell (bilaga A, 2.1, *att* ~ *å*). Vid denna analys utelämnades materialet från Korpo och Västansfjärd eftersom det inte finns några belägg på *å* från dessa orter. Varken ålder eller kön som faktorer är av betydelse för variationen mellan *att* och *å*. Mellan orterna finns det däremot signifikanta skillnader. Från de flesta orter finns det fler belägg på *att* än på *å*. Houtskär utgör ett undantag med betydligt fler belägg på *å* än på *att*.

Vilket verb som följer efter infinitivmärket verkar inte vara avgörande för distributionen mellan *att* och *å* (likelihood-kvottest, $p = 0,557$). Däremot har ordet före infinitivmärket betydelse (likelihood-kvottest, $p = 0,001$). Att uttrycket *så att säga* alltid har *att* som infinitivmärke har redan nämnts. Ett annat frekvent ord i materialet som favoriserar valet av *att* som infinitivmärke framför *å* är ordet *för*. Det finns många belägg på *för att* (och även *för ti*) men inga *för å*. Som exempel på en informant som allmänt sett verkar favorisera *å* som infinitivmärke men som ändå använder *att* efter *för* kan Pa.ywo1.sp nämnas. Hon har sammanlagt tretton belägg på infinitivmärket varav åtta *å* och fem *att*. Vid två av de fem beläggen på *att* är det föregående ordet *för*.

(3.32) *ja: hadd int (.) tirekkit me arbe:tsä:rfa:renhe:t (.) för att kona sö:ka di:t* (Pa.ywo1.sp)

jag hade inte (.) tillräckligt med arbetserfarenhet (.) för att kunna söka dit

Som redan har framgått är varianten *att* överlag vanligare än *å*. Bland de mer frekventa orden som finns belagda med efterföljande infinitivmärke finns det inget där *å* skulle vara den dominerande varianten. Vid de ord som har de högsta andelarna av *å* är *å* och *att* ungefär lika vanliga. Detta gäller t.ex. ordet *om* med nio belägg på efterföljande *att* (som i exempel 3.31 ovan) och nio efterföljande *å* (exempel 3.33).

(3.33) *ja tykt inga åm å gå: i skola he va inga als ro:lit* (Pa.omo3.sp)

jag tyckte inte om att gå i skola det var inte alls roligt

Figurerna I.12 och I.13 visar på en mycket stor variation vad gäller infinitivmärket. Denna variation kan vid första anblicken tyckas godtycklig. Men en närmare analys har visat att variationen mellan de tre varianterna åtminstone delvis är systematisk och styrs av både utomspråkliga och inomspråkliga faktorer. Analysen av kontextens betydelse för valet av infinitivmärke har visat att *ti* som infinitivmärke är särskilt vanligt då infinitivmärket föregås av ett adjektiv, medan *ti* är ovanligt i uttrycket *tycka om att*. Dessa resultat påminner om de resultat som Hagren (2008) har redovisat för dialekterna i Sverige. Hennes resultat visar att *ti* dominerar bl.a. efter adjektiv och prepositioner, medan *å* (Hagren gör ingen skillnad mellan *å* och *att*) är vanligare än *ti* vid vissa partikelverb (Hagren 2008, 234). I dialekter med hög benägenhet för *ti* är *å* enligt Hagren (2008, 230) vanligt bara vid partikelverb.

Eftersom resultaten av föreliggande undersökning pekar åt samma håll som Hagrens resultat kan man mycket väl tänka sig att variationen mellan *ti* och *å* i de traditionella dialekterna i Åboland har varit liknande som för dialekter med hög benägenhet för *ti* i Sverige (Sveamål, främst uppsvenska). Varianten *ti* har då möjligen inte varit allenarådande som infinitivmärke, utan framför allt vid partikelverb kan även *å* ha använts. Vid dialektutjämning införs den skriftspråkliga varianten *att* som komplicerar variationsmönstren ytterligare.

En fullständig syntaktisk analys av variationen vad gäller infinitivmärke skulle kunna säga betydligt mer om variationen i Åboland än vad analysen ovan har möjliggjort. Med hjälp av ett större dialektmaterial

än det som har använts här kunde definitionen av kontext också utökas till bigram eller trigram (dvs. kombinationer av två eller tre ord) i stället för bara ett ord till höger och vänster om infinitivmärket. På så vis skulle fler systematiska mönster kunna upptäckas, även vad gäller variationen mellan *å* och *att*.

3.4.2 Infinitiv

Infinitiven har i Åboland två olika former. Den kan stå antingen med eller utan *-a*. Reglerna för när infinitiv-*a*:et faller bort och när det bibehålls har varit olika på olika håll i Åboland. Socknarna från Pargas österut har haft villkorlig apokope, vilket innebär att ändelsebortfallet är beroende av verbets ställning i satssammanhanget. Ändelsevokalen faller då ofta framför ett ord som är nära förbundet med verbet (Ahlbäck 1945, 29–30). Fristående, tryckstarka infinitiver behåller däremot slutvokalen. I motsats till östra Åboland har apokoperingen i västra Åboland varit beroende av kvantitet. Dialekterna i västra Åboland har haft ovillkorlig apokope av långstaviga verb, dvs. ändelsevokalen har fallit bort i alla ställningar. Men för kortstaviga ord har reglerna för villkorligt bortfall gällt också i västra Åboland.

Apokopen har också påverkat kvantiteten hos kortstaviga infinitiver. I västra Åboland har kortstaviga ord förlängts vid vokalbortfall. Variationen för t.ex. verbet *fara* har alltså skett mellan *fa:r* och *fara* i västra Åboland, men mellan *far* och *fara* i östra Åboland (Harling-Kranck 1998, 143).

I Houtskär har *-a* gått över till *-å* i kortstaviga ord (Zilliacus 1992, 57), som i utdraget nedan:

- (3.34) *ja:g ha no: mest bakå veitbrö:d ja å tå: sen nå vi va li:te såm största hu:shålle så kond e ga: en tjugufem brö:d i vikun så (.) va no: ti bakå he ibla:nd kond e he:nd att man bakå på måndain å så va e slu:t på fri:dain så fikk man å:tår bör bakå* (Ho.owo1.sw)

jag har nog mest bakat vetebröd jag och då sedan när vi var lite som största hushållet så kunde det gå en tjugofem bröd i veckan så (.) var nog att baka det ibland kunde det hända att man bakade på måndagen och så var det slut på fredagen så fick man åter börja baka

Följande två utdrag visar exempel på hur reglerna för villkorlig apokope fungerar i östra Åboland. I exempel 3.35 står infinitiverna *lyda*, *niga* och *bocka* i slutet av en sats och behåller då slutvokalen. Infinitiven *säga* som följs av ett objekt har däremot vokalbortfall. Exempel 3.36 visar en typisk konstruktion där ändelsevokalen faller bort. Den relativt svagt betonade infinitiven *hitta* följs av partikeln *på* som har betoning. Infinitiven står då i s.k. upptaktsställning till partikeln och förlorar sin ändelsevokal. Trots att det går att ställa upp vissa regler för den villkorliga apokopen är det enligt Ahlbäck (1945, 29) inte ovanligt med avvikelser från reglerna.

- (3.35) *ja: no sko man ly:da (.) int va he ti sejj någå du: å:t å:t lä:rarn int at no sko man ni:ga å båkka* (Ki.owo1.sp)

ja nog skulle man lyda (.) inte var det att säga något du åt åt läraren inte att nog skulle man niga och bocka

- (3.36) *nu bru:kar nu påjkana hitt på: någå sårts hyss alti dām*
(Ki.owo1.sp)

nu brukar nu pojkarna hitta på någon sorts hyss alltid de

I västra Åboland har långstaviga ord som nämnts ovillkorlig apokope. Följande utdrag från Houtskär ger exempel på ett långstavigt verb (*spring*) före ett adverbial som uppbär betoningen samt ett belägg på långstavigt verb i satsslut (*ro:p*):

- (3.37) *nä:r in fikk sy:n på na:in så måsta in spring tiba:k ti tre:de å ro:p*
(Ho.ymo1.sp)

när en fick syn på någon så måste en springa tillbaka till trädet och ropa

Medan reglerna för apokopering av långstaviga ord är relativt entydiga, med ovillkorlig apokope i västra Åboland och villkorlig i östra, är förhållandena för kortstaviga ord som ovan framgått mer komplicerade. Förutom variation mellan apokoperade och icke-apokoperade former finns det också en variation i huruvida stamvokalen förlängs eller inte vid apokopering av kortstavingar. Enligt Zilliacus (1992, 62) är situationen vad gäller apokopering ”oklar för många ordtyper och

särskilt när det gäller kortstavingarna”. För 2000-talsmaterialets del kompliceras situationen ytterligare av att ursprungligt kortstaviga ord kan uppträda i långstavig form som resultat av dialektutjämning (se avsnitt 2.3). Eftersom variationen hos de kortstaviga infinitiverna är så mångfasetterad skulle ett betydligt större dialektmaterial krävas för en systematisk analys. Redovisningen nedan fokuserar därför på ursprungligt långstaviga verb.

Infinitiverna har excerperats manuellt i materialet. Ett problem uppstår vid pseudosamordningar av verb, där det inte alltid är möjligt att avgöra om det senare ledet verkligen är ett samordnat finit verb eller en infinitiv. I dessa konstruktioner används enligt Bergroth (1928, § 162) i finlandssvenskan ofta infinitiv där standardsvenskan har en finit form. För att ta ett av Bergroths exempel motsvaras det finlandssvenska *han va ute å ro* av *han va ute å rodde* i standardsvenskan. I exemplet med verbet *ro* är det entydigt att verbet står i infinitiv i den förstnämnda satsen. För verb som böjs enligt första konjugationen, där de finlandssvenska dialekterna och finlandssvenskan i allmänhet har bortfall av *-de* i preteritum och av *-t* i supinum (och preteritum- och supinumformerna därmed är identiska med infinitiven), kan det däremot vara svårt att avgöra vilken verbform det är frågan om. Problemet förstärks av att infinitivmärket *att* och konjunktionen *och* bägge kan ha uttalet *å*.

I många finlandssvenska dialekter används liksom i standardsvenskan pseudosamordning av verb (se t.ex. Lundström 1939, 107; Ivars 2010, 133–136). I det åboländska materialet finns det belägg både på pseudosamordnade verb och konstruktioner där pseudosamordningens andra led motsvaras av en infinitivfras. Exempelen 3.38–3.41 nedan, som alla visar konstruktioner med verbet *fara*, belyser problemet med att avgöra vilken av de två konstruktionerna det handlar om. I exempel 3.38 och 3.39 är det andra verbet (*sova* resp. *följa*) en tydlig infinitiv, och i exempel 3.40 står verbet *se* i preteritum. Problemet uppstår i exempel 3.41 där verbet *simma* kan tolkas antingen som en infinitiv eller en preteritumform.

(3.38) *de ha:r int nå sjilna åm ja e på jäbb elä int så int **fa** ja å såva före tålv* (Ki.ymo3.sp)

det har inte någon skillnad om jag är på jobb eller inte så inte far jag och sova före tolv

- (3.39)
- så fo:r vi å följ henne*
- (Ki.owo1.sp)

så for vi och följa henne

- (3.40)
- nå:gån eftermidda: elär va de va så fo:r vi (.) fo:r vi å så: på de*
-
- (Df.omo2.sp)

någon eftermiddag eller vad det var så for vi (.) for vi och såg på det

- (3.41)
- sen hadd vi ju alti nångång at man for in ti ti å:bo å simma*
-
- (Na.ywo1.sp)

sedan hade vi ju alltid nångång at man for in till till Åbo och simma/simmade

För att undvika att exciperera belägg som inte säkert är infinitiver (eller bara exciperera apokoperade belägg som avslöjar att det handlar om en infinitiv, vilket skulle ge en skev bild av distributionen) skulle man kunna undvika alla verb av första konjugationen i konstruktioner som kan tolkas som pseudosamordningar. Men verben följer inte alltid samma konjugationsmönster i dialekterna som i standardspråket (se t.ex. utdrag 3.44 nedan). Därför har samtliga verb i denna typ av syntaktisk konstruktion utelämnats vid exciperingen, dvs. inget av verben i exempel 3.38–3.41 ingår i analysen av infinitivapokope. Endast då även det första verbet i samordningen också är en infinitiv (som i 3.42 nedan) kan man med säkerhet veta att även det senare verbet är en infinitiv, och i dessa fall har båda infinitiverna exciperats.

- (3.42)
- de flyttades ba: ti e:n ö: (.) dä:r riktigt mitt emo:t så at de va no:g*
-
- nä:ra att fara å mjölka dåm dä:r ju:ren*
- (Ko.owo3.sp)

det flyttades bara till en ö (.) där riktigt mitt emot så att det var nog nära att fara och mjölka de där djuren

Efter den manuella exciperingen delades infinitiverna in i två grupper utifrån följande ord i kontexten. De två grupperna utgörs av infinitiver i *satsfinal ställning* och infinitiver i *upptaktsställning* och är tänkta att motsvara reglerna för villkorlig apokope. En syntaktisk taggning av dialektmaterialet hade möjliggjort en betydligt mer tillförlitlig indelning av infinitiverna. Eftersom inget syntaktiskt taggat material förelåg är

indelningen rudimentär och en stor del av infinitiverna har inte kunnat hänföras till någondera gruppen. Det ord eller tecken i transkriptionen som följer efter infinitiven har fått stå som grund för indelningen. En lista över ord och tecken som definierar vardera kontexten gjordes upp manuellt, och därefter skedde klassificeringen automatiskt. En infinitiv har klassats som satsfinal om följande ord är en konjunktion eller subjunktion eller om infinitiven står i slutet av en tur eller före ett paustecken. Vissa adverb som oftast är obetonade och kan användas som konjunktioner (som *därför*, *sedan*, *så*) har också inkluderats i definitionen av satsfinal kontext.

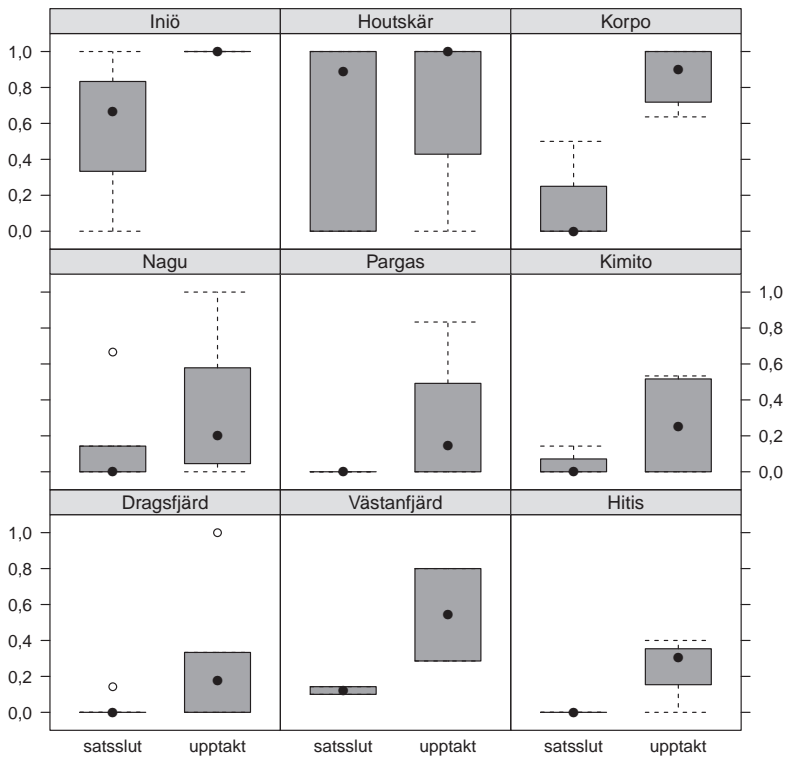
Enligt SAG (3, 417) är partikeladverbialen den av verbets bestämmningar ”som är närmast knuten till verbet”. Partikeladverbialen utgörs enligt SAG nästan alltid av antingen ett kort adverb som anger riktning eller en preposition utan rektion. Partikeladverbialen följer direkt efter det infinitiva verbet och har stark betoning medan verbet är svagt betonat. Adverb och prepositioner som kan fungera som partikeladverbial enligt SAG (kap. 20 Verbfraser: Adverbial § 7–8) har fått stå som definition för infinitiver i upptaktsställning. I gruppen av ord som definierar upptaktsställningen har dessutom reflexiva pronomen inkluderats (jfr Ahlbäck 1956, 37). Klassificeringen har skett automatiskt vilket innebär att ingen hänsyn har tagits till om prepositionerna de facto fyller funktionen som partikeladverbial eller om de står i en prepositionsfras. Definitionen av kontext är således grov.

Endast tvåstaviga verb (eller sammansatta verb där den senare leden är tvåstavig) har fått ingå i analysen. Motiveringen till detta är att antalet stavelser påverkar rytmförhållandena vilket kan inverka på apokoperingen. På så vis har t.ex. verb som slutar på *-era* (*koncentrera*, *studera*) uteslutits. Enstaviga verb som även i standardspråket saknar *-a* (*dö*, *ha*, *ge*, *stå* etc. samt kortformer som *va* av *vara*) har givetvis också uteslutits. Det totala antalet tvåstaviga, långstaviga infinitiver i materialet är 852. Av dessa har 181 klassificerats som infinitiver i satslut och 300 som infinitiver i upptaktsställning (371 har alltså inte kunnat hänföras till någon av grupperna). Medianvärdet för antal belägg per informant är 6 i upptaktsställning och 4 i satslut. Två informanter saknar helt belägg på infinitiver i satsfinal kontext.

Figur I.14 visar lådagram över andelen apokoperade belägg i de båda kontexterna för varje ort. Utifrån beskrivningarna av de traditionella

sockendialekterna i området skulle man förvänta sig en mycket hög andel apokope i bägge kontexterna för Iniö, Houtskär, Korpo och Nagu. För de östligare orterna skulle man däremot förmoda låga andelar apokope i satsslut och höga andelar apokope i upptaktsställning. Denna bild bekräftas bara delvis av figur I.14.

En regressionsanalys med informant som stokastisk effekt (se bilaga A, 2.2) tillämpades på materialet. Analysen visar att det finns ett starkt samband mellan andelen apokoperade former och kontexten. Andelen apokoperade former är signifikant lägre i satsslut än i upptaktsställning, vilket syns tydligt i figur I.14. Det finns också en stark effekt för ort, vilket innebär att andelen apokoperade former är generellt högre på vissa orter än på andra. Figuren visar att de högsta genomsnittliga andelarna apokope finns i Iniö och Houtskär. Däremot finns det ingen



Figur I.14. Lådagram över andelen dialektvarianter (apokoperade former) av infinitiver i satsslut och i upptaktsställning per ort.

signifikant interaktion mellan ort och kontext. Utifrån det vi vet om de traditionella dialekterna i Åboland skulle en sådan interaktion vara att vänta eftersom kontexten har varit avgörande i östra Åboland men inte i västra. Men i 2000-talsmaterialet ser mönstret lika ut för alla orter, dvs. andelen apokoperade former är högre i upptaktsställning än i satsslut också i västra Åboland. Att slutvokalen införs för långstaviga infinitiver i västra Åboland kan ses som en utveckling i riktning mot standardspråket. Men det är anmärkningsvärt att vokalen införs enligt reglerna för villkorlig apokope. Man kan tänka sig två olika förklaringar till detta. Antingen förs reglerna om villkorlig apokope över från de kortstaviga verben till att gälla även långstaviga, eller också har den villkorliga apokopen av långstavingar införts genom inflytande från öståboländska dialekter.

Den kvantitativa analysen visar inga signifikanta skillnader för faktorerna ålder och kön.

Om man ser till de enskilda informanterna kan man konstatera att i upptaktsställningen, där alla sockendialekter har haft apokope, har sammanlagt elva informanter inga belägg alls på apokope utan endast belägg med slutvokal. Dessa informanter kommer huvudsakligen från östra Åboland. Det finns också elva informanter som apokoperar infinitiverna i 100 procent av fallen i denna ställning. De flesta av de informanter som apokoperar konsekvent kommer från Iniö, Houtskär och Korpo. I satsslut däremot har bara fyra informanter endast apokoperade belägg. Dessa är de två äldre informanterna från Björkö (Houtskär) och de två yngre kvinnorna från Iniö och Houtskär. Hela 28 informanter har inga belägg på apokope i satsslut.

Som framkommit ovan är antalet belägg i vardera kontexten relativt litet per informant, om man räknar bara ursprungligt långstaviga ord. Samtidigt ser det ut som att långstaviga verb även i västra Åboland delvis styrs av reglerna för villkorlig apokope, och den förmodade skillnaden mellan långstaviga och kortstaviga verb blir därmed mindre än väntat. För den kvantitativa analysen som redovisas i kapitel 4 har därför även de kortstaviga verben inkluderats. Det sammanlagda antalet belägg blir på så vis 996, varav 521 i upptaktsställning och 475 i satsslut. Medianvärdet för antal belägg per informant är 12 i upptaktsställning och 10 i satsslut, och ingen informant saknar helt belägg i någon av de två kontexterna. Produktmomentkorrelationen mellan en analys som

inkluderar bara långstavingar och en analys som inkluderar alla typer av verb vid beräkningen av andelen apokoperade infinitiver är hög: 0,837 ($n = 41, p < 0,001$) i satsslut och 0,887 ($n = 43, p < 0,001$) i upptaktsställning. De höga korrelationskoefficienterna antyder att en analys av alla verb i stort sett mäter samma sak som en analys av bara långstavingar.

3.4.3 Preteritum

Såsom finlandssvenskan i allmänhet saknar de åboländska dialekterna ändelsen *-de* i preteritum av första konjugationens verb. I materialet finns ytterst få belägg på att preteritumändelsen uttalas i denna verbklass.

- (3.43) *vi **starta** vel fy:rati:dn på mårån från å:bo å (..) å ja tenkt at nu sku ja no ha bo:rda narr dâm dä:r å aprill no: (.) å dâm satt no å så:v i bossn [...] så bakâm en korva så (..) så så **bråmsa** ja in rikit haståt å (..) å så reiv ja ta:g i: i yksen sâm e i bossn de måsta alti förr i bossana va: en sân dä:r yks [...] ne men så:g ni inga då han där han där älgin sâm va: va: (.) på:tjö:rd hä:r bakâm bakâm den där korvan (.) att den lå:g å **sparka** no måst man fara å slå: ta: li:ve å:v den int kan man lemn den dä:r ti pi:nas [...] så **ru:sa** dâm alliho:p u:t [...] å ja sto: i bossdören sen ja när när de var halvve:gs di:t å så **ro:pa** ja aprill aprill (Na.omo2.sp)*

vi startade väl fyratiden på morgonen från Åbo och (..) och jag tänkte att nu skulle jag nog ha bort narra dem där också april nog (..) och de satt nu och sov i bussen [...] så bakom en kurva så (..) så så bromsade jag in riktigt hastigt och (.) å så rev jag tag i i yxan som är i bussen det måste alltid förr i bussarna vara en sådan där yxa [...] nej men såg ni inte då den där den där älgan som var var (.) påkörd här bakom bakom den här kurvan [...] att den låg och sparkade nog måste man fara och slå ta livet av den inte kan man lämna den där att pinas [...] så rusade de allihop ut [...] och jag stod i bussdörren sedan jag när när de var halvvägs dit och så ropade jag april april

I motsats till apokopering av första konjugationens verb, som är allmän i all finlandssvenska, är apokopering i preteritum av andra och tredje konjugationens verb kännetecknande framför allt för dialekterna (Ahl-

bäck 1945, 5, 28). I städerna förekommer denna typ av apokopering framför allt bland personer med lägre utbildning (Nyholm 1984, 71; Gullmets-Wik 2004, 201–202). Apokoperingen i preteritum är inte underkastad reglerna för villkorlig apokope, utan preteritum apokoperas ovillkorligt också i dialekter som har villkorlig apokope i infinitiv (Ahlbäck 1945, 28–29; Nyholm 1984, 69–70; Gullmets-Wik 2004, 143).

Variationen hos variabeln 'apokope i preteritum av andra och tredje konjugationens verb' sker mellan den dialektala formen med apokoperad ändelsevokal (*bodd, hadd, tykt*) och den standardsvenska icke-apokoperade formen (*bodde, hade, tyckte*). För variabeln har alla andra och tredje konjugationens verb och därmed jämförbara oregelbundna verb (*ha, göra, kunna, veta* etc.) excerperats. Men verbkonjugationen i de åboländska dialekterna är inte alltid identisk med den standardsvenska. Verb som böjs enligt andra eller tredje konjugationen i standardspråket behöver inte nödvändigtvis göra det i dialekterna. Vid excerperingen har först alla verb som i standardspråket har preteritumallomorfen *-de* eller *-te* excerperats automatiskt. Belägg som av informanten i fråga realiseras med en annan allomorf har i följande steg uteslutits manuellt. Exempel på detta är former som *kona* 'kunde' och *å:ka* 'åkte'. Vissa informanter växlar mellan olika paradigm, som den äldre mannen från Dragsfjärd i följande två exempel. Han använder först preteritumformen *å:ka* och lite senare under intervjun *å:kte*.

(3.44) *nå: de fandes int mykky bi:lar ne: hä:r de va (.) de va nu bussar såm såm vi: (.) såm såm å:ka då* (Df.omo2.sp)

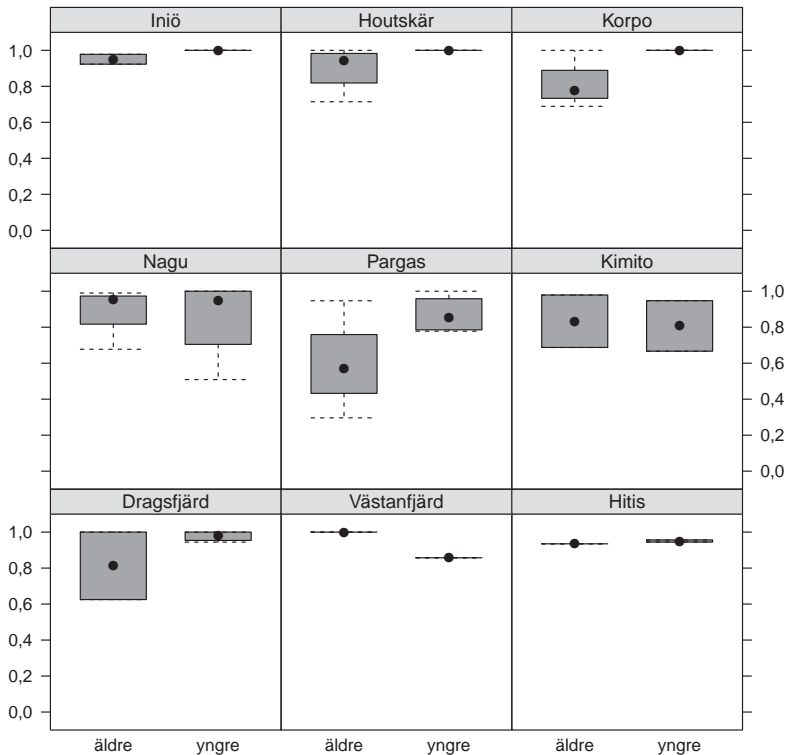
nå det fanns inte mycket bilar nej här det var (.) det var nu bussar som som vi (.) som som åkte då

(3.45) *så att ja: ja: å:kte in me me me buss di:t senn* (Df.omo2.sp)

så att jag åkte in med med med buss dit sedan

Av beläggen på verbet *åka* i utdragen ovan har det i exempel 3.45 räknats som ett belägg på variabeln 'preteritum av andra och tredje konjugationens verb', medan det i exempel 3.44 inte har medräknats.

Det totala antalet belägg på variabeln i materialet är 1 600 och medianvärdet per informant är 29. Den informant som har minst belägg har 6 stycken och informanten med flest belägg har 98. Figur I.15 visar



Figur 1.15. Lådagram över andelen dialektvarianter (apokoperade former) i preteritum av andra och tredje konjugationens verb efter ort och ålder.

lådagram över andelen apokoperade former av alla belägg på variabeln efter ort och åldersgrupp. Figuren visar mycket höga andelar av apokoperade former för alla orter och åldersgrupper. Hela tretton informanter har apokoperad ändelsevokal i samtliga belägg. Den informant som har den lägsta andelen apokope (en äldre kvinna från Pargas) apokoperar ändelsevokalen i 30 procent av fallen. Detta är det bäst bevarade av alla dialektdrag i det undersökta materialet.

Som figuren visar är skillnaden mellan olika orter och åldersgrupper liten. En logistisk regressionsanalys med blandade effekter (se bilaga A, 2.3) med informant som stokastisk effekt ger signifikanta effekter för ort och åldersgrupp men ingen signifikant effekt för kön. Pargas har den lägsta andelen dialektvarianter, och yngre informanter använder apokoperade former i något större utsträckning än de äldre. Modellen

förbättras signifikant av att man lägger till ord som stokastisk effekt (likelihood-kvottest, $p < 0,001$). Stark benägenhet för att apokoperas har orden *hade, kunde, lärde, satte* och *skötte*, medan orden *hörde, köpte* och *ringde* har en relativt svag benägenhet för apokopering. Något tydligt mönster kan inte ses i fråga om vilka ord som apokoperas oftare än andra och varför.

Samtalsämnet verkar spela en roll åtminstone för vissa informanter. En av de äldre kvinnorna från Korpo apokoperar preteritumformerna ganska konsekvent när hon talar om livet ute i skärgården, men när hon berättar om arbetslivet på Kyrklandet använder hon betydligt fler standardformer. Som exempel kan nämnas att hon har sju belägg på verbet *tycka* i preteritum, av vilka tre är apokoperade och fyra är fullformer. De tre apokoperade beläggen produceras alla när det är tal om hemön, medan de fyra icke-apokoperade produceras på tal om skolgång och arbetsliv.

3.5 Negationen

Som negation används vid sidan av *inte* även *inga*. De båda negationsorden förekommer sida vid sida även i äldre dialektmaterial (se t.ex. uppslagsorden *inte* och *inga* i FO och språkprov hos Zilliacus 1992). Om det finns någon syntaktisk eller semantisk skillnad i användningen av de båda varianterna är outrett.

Ordet *inga* verkar vara ett iöronfallande dialektdrag som utomstående reagerar på, vilket kommer fram i intervjun med Df.ywo2.sp när intervjuaren frågar om det finns drag i Dragsfjärdsdialekten som folk utifrån fäster sig vid:

- (3.46) *just de hä:r inga* (.) *så de e många såm no:g li:te så hä skrattar å:t*
(.) *elä: såm tykkär att de lå:ter fe:l å fremande* (.) *nä man sejä inga*
(.) *i ställe fö inte* (Df.ywo2.sp)

just det här inga (.) *så det är många som nog lite så här skrattar*
åt (.) *eller som tycker att det låter fel och främmande* (.) *när man*
säger inga (.) *i stället för inte*

Det är mycket vanligt att negationen dubblas i slutet av satsen (jfr Ivars 2009, 61–64). Det förekommer både att samma variant upprepas

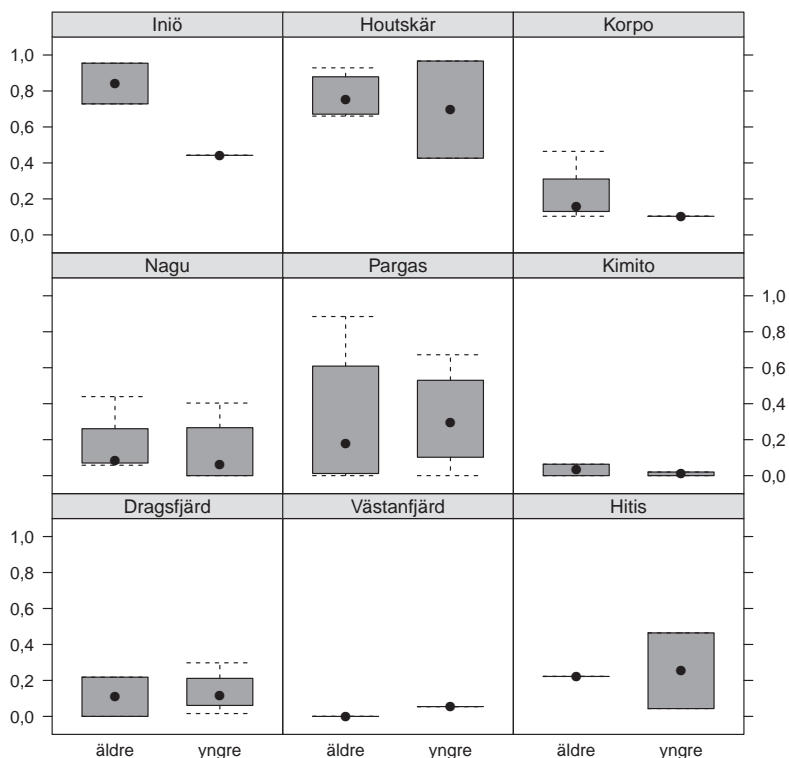
och att *inte* dubbleras med *inga* eller tvärtom. Hos vissa informanter är finaldubbling betydligt vanligare än hos andra.

- (3.47) *man kan ju **int** kritise:ra alt **int*** (Ki.ymo3.sp)
man kan ju inte kritisera allt inte
- (3.48) *he va **inga** he:ft **int*** (Pa.omo3.sp)
det var inte hävt inte
- (3.49) *vi e **inga** u:tan musi:k **int*** (Ho.omo1.sp)
vi är inte utan musik inte
- (3.50) *he va: **inga** någå vi:dare de **inga** vett do* (Pa.omo3.sp)
det var inte något vidare det inte vet du
- (3.51) ***inga** sku ja tjenna så: **inga*** (Ho.ymo1.sp)
inte skulle jag känna så inte
- (3.52) ***int** va ja någå specielt bra: **inga*** (Pa.omo3.sp)
inte var jag något speciellt bra inte
- (3.53) *hallåve:n brukar vi **int** jöra så mykke **inga*** (Hi.ywo1.sp)
Halloween brukar vi inte göra så mycket inte

Vid finaldubbling reduceras negationen ibland till bara ett *i* eller *e* (jfr FO bd III, s. 150):

- (3.54) *då ja va ba:rn så fans hä:r no **ingga** riktig nå sje:l hellär e*
(Io.owo1.sp)
då jag var barn så fanns här nog inte riktigt någon säl heller inte

Vid kvantifieringen har alla ord som motsvaras av *inte* i den standardiserade transkriptionen excerperats automatiskt. Varianten *inga* har uttalsvarianterna *inga* och *ingga*. Varianten *inte* uttalas oftast *int*. Uttalet *inte* utan apokoperad slutvokal förekommer i enstaka fall, främst bland de äldre informanterna.



Figur I.16. Lådagram över andelen dialektvarianter (*inga*) av negationen efter ort och ålder.

Det totala antalet belägg i materialet är 2 141, varav 1 525 *inte* och 616 *inga*. Medianvärdet för belägg per informant är 51. Den informant som har det minsta antalet har 18 belägg, medan det största antalet belägg är 99.

Figur I.16 visar lådagram över andelen *inga* som negation efter ort och åldersgrupp. Figuren visar att *inga* används nästan överallt men är mest frekvent i Iniö och Houtskär. Minst frekvent verkar *inga* vara på Kimitoön.

4

Kvantitativ analys

I detta avsnitt behandlas de tretton språkliga variabler som har kvantifierats i en sammanlagd analys. Variablerna listas i tabell I.4, som ger exempel på varje variabel och anger antalet belägg och andelen dialektvarianter. Varje variabel har getts en förkortning för att underlätta behandlingen av variablerna i figurerna och i den löpande texten.

Den mest frekventa variabeln i materialet är NeutrPron med sammanlagt 4 778 belägg i materialet. De minst frekventa variablerna är NorrlF och MaskB som bägge har kring 400 belägg i materialet. Medianvärdet för antal belägg per informant är 8 för de minst frekventa variablerna (NorrlF och MaskB) och 99 för den mest frekventa variabeln (NeutrPron). För sju av variablerna har alla informanter minst fem belägg. För de sex variabler som har ett minimiantal belägg under fem gäller att de informanter som har färre än fem belägg är mycket få till antalet.

Andelen dialektvarianter i tabell I.4 är uträknad per informant. Noll procent innebär bara standardvarianter medan hundra procent innebär bara dialektvarianter. Flera av variablerna har ett minimivärde på 0 procent och ett maximivärde på 100 procent, vilket visar på stor variation bland informanterna. Det finns alltså både informanter som bara använder standardvarianten av variabeln i fråga och sådana som bara använder dialektvarianten. Den enda variabeln för vilken samtliga informanter har belägg på dialektvarianten är PretApo. Den stora skillnaden mellan medeltal och medianvärde för vissa av variablerna visar på en skev distribution.

Tabell 1.4. Variabler i den kvantitativa analysen. Förutom det totala antalet belägg i materialet anges minimum, median och maximum för antalet belägg per informant. Andelen dialektvarianter anges i procent uträknat per informant. För andelen dialektvarianter anges medeltal samt minimum, median och maximum.

	VARIABEL		
	FÖRKORTNING	VARIANTER	EXEMPEL
<i>ei</i> -diftongen	PrimDift	<i>ei</i> ~ <i>e:</i>	<i>veit</i> ~ <i>ve:t</i>
bevarat <i>a</i>	LangtA	<i>a(:)</i> ~ <i>å(:)</i>	<i>ga(:)ng</i> ~ <i>gång</i>
kort <i>u</i>	KortU	<i>o</i> ~ <i>u</i>	<i>ront</i> ~ <i>runt</i>
<i>ng</i> -ljudet	Ngg	<i>ngg</i> ~ <i>ng</i>	<i>penggar</i> ~ <i>pengar</i>
norrländsk förmjukning	NorrlF	<i>tj</i> ~ <i>k(k)</i>	<i>mytjy</i> ~ <i>myky/</i> <i>mykke</i>
kortstavighet	Kortst	VC ~ V:C/VC:	<i>dagar</i> ~ <i>da:gar</i>
bestämd slutartikel, maskuliner	MaskB	<i>-i(n)/-n</i> ~ <i>-en</i>	<i>bå:ti(n)/</i> <i>bå:tn</i> ~ <i>bå:ten</i>
neutralt pronomen	NeutrPron	<i>het</i> ~ <i>det</i>	<i>(h)e</i> ~ <i>de</i>
infinitivmärket	InfM	<i>ti</i> ~ <i>att/å</i>	
infinitiver, satsfinal ställning	InfApoFi	<i>-∅</i> ~ <i>-a</i>	<i>spring</i> ~ <i>springa</i>
infinitiver, upptaktsställning	InfApoUp	<i>-∅</i> ~ <i>-a</i>	<i>spring</i> ~ <i>springa</i>
preteritum av 2:a och 3:e konjugationens verb	PretApo	<i>-∅</i> ~ <i>-e</i>	<i>hett</i> ~ <i>hette</i>
negation	Neg	<i>inga</i> ~ <i>inte</i>	<i>ing(g)a</i> ~ <i>int(e)</i>

	ANTAL BELÄGG				ANDEL (%)				BESKRIVS I AVSNITT
	TOTALT	MIN.	MEDIAN	MAX.	MEDELTAL	MIN.	MEDIAN	MAX.	
	962	7	21	48	14,7	0	6,3	77,8	2.1.1
	1 425	10	30	97	11,7	0	0	97,8	2.1.4
	637	5	14	44	49,4	0	50,0	100,0	2.1.5
	1 992	23	42	101	8,1	0	0	60,9	2.2.1
	392	3	8	28	6,4	0	0	85,7	2.2.6
	474	0 (3)	10	30	54,6	0	63,6	100,0	2.3.1
	407	1	8	24	53,4	0	60,0	100,0	3.1
	4 778	59	99	233	17,9	0	3,7	98,0	3.3.1
	519	1	11	28	48,4	0	50,0	100,0	3.4.1
	475	2	10	23	22,7	0	11,8	100,0	3.4.2
	521	2	12	23	40,3	0	36,4	100,0	3.4.2
	1 600	6	29	98	86,9	29,6	94,7	100,0	3.4.3
	2 141	18	51	99	29,5	0	15,7	96,7	3.5

I avsnitt 4.1 jämförs de olika variablerna med varandra och i 4.2 jämförs de enskilda informanterna. I avsnitt 4.3 redovisas resultaten av en principalkomponentanalys som visar relationen mellan variabler och informanter.

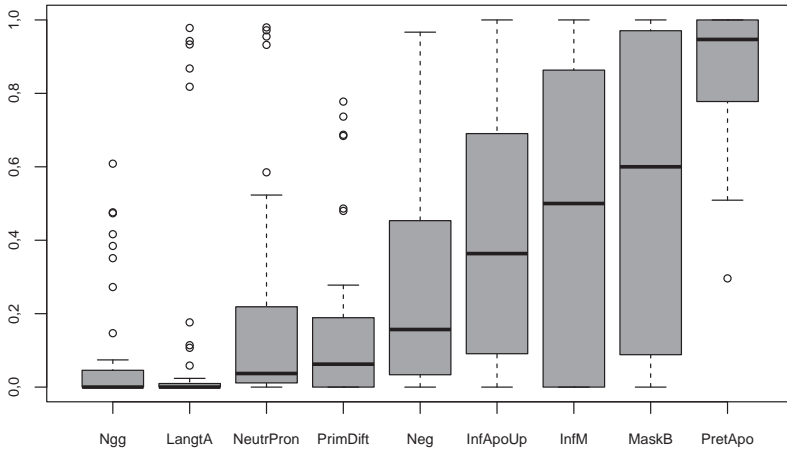
4.1 Jämförelse av variablerna

Figurerna I.17, I.18 och I.19 visar fördelningen av andelen dialektvarianter per variabel. I figur I.17 visas de nio dialektdrag som har kvantifierats och som enligt Zilliacus (1992) tidigare har funnits i samtliga åboländska sockenmål. Lådorna i detta diagram bygger på andelar dialektvarianter hos alla 43 informanter. Ett högt medianvärde (det svarta strecket i varje låda) anger alltså en hög genomsnittlig användning av dialektvarianten. Lådorna är ordnade enligt stigande medianvärde, och dialektdragen längst till vänster i figuren används följaktligen minst av informanterna, medan dialektdragen längst till höger är de bäst bevarade i 2000-talsmaterialet.

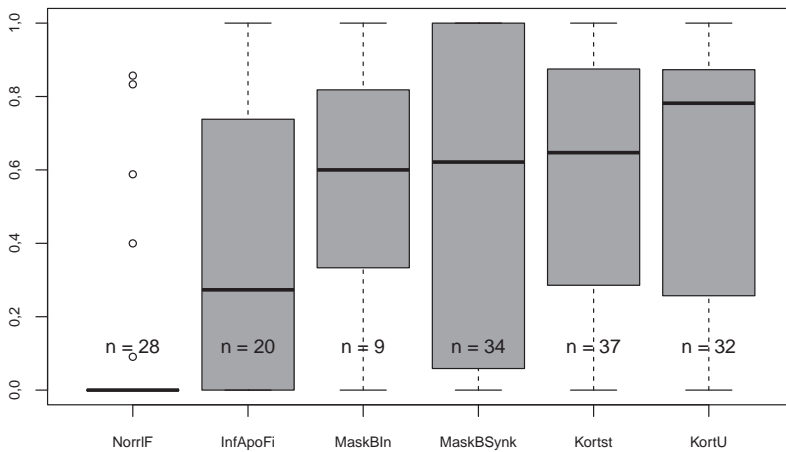
Dialektdrag som är på väg att försvinna ur de åboländska dialekterna är *ng*-ljud med hörbart *g*, bevarat *a* för standardspråkets *å*, *het* som neutralt pronomen samt *ei*-diftongen. Bara ett fåtal informanter har en hög andel dialektvarianter av dessa variabler. Drag som står sig väl är däremot apokope i preteritum av andra och tredje konjugationens verb, den traditionella maskulina slutartikeln samt infinitivmärket *ti*. Medelprocenten är högst för variabeln PretApo; alla utom en informant apokoperar preteritumändelsen i mer än 50 procent av fallen. För variabeln MaskB har andelen traditionella slutartiklar för starka maskuliner räknats. Variationen sker alltså mellan dialektvarianten *-i(n)* (Iniö och Houtskär) eller *-n* (övriga orter) och standardvarianten *-en* (se närmare avsnitt 3.1).

Man kan konstatera att de väl bevarade dialektdragen i figur I.17 hör till kategorin formlära och formord, medan de rent fonologiska dragen verkar ha en svag ställning. En motsvarande tendens har noterats för finlandssvenska stadsmål, som har tagit upp morfologiska särdrag ur de omkringliggande dialekterna, medan de fonologiska dragen i dialekterna är starkt markerade och saknas i stadsmålen (Ivars 2005, 1053).

Figur I.18 visar lådagram över de variabler vilkas dialektvariant har haft en lokal utbredning inom Åboland, dvs. dragen har funnits på

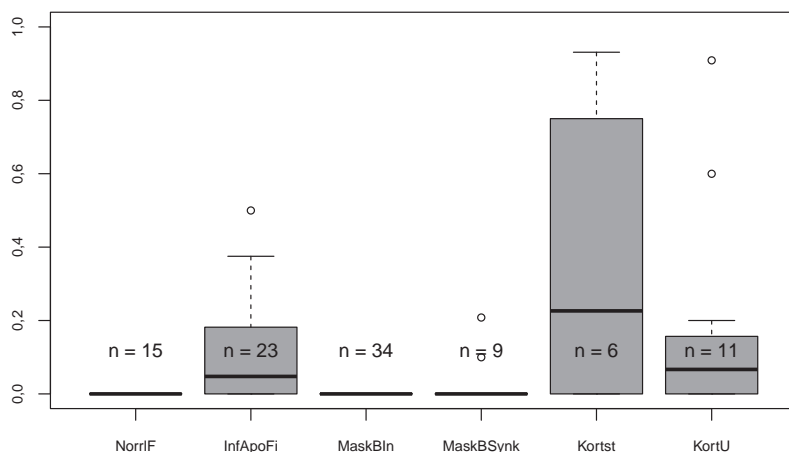


Figur I.17. Lådagram över variabler vilkas dialektvariant tidigare funnits i alla åboländska dialekter. Figuren visar andelen dialektvarianter för varje variabel uträknad per informant ($n = 43$).



Figur I.18. Lådagram över variabler vilkas dialektvariant har haft en lokal utbredning i Åboland. Värden enbart för de informanter som kommer från orter där dialektdraget tidigare har funnits. Antal informanter för varje variabel anges i grafen.

vissa orter men inte på alla. Lådorna är ritade så att bara värden för de informanter som kommer från en ort där dialektdraget har funnits tidigare enligt Zilliacus (1992) har beaktats. Antalet informanter varierar alltså från låda till låda i diagrammet. För variabeln NorrlF



Figur I.19. Lådagram över dialektdrag som har haft en lokal utbredning i Åboland. Värden enbart för de informanter som kommer från orter där dialektdraget *inte* har varit utbrett tidigare. Antal informanter för varje variabel anges i grafen.

t.ex. har informanterna från Nagu och Pargas, som inte har hört till förmjukningsområdet, inte räknats med och lådan bygger därmed på värden för 28 informanter. För variabeln Kortst har informanterna från Hitis och den yttre skärgården i Korpo uteslutits.

Variabeln MaskB har delats upp i två: MaskBIn som står för varianten *-i(n)* (som funnits i Iniö och Houtskär) och MaskBSynk som står för den synkoperade varianten *-n* (övriga orter). Lådan för variabeln MaskBIn bygger alltså på endast nio informanter: tre från Iniö och sex från Houtskär.

I jämförelse med figur I.17 kan man säga att många av dragen med lokal utbredning används i relativt stor utsträckning av informanterna. Endast ett drag, NorrlF, verkar vara på väg att försvinna. För de övriga dragen är variationen stor. Det finns både informanter som helt undviker dialektvarianten och sådana som använder den i 100 procent av fallen. Fyra av variablerna har ett medianvärde över 50 procent. Både den östliga och den västliga varianten av MaskB används i ungefär lika stor utsträckning av informanterna i respektive område.

Figur I.19 visar samma variabler som figur I.18, men bygger enbart på värden för informanter som kommer från orter där man *inte* skulle vänta sig att finna dialektdraget enligt Zilliacus (1992). Lådan för NorrlF byg-

ger alltså på de sju informanterna från Nagu och de åtta informanterna från Pargas. Figuren visar på så vis vilka dialektdrag som har spridits till områden där de tidigare inte har funnits. Figuren visar att det framför allt är kortstavigheten som har brett ut sig till informanterna från den yttre skärgården, som tidigare har haft genomförd stavelseförlängning.

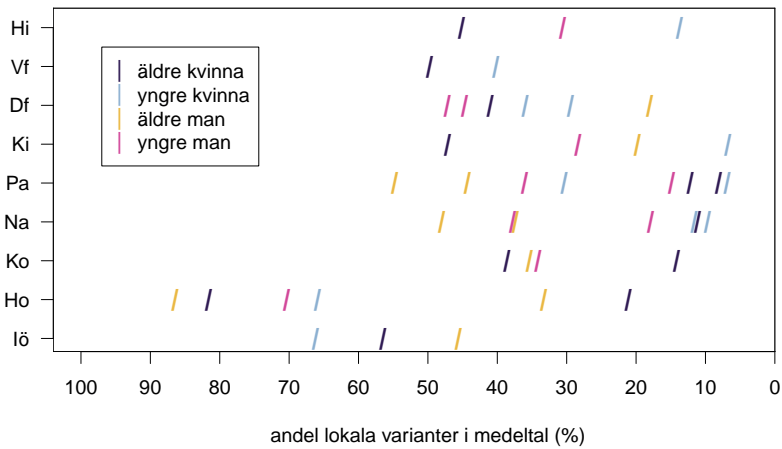
Ytterligare två dialektdrag som finns belagda även hos informanter som inte kommer från det ursprungliga utbredningsområdet för draget i fråga är InfApoFi och KortU. Men som redan tidigare har konstaterats finns det vissa svårigheter med att tillförlitligt transkribera det korta *u*-ljudet (2.1.5), och det är inte ovanligt med undantag från reglerna som gäller den villkorliga apokoperingen av infinitiver (3.4.2). För dessa två drag kan man alltså inte utifrån resultaten med säkerhet påstå att dragen skulle ha spridits inom Åboland.

4.2 Jämförelse av informanterna

Figur I.20 visar andelen lokala varianter i genomsnitt hos varje informant. Andelen är först uträknad per variabel, och den sammanlagda andelen per informant utgörs av medeltalet för de variabler vilkas dialektvariant traditionellt funnits på orten. Antalet variabler som medeltalet är uträknat på varierar mellan elva (t.ex. Nagu som saknat norrländsk förmjukning och *o*-uttal av kort *u*) och tretton.

I figuren skiljer sig Iniö och Houtskär tydligt från de övriga orterna. En stor del av informanterna på dessa två orter har en genomsnittlig andel dialektvarianter på över 50 procent. På de övriga orterna har nästan alla informanter en andel under 50 procent. Informanterna med de lägsta andelarna (kring 10 procent) kommer från Nagu, Kimito och Pargas. Färgkoderna i figuren visar att det är kvinnor på dessa orter som har de allra lägsta andelarna. För Nagu och Pargas ser man dessutom en tydlig åldersfördelning med lägre andelar dialektvarianter hos yngre än hos äldre talare. De informanter på dessa två orter som har de högsta andelarna dialektvarianter är äldre män.

För de flesta orters del kan man säga att variationen mellan informanterna är kontinuerlig, dvs. informanterna befinner sig på en någorlunda jämn skala med högre eller lägre andelar dialektvarianter. En tendens till två grupper av informanter kan man kanske se i Nagu och Pargas som har ett tomrum mellan 20 och 30 procent i figuren.



Figur I.20. Den genomsnittliga andelen lokala varianter på individnivå. Varje lodrätt streck i figuren representerar en informant.

Endast i Houtskär är fördelningen i två grupper riktigt tydlig. Informanterna från Hönsnäs har relativt låga andelar dialektvarianter medan informanterna från Norrbyarna har höga andelar.

4.3 Principalkomponentanalys

En principalkomponentanalys (PCA) utfördes för att undersöka hur de enskilda informanterna står i relation till de olika språkliga variablerna. PCA är en datareduktionsmetod som gör det möjligt att utforska huruvida en uppsättning variabler kan reduceras till ett mindre antal komponenter genom att variabler som uppvisar en hög korrelation kombineras med varandra. I detta fall är målet att se om det finns dialektdrag som uppvisar en liknande distribution bland informanterna. Samtidigt visar metoden om informanterna kan delas in i olika grupper vad gäller språkligt beteende eller om det är frågan om en kontinuerlig variation.

De 43 informanterna utgör objekt (datarader) i analysen och de fjorton dialektdragen utgör variabler. Variabeln MaskB har delats upp i MaskBIn ($-i(n) \sim -n/-en$) och MaskBSynk ($-n \sim -i(n)/-en$). På så vis ingår formvariationen i analysen, vilket ger en variabel mer än i tabell I.4. Variablerna anger andelen dialektvarianter för varje informant.

Analysen resulterar i en uppsättning laddningar och en uppsättning poäng. Laddningarna visar i hur stor utsträckning var och en av de språkliga variablerna korrelerar med de extraherade komponenterna, medan poängen ges till informanterna. Efter analysen kan varje informants språkliga beteende beskrivas genom poäng på ett färre antal komponenter i stället för genom de fjorton ursprungliga variablerna.

Två komponenter extraherades, vilka tillsammans förklarar 71,7 procent av variansen i materialet. Genom att använda bara två komponenter kan man alltså säga lika mycket om informanternas språkliga beteende som genom att använda 71,7 procent av den information som uttrycks med hjälp av fjorton variabler.

Figur I.21 visar en s.k. biplot (tvådimensionellt punktdiagram) av analysen. Signumet anger varje informants position i diagrammet. Positionerna bestäms av informantens poäng på de två komponenterna. Pilarna bygger på laddningarna och anger den riktning i diagrammet i vilken andelen dialektvarianter för varje variabel ökar. Informanter som befinner sig nära varandra i diagrammet har ett liknande språkligt beteende, medan informanter som befinner sig på långt avstånd från varandra är språkligt sett olika.

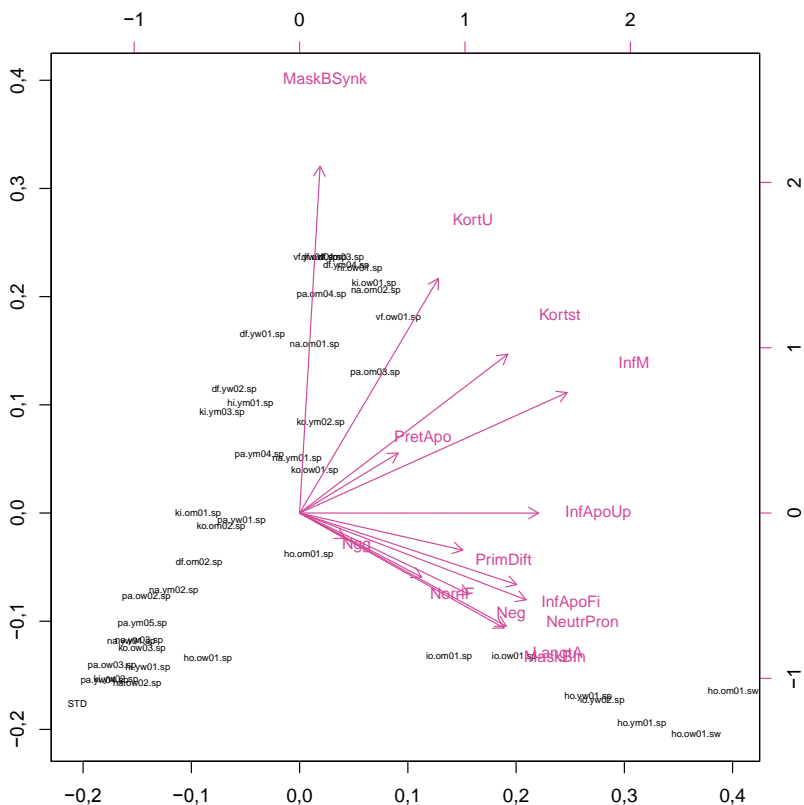
Utifrån figur I.21 kan man grovt dela in informanterna i tre grupper: en grupp nere i vänstra hörnet, en grupp nere i högra hörnet och en tredje grupp högre upp i figuren. Förutom de 43 informanterna har också en kod för standardsvenska (STD) lagts in i analysen. Denna kod finns längst nere till vänster i figuren. En informant som inte hade några belägg alls på dialektvarianter för de fjorton variablerna skulle hamna på detta ställe i diagrammet. Ingen informant saknar helt belägg på dialektvarianter, men en grupp informanter befinner sig mycket nära detta ställe i figuren. Dessa är de mest utjämnade informanterna som bara har få belägg på dialektdrag. I denna grupp befinner sig främst kvinnor (både äldre och yngre) från mellersta och östra Åboland (jfr figur I.20).

Längst upp i figuren hittar vi informanter från Kimitoön och Hitis och även äldre män från Nagu och Pargas. Det dialektdrag som ökar i användning ju högre upp i figuren en informant befinner sig är synkoperingen av starka maskuliner i bestämd form.

Nere till höger i figuren finns informanter från Iniö och Houtskär. En mängd olika pilar visar i denna riktning. De längsta pilarna har

variablerna MaskBIn, LangtA och NeutrPron. Dessa drag är alltså kännetecknande för informanterna från Iniö och Houtskär. Men även flera andra drag, om än med kortare pilar, ökar i denna riktning. Dialektvarianter av variabler som NorrIF, Neg och PrimDift används även på övriga håll i Åboland i olika utsträckning, men de är vanligast bland informanterna från Iniö och Houtskär.

De tre grupper av informanter som kan urskiljas i figuren är alltså en östlig grupp, en västlig grupp och en standardnära grupp. Flera



Figur I.21. Diagram över resultatet av principalkomponentanalysen. Informanterna kan grovt delas in i en östlig, en västlig och en standardnära grupp utifrån andelen dialektvarianter av fjorton språkliga variabler.

informerter placerar sig också mellan den östliga gruppen och den standardnära gruppen. Här kan man alltså inte tala om två strikt avskilda grupper, utan informanterna verkar finnas på en glidande skala mellan dialekt och standard. Den västliga gruppen är däremot ganska väl avskild från de två övriga grupperna, men även här befinner sig informanterna på en glidande skala där Ho.omo1.sw och Ho.owo1.sw är de mest dialektala informanterna, följda av de tre yngre informanterna. De två äldre informanterna från Iniö som under intervjun säger att de i inspelningssituationen inte talar lika dialektalt som de skulle göra i en naturlig talsituation med andra ortsbor befinner sig tydligt i den västliga informantgruppen, om än inte lika långt till höger som den yngre Iniö-informanten och informanterna från Norrbyarna i Houtskär. De två informanterna från Hönsnäs (Houtskär) återfinns i den standardnära gruppen, men ändå med dragning åt det västliga hållet.

Ett antal pilar visar i riktning från de standardnära informanterna mitt emellan den östliga och den västliga gruppen. Dessa dialektdrag är gemensamma för både de öståboländska och de väståboländska informanterna, men de standardnära informanterna använder dem endast i liten utsträckning eller inte alls. PretApo har den kortaste pilen och är det enda drag som används i stor utsträckning också av de standardnära informanterna. Variablerna KortU och InfApoUp har en hög andel dialektvarianter bland informanter både i öster och väster, men KortU har lite högre andelar i öster och InfApoUp har högre andelar i väster (jfr figur I.5 på s. 66 och figur I.14 på s. 117). Variablerna Kortst och InfM har ungefär lika höga värden bland de mest dialektala informanterna i östra och västra Åboland.

Analysen visar att det fortfarande finns en tydlig öst-västdistinktion i Åboland. Denna distinktion beror delvis på att vissa dialektdrag med lokal utbredning i Åboland är relativt väl bevarade (se figur I.18). Men den beror också på att ett antal drag som tidigare har varit utbredda i hela Åboland numera verkar vara livskraftiga bara längst i väster.

5

Slutdiskussion

Dialektutjämningsen i Åboland har i denna studie analyserats genom granskning av ett antal fonologiska och morfologiska dialektdrag i ett material insamlat i början av 2000-talet. I jämförelse med de traditionella dialekterna i Åboland som de beskrivs av Zilliacus (1992) har dialektutjämningsen varit stor. Dialektvarianter konkurrerar med standardsvenska varianter i fråga om alla undersökta variabler. Dialektutjämningsen har dock fortskridit över en lång period. Man bör hålla i minnet att Zilliacus samlade in dialektmaterial i Åboland främst under 1950-talet men hans mål var att beskriva ännu äldre stadier av dialekterna än de som talades vid insamlingstiden. Hans syfte var att beskriva dialekterna så som de talats i början av 1900-talet (se 1.2).

Trots att dialektutjämningsen varit stor finns det ett antal dialektdrag som ännu används relativt ofta av de åboländska informanterna. Dialektdrag som fortfarande verkar ha en ganska stark ställning i Åboland är:

- apokopering i preteritum av andra och tredje konjugationens verb (*tykt* 'tyckte')
- *o*-uttal av kort *u* (*hondra* 'hundra')
- kortstavighet (*kåma* 'komma')
- traditionell bestämd slutartikel i singular för maskulina substantiv (*bå:tin/bå:tn* 'båten')
- *ti* som infinitivmärke (*de e ba: ti tjö:r åmbo:rd* 'det är bara att köra ombord')
- apokopering i infinitiv (*hitt* 'hitta')

- *inga* som negation (*he va inga he:ft* 'det var inte hävt')
- supinumformer av typen *slaji* 'slagit'
- pronomenet *han* som objektsform (*så slo: ja han ra:kt på nesan* 'så slog jag honom rakt på näsan').

De ovannämnda dragen är relativt starkt representerade i det undersökta materialet. Men för alla drag utom apokopering i preteritum gäller att det också finns informanter som inte har några belägg alls på dialektvarianten utan bara använder standardvarianter. Variationen mellan informanterna är alltså stor.

Exempel på dialektdrag som förekommer i det undersökta materialet, men som hos de flesta informanter till stor del har ersatts av standardspråkliga former är:

- *ng*-ljud med hörbart *g*
- bevarat *a* i stället för *å* (*la:ngt* 'långt')
- *het* som neutralt pronomen (*he va ro:lit* 'det var roligt')
- primära diftonger (*veit* 'vet', *röyken* 'röken')
- sekundära diftonger (*boån* 'barn', *jåort* 'gjort')
- norrländsk förmjukning (*trafi:tjin* 'trafiken')
- oförmjukat *g*-, *k*- och *sk*- initialt före främre vokal (*kö:ra* 'köra', *gä:r* 'gör')
- suffixet *-og* för adjektiv (*to:kå* 'tokig')
- *tocken/tolin* i betydelsen 'sådan' (*tåkna i:svintrar* 'sådana isvintrar').

Den största språkliga skillnaden mellan äldre och yngre informanter finns i mellersta Åboland (Nagu, Pargas, Kimito) där flera av de yngre informanterna och av kvinnorna har en låg andel dialektvarianter av de undersökta dragen. Men skillnaden mellan de äldre informanterna (födda på 30- och 40-talen) och de yngre (födda på 70- och 80-talen) är för de flesta undersökta variabler överlag inte påfallande stor. Det är alltså inte fråga om någon stor *pågående* dialektutjämning. Snarare verkar det som att en stor del av dialektutjämningen skulle ha skett mellan den äldre generationen av informanter i det aktuella materialet och en ännu äldre generation av dialekttalare. Denna bild bekräftar av Karlssons (2006) fallstudie från Nagu. Karlssons äldsta informant är född 1903 (inspelad i slutet av 70-talet), medan den mellersta informan-

ten är identisk med Na.omo2.sp i föreliggande undersökning och den yngsta informanten är samma person som Na.ymo1.sp. Av Karlssons informanter talar den äldsta traditionell dialekt medan den mellersta informanten använder både dialektvarianter och standardvarianter. Den yngsta informantens språk är utjämnat och saknar i stort sett markerade dialektdrag.

Ett av de få drag som visar en tydlig förändring i skenbar tid (dvs. mellan de två åldersgrupperna) i materialet är *ng*-ljudet. Bland de äldre informanterna finns det belägg med hörbart *g*, men bland de yngre saknas detta uttal så gott som helt. *Ng*-ljudet med hörbart *g* verkar alltså vara på väg att försvinna med den generation talare som är födda på 30- och 40-talen.

I introduktionen (1.3) beskrevs de olika tendenser som dialektutjämning i europeiska språk vanligen följer och som hypotetiskt kan ha kännetecknat dialektutvecklingen även i Åboland. Med *vertikal dialektutjämning* avses utjämning i riktning mot standardspråket medan *horisontell dialektutjämning* innebär att olika dialekter inom ett geografiskt område influerar varandra så att de närmar sig språkligt. Vid *regionalisering* uppstår språknivåer mellan traditionell dialekt och standardspråk. Man kan också skilja mellan *urban* och *rural regionalisering* (Akselberg 2003). En regional dialekt kan, men behöver inte, ha status av varietet (Auer 2011, 492). För att kunna klassificeras som varietet behöver en språknivå i regel kännetecknas av en viss intern koherens (dvs. variationen inom en varietet är mindre än variationen mellan varieteter) och av tydligt formulerade pragmatiska användningskriterier som bl.a. reglerar i vilken typ av talsituationer varietet används (Thelander 1979; Lenz 2003, 388–395). Regional dialekt kan däremot enligt Auer (2011, 492) också innebära en glidande skala mellan ett mer dialektalt och ett mer standardnära sätt att tala. Kristiansen (2003) problematiserar definitionen av regional dialekt som endast en språknivå mellan lokal dialekt och standardspråk och argumenterar för att existensen av ett regionalt språk ska definieras på normnivån och inte på språkbruksnivån. Enligt Kristiansens synsätt kan man motivera bruket av termen regionalt språk (eller regional dialekt) bara i den mån språkbrukarna själva förknippar det regionala språkbruket med en positiv regional identitet och uppvärderar det regionala språket i jämförelse med standardspråket.

I Åboland verkar vertikal dialektutjämning spela en betydligt större roll än horisontell dialektutjämning. Alla undersökta variabler visar på en språkförändring i riktning mot standardspråket. En liten spridning av dialektdrag som funnits i många, men inte alla, sockenmål i området har noterats – detta verkar gälla kortstavighet och möjligen också *o*-uttal av kort *u* – men det faktum att flera språkdrag som tidigare funnits bara i delar av Åboland står sig väl jämfört med dialektdrag som varit gemensamma för hela regionen (se 4.1) visar ändå att betydelsen av horisontell dialektutjämning inte är särskilt stor. Resultatet av principalkomponentanalysen i avsnitt 4.3 visar att det fortfarande finns en tydlig skillnad mellan språket i östra och västra Åboland. En rural regionaliseringsprocess som skulle omfatta hela Åboland och göra dialektområdet mer homogent verkar alltså inte vara förhånden. Genom att utjämningen verkar ha gått längst på de orter som ligger närmast Åbo kan man snarare tänka sig en urban regionaliseringsprocess med utgångspunkt i Åbo. Eftersom språket i Åbo fonologiskt och morfologiskt motsvarar det finlandssvenska riksspråket (se t.ex. Ivars 1996, 34–37 som ger en beskrivning av det finlandssvenska riksspråket som referensspråk) går det ändå inte att avgöra om standardspråket sprids från just Åbo eller t.ex. via finlandssvenska massmedier.

Dialektutjämning kan hypotetiskt också ske under inflytande av skriftspråket eller sverigesvenska. Mot sverigesvenskt inflytande talar att typiska finlandssvenska talspråksdrag används av de åboländska informanterna. Exempel på sådana drag är konsekvent utelämnande av vissa ändelser (t.ex. *-t* vid neutrala substantiv och *-de* i preteritum av första konjugationens verb) samt former som *si*: 'se' och *tri*: 'tre'. Vad gäller dessa drag används den finlandssvenska varianten av alla informanter i så gott som hundra procent av fallen. Bilden av att det är det finlandssvenska standardspråket och inte det sverigesvenska som utgör normen stöds ytterligare av att hörbart *d* i initialt *dj*- återinförs på vissa håll i Åboland där dialekterna tidigare haft bortfall av *d*. Skriftspråkligt inflytande kan anas i enstaka uttal av *inte* med hörbart *e* och några få fall av tydligt uttalade ändelser. Dessa belägg finns främst hos de äldre informanterna som kanske uppfattar inspelningssituationen som mer formell än de yngre.

Bäst bevarade är dialekterna i Iniö och i Norrbyarna i Houtskär. På dessa orter har både de äldre och de yngre informanterna en hög

andel dialektvarianter av de flesta av de undersökta dragen. Iniö och Houtskär ligger längst i väster i Åboland och nås enbart med flera färjor. Det perifera läget har antagligen bidragit till att dialekterna har bevarats särskilt väl i dessa områden. Flera dialektdrag som tidigare varit allmänna i hela Åboland är i början av 2000-talet vanliga bara i Iniö och i Houtskärs Norrbyar. Detta gäller t.ex. *het* som neutralt pronomen, *ei*-diftongen, bevarat gammalt långt *a*-ljud (som i standardsvenska gått över till *å*), norrländsk förmjukning samt förlängt *a* och *ä* före konsonantkombinationerna *ld*, *mb*, *nd* och *ng*. Den som gjorde intervjuerna i Iniö och Houtskär talar inte själv den lokala dialekten, och vissa av informanterna nämner att detta påverkar deras sätt att tala under intervjun. Trots denna omständighet, som kan tänkas leda till att informanterna talar mer standardnära, är informanterna från Iniö och Norrbyarna de som använder flest dialektdrag i den åboländska delen av *Spara talet*-korpusen.

I de övriga delarna av Åboland kan vissa av de dialektdrag som är infrekventa i materialet ha fått nya funktioner. Drag som är infrekventa i 2000-talsmaterialet kan användas som stilmedel eller för att visa ”var man har sina rötter” (SLS 1746: 466). Exempel på sådana drag är *het* som neutralt pronomen, hårt *g*-, *k*- och *sk*- initialt före främre vokal (i Nagu och Pargas), *åo*-diftongen och *og*-suffixet för adjektiv. Dessa dialektdrag är inte helt försvunna ur språket trots att de till stor del har trängts undan av de standardspråkliga motsvarigheterna. De finns kvar som ett skikt i dialekttalarnas medvetenhet som man kan ta till för att uttrycka lokal identitet eller ge extra färg åt språket (jfr Sandström 2010, 368ff.). I en annan typ av talsituation än intervjun kan det tänkas att denna typ av drag skulle uppvisa en högre frekvens. Inspelningssituationen kan ha känts formell vilket kan leda till ett mer standardnära språkbruk. Trots det säger det ringa antalet dialektvarianter i det inspelade materialet en hel del om dialektsituationen i Åboland. En jämförelse kan göras med Österbotten där precis samma typ av intervjuer spelades in i *Spara talet*-projektet. För de österbottniska informanterna var det naturligt att tala den lokala dialekten också i inspelningssituationen (se t.ex. Ivars 2012).

Regionalisering har skett i Åboland i bemärkelsen att det uppstått språknivåer som ligger mellan traditionell dialekt och standardspråk. Men utvecklingen har varit olika på olika håll i Åboland. I de väst-

ligaste delarna har språksituationen fortfarande karaktär av *diglossi* (typ B-repertoar enligt Auers, 2011, typologi). Detta framkommer av informanternas egna uttalanden om språkbruket (se 1.5), men framgår också av principalkomponentanalysen (4.3) där informanterna från Iniö och Houtskärs Norrbyar bildar en tydligt avgränsad grupp. Övriga delar av Åboland kännetecknas snarare av en *diaglossisk* repertoire (typ C enligt Auers typologi). I en diaglossisk repertoire bildar dialekter och standardspråk inte strukturellt avskilda system, utan det finns en glidande språklig skala mellan dialekt och standardspråk som talarna kan utnyttja i olika grad i olika situationer. I principalkomponentanalysen bildar informanterna från östra och mellersta Åboland ett sådant kontinuum (se figur I.21, s. 134). En viss polarisering finns trots det med flera informanter mycket nära standard och i den dialektala ändan av kontinuumet och färre informanter i mitten av kontinuumet.

Frågan om det har bildats en regional åboländska som har status av en språklig varietet är svårare att besvara. Det material som analyserats här fyller inte kriteriet intern koherens, eftersom man kan se en tydlig indelning i öst och väst. För en fullständig analys av uppkomsten av en regional varietet skulle dock informanternas språkbruk behöva observeras i olika talsituationer, och även språkbrukarnas attityd till och identifikation med den egna regionala dialekten i jämförelse med standardspråket skulle behöva analyseras. Med regionalisering avses ofta också minskad differentiering på den lokala nivån till förmån för differentiering på den regionala nivån (Hinskens et al. 2005, 12; Kerswill & Trudgill 2005, 202). Differentieringen på den lokala nivån har minskat i Åboland, eftersom man med hjälp av de variabler som har analyserats här inte kan skilja mellan enskilda sockenmål utan endast kan göra en grov indelning i öst och väst. För en analys av differentieringen på den regionala nivån, dvs. om det finns en åboländsk regional varietet som är tydligt avgränsad från exempelvis en nyländsk, skulle talspråksmaterial från ett större område än endast Åboland behövas.

I den mån man för de mest utjämnade informanternas del kan tala om regional åboländska, handlar det om en varietet som förutom allmänfinlandssvenska talspraksdrag omfattar apokopering i preteritum, men annars inte har särskilt många avvikelser från standardspråket i fråga om fonologi och morfologi. Apokopering i preteritum av andra och tredje konjugationens verb är inte någon specifikt åboländsk företeelse,

utan finns i alla finlandssvenska dialekter och förekommer även i städerna. Eventuella specifikt åboländska regionala drag kanske ska sökas på andra språkliga nivåer än inom fonologi och morfologi. Fonetiska särdrag bevaras ofta bättre än fonologiska och morfologiska. Aho (2010, 17) menar att en stigande ton i den betonade stavelsen i fokuserade ord är ett typiskt åboländskt drag ”som ger dialekten sitt särdrag även hos sådana informanter som annars talar nästan standardspråk”. En annan nivå i språket där regionala drag kanske kan hittas är fonetisk kvantitet. Schaeffler (2005, 88) har visat att de finlandssvenska dialekterna kan indelas i åtminstone tre grupper vad gäller vokal- och konsonantlängd i stavelsetyperna V:C och VC:. De åboländska dialekterna hör till ett sydfinländskt kvantitetsområde som förutom åboländska dialekter även omfattar västnyländska och östnyländska. Även inom syntax kan regionala drag tänkas förekomma.

Undersökningen har gett en bild av förändringstendenser och geografisk variation i åboländskan i början av 2000-talet. Allmänna tendenser har kunnat påvisas, som att närheten till Åbo spelar en roll vid dialektutjämningen och att en tydlig öst–väst-distinktion upprätthålls även när dialekterna utjämnas. Samtidigt har bara ett begränsat antal frekventa språkdrag analyserats, och från vissa orter ingår ytterst få informanter i analysen. Mycket kvarstår alltså att undersöka. De talrika språkproven i texten ger förutom exempel på de undersökta språkdragen även exempel på drag som inte har undersökts systematiskt här, men som bidrar till att ge språket lokal färg (former som *li:tn* ’liten’, *lå:do* ’låda’, *å:v* ’av’ osv.). Förhoppningsvis kommer framtida undersökningar som behandlar sådana språkdrag och språknivåer som inte har analyserats här att bidra med ytterligare information om den dialektala variationen i Åboland. Vad gäller uppkomsten av regionala språkvarieteter skulle en empirisk undersökning som omfattar hela södra Finland eller t.o.m. hela det svenska språkområdet i Finland betydligt bättre kunna belysa de mekanismer som ligger bakom regionaliseringen av språket. Enligt översikten hos Ivars (2003) bevarar de öståboländska dialekterna vid utjämning framför allt drag som är gemensamma med västnyländska dialekter. De historiska landskapsgränserna verkar inte sammanfalla med de skarpaste dialektgränserna i dagens läge, och en undersökning som omfattar ett större område än ett enda landskap kunde kanske ge fler svar på frågor om hur den dialektala identiteten

formas på 2000-talet. Att landskapsgränserna trots allt fortfarande upplevs spela roll för den lokala identiteten kan avslutningsvis få exemplifieras med följande utdrag:

- (5.1) *no e man ju en u:tbö:ling dä:r i ny:land* (Df.ywo1.sp)
nog är man ju en utböling där i Nyland

Del II

FRÅN *DI* OCH *DEM* TILL *DÅM* – VARIATION OCH FÖRÄNDRING I FYRA STÄDER I SVENSKFINLAND

1 Introduktion

I talspråket i Sverige används nästan uteslutande uttalet /dåm/ för pronomenen *de* och *dem* och artikeln *de*.¹ Skriftens skillnad mellan *de* som subjektform och *dem* som objektform saknas alltså i talspråket. Formen /dåm/ är en ursprunglig objektform som först blivit vanlig också i subjektställning i norra och mellersta Sverige och sedan spridit sig allt längre söderut och ersatt den äldre formen /di/ (Grass 1970, 1–3). Enligt Bergman (1942, 151–153) blev /dåm/ på 1930-talet allt vanligare även i södra Sverige.

Att skriftspråket upprätthåller en grammatisk distinktion som många svenskar saknar språkkänsla för kan ställa till med problem när man ska skriva korrekt. Det är inte bara skolelever som har svårt att lära sig att använda *de* och *dem* korrekt i skrift, utan även vuxna skribenter känner sig osäkra. Detta framgår av att råd om korrekt användning av *de* respektive *dem* regelbundet ges i språkspalter till allmänheten.²

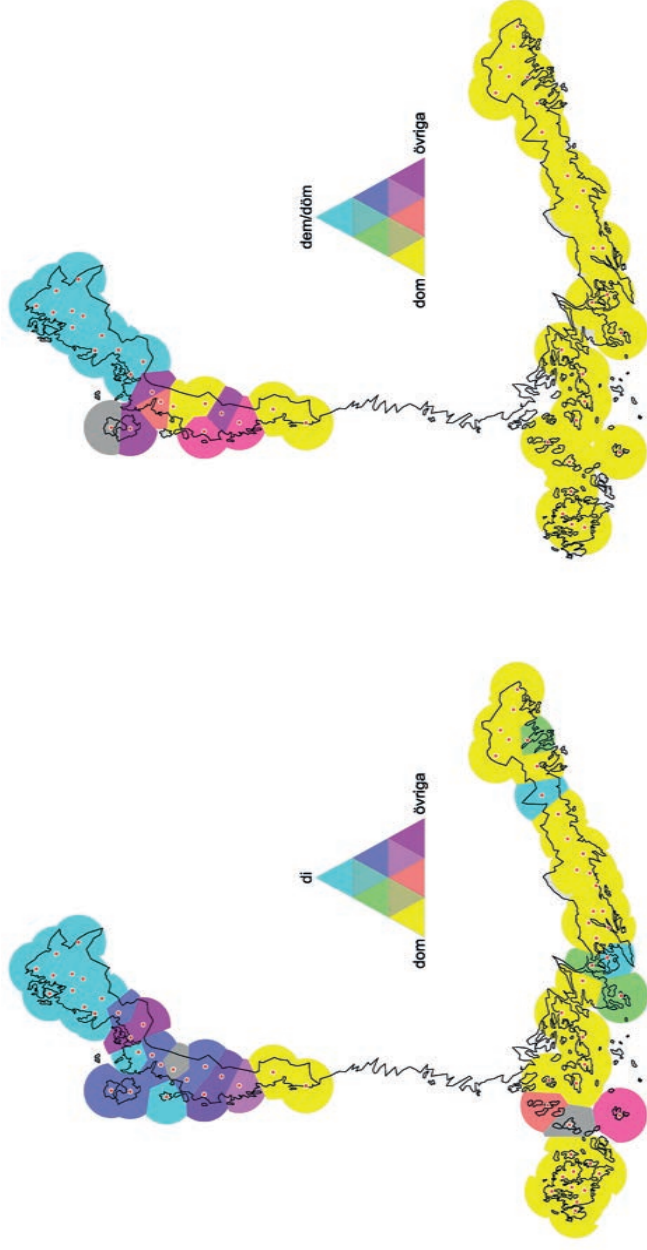
Trots att många har svårt att skilja mellan *de* och *dem* har språkvårdarna fortsatt rekommendera att den grammatiska distinktionen upprätthålls i skrift, i stället för att man skulle övergå till enbart *dom* enligt mönster från talspråket. Motiveringen har varit att en *dom*-reform skulle vara alltför radikal eftersom det rör sig om ett frekvent ord och texter skrivna före reformen snabbt skulle få en stämpel som ålderdomliga. En *dom*-reform skulle även kunna göra det svårt för kommande generationer att läsa äldre texter (*Språkriktighetsboken*, 207). Ett annat ofta återkommande argument mot att övergå till *dom* i skrift har varit att en del svenskspråkiga fortfarande upprätthåller distinktionen mellan subjekt och objekt också i talspråket. Som exempel nämns i dessa fall ofta finlandssvenskar och (äldre) skåningar.³

1. I denna text används kursiv för skriftspråkets former medan uttalsformer står mellan snedstreck.
2. Se t.ex. Catharina Grünbaum i DN 1/4 2006, "Får man skriva 'dom' i DN?", Siv Strömquist i SvD 21/3 2009, "Rätt val kan ge irritation", Olle Josephson i SvD 29/11 2010, "Det finns en utväg ur dem-problemet", Ylva Byrman i SvD:s Språkblogg 2/6 2011, "De eller dem?", samt Språkrådets frågelåda (<http://www.sprakradet.se/frågelådan>) sökord *dem*.
3. T.ex. i Språkrådets frågelåda, fråga: "Varför skriver man inte dom i stället för de och dem? Det uttalas ju så."

På 1960- och 1970-talen ökade användningen av *dom* något i skrift, men detta bruk slog aldrig igenom utan användningen av *dom* i skrift har snarare minskat under senare tid (*Språkriktighetsboken*, 206; Ewerth 2008). Man tycks alltså hellre acceptera ett felanvänt *de* eller *dem* då och då än en konsekvent övergång till *dom*. I SvD:s Språkspalt 29/11 2010 konstaterar dock Olle Josephson att ”vi vill befria svensklärare och skribenter från de–dem-grubblrier”. I stället för en övergång till *dom* i skriftspråk föreslår Josephson en mellanväg med *de* som enda skriftspråksform. Detta skulle enligt Josephson innebära en mindre förändring än en övergång till *dom* eftersom *de* redan finns i skriftspråket och *de* förekommer betydligt oftare än *dem*. Efter att Ylva Byrman hänvisat till Josephsons språkspalt i sin Språkblogg i juni 2011 utbröt en webbdiskussion där många förfasade sig över rekommendationen att använda *de* i alla lägen. Diskussionen fortsatte i *Språktidningen* i form av ett debattinlägg för en övergång till *dom* (Ola Karlberg, ”Dags att skrota *de* och *dem*!”, nr 2/2012) samt några läsarbrev. I en replik till Karlberg i *Språktidningen* 3/2012 är Josephson försiktigare och menar att en nackdel med ett enhets-*de* trots allt kunde vara att ordet *det* uttalas just *de*, vilket skulle kunna leda till ytterligare sammanblandning. I repliken öppnar Josephson försiktigt för en *dom*-reform.

I *de/dem/dom*-debatterna nämns alltså ibland att finlandssvenskar skiljer mellan subjekt och objekt även i talspråk. Men detta gäller inte alla svenskspråkiga i Finland. I dialekterna i södra Finland (dvs. Åland, Åboland och Nyland) är /däm/ den vanligaste formen både som subjekt och objekt, medan /di/ som subjektsform förekommer i österbottniska dialekter (Ahlbäck 1956, 41; FO). Kartorna i figur II.1 visar distributionen av olika former av pronomenet i de traditionella dialekterna enligt FO.⁴ I stora delar av språkområdet har man inte gjort någon skillnad mellan subjektsform och objektsform (området som är gult på bägge kartorna). Tydligast är distinktionen i norra Österbotten (cyanblått), där /di/ används som subjektsform och /dem/ eller /döm/ som objektsform. Förutom i norra Österbotten har formerna /di/ som subjekt och /dem/ som objekt också varit kännetecknande för högspråket i städerna Åbo och Helsingfors. Dessa former kom

4. För tillverkning av kartorna användes Gabmap (Nerbonne et al. 2011) och RuG/Lo4 (Kleiweg 2013).



Figur II.1. Distributionen av olika former av pronomenet *de* i de svenska dialekterna i Finland. Kartorna bygger på belägg i FO (*den*, plur.). Beläggorterna, som är något färre för objektsformen än för subjektsformen, är markerade med röda punkter. Gult står för /däm/ (eller /täm/) medan cyanblått står för /di/ (eller /ti/) respektive /dem/ (eller /döm/, inkl. t-former). Övriga former är markerade med magenta. Övriga former finns främst i Österbotten, där /dem/ och /döm/ används som subjektsform i Oravais, Vörå och Maxmo, och former som /täij/, /tär/, /där/ förekommer som både subjekts- och objektsform i mellersta och södra Österbotten. Även från östra Åland finns belägg på /täij/. Orter med belägg på flera olika former har färgnyanser mellan de tre grundfärgerna. Grön färg för vissa orter i södra Finland innebär t.ex. att det finns belägg på både /di/ och /dom/, medan olika nyanser av violett betyder att belägg finns på både /di/ och övriga former.

antagligen till Åbo med inflyttare från Mälardalen under 1600- och 1700-talen och till Helsingfors från Åbo i och med att Helsingfors blev huvudstad 1812 (Nyholm 1976, 32).

Ahlbäck förutspådde 1956 att användningen av /dåm/ antagligen skulle komma att öka i Finland eftersom formen har stöd både i sverigesvenskan och i dialekterna på landsbygden som omger Helsingfors och Åbo. I en sociolingvistisk analys av Helsingforssvenskan kunde Nyholm (1976, 36; 1984, 41) konstatera att /di/ fortfarande var prestigeformen på 1970-talet medan /dåm/ framför allt användes av talare i lägre socialgrupper, där många var inflyttade från den nyländska landsbygden. Informanterna frågades också om åsikter om de båda varianterna: /di/ uppfattades allmänt som den ”goda varianten” medan /dåm/ av informanterna uppfattades som dialektalt och sämre (Nyholm 1984, 42). Nyholms material visar dock på en stigande andel /dåm/ med sjunkande ålder, vilket stödjer Ahlbäcks antagande. Av de yngsta informanterna i Nyholms material (skolelever i åldern 15–17 år) använde 29 procent i den högsta socialgruppen och 72 procent i den lägsta socialgruppen övervägande /dåm/ (Nyholm 1976, 36).

Eftersom /dåm/ ursprungligen är en objektsform och pronomenet mångdubbelt oftare förekommer som subjekt än som objekt, kan det kanske tyckas oväntat att objektsformen ersatt subjektsformen. Övergången till enbart /dåm/ i talspråket verkar ändå följa vissa mönster. I en studie av övergången från /di/ till /dåm/ i Eskilstuna visar Grass (1970) att faktorerna betoning och grammatisk kategori spelar roll vid övergången. Formen /dåm/ är vanligare i betonad ställning än i obetonad, och /dåm/ är vanligare som personligt och demonstrativt pronomen än som bestämd artikel och i konstruktionen *de här/där* (Grass 1970, 19–20). Vidare är /dåm/ också vanligare i omvänd ordföljd än i rak ordföljd (Grass 1970, 40). Formen /dåm/ gynnas alltså när pronomenet är betonat och står självständigt, medan /di/ gynnas när pronomenet står obetonat och förenat.

Eftersom /dåm/ är längre är det lättare att framhäva eller ge prominens än det kortare /di/, som dessutom lätt kan sammanblandas med *vi*, *ni* och *det*. Ett behov av tydlighet verkar alltså ha drivit språkförändringen, och eftersom pronomenet förekommer betydligt oftare självständigt än som bestämd artikel eller i *de här/där* har /dåm/ till slut spridits till alla funktioner (Grass 1970, 42). Variationen efter gram-

matisk funktion verkar inte finnas enbart i talspråk utan återfinns ibland även i skrift (Ewerth 2008).

I Nyholms (1976, 32–33, 1984, 39) analys av variationen /di/ ~ /dåm/ i Helsingforsvenskan på 1970-talet visade sig variabeln vara polariserad så att de flesta informanter ganska konsekvent använde antingen /di/ eller /dåm/. Variationen var alltså snarare social än lingvistisk, och eftersom endast få informanter växlade mellan de båda formerna var det svårt att dra slutsatser om eventuella lingvistiska faktorerens betydelse för variationen. Precis som Grass fann Nyholm ändå att /dåm/ var vanligare i omvänd ordföljd än i rak. Enligt Nyholm (1984, 44) kan användningen av /dåm/ i omvänd ordföljd stödjas av att pronomenet står i den position där objektet brukar stå.

Syftet med denna studie är att undersöka variationen mellan /di/ och /dem/ å ena sidan och /dåm/ å andra sidan i de fyra så kallade regionhuvudstäderna i Svenskfinland: Helsingfors, Åbo, Mariehamn och Vasa. En frågeställning för undersökningen är om distinktionen mellan subjektform och objektform upprätthålls eller om /dåm/ blivit den vanligaste formen även i svenskan i Finland. En ytterligare frågeställning är om förändringen i Finland följer liknande språkliga mönster som i Sverige. Materialet för analysen omfattar en äldre och en yngre generation av informanter, vilket möjliggör analys av språkförändring i skenbar tid. Eftersom fyra städer ingår i analysen kan städerna dessutom jämföras med varandra vad gäller de olika varianternas frekvens.

2 Material

Informanterna spelades in i *Spara talet*-projektet. Fördelningen efter stad och åldersgrupp ses i tabell 1 på s. 15 i inledningen till boken. Därtill kommer en yngre man från Esbo som ingår i materialet eftersom han intervjuades tillsammans med en av de yngre männen från Helsingfors. Esbobon har fått ingå i Helsingforsgruppen i analysen och det totala antalet informanter är således 48. I denna delstudie studeras även intervjuarnas språkbruk. Antalet intervjuare i stadsmaterialet är elva, av vilka samtliga är kvinnor och alla utom en hör till den yngre åldersgruppen.

I Nyholms studier av Helsingforsvenskan (Nyholm 1976 och 1984) ingick ett stort antal informanter ur både lägre och högre socialgrupper, såväl inflyttade till som födda och uppvuxna i Helsingfors. En

sådan bredd finns inte i det aktuella materialet. I *Spara talet*-projektet koncentrerade man sig på informanter som var födda och uppvuxna på den ort de representerar i materialsamlingen och även hade minst en förälder som också var det (Ivars & Södergård 2007, 204). Därmed faller gruppen inflyttade bort. Vid insamlingen eftersträvades en jämn spridning med avseende på informanternas utbildningsnivå. I praktiken kom dock informanter med hög utbildningsnivå att vara överrepresenterade i materialet från städerna. Detta sammanhänger delvis med att de flesta intervjuare var universitetsstuderande och rekryterade informanter i sin egen bekantskapskrets, men även med det faktum att den svenska bosättningen i Helsingfors är koncentrerad till de södra och västra stadsdelarna där utbildningsnivån är hög (Höckerstedt 2010, 9–19; *Helsingin seudun aluesarjat*, Väestön koulutus rakenne).⁵

Vid insamlingen av *Spara talet*-materialet gjordes noteringar om informanternas egen utbildningsnivå, men inte om föräldrarnas. Flera av de unga informanterna går fortfarande på gymnasiet eller har nyligen inlett högskolestudier, vilket gör att det totala antalet utbildningsår inte kan räknas ut för de unga. Men det faktum att nästan alla unga har tagit studentexamen eller siktar mot studentexamen kategoriserar automatiskt de unga informanterna som högutbildade. Bland de äldre informanterna finns det en något större spridning vad gäller utbildningsnivå, men eftersom informanterna är så få till antalet har ingen indelning enligt socialgrupp gjorts utan analysen koncentreras till variablerna stad och ålder.

Excerpteringen skedde genom sökning på orden *de* och *dem* i de ortografiska transkriptionerna. Den ljudenliga transkriptionen kontrollerades och ändrades i vissa fall. Det kan t.ex. vara mycket svårt att höra om en informant säger /dem/ eller /döm/. Vid osäkerhet mellan dessa två varianter valdes uttalet /dem/ medan /döm/ bara valdes när vokalen entydigt kunde identifieras som /ö/.

5. Att andelen svenskspråkiga är låg i de lägre samhällsklasserna i Helsingfors hänger också samman med att språkbytet till finska varit omfattande just i arbetarklassen, där tvåspråkigheten är mycket utbredd och arbetslivet vanligen är finskdominerat (Tandefelt 1986, 94 och 1996, 42–43). För Nyholms (1984) studie var de lågutbildade informanterna handplockade ur mantalslängderna, vilket gjorde att ett tillräckligt stort antal lågutbildade informanter kunde hittas. Något sådant förfarande tillämpades inte i *Spara talet*-projektet.

Antalet belägg per informant är i medeltal 31,6 och medianvärdet är 27. Den informant som har minst antal belägg har endast 10 medan den som har flest har 82.

Av de totalt 48 informanterna i materialet använder 12 genomgående varianten /dåm/, medan 13 informanter helt saknar belägg på /dåm/ och i stället använder /di/ och /dem/ eller /döm/ och därmed skiljer mellan subjektsform och objektsform. Sammanlagt 23 informanter använder både /dåm/ och (någon av) de andra formerna. Dessa informanter, som växlar mellan /dåm/ å ena sidan och /di/, /dem/ eller /döm/ å andra sidan, är särskilt intressanta för studiet av övergången till enhetsformen /dåm/.

3 Grammatisk indelning

Vid excerperingen delades beläggen in enligt grammatisk kategori. Därtill noterades ordföljd samt huruvida pronomenet var betonat eller inte. Vid den grammatiska indelningen följdes SAG. En uppdelning gjordes i följande grammatiska kategorier:

- *personligt pronomen*, kan fungera som subjekt (*de* i skriftspråket) eller objekt (*dem* i skrift); ex. **de** såg **dem**
- *bestämd artikel*, förekommer i ”definit nominalfras med deskriptivt framförställt attribut om nominalfrasen saknar annat definit attribut” (SAG 2, s. 301); ex. **de** röda byggnaderna
- *de här/där*, kan fungera antingen nominalt (ex. *de var bra de där*) eller som definit attribut i en nominalfras (ex. *man lärde sig ganska snabbt att göra de här anteckningarna*)
- *demonstrativt pronomen*, normalt starkare betonat än den bestämda artikeln, kan (men behöver inte) vara skilt från det substantiviska huvudordet av ett adjektiv; ex. *svenskspråkiga pojkar var helt integrerade delar av de pojkgängen*
- *determinativ användning*, korrelerat till bisats; ex. *det har varit intressant nu under de månader som jag har jobbat där.*

Pronomenet står självständigt när det används som personligt pronomen. I övriga funktioner står det förenat.⁶ När pronomenet efterföljs av en *som*-sats vacklar bruket mellan subjekts- och objektsform såväl i tal som i skrift (*Språkriktighetsboken*, 209–210).⁷ Eftersom syftet med undersökningen var att analysera om skillnaden mellan subjektsform och objektsform upprätthålls var det inte meningsfullt att inkludera belägg i en kontext där bruket mellan de två formerna vacklar, och belägg i konstruktionen *de/dem som* utelämnades följaktligen.

Antalet belägg i de olika grammatiska kategorierna är följande: *personligt pronomen som subjekt* 900 (59,3 %), *personligt pronomen som objekt* 240 (15,8 %), *bestämd artikel* 65 (4,3 %), *de här/där* 282 (18,6 %), *demonstrativt pronomen* 18 (1,2 %), *determinativt* 13 (0,9 %). Vanligast är alltså det personliga pronomenet. På demonstrativt pronomen och determinativ användning finns bara ströbelägg i materialet.

Tabell II.1 visar fördelningen på de olika grammatiska kategorierna efter ort. Anmärkningsvärd är den mycket låga andelen av *de här/där* i Mariehamn jämfört med de tre övriga städerna. Av de sex informanterna från Mariehamn saknar fyra stycken helt belägg på *de här/där*. I Helsingfors, Åbo och Vasa har samtliga informanter minst ett belägg på *de här/där*, och medelantalet belägg på *de här/där* per informant ligger på 6,6 för dessa tre städer.

Enligt SAG (2, s. 314) tillhör pronomenen *den här* och *den där* ”i första hand talspråk och ledigt skriftspråk”. I finlandssvenskt talspråk används dessa pronomen mer frekvent än i sverigesvenskan (Bergroth 1928; Saari 2000; Wide 2011b). Förutom som demonstrativa pronomen används *den här* och *den där* som planeringspartiklar⁸ i svenskan i Finland (Wide 2011a), men även om man bortser från funktionen som planeringspartikel uppvisar både *den här* och *den där* högre frekvenser i svenskan i Finland än i Sverige (Wide 2011b). Den höga frekvensen

6. Självständigt *de* kan även tolkas som demonstrativt pronomen när det används betonat (SAG 2, s. 271–272). Här har dock i enlighet med SAG (2, s. 306) alla självständiga *de* tolkats som personligt pronomen, och huruvida belägget är betonat eller inte har noterats separat.
7. Ex. *en utav de/dem som satt i stadsstyrelsen*.
8. Ex. *de e nu int rikit min grej (.) men at **de hä:r** nu gå:r man ju allti ibland på på (.) på nå: sänhäna ka:ssajatkån* det är nog inte riktigt min grej (.) men att **det här** nog går man ju alltid ibland på på (.) på några sådana-här Casajatkon (= efterfester på Casa) (Hf.ymo4).

Tabell II.1. Fördelningen på olika grammatiska kategorier efter stad. Andelen anges som procent av det totala antalet belägg *de/dem* per stad.

STAD	PERS.PRON. (SUBJ.)	PERS.PRON. (OBJ.)	BEST.ART.	DE HÄR/ DÄR	DEM.PRON.	DETERM.
Helsingfors	56,2	15,8	2,7	23,3	1,3	0,7
Mariehamn	69,5	16,9	6,5	3,9	1,3	1,9
Vasa	67,0	11,2	3,1	17,5	0,5	0,7
Åbo	51,5	20,2	6,8	19,0	1,8	0,8

av demonstrativa pronomen i finlandssvenskan anses bero på kontakt med finskan (Saari 2000; Sundman 2000). Partikelfunktionen är sällsynt på Åland enligt Wide (2011a).

Som planeringspartikel används neutrumformen av pronomenen. Eftersom föreliggande material endast består av pluralformer, fungerar *de här/där* här uteslutande som demonstrativt pronomen. En mer exakt bild av frekvensen av *de här/där* ges i tabell II.2 som visar de relativa andelarna *de här* och *de där* av det totala antalet ord i materialet per stad. Medan de tre övriga städerna har andelar mellan 0,2 och 0,3 procent är andelen i Mariehamn endast 0,05 procent. Mariehamn skiljer sig alltså klart från de övriga städerna. Detta gäller såväl *de här* som *de där*. Man kan fråga sig om detta har med närheten till Sverige att göra och om siffrorna från Mariehamn påminner mera om motsvarande siffror från Sverige. En sökning i materialet från Sverige i *Nordic Dialect Corpus* (Johannessen et al. 2009) ger en andel på 0,097 procent *de här/där* av

Tabell II.2. Andel *de här/där* av det totala antalet ord i materialet per stad.

	DE HÄR		DE DÄR		DE HÄR/DÄR		ANTAL ORD
	%	ANTAL	%	ANTAL	%	ANTAL	
Helsingfors	0,148	76	0,105	54	0,254	130	51 186
Mariehamn	0,022	3	0,030	4	0,052	7	13 472
Vasa	0,138	41	0,135	40	0,273	81	29 683
Åbo	0,087	34	0,113	44	0,200	78	38 975
Totalt	0,116	154	0,107	142	0,222	296	133 316

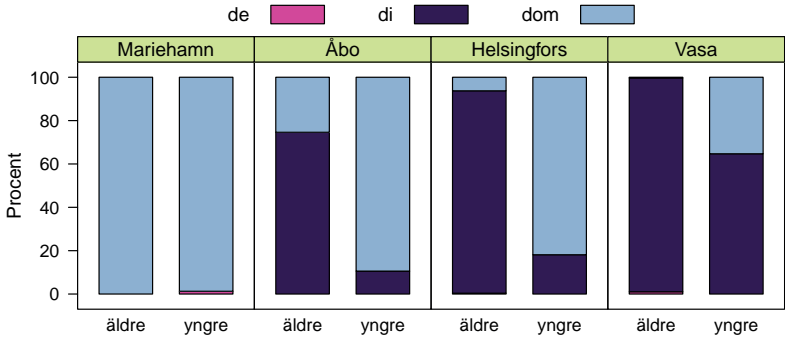
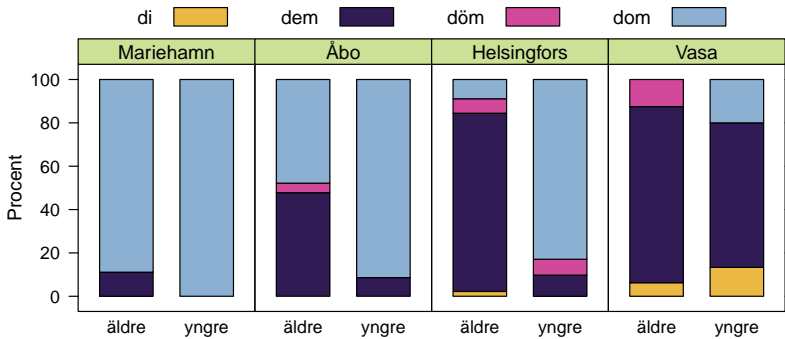
det totala antalet ord.⁹ Siffrorna från Mariehamn är alltså t.o.m. lägre än i talspråk från Sverige, medan siffrorna från de tre övriga städerna klart överstiger de sverigesvenska siffrorna.

4 Sociala variabler

Figurerna II.2 och II.3 visar andelen belägg på olika former av skriftspråkets *de* respektive *dem* i de fyra städerna. Figur II.2 visar belägg på skriftspråkets *de*, dvs. alla grammatiska kategorier utom personligt pronomen som objekt. Det finns några enstaka belägg på uttalet /de/ (så få att de knappt syns i staplarna), men vanligen är uttalet antingen /di/ eller /däm/. Staplarna visar både på en geografisk och en åldersmässig variation. I Mariehamn förekommer uttalet /di/ inte alls. I de tre övriga städerna är /di/ vanligare bland äldre informanter än bland yngre. Ovanligast är formen /däm/ bland de äldre informanterna från Vasa – i denna grupp finns bara ett enda belägg på /däm/. Figur II.3 visar fördelningen av varianter för skriftspråkets *dem*, dvs. personligt pronomen som objekt. Fördelningen är liknande som för *de*, med den högsta andelen /däm/ i Mariehamn och den lägsta i Vasa samt högre andelar /däm/ bland yngre än bland äldre informanter. Varianten /döm/ förekommer framför allt i Helsingfors och Vasa, men /döm/ är betydligt ovanligare än /dem/. Om formen /di/ som objekt, se avsnitt 5.2 nedan.

Faktorerna stad och ålder (äldre eller yngre) som illustreras i figurerna II.2 och II.3 testades för statistisk signifikans med logistisk regressionsanalys med blandade effekter (för en beskrivning av metoden se I.1.4.4). Den beroende variabeln var pronomenet *de/dem*, som behandlades som en binär variabel med varianterna *dom* eller *annan form* (dvs. /di/, /de/, /dem/ eller /döm/). Informant behandlades som stokastisk effekt i modellen. Faktorn ålder har hög signifikans ($z = 6,659, p < 0,001$) och visar att *dom* håller på att bli den vanligaste formen åtminstone i södra Finland. Även faktorn stad är signifikant¹⁰:

9. Sökningen gjord 24/9 2013. Antalet svenska informanter i korpusen var då 126 (från 38 olika orter) och det totala antalet ord 307 861.
10. Likelihood-kvottest av en modell med endast ålder som förklarande variabel jämfört med en modell med ålder och stad som variabel: $p < 0,001$. Interaktionen mellan stad och ålder testades också men förbättrade inte

Figur II.2. Stapeldiagram över *de*.Figur II.3. Stapeldiagram över *dem*.

dom är vanligast i Mariehamn och ovanligast i Vasa, däremot föreligger ingen statistiskt signifikant skillnad mellan Åbo och Helsingfors utan varianten *dom* är ungefär lika vanlig i dessa båda städer ($z = -0,927$, $p = 0,354$).

5 Språkliga variabler

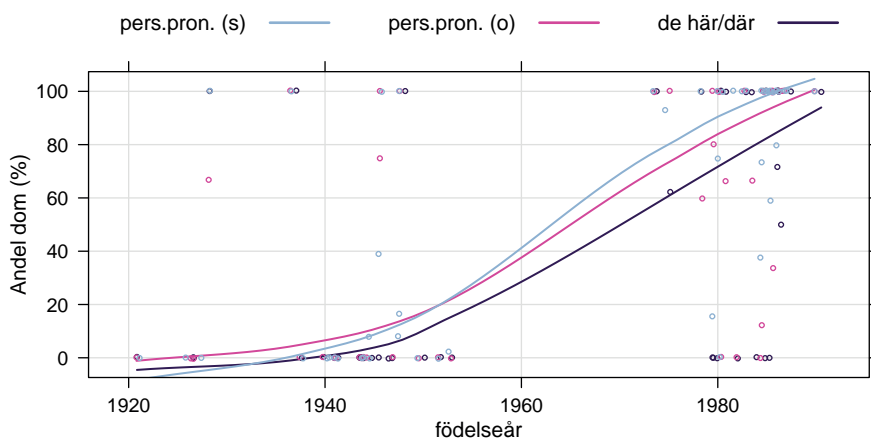
De lingvistiska faktorer som är av intresse för analysen av variationen $de/dem \sim dom$ är grammatisk kategori, betoning och ordföljd.

modellens anpassningsgrad (likelihood-kvottest av modell utan interaktion jämfört med en modell med interaktion: $p = 0,331$). Den slutgiltiga modellen med två fixa effekter och en stokastisk effekt har en anpassningsgrad på Somers $D_{xy} = 0,961$. Det totala antalet belägg på variabeln i analysen var 1 518.

5.1 Grammatisk kategori och betoning

Grammatisk kategori och betoning är starkt beroende av varandra eftersom *de* som obestämd artikel och i *de här/där* så gott som alltid är obetonat. Det personliga pronomenet kan vara både obetonat och betonat. I analysen av de språkliga variablerna inkluderades endast de 23 informanter som uppvisar variation, dvs. varierar mellan *de/dem* och *dom*. Demonstrativa pronomen (förutom demonstrativt *de här/där*) och determinativ användning lämnades utanför analysen eftersom antalet belägg i dessa kategorier är så litet. Efter denna reducering av materialet var antalet belägg på variabeln 703. Materialet analyserades med logistisk regressionsanalys med blandade effekter. Informant behandlades som stokastisk effekt. De lingvistiska faktorernas signifikans testades med likelihood-kvottest.

Analysen visade att effekten av grammatisk kategori är större än effekten av betoning.¹¹ Varianten *dom* är vanligast som personligt pronomen i subjektställning, följt av objektsform och bestämd artikel,



Figur II.4. Förändringen från *de/dem* till *dom* över tid efter grammatisk kategori.

11. Log-likelihood för modellen med grammatisk kategori som förklarande variabel var $-266,31$ (likelihood-kvottest $p < 0,001$), medan den för modellen med betoning som förklarande variabel var $-289,03$ (likelihood-kvottest $p = 0,001$). Modellen med grammatisk kategori som fix effekt och informant som stokastisk effekt har en anpassningsgrad på Somers $D_{xy} = 0,867$.

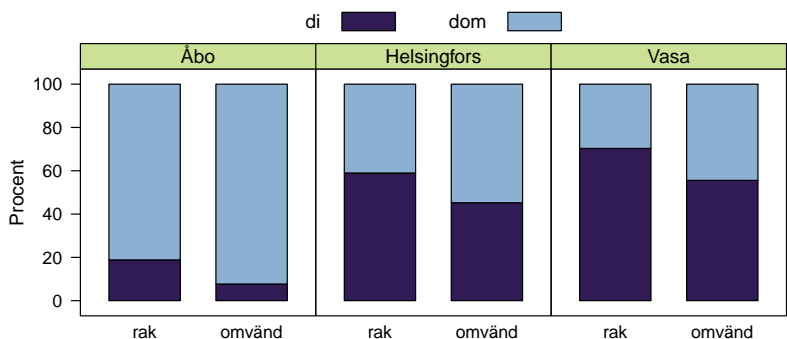
och ovanligast i *de här/där*. Betoning förbättrar inte en regressionsmodell där man redan beaktat grammatisk kategori (likelihood-kvottest $p = 0,758$).

Betydelsen av grammatisk kategori i förändringsprocessen illustreras i figur II.4. De tre största grammatiska kategorierna, dvs. personligt pronomen som subjekt och objekt samt *de här/där*, ingår i figuren. Andelen *dom* i dessa kategorier hos varje informant är inprickad med en boll i figuren. Varje informant representeras alltså av tre bollar i figuren, en av varje färg.¹² Informantens födelseår anges på x -axeln och andelen *dom* på y -axeln. Linjerna i figuren visar utvecklingstendensen över tid.¹³ Tendensen är överlag stigande, vilket visar att användningen av *dom* ökar med stigande födelseår, men vi ser också en skillnad mellan de olika grammatiska kategorierna. Linjen som representerar personligt pronomen som subjekt ligger överst, vilket antyder att förändringen börjar i denna kategori. Linjen som representerar *de här/där* ligger underst. Utvecklingen verkar stämma överens med förändringsprocessen i Eskilstuna beskriven av Grass (1970; se avsnitt 1).

5.2 Ordföljd

Vid excerperingen noterades ordföljd för alla personliga pronomen. Det totala antalet belägg på det personliga pronomenet hos de informanter som överhuvudtaget uppvisar variation i fråga om variabeln är för litet för att testa statistisk signifikans. Figur II.5 åskådliggör skillnaden i andelen *dom* i rak respektive omvänd ordföljd för personligt pronomen som subjekt ($n = 362$). Formen */dåm/* verkar gynnas av omvänd ordföljd medan */di/* gynnas av rak ordföljd. Skillnaden är inte påfallande stor, men den är konsekvent för de tre städerna, och tendensen är densamma som tidigare påvisats av Grass (1970) och Nyholm (1984).

12. Undantag utgör de fyra yngre informanterna från Mariehamn som saknar belägg på *de här/där* samt fyra informanter (en från Åbo, två från Helsingfors och en från Mariehamn) som saknar belägg på personligt pronomen som objekt.
13. Linjerna är utjämnade med funktionen *loess* i programmet R (R Core Team 2013).



Figur II.5. Stapeldiagram över ordföljdens betydelse för variationen /di/ ~ /dåm/ hos personligt pronomen som subjekt. Varianten /dåm/ förekommer något oftare i omvänd ordföljd än i rak ordföljd i alla tre städer.

För skriftspråkets *dem* används även varianten /di/ i talspråket. Detta sker ibland när objektet inleder en sats, som i följande exempel:

di ha:r ja i kru:kor (va.owo2)

dem har jag i krukor

di minns man jettebra: (va.ym10)

dem minns man jättebra

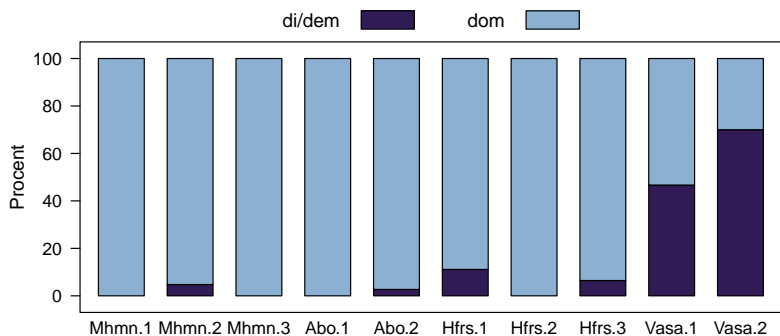
Att objektet inleder en sats är en ovanlig ordföljd; det vanliga är att subjektet står i denna position. Därmed kan en generalisering ha skett som innebär att pronomenet har formen /di/ när det inleder en sats, oberoende av om det är frågan om ett subjekt eller objekt.

Antalet objekt i början av en sats är inte så stort i materialet att långtgående slutsatser kan dras ($n = 19$). Från Vasa finns det fem belägg på objekt i denna ställning, varav fyra uttalas /di/ och ett /dem/. Eventuellt kan alltså en tendens skönjas till att /di/ uppfattas som den normala varianten i denna position i Vasa. Från Helsingfors finns åtta belägg varav fem /dåm/, två /dem/ och ett /di/. Från Åbo och Mariehamn finns bara belägg på varianten /dåm/ i denna ställning.

6 Intervjuarna och deras inflytande på informanterna

I Nyholms (1976, 39) studie av formväxlingen i Helsingforssvenskan använde de tre intervjuarna som utförde de sociolingvistiska intervjuerna nästan uteslutande varianten /di/ av variabeln /di/ ~ /dåm/. Intervjuarna var nordister (Nyholm 1976, 18), alltså högutbildade, språkmedvetna personer. Den allmänna attityden till /di/ respektive /dåm/ bland informanterna var att /di/ var den "goda varianten" (Nyholm 1984, 42). Intervjuarna använde alltså den variant som i 70-talets Helsingfors uppfattades som prestigeform.

Även i materialet för denna studie är de flesta av intervjuarna nordister. Intervjuarna i de två undersökningarna är alltså jämförbara med varandra. Men vad gäller användningen av *de/dem* ~ *dom* skiljer sig de två grupperna av intervjuare markant. Figur II.6 visar staplar för var och en av intervjuarna.¹⁴ Någon indelning efter grammatisk kategori har inte gjorts utan i materialet ingår alla belägg på *de/dem*. Varianten /dåm/ dominerar bland intervjuarna, utom i Vasa där andelen /dåm/ är mellan 30 och 50 procent. Intervjuarna (utom en av de åländska) hör till den yngre åldersgruppen och betar sig på samma sätt som de yngre informanterna i respektive stad vad gäller denna språkliga variabel.



Figur II.6. Stapeldiagram över andelen /dåm/ bland intervjuarna.

14. Utom en av intervjuarna från Helsingfors som bara utförde en enda intervju där belägg på variabeln helt saknas.

Intervjuarna och informanterna kan tänkas influera varandra språkligt (Giles et al. 1991). I en sociolingvistisk analys är man i synnerhet intresserad av att veta om informanten anpassar sig till intervjuaren eller en annan samtalspartner, t.ex. så att informanten i ett svar använder en viss variant för att intervjuaren använt denna i sin fråga.

Det går att hitta exempel på denna typ av språklig anpassning i materialet. Det enda belägget på /dåm/ hos äldre informanter i Vasa hittar man t.ex. hos en av de äldre kvinnorna i en direkt upprepning av intervjuarens replik (*recycling of phrase*, jfr Nordberg 2012). Men flera av informanterna låter sig inte alls påverkas. Av de äldre informanterna från Åbo och Helsingfors använder t.ex. sju stycken (fyra från Åbo och tre från Helsingfors) inte en enda gång varianten /dåm/ trots att denna form dominerar hos intervjuarna.

En viss språkforms varaktighet (dvs. förmåga att fortplanta sig) inom ett samtal kan bero på många olika faktorer (bl.a. psykolingvistiska, diskursbundna och inomspråkliga, för en översikt se Szmrecsanyi 2006, 9–42). I hur stor utsträckning (brist på) språklig variation kan förklaras genom denna typ av varaktighet har visat sig vara olika för olika typer av språkdrag (Szmrecsanyi 2006, 182). Hypotetiskt skulle man kunna testa varaktighetens betydelse för variationen *de/dem* ~ *dom*. En möjlighet till direkt inflytande från intervjuaren föreligger när föregående belägg på variabeln i kontexten har uttalats av intervjuaren. Eftersom det handlar om intervjuer där intervjuaren ställer relativt korta frågor och informantens talturer är betydligt längre, visade det sig dock att det föregående belägget på variabeln i mer än 90 procent av fallen uttalas av informanten själv. Eftersom informanten och intervjuaren i flera av intervjuerna dessutom konstant använder en identisk variant, är fallen där inflytande från intervjuaren överhuvudtaget kan förekomma och mätas så få att ingen systematisk analys är möjlig.

7 Slutsatser

Undersökningen pekar mot en övergång från separata subjekts- och objektsformer till enhetsformen /dåm/ i de finlandssvenska regionhuvudstäderna. Förändringen kan ses såväl i skenbar tid i det aktuella materialet som i verklig tid i jämförelse med Nyholms (1976 och 1984) Helsingforsundersökningar från 1970-talet. På 2000-talet är /dåm/

betydligt vanligare hos yngre personer än hos äldre personer, och i Helsingfors använder unga personer på 2000-talet /dåm/ oftare än unga personer gjorde på 1970-talet. I jämförelse med Nyholms material kan man också konstatera att unga högutbildade språkmedvetna personer (dvs. intervjuarna) på 1970-talet uteslutande använde formerna /di/ och /dem/, medan /dåm/ är klart dominerande hos motsvarande språkbrukare på 2000-talet.

Det finns också en skillnad mellan de fyra städerna i undersökningen. För Mariehamns del finns varianten /di/ inte belagd i stadsspråket vare sig i detta material eller tidigare. De få beläggen på objektsformen /dem/ i Mariehamn kan bero på inflytande från skriftspråket. De tre övriga städerna visar alla ett liknande mönster där formerna /di/, /dem/, /döm/ och /dåm/ förekommer sida vid sida, men andelen /dåm/ ökar med sjunkande ålder. Formen /döm/ är relativt ovanlig, men förekommer i alla tre städer i synnerhet bland de äldre. Medan förändringen i Åbo och Helsingfors verkar ske ungefär samtidigt, ligger Vasa något efter och har alltså lägre andelar /dåm/ i båda åldersgrupperna än Åbo och Helsingfors. Att förändringen i Vasa sker långsammare kan bero på att staden omges av landsbygd där dialekterna traditionellt upprätthållit skillnaden mellan subjektsform och objektsform, medan landsbygdsdialekterna kring Åbo och Helsingfors har /dåm/ både som subjekt och objekt.

Övergången till /dåm/ följer vissa språkliga mönster, och dessa verkar vara desamma i Finland som de varit i Sverige. Grammatisk kategori och ordföljd är faktorer som inverkar på valet av variant. Den längre och fonetiskt mer distinkta formen /dåm/ är vanligare när pronomenet används självständigt än när det står förenat. Formen /dåm/ är ovanligare i det demonstrativa pronomenet *de här/där* än som personligt pronomen. Samtidigt är *de här/där* betydligt mer frekvent i finlandssvenskan än i svenskan i Sverige. Eventuellt kan den mer frekventa användningen av *de här/där* delvis förklara att övergången till /dåm/ skett långsammare i Finland än i Sverige. Ordföljden inverkar på valet av variant i båda riktningarna: den ursprungliga objektsformen /dåm/ är vanligare i omvänd än i rak ordföljd, och subjektsformen /di/ används ibland för objekt när objektet inleder en sats.

Resultaten antyder att /dåm/ liksom i sverigesvenskan så småningom kommer att bli den allenarådande formen också i de finlandssvenska

stadsspråken. Det faktum att finlandssvenskar skiljer mellan /di/ och /dem/ i talspråket har använts som argument mot en *dom*-reform i svenskt skriftspråk. Men de finlandssvenskar som upprätthåller distinktionen ser ut att bli färre och argumentet blir därmed förlegat. Om någon generation har kanske även finlandssvenska skribenter tappat språkkänslan för *de* och *dem* i skrift.

Del III

FONETISK KVANTITET
I FINLANDSSVENSKA STADSMÅL

1 Introduktion

Svenskan är ett språk med s.k. komplementär kvantitet, vilket innebär att varje betonad stavelse har antingen en lång vokal /V:(C)/, t.ex. *mata* /ma:ta/, eller en kort vokal följt av en lång konsonant /VC:/, t.ex. *matta* /matta/ (Elert 1964, 39; Riad 2014, 10). Detta system är ett resultat av kvantitetsskiftet som ägde rum i äldre fornsvensk tid (med början kring 1300; Riad 1992, 235ff.). I fornnordiskan fanns fyra möjliga vokal-konsonant-sekvenser i betonade stavelser: VC, V:C, VC: och V:C:. Kvantitetsskiftet ledde till en förlust av korta stavelser (VC) och överlånga stavelser (V:C:); ord med någon av dessa två stavelsetyper övergick i de två andra typerna.

Dialekter i olika delar av språkområdet påverkades i olika grad av kvantitetsskiftet. Alla svenska dialekter i Finland utom de åländska bevarade den fonologiska kortstavigheten i ord som t.ex. *gräset* /grese/ (Ahlbäck 1971, 10). Även fonologiskt överlånga stavelser bevarades, åtminstone delvis, i många dialekter i Finland; de försvann endast i de västligaste delarna av dialektområdet (dvs. Åland, västra Åboland och södra Österbotten; Ivars 1988, 67). Ett exempel på ett fonologiskt långstavigt ord i dialekterna är *kryssa* /kry:ss(a)/.

Det svenska standardspråket i Finland saknar överlånga stavelser. Kortstavigheten finns däremot bevarad, om än i ett mer begränsat antal ord än i dialekterna. Framför allt handlar det om högfrekventa ord som normalt är obetonade i sin kontext, t.ex. pronomenet *mina* /mina/, adverbet *bara* /bara/, och vissa typer av lånord, t.ex. *kamera* /kamera/ (Reuter 1980; Kiparsky 2008).

De bevarade kvantitetstyperna i svenskan i Finland som saknas i Sverige är enligt Reuter (1992, 108) delvis ett arkaistiskt drag, men antagligen i ännu högre grad ett resultat av kontakt med finskan. I motsats till svenskan är finskan ett typiskt kvantitetsspråk där två längdgrader är möjliga för nästan alla språkljud i såväl betonad som obetonad ställning. Ytterligare en likhet mellan finlandssvenskan och finskan är att den kvalitativa skillnaden mellan motsvarande korta och långa vokaler är liten, medan sverigesvenskan utnyttjar kvalitativa skillnader vid sidan av kvantitativa skillnader för den fonematiska distinktionen (Reuter 1992, 108). Uttalet av vokalerna i svenskan i Finland ligger mellan det finska och det sverigesvenska uttalet (Reuter 1971).

Kvantitetssystemen i olika varieteter av svenska i Finland skiljer sig inte åt bara fonologiskt utan även i fråga om den exakta fonetiska längden hos fonologiskt långa och korta vokaler och konsonanter. Fonologiskt korta men fonetiskt relativt långa tonlösa obstruenter (dvs. /f/, /k/, /p/, /t/, /s/ och /ʃ/) i V:C-sekvenser (i ord som *läsa*, *veta*) har varit ett karakteristiskt särdrag för den s.k. bildade svenskan i Helsingfors, och kvantitet har setts som en av de tydligaste sociala markörerna i svenskan i Finland (Reuter 1982, 197; Höckerstedt 2010, 61). I motsats till de långa konsonanterna i Helsingfors, karakteriseras dialekter på landsbygden av mycket korta konsonanter i denna ställning, klart kortare än motsvarande konsonanter i sverigesvenska dialekter (Reuter 1982, 194; Schaeffler 2005, 80–82).

Lyssnare som är vana vid ett fonologiskt system med både korta och överlånga stavelser (t.ex. talare av finlandssvenska dialekter eller av finska) uppfattar lätt det helsingforssvenska uttalet av ett ord som *bita* som överlångt /bi:tta/ trots att det inte handlar om ett fonologiskt långstavigt ord, utan den fonologiska formen är /bi:ta/. Orsaken till detta är alltså att konsonanten i Helsingforssvenskan kan uttalas betydligt längre än i dialekterna, där den konkreta längdskillnaden mellan vokalen och konsonanten är större.

Flera forskare har studerat kvantiteten i svenskan i Finland (se nedan). Även ett litet antal akustiska studier föreligger, men dessa har fokuserat antingen på högutbildade talare från Helsingfors (Reuter 1982) eller på landsbygdsdialekter (Schaeffler 2005). Studier som jämför olika stadsspråk saknas, och syftet med denna undersökning är att studera fonetisk kvantitet i V:C-sekvenser i fyra städer: Helsingfors, Åbo, Mariehamn och Vasa.¹ Undersökningen är begränsad till V:C-sekvenser där konsonanten är en tonlös obstruent i mellanvokalisk ställning, dvs. ord av typen *läsa*, *veta*.

2 Tidigare forskning

Det finlandssvenska standardspråket uppstod i Åbo och vidareutvecklades sedan i Helsingfors i och med att Helsingfors blev huvudstad 1812 (Ahlbäck 1971, 18–30). Helsingfors- och Åbosvenskan har därmed

1. Delar av resultaten har tidigare publicerats på engelska i Leinonen 2013.

traditionellt varit prestigeformen av finlandssvenska och varit den varietet som använts t.ex. av nyhetsuppläsare i radio och tv. På senare år verkar det ändå ha skett en förändring i vilket språk som finlandssvenskar uppfattar som standard eller ”det bästa språket”. Ivars (2003) och Östman & Mattfolk (2011) har rapporterat att Helsingforssvenskan inte automatiskt uppfattas som standardspråk av finlandssvenskar utanför Helsingforsregionen, i synnerhet i Österbotten. Stenberg-Sirén & Östman (2012) och Stenius (2012) har också noterat en minskad användning av typiska Helsingforsdrag – inklusive långa mellanvokaliska konsonanter i V:C-sekvenser – i massmedierna.

Diskussionen nedan begränsar sig till kvantiteten i V:C-sekvenser där konsonanten är en tonlös obstruent (/f/, /k/, /p/, /t/, /s/, /ʃ/). Tonlösa obstruenter har i tidigare forskning visat sig uppvisa stor variation å ena sidan mellan stadsspråk och dialekter i Finland och å andra sidan mellan dialekter i Finland och Sverige. I svenskan är de tonande konsonanterna i allmänhet kortare än de tonlösa (Elert 1964, 148; Reuter 1982, 88), och den regionala variationen är mindre för de tonande konsonanterna än för de tonlösa. Av relevans i sammanhanget är också att finskan endast har tonlösa obstruenter.²

Kvantitet är ett språkdrag med starka regionala och sociala associationer i svenskan i Finland. Långa mellanvokaliska obstruenter i V:C-sekvenser har varit typiskt för de högre socialgruppernas språk i Helsingfors men också i Åbo (Reuter 1982, 100). Reuter (1982, 113–131) påvisade dessutom att variabeln är socialt stratifierad i Helsingfors: högre socioekonomisk status innebär i regel längre konsonantlängder. Reuter kunde även konstatera att talare uppvuxna i Helsingfors hade längre konsonanter än inflyttade helsingforsare. Kvinnor hade längre konsonanter än män i alla socialgrupper i Reuters studie och detta gällde såväl infödda som inflyttade helsingforsare, vilket visar på det mycket vanligt förekommande sociolingvistiska mönstret där kvinnor använder en högre andel prestigeformer än män.

Variabeln kan kallas en sociolingvistisk stereotyp i Labovs (1972) termer, eftersom den ofta imiteras och kommenteras av icke-språkvetare

2. Utom /d/, som förekommer i samband med stadiesväxling och som fonetiskt befinner sig mellan en klusil (dvs. en obstruent) och en flappaktig konsonant (Suomi et al. 2008, 33). Konsonanterna /b/ och /g/ finns i lånord i finskan men uttalas inte tonande av alla finskspråkiga.

re (Reuter 1982, 100; af Hällström-Reijonen 2010). I en folklingvistisk studie av Haapamäki & Kvarnström (2014) nämnde flera av deltagarna från Helsingfors och Åbo kvantitet som ett särskiljande drag för svenskan i Helsingforsregionen. Åboborna i studien ansåg att kvantitet utgör en skillnad mellan Åbosvenskan och Helsingforssvenskan, som Åboborna överlag hade mycket negativa attityder till (s. 112).

En av informanterna i *Spara talet*-materialet nämner språkdraget under intervjun. Det är en yngre man från Helsingfors som berättar om vilken typ av svenska som talas på den yrkeshögskola där han studerar:

*dä: fösö:kä: mensjå: (.) pra:ta me:r de hä närma:la (.)
helsingfårssprå:ke men de e ju int de dä: "ha:tta: må:ssa:" de hä
granigrejjen (.) u:tan de e ju de e ju ganska så hä: (.) plein
(Hf.ymo8)*

där försöker mänsikor (.) prata mer det här normala (.)
Helsingforsspråket med det är ju inte det där "hatar måsar" det
här Granigrejen (.) utan det är ju det är ju ganska så här (.) plain

Informanten imiterar det han kallar "Granigrejen"³ och använder två ord med V:C-sekvenser med tonlösa obstruenter (*hatar måsar*). Vid imitationen använder han nasal röst och överdriver längden på de mellanvokaliska konsonanterna. Imitationen är klart nedsättande och enligt informanten används ett neutralare uttal på hans skola. Informanten själv använder de facto ett ganska neutralt uttal utan extremt låga eller extremt höga V:/C-värden (se figur III.2 och tabell III.3).

Flera forskare har försökt förklara kvantitetsmönstren i finlandssvenskan genom språkkontakt med finskan. Konsonantlängden i V:C-sekvenser i mellansvenskt standardspråk är längre än i finska V:C-sekvenser, och finskspråkiga tenderar att uppfatta svenska V:C-sekvenser som V:C: (Hakulinen 1979, 28). En del forskare har förklarat de långa konsonantlängderna i V:C-sekvenser i Helsingforssvenskan som ett resultat av interferens hos tvåspråkiga talare (Hakulinen 1979, 28; Itkonen 1965, 261). Konsonanterna skulle alltså ha förlängts som en

3. Grani syftar på staden Grankulla ca 15 km väster om Helsingfors centrum som traditionellt haft en stor andel invånare med höga inkomster.

substratumeffekt genom att V:C uppfattas som V:C: av personer med finska som första eller starkaste språk.

Jämförbara mätningar av segmentlängder i rikssvenska och Helsingforsvenska är sällsynta, och det förblir oklart i vilken grad Helsingforsvenska konsonanter är längre än konsonanterna i mellansvenskt standardspråk. På grund av olika experimentella upplägg och olika segmenteringsregler är resultaten från olika studier sällan direkt jämförbara. Men åtminstone under vissa omständigheter kan man observera förlängning av postvokala korta konsonanter även i rikssvenskan. Elert (1964, 166) skriver:

in utterances consisting of isolated words, or lists thereof, all phonemes are lengthened, especially the consonants following after a stressed vowel. Even phonemically short consonants following long vowels have longer duration than the vowels.

Helgason et al. (2010 och 2013) mätte segmentlängder hos sex talare från Stockholm och tolv finlandssvenskar. Åtta av talarna från Finland producerade mycket korta mellanvokaliska tonlösa obstruenter i V:C-sekvenser (ruralt uttal [min term]), medan fyra talare producerade signifikant längre mellanvokaliska konsonanter (helsingforsvenskt uttal [likaså min term]). Av de fyra talarna med helsingforsvenskt uttal i studien härstammade de facto endast en från Helsingfors, medan två var födda i Åbo och en i Ekenäs. Talarna med ruralt uttal härstammade från Österbotten eller mindre orter i södra Finland. Alla informanterna var bosatta i Åbo då studien genomfördes. Medelvärdena för den procentuella andelen V: av V: + C (dvs. vokalens procentuella andel av den totala längden på V:C-sekvensen) visar att det rurala finlandssvenska uttalet ($M = 70,7\%$) avviker klart mer från Stockholmssvenskan ($M = 55,3\%$) än vad det helsingforsvenska uttalet gör ($M = 50,1\%$) (Helgason et al. 2013, 539). Variationen mellan olika talare och hos varje talare rapporteras inte. Dessa forskares resultat motsvarar hur som helst Reuters resultat med avseende på att det stockholmska uttalet ligger mitt emellan de två olika finlandssvenska uttalen men närmare det helsingforsiska uttalet än det rurala finlandssvenska uttalet. Reuter (1982, 108–112) mätte en aning högre V:/C-värden hos två talare från Stockholm än för talare från Helsingfors.

Inte bara de långa konsonantlängderna i Helsingforssvenskan, utan också de exceptionellt korta konsonantlängderna i finlandssvenska dialekter har förklarats genom finskt inflytande (Itkonen 1965, 261–262). Segmentlängderna i V:C-sekvenser i finlandssvenska dialekter är påfallande lika de finska V:C-sekvenserna och avviker klart mer från V:C-sekvenser i sverigesvenska. Men de korta konsonanterna i finlandssvenska dialekter kan också förklaras inomspråkligt. Schaeffler (2005, 121–127) bortser från allt eventuellt finskt inflytande och menar att variationen i segmentlängd bland svenska dialekter kan förklaras genom de olika fonologiska systemen. I dialekter där man har, eller intill nyligen haft, ett 4-vägssystem (dvs. ett system med VC-, V:C-, VC:- och V:C:-sekvenser i betonad ställning) görs tydliga längddistinktioner både i fråga om vokallängd och konsonantlängd. I dialekter med ett 3-vägssystem (dvs. dialekter där överlånga stavelser saknas) är konsonantlängden i V:C-sekvenser inte längre kontrastiv eftersom V:C:-sekvenser saknas i dialekten, vilket tillåter längre konsonanter i V:C-sekvenser. I 3-vägssystemet, liksom i 4-vägssystemet, är C: i VC: fortfarande lång eftersom konsonantlängden kontrasterar VC och VC:. I det sista skedet, 2-vägssystemet, har VC försvunnit vilket tillåter kortare konsonantlängd överlag, eftersom den enda kvarvarande kontrasten är den mellan V:C och VC: och dessa två sekvenser kan åtskiljas enbart genom olika vokallängd.

Dialekterna i södra Sverige har nått det sista stadiet, 2-vägsstadiet, där kvantitetsskillnaden huvudsakligen är begränsad till en skillnad i vokallängd. Motsvarande system återfinns för övrigt i andra germanska språk som danska, nordtyska och nederländska (Bruce 2012, 110). Dialekterna i nordligare delar av Sverige har segmentlängder som motsvarar 3-vägssystemet, vilket sammanfaller med att kortstavigheten finns eller åtminstone ännu under 1900-talet funnits bevarad på flera håll i Norrland (Schaeffler 2005, 43). I många svenska dialekter i Finland har 4-vägssystemet bevarats betydligt längre än i Sverige, vilket enligt Schaeffler (2005) förklarar de mycket korta konsonanterna i V:C-sekvenser i dessa dialekter. Här kan man inflika att det faktum att 4-vägssystemet överhuvudtaget har bevarats i Finland mycket väl kan bero på att kontakt med kvantitetsspråket finska haft en konserverande verkan, och språkkontakt som en åtminstone delvis förklarande faktor är därmed inte utesluten (jfr Reuter 1992, 108).

I en optimalitetsteoretisk analys av olika varieteter av svenska i Finland utgår Kiparsky (2008) precis som Schaeffler (2005) från att kort C-längd i finlandssvenska dialekter kan förklaras som en kontrastbevarande effekt eftersom konsonantlängden behövs för att åtskilja V:C- och V:C:-sekvenser. Mellanvokaliska obstruenter i V:C-sekvenser i Helsingforssvenskan analyseras som geminator, dvs. långa konsonanter, av Kiparsky.

Tidigare akustiska mätningar av kvantitet i olika svenska varieteter i Finland har byggt på omsorgsfullt eliciterat material där målsekvenserna har uttalats i isolerade ord eller i korta exempelmeningar (Reuter 1982; Schaeffler 2005; Helgason et al. 2013). Faktorer som kontext, taltempo och prominens är välkontrollerade i denna typ av experiment. Den verkliga variationen som finns i spontant tal kommer ändå inte nödvändigtvis fram när informanter läser ord från en ordlista, och faktorer som final förlängning kan påverka resultaten (jfr Elert 1964, 166). Särskilt när det handlar om ett språkdrag som är socialt markerat kan talare också påverkas av att inspelningssituationen känns formell. Genom en auditiv analys av ett socialt skiktat intervjumaterial från Helsingfors kunde Reuter (1982, 127) dra slutsatsen att påfallande långa och påfallande korta mellanvokaliska konsonanter är undantag snarare än regel och att de flesta av talarna för det mesta producerade mellanlånga konsonanter.

Undersökningar av Heldner & Strangert (2001) gällande svenska och av Pind (1999) gällande isländska (som liksom svenskan har komplementär kvantitet) har visat att förhållandet mellan vokallängd och konsonantlängd inte är konstant utan tenderar att variera med talhastighet och betoningsgrad. Bägge studierna har visat att det segment som varierar minst i längd är vokalen i VC:-sekvenser. När VC:-sekvenser förlängs, på grund av långsammare taltempo eller starkare betoning, är det den fonologiskt långa konsonanten som förlängs. Vid V:C-sekvenser förlängs både vokalen och konsonanten när taltempot minskar, men variationen i vokallängd är något större än i konsonantlängd. I bägge fallen är det alltså det fonologiskt långa segmentet som förlängs mest, vilket maximerar kvantitetskontrasten såväl mellan det långa och korta segmentet inom sekvensen som mellan de båda sekvenstyperna. Vad gäller typiska kvantitetsspråk som finska och estniska, har Engstrand & Krull (1994) påvisat att talare av dessa språk i spontant tal i större

utsträckning verkar upprätthålla fonetiska längddistinktioner än vad svenskspråkiga gör.

Ytterligare en faktor som påverkar segmentlängd är vokalhöjd. Slutna vokaler har en inneboende egenskap att vara kortare än öppna vokaler (Lehiste 1970, 18). Detta är en universell tendens i världens språk som också har påvisats för rikssvenska (Elert 1964), finska (Lehtonen 1970) och helsingforssvenska (Reuter 1982). Vad gäller konsonanter så är tonande konsonanter i regel kortare än tonlösa (Elert 1964, 148), men i övrigt verkar inneboende konsonantlängd variera mellan olika språk (Lehiste 1970, 27–30).

I denna artikel analyseras variabeln 'fonetisk längd hos mellanvokaliska tonlösa obstruenter i V:C-sekvenser' bland svenskspråkiga i fyra städer i Finland. För att studera den socialt markerade variabeln i en naturlig kontext användes spontant talmaterial.

3 Material

Materialet består av informanterna från Helsingfors, Åbo, Mariehamn och Vasa i *Spara talet*-databasen. Antalet informanter från de olika städerna och åldersgrupperna ses i tabell 1 på s. 15 i inledningen till boken.⁴

Undersökningen koncentrerades till V:C-sekvenser där konsonanten är en tonlös obstruent i mellanvokalisk ställning. Detta har en praktisk anledning eftersom konsonantslut vid mellanvokaliska konsonanter kan markeras vid början av följande vokal, vilket underlättar segmenteringen i synnerhet av spontant tal i varierande inspelningsmiljöer. Resultaten går att jämföra med Reuters (1982) resultat eftersom han analyserade vokaler i såväl enstaviga som tvåstaviga ord. Schaeffler (2005) analyserade däremot kvantiteten endast i enstaviga ord. Reuter (1982, 89) kunde konstatera att de tonlösa obstruenterna var längre i tvåstaviga än i enstaviga ord.⁵ Tvåstaviga substantiv, verb och adjektiv med betoning på den första stavelsen användes i analysen.

4. En av de äldre männen från Åbo måste dock uteslutas ur kvantitetsstudien på grund av en röstegenskap där stämbandspåslagningen är försenad, vilket omöjliggjorde mätningen av konsonant- och vokallängd.
5. Schaefflers (2005) V:/C-värden för landsbygdsdialekterna översteg ändå klart Reuters (1982) mätvärden för Helsingforssvenskan i både de enstaviga och de tvåstaviga orden.

Även trestaviga verb med vanliga obetonade prefix (*be-*, *för-*) ingick i materialet. Transkriptionerna av *Spara talet*-materialet kunde användas för att identifiera de ord i materialet som fyllde dessa kriterier.

De tonlösa obstruenter som förekommer i mellanvokalisk ställning i svenskan är /f/, /k/, /p/, /t/, /s/ och /ʃ/. Av de möjliga konsonanterna är /f/ och /ʃ/ dock mycket ovanliga och förekommer bara i lånnord. Dessa två konsonanter segmenterades därför inte. Efter segmenteringen visade det sig att så många som 19 av informanterna helt saknade belägg med /p/. Belägg med /p/ som mellanvokalisk konsonant togs därför inte heller med i analysen.⁶ De konsonanter som kvarstod var alltså /k/, /t/ och /s/. Exempel på vanliga ord i materialet är olika former av verben *läsa*, *försöka*, *sluta* och *prata* och substantiv som *saker*, *dator* och *huset*.

I svensk prosodisk notation urskiljer man vanligen fyra nivåer av prominens: obetonad (ingen framhävninɡ), betonad (svagt framhävd, bibetonad), accentuerad (starkt framhävd) och fokuserad (mycket starkt framhävd) (Bruce 1994; Bruce 2012, 183–185). Att skilja mellan de olika prominensnivåerna i spontant tal har ändå visat sig vara svårt t.o.m. för tränade fonetiker (Strangert & Heldner 1995). I analysen ingår endast ord med någon av de två högsta prominensnivåerna (dvs. orden har huvudbetoning, enligt författarens perception), men ingen uppmärkning gjordes av distinktionen mellan accentuering och fokusering, eftersom i synnerhet skillnaden mellan de två högsta prominensnivåerna har visat sig vara svår att avgöra (Strangert & Heldner 1995). Segmenteringsreglerna från *SweDia*-projektet, beskrivna av Schaeffler (2005, 60–61), tillämpades på materialet. Schaefflers studie gäller dock enstaviga ord, vilket gjorde det ändamålsenligt att markera slutet av klusiler vid början av klusilexplosionen. Eftersom den aktuella studien gäller tvåstaviga ord har både det tysta intervallet och explosionsbruset räknats in i konsonantlängden.

Spara talet-inspelningarna gjordes i relativt tysta rum i informanternas hem. Inspelningskvaliteten skiljer sig ändå i viss mån mellan intervjuerna, och vissa inspelningar har mer bakgrundsljud än andra. Före segmenteringen reducerades brusnivån i inspelningarna med

6. /p/ är kortare än andra klusiler i många språk (Lehiste 1970), bl.a. i svenska (Elert 1964).

programmet Audacity. I vissa inspelningar finns rumseko som kan göra det svårt att identifiera var en vokal slutar och en klusil börjar. Bara belägg med tydliga övergångar mellan segmenten segmenterades. Segmenteringen och de akustiska mätningarna utfördes i Praat (Boersma & Weenink 2013).

4 Förhållandet V:/C

V:C-sekvenserna analyserades med regressionsanalys med blandade effekter (för en beskrivning av metoden se I.1.4.4) för att testa vilka språkliga och utomspråkliga faktorer som påverkar segmentlängderna. Den beroende variabeln var V:/C-värdet, alltså vokallängd dividerat med konsonantlängd, dvs. relativa segmentlängder. De språkliga fixa faktorer som testades var *vokalhöjd*, *konsonant* och *lokalt taltempo* (kvantifierat som den totala längden på V:C-sekvensen). Utomspråkliga fixa faktorer var *stad* och *ålder*. Informant behandlades som stokastisk faktor i modellen.⁷ Mariehamn togs inte med i denna analys eftersom antalet informanter från Mariehamn är litet och det för de övriga städernas del visade sig behövas en interaktionseffekt mellan stad och ålder.⁸ Tabell III.1 visar estimat och *t*-värden för de fixa effekterna i modellen som beskrivs närmare nedan. Ett absolut *t*-värde högre än 2 indikerar statistisk signifikans på 5-procentsnivån (Baayen 2008, 248).

7. Analysen gjordes med *lmer*-funktionen i paketet *lme4* (Bates et al. 2011) i R (R Core Team 2013). Som basmodell användes en modell med enbart den stokastiska faktorn informant för att förklara V:/C-värdet. Övriga faktorer lades till steg för steg och likelihood-kvottest användes för att identifiera signifikanta faktorer (Baayen 2008, 253). Den beroende variabeln V:/C-värde transformerades logaritmiskt före analysen för att uppnå en normal distribution (Kolmogorov–Smirnov-test: V:/C $D = 0,096$, $p < 0,001$; $\log(V:/C)$ $D = 0,031$, $p = 0,612$). Även den numeriska förklarande variabeln taltempo transformerades logaritmiskt för att minska effekten av extremvärden; dessutom centerades värdena eftersom en slumpnotning (eng. *random slope*) per informant krävdes för taltempo i modellen (jfr Baayen 2008, 254–255). Det totala antalet belägg i analysen var 647. Medelantalet belägg per informant var 16,2 ($SD = 6,5$, $min. = 6$, $max. = 41$).
8. Genom att utelämna Mariehamn kunde antalet faktornivåer i förhållande till antalet informanter hållas på en rimlig nivå.

Tabell III.1. Koefficienter för fixa effekter i regressionsmodellen med blandade effekter med log(V:/C) som beroende variabel.

	ESTIMAT	MEDELFEJ	T-VÄRDE
Intercept	0,299	0,090	3,314
log(V:/C längd)	0,167	0,093	1,807
Konsonant = <i>klusil</i>	-0,145	0,026	-5,578
Vokalhöjd	-0,045	0,016	-2,849
Stad = <i>Åbo</i>	-0,154	0,132	-1,172
Stad = <i>Vasa</i>	0,369	0,125	2,942
Ålder = <i>yngre</i>	-0,106	0,113	-0,937
log(V:/C längd): <i>Åbo</i>	0,490	0,143	3,438
log(V:/C längd): <i>Vasa</i>	0,301	0,148	2,039
<i>Åbo:yngre</i>	0,514	0,173	2,970
<i>Vasa:yngre</i>	0,144	0,188	0,765

4.1 Språkliga variabler

Både konsonant och vokalhöjd förbättrar modellens anpassningsgrad. En första analys visade att ingen signifikant skillnad fanns mellan konsonanterna /k/ och /t/, vilket motiverade att dessa två konsonanter slogs ihop till en faktornivå. Faktorn konsonant hade därmed två nivåer, vilka motsvarar artikulationssätt: frikativ (/s/) och klusil (/k/ och /t/).

Tabell III.1 visar en signifikant huvudeffekt för konsonant. Nivån frikativ har mappats till interceptet och modellen visar ett negativt estimat för klusil, vilket innebär att V:/C-värdet är lägre för klusiler än för den frikativa konsonanten. Effektstorleken är ändå mycket liten (-0,145). Ett lägre V:/C-värde kan innebära antingen längre vokaler eller kortare konsonanter, eller både och. Medellängden för vokaler före /s/ är något längre ($M = 136$ ms, $SD = 45$, $n = 148$) än före klusiler ($M = 125$ ms, $SD = 45$, $n = 499$), medan konsonantlängden är lika (/s/: $M = 103$ ms, $SD = 29$; klusiler: $M = 103$ ms, $SD = 33$). Det mänskliga örat sägs kunna uppfatta längdskillnader bara om skillnaden är minst 10 procent (Rietveld & Van Heuven 2009, 221), vilket gör att effekten av konsonant på V:/C-värdet närmast är av teoretiskt intresse. Att

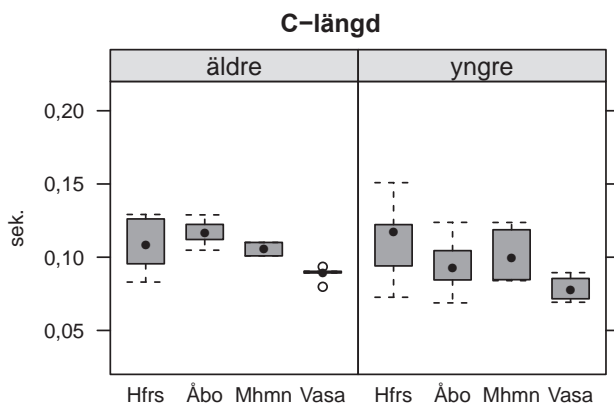
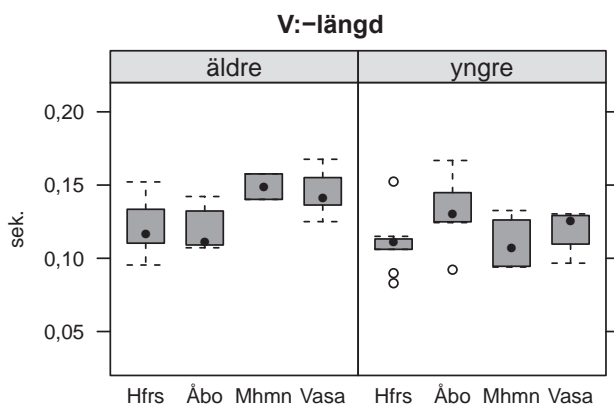
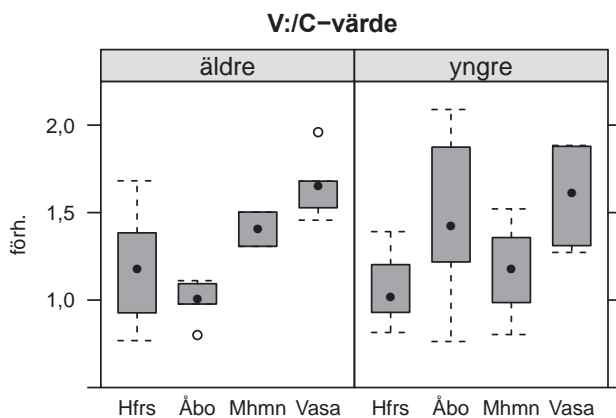
vokaler är längre framför frikativter än framför klusiler är en universell tendens (Davis 2011, 104).

Vokalhöjd behandlades som en numerisk variabel med värdena 1 (= sluten: /i:/ /y:/ /u:/ /o:/; $n = 197$), 2 (= mellan: /e:/ /ö:/ /å:/; $n = 330$) och 3 (= öppen: /a:/; $n = 120$). Effekten av vokalhöjd är signifikant, men ännu mindre än för konsonant. Öppna vokaler är längre än slutna vokaler, vilket också tidigare forskning har visat (Lehiste 1970, 18). Medelvokallängderna i materialet är 121 ms ($SD = 42$) för slutna vokaler, 125 ms ($SD = 48$) för mellanvokaler (skillnaden mellan slutna vokaler och mellanvokaler är inte statistiskt signifikant) och 143 ms ($SD = 40$) för öppna vokaler. Konsonantlängden påverkas inte av vokalhöjden.

4.2 Sociala variabler

Av de tre städerna mappades Helsingfors till interceptet (dvs. det ställe där regressionslinjen skär y-axeln). Skillnaden mellan Helsingfors och Åbo är inte statistiskt signifikant, men Vasa har ett signifikant högre V:/C-värde. I modellen ingår också en signifikant interaktion mellan stad och ålder. I Helsingfors och Vasa finns ingen skillnad mellan åldersgrupperna, men i Åbo har de yngre informanterna signifikant högre V:/C-värden än de äldre.

Figur III.1 visar lådagram över de tre värdena V:/C-värde (utan logaritmisk transformation), V:-längd och C-längd (i I.1.4.3 ges en närmare beskrivning av hur diagram av denna typ ska läsas). Diagrammen har ritats utifrån medianvärden per informant. Ett V:/C-värde på 1 innebär att vokalen och konsonanten i V:C-sekvenser är lika långa. Många talare från Helsingfors och Åbo har värden kring 1, och en del talare har t.o.m. medianvärden mindre än 1, vilket innebär att V: är kortare än C (trots att det fonologiskt sett alltså är vokalen som är lång). I Vasa är V:/C-värdena klart högre än 1 och V: är alltså betydligt längre än C. Flera av de yngre informanterna från Åbo har värden som påminner mer om värdena från Vasa än om värdena hos Helsingforsinformanter och äldre Åboinformanter. Variationen är också påfallande stor bland de yngre Åboborna. Den stora skillnaden mellan äldre och yngre Åbobor tyder på en pågående språkförändring. I figuren ingår även värdena från Mariehamn. Tendensen i Mariehamn är snarare den motsatta



Figur III.1. Lådagram över V:/C-värde, V:-längd och C-längd efter stad och åldersgrupp. Diagrammen baseras på medianvärden per informant.

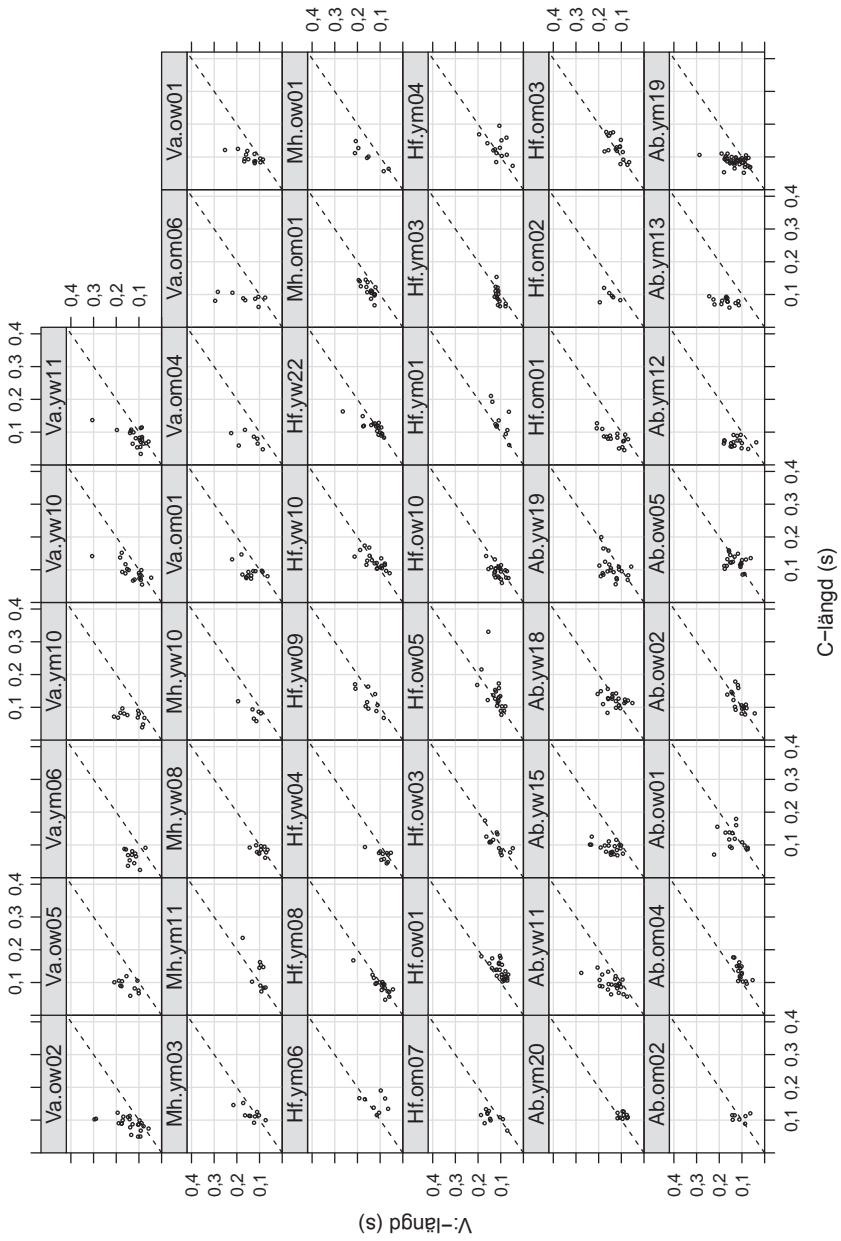
än i Åbo, dvs. de äldre informanterna har värden som påminner om värdena från Vasa, medan de yngre informanterna har värden som liknar Helsingforsvärdena.

De två understa diagrammen i figur III.1 visar att såväl vokallängd som konsonantlängd bidrar till de stora skillnaderna i V:/C-värde mellan informantgrupperna.

4.3 Taltempo

I modellen ingår faktorn V:C-längd (logaritmiskt transformerad) som har en signifikant interaktion med stad. V:/C-värdet ökar mera med sjunkande sekvenslängd i Åbo och Vasa än i Helsingfors. Skillnaden mellan städerna handlar alltså inte bara om en skillnad i medelvärde för förhållandet V:/C, utan det finns också en skillnad i hur V:/C-värdet påverkas av långsammare taltempo eller prominens. Talarna från Åbo och Vasa gör en tydligare längdskillnad mellan fonologiskt långa och korta segment vid långsammare lokalt taltempo än vid snabbare taltempo. Detta gäller inte för Helsingforsinformanterna.

En slumpplutning för V:C-längd för talare förbättrade också anpassningsgraden, vilket innebär att det finns individuella skillnader i hur lokalt taltempo påverkar V:/C-värdet. Figur III.2 visar spridningsdiagram där vokallängden och konsonantlängden har prickats in separat för varje informant. Den streckade linjen i diagrammen anger det ställe där V:/C är 1, dvs. vokalen och konsonanten är lika långa. Många informanternas (så som Ab.y_m13, Hf.o_m01 och de flesta från Vasa) uttal ligger tydligt ovanför denna linje, vilket innebär att V: alltid är längre än C. Men det finns också talare (såsom Ab.o_m04 och Hf.o_w01) vars punkter ligger under linjen – konsonanten är alltså längre än vokalen i V:C-sekvenserna. Skillnaden i hur taltempo inverkar på värdena är också tydlig. Talare som Ab.y_m19 och Va.o_m06 ökar helt tydligt på vokallängden när den totala sekvenslängden ökar, medan andra (som Hf.o_w05 och Hf.y_m01) varierar konsonantlängden mer än vokallängden. Några talarers punkter (t.ex. Hf.y_m08) finns centrerade kring den streckade linjen, vilket innebär att vokalen och konsonanten är ungefär lika långa och förlängs lika mycket när taltempot minskar, vilket håller V:/C-värdet konstant.



Figur III.2. Punktdiagram över V-längd och C-längd för varje informant. Den streckade linjen anger $V/C = 1$.

4.4 Slumpintercept

De enda sociala variablerna i modellen är stad och åldersgrupp. Ytterligare faktorer som vore av intresse, utifrån tidigare forskning, är kön och socioekonomisk nivå. Eftersom antalet informanter i varje grupp är så litet, var det inte möjligt att testa dessa faktorer kvantitativt. När det gäller socioekonomisk nivå har databasen dessutom inte tillräckligt med information om alla informanter. De flesta av de yngre informanterna har gått eller går på gymnasiet. Information om föräldrarnas bakgrund saknas. Bland de äldre informanterna varierar utbildningsnivån mer än bland de yngre.

För att i alla fall bilda sig en uppfattning om hur faktorer som kön och utbildningsnivå påverkar kvantiteten, kan regressionsmodellens slumpintercept studeras. Drager & Hay (2012) visade att slumpintercept inte enbart behöver ses som en biprodukt av regressionsmodeller med blandade effekter, utan att de kan tolkas kvalitativt. Slumpinterceptet visar hur en persons språkbruk förhåller sig till andra talares i språkgemenskapen. Slumpinterceptet estimerar i vilken grad en person avviker från gruppvärdena i den fixa delen av regressionsmodellen.

Ett sätt att studera slumpintercept är att använda slumpintercepten från den fullständiga modellen med de fixa effekter som visas i tabell III.1. Ett annat sätt är att helt utelämna de sociala variablerna och låta slumpintercepten omfatta all informantbunden variation, inklusive stad och ålder. Bägge metoderna är av intresse. Tabell III.2 visar slumpintercepten från den fullständiga modellen (modellen i tabell III.1). Tabellen är grupperad enligt stad och åldersgrupp eftersom denna variation förklaras av de fixa faktorerna. Informanterna i varje grupp är ordnade efter stigande slumpintercept. I praktiken innebär detta att talarna med det högsta V:/C-värdet står längst ner i varje grupp.

Tabell III.3 visar slumpintercepten från en analys som är identisk med den i tabell III.2, utom att faktorerna stad och åldersgrupp saknas som fixa faktorer. Informanterna är ordnade efter stigande slumpintercept, så alla eventuella mönster som uppstår som har med stad, ålder eller kön att göra är bara ett resultat av slumpintercepten. I denna analys har även informanterna från Mariehamn inkluderats.

Tabell III.2. Slumpintercept för regressionsmodellen i tabell III.1.

Hf.ym06	-0,319	Ab.ym20	-0,408	Va.yw10	-0,181
Hf.ym01	-0,192	Ab.yw18	-0,381	Va.yw11	-0,163
Hf.ym04	-0,159	Ab.yw19	-0,062	Va.ym10	0,155
Hf.yw10	-0,069	Ab.yw11	-0,011	Va.ym06	0,223
Hf.ym08	0,063	Ab.ym19	0,056		
Hf.yw22	0,082	Ab.yw15	0,125		
Hf.ym03	0,164	Ab.ym13	0,310		
Hf.yw09	0,188	Ab.ym12	0,342		
Hf.yw04	0,239				
Hf.ow01	-0,399	Ab.om04	-0,165	Va.om01	-0,102
Hf.ow05	-0,165	Ab.ow05	-0,015	Va.ow01	-0,099
Hf.om03	-0,157	Ab.ow02	0,029	Va.ow05	-0,033
Hf.ow03	-0,031	Ab.om02	0,070	Va.ow02	-0,009
Hf.ow10	0,041	Ab.ow01	0,109	Va.om06	0,033
Hf.om07	0,074			Va.om04	0,176
Hf.om02	0,269				
Hf.om01	0,372				

De flesta informanterna från Helsingfors finns högst upp i tabell III.3, medan informanterna från Vasa finns längst ner. De äldre informanterna från Åbo finns bland Helsingforsinformanterna uppe i tabellen, medan en del av de unga Åboborna återfinns långt nere i tabellen. Detta återspeglar resultaten av de fixa variablerna i den ursprungliga modellen. Det intressanta med tabell III.3 är de informanter som avviker från mönstret.

Tabell III.2 visar att det verkar finnas könsbunden variation. I Helsingfors har kvinnorna de lägsta V:/C-värdena bland de äldre informanterna, men bland de yngre är det tvärtom. Värdena är jämförbara bara inom varje grupp i tabell III.2, men i tabell III.3 ser man att yngre män och äldre kvinnor från Helsingfors har liknande värden. Bland de yngre Åboinformanterna finns främst kvinnor högst upp i tabell III.2 medan de flesta män finns längre ner. Söker man efter dessa informan-

Tabell III.3. Slumpintercept för en regressionsmodell utan stad och ålder som fixa effekter.

STAD	ÅLDER	KÖN	INFORMANTNR	INTERCEPT
Helsingfors	yngre	man	6	-0,523
Helsingfors	äldre	kvinn	1	-0,505
Åbo	äldre	man	4	-0,405
Helsingfors	yngre	man	1	-0,404
Mariehamn	yngre	man	11	-0,364
Helsingfors	yngre	man	4	-0,361
Åbo	yngre	man	20	-0,314
Helsingfors	yngre	kvinn	10	-0,277
Helsingfors	äldre	kvinn	5	-0,275
Helsingfors	äldre	man	3	-0,264
Åbo	äldre	kvinn	5	-0,238
Åbo	yngre	kvinn	18	-0,229
Åbo	äldre	kvinn	2	-0,222
Åbo	äldre	man	2	-0,156
Helsingfors	äldre	kvinn	3	-0,126
Helsingfors	yngre	man	8	-0,121
Helsingfors	yngre	kvinn	22	-0,121
Mariehamn	yngre	man	3	-0,118
Åbo	äldre	kvinn	1	-0,108
Helsingfors	äldre	kvinn	10	-0,051
Helsingfors	yngre	man	3	-0,037
Helsingfors	äldre	man	7	-0,025
Helsingfors	yngre	kvinn	9	-0,012
Mariehamn	yngre	kvinn	8	0,004
Mariehamn	äldre	man	1	0,067
Åbo	yngre	kvinn	19	0,084
Vasa	yngre	kvinn	10	0,107
Vasa	yngre	kvinn	11	0,118
Helsingfors	yngre	kvinn	4	0,130
Åbo	yngre	kvinn	11	0,137
Vasa	äldre	man	1	0,142
Vasa	äldre	kvinn	1	0,144

STAD	ÅLDER	KÖN	INFORMANTNR	INTERCEPT
Mariehamn	äldre	kvinnna	1	0,153
Mariehamn	yngre	kvinnna	10	0,172
Helsingfors	äldre	man	2	0,181
Åbo	yngre	man	19	0,203
Vasa	äldre	kvinnna	5	0,210
Vasa	äldre	kvinnna	2	0,242
Åbo	yngre	kvinnna	15	0,274
Vasa	äldre	man	6	0,277
Helsingfors	äldre	man	1	0,290
Vasa	äldre	man	4	0,416
Vasa	yngre	man	10	0,450
Åbo	yngre	man	13	0,467
Vasa	yngre	man	6	0,489
Åbo	yngre	man	12	0,502

ter i tabell III.3, återfinns de unga männen från Åbo längst nere bland Vasainformanterna, medan kvinnorna finns mer i mitten av tabellen. De unga kvinnorna från Åbo och Helsingfors har de facto ganska liknande värden, medan det är de unga männen från dessa två städer som gör att grupperna som helhet skiljer sig så mycket. De unga männen från Helsingfors har mycket låga V:/C-värden medan de unga männen från Åbo har höga V:/C-värden. Endast en av de unga männen från Åbo (Ab.y_{m20}) avviker klart från detta mönster.

De informanter som har de lägsta värdena i tabell III.3 är Hf.y_{m06}, Hf.o_{w01}, Ab.o_{m04} och Hf.y_{m01}. Av dessa har Hf.o_{w01} och Ab.o_{m04} universitetsutbildning och har arbetat med undervisning i svenska. Hf.y_{m01} är utbildad skådespelare. Alla tre har alltså arbetat på en professionell nivå med svenska språket och är troligen mycket medvetna om sitt uttal. De låga V:/C-värdena (dvs. lång konsonant i förhållande till vokalen) hos dessa talare reflekterar antagligen ett uttal som (åtminstone tidigare?) setts som prestigeformen. Hf.y_{m06} har däremot

en helt annan bakgrund. Han har den lägsta utbildningsnivån av alla unga Helsingforsinformeranter (yrkesskola) och hör därmed inte till den grupp där man normalt skulle förvänta sig ett prestigeuttal.

En av de äldre Helsingforsborna (Hf.omo1) finns bland Vasainformeranterna långt nere i tabell III.3. Han har universitetsutbildning och man skulle därför kunna förvänta sig ett lägre V:/C-värde. Men han är uppvuxen i östra Helsingfors, som till stora delar bestod av landsbygd under informantens uppväxttid, vilket antagligen förklarar hans uttal. Den av de äldre kvinnorna i Helsingfors som har det högsta V:/C-värdet har den lägsta utbildningsnivån i gruppen.

Informanterna från Vasa och de äldre från Åbo är så få att det är svårt att dra slutsatser av materialet. Man kan ändå konstatera att det verkar finnas en korrelation mellan V:/C-värdet och utbildning bland de äldre Åboinformeranterna, så att talare med den högsta utbildningsnivån också har de lägsta V:/C-värdena.

Av de yngre Vasainformeranterna uttrycker den som har det högsta V:/C-värdet en mycket stark tillhörighet till Vasa under intervjun och negativa attityder gentemot Helsingfors:

ja vill int flytt di:t fö ja tykkä int åm han sta:n (Va.ymo6)

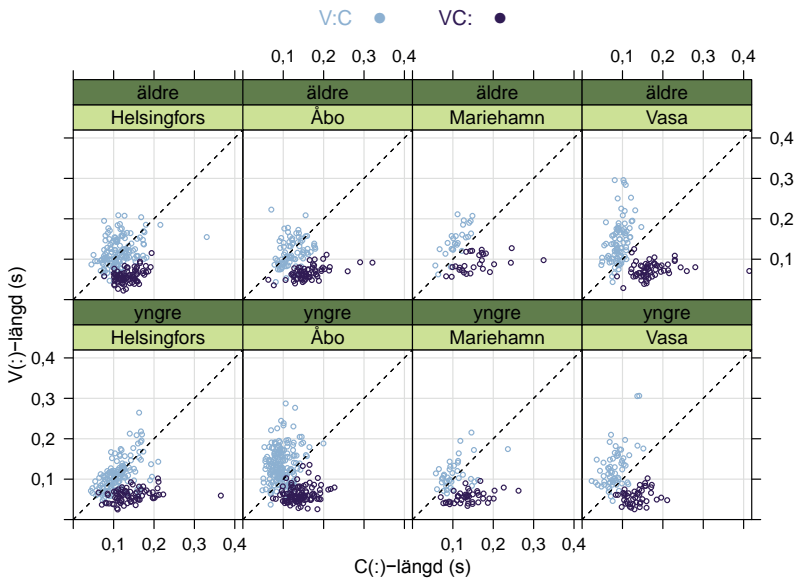
jag vill inte flytta dit för jag tycker inte om den staden

Det höga V:/C-värdet skulle här kunna tolkas som en identitetsmarkör som används för att understryka gränsdragningen mot Helsingforssvenskan. Jämför Larsens (1917) term *naboopposition*, som innebär att det lokala språkets särdrag överdrivs för att markera en klar skillnad mot "de andras" språk.

För att kunna dra närmare slutsatser om faktorerna kön och socialgrupp borde dessa testas kvantitativt med ett större material än vad som var tillgängligt för denna studie. Den kvalitativa analysen av slumpintercepten antyder att dessa faktorer, som Reuter (1982) fann vara signifikanta för konsonantlängd i Helsingfors på 1970-talet, spelar roll också på 2000-talet. Den största skillnaden i förhållande till Reuters resultat är att unga män i Helsingfors har lägre V:/C-värde än unga kvinnor. De äldre Helsingforsinformeranterna betar sig språkligt mer i enlighet med Reuters resultat.

5 Distinktionen V:C–VC:

Så som figurerna III.1 och III.2 visar har vissa talare längre konsonant än vokal i fonologiska V:C-sekvenser. Den fråga som inställer sig är då hur V:C-sekvenserna skiljs åt från VC:-sekvenserna. I motsats till mellansvenskt standardspråk⁹ har finlandssvenskan bara minimala kvalitativa skillnader mellan långa och korta vokaler, vilket gör att kvantitet är den enda ledtråden för att skilja mellan långa och korta segment (Reuter 1992, 108). Man förväntar sig alltså längdskillnader mellan V:C- och VC:-sekvenser.



Figur III.3. Punktdiagram över spridningen av vokallängd och konsonantlängd i V:C-sekvenser (ljusblått) resp. VC:-sekvenser (mörkblått) i de olika informantgrupperna. Mellan grupperna finns tydliga skillnader i hur väl de båda sekvenstyperna separeras i fråga om kvantitet.

9. I mellansvenskt standardspråk finns kvalitativa skillnader mellan långa och korta vokaler utöver de kvantitativa skillnaderna. Vokaler med stor skillnad i vokalkvalitet mellan den långa och den korta motsvarigheten är t.ex. *a* och *u*. Trots detta verkar kvantitet vara den viktigaste ledtråden för den fonematiska skillnaden, och vokalkvalitet är en tilläggsledtråd som finns för vissa vokalar men inte för alla (Hadding-Koch & Abramson 1964; Behne et al. 1996).

Tabell III.4. Somers D_{yy} -värden för fem olika regressionsmodeller körda separat för sju informantgrupper. Värdet anger korrelation mellan modellerad sannolikhet och observerade värden. 0 = slumpmässigt samband, 1 = perfekt korrelation.

STAD	ÅLDERSGRUPP	ANTAL BELÄGG				OBEROENDE VARIABEL/VARIABLER					
		TOTALT	V:C	VC:	VC:	V(:)-LÄNGD	C(:)-LÄNGD	V(:)C(:)-LÄNGD	V(:)/C(:)-VÄRDE	V(:)-LÄNGD o. C(:)-LÄNGD	
Helsingfors	äldre	220	141	79	79	0,954	0,606	0,465	0,990	0,993	
	yngre	209	136	73	73	0,886	0,560	0,396	0,962	0,963	
Åbo	äldre	126	82	44	44	0,902	0,590	0,302	0,966	0,968	
	yngre	286	183	103	103	0,930	0,880	0,323	0,991	0,997	
Mariehamn	yngre	84	39	45	45	0,969	0,578	0,201	0,993	0,990	
Vasa	äldre	155	93	62	62	0,905	0,978	0,344	0,999	1,000	
	yngre	109	63	46	46	0,936	0,875	0,295	1,000	1,000	

Figur III.3 visar spridningsdiagram där både V:C- och VC:-sekvenser ingår. Figuren är uppdelad efter stad och åldersgrupp och man kan se tydliga skillnader mellan grupperna. De äldre informanterna från Vasa gör den allra tydligaste distinktionen mellan V:C och VC:, medan de båda sekvenstyperna överlappar varandra betydligt mer i diagrammen över de övriga informantgrupperna. De äldre Vasainformanterna använder såväl vokallängd som konsonantlängd för att skilja åt sekvenstyperna, medan bara vokallängden verkar skilja sig bland de äldre från Helsingfors och Åbo.

För att testa vilka akustiska ledtrådar som används för att skilja de båda sekvenstyperna åt gjordes ett antal logistiska regressionsmodeller med blandade effekter. Analyserna kördes separat för de olika informantgrupperna. Den beroende variabeln var densamma i alla analyser, nämligen den kategoriska variabeln sekvenstyp med de två kategorierna V:C och VC:. I alla analyser fanns också en stokastisk effekt för informant. Fem olika modeller med olika oberoende variabler testades:

1. vokallängd
2. konsonantlängd
3. sekvenslängd, dvs. den totala längden vokal + konsonant
4. det relativa värdet vokal/konsonant
5. segmentlängder, dvs. två oberoende variabler: vokallängd och konsonantlängd.

Som mått på anpassningsgraden, alltså hur väl modellen förklarar observationerna, användes Somers D_{xy} (jfr Baayen 2008, 204, 281). Resultaten, som kan ses i tabell III.4, visar att vokallängd är en mycket god förklarande variabel för sekvenstyp i alla informantgrupper. Konsonantlängd är en dålig ledtråd för sekvenstyp i Helsingfors och bland äldre talare i Åbo och yngre i Mariehamn¹⁰, men en bra ledtråd i Vasa och bland yngre Åboinformanter.

Sekvenslängd är överlag en dålig ledtråd, vilket innebär att det finns en tendens till att hålla den totala längden av kombinationen vokal-konsonant konstant oberoende av sekvenstyp. Sekvenslängd fungerar

10. De äldre talarna från Mariehamn ingår inte i analysen eftersom de är bara två till antalet.

något bättre som ledtråd i Helsingfors än på de andra orterna. Bland i synnerhet äldre Helsingforsinformanter är V:C-sekvenser aningen längre än VC:-sekvenser, på grund av förlängd konsonant.

För alla informantgrupper ger det relativa värdet $V(:)/C(:)$ en något högre anpassningsgrad än när bara vokallängd används som oberoende variabel. Modeller med två oberoende variabler (vokallängd och konsonantlängd) är lika bra som modeller med ett enda relativt värde (vokallängd/konsonantlängd), vilket tyder på att det är det relativa värdet som är den viktigaste ledtråden. För Vasatalarna ger både modellen med relativt värde och modellen med två oberoende variabler en perfekt anpassningsgrad.

6 Diskussion

Resultaten visar på uppenbara skillnader i fonetisk kvantitet mellan de fyra undersökta städerna. I Vasa är V:/C-värdet klart högre än i Helsingfors. I Åbo har de äldre talarna och vissa av de yngre talarna värden som motsvarar dem från Helsingfors, medan vissa yngre talare har värden som påminner mer om dem från Vasa. De yngre talarna från Mariehamn har värden som motsvarar Helsingforsvärdena.

Distinktionen mellan V:C- och VC:-sekvenser signaleras huvudsakligen med hjälp av vokallängd. Vokallängd används för att markera den fonologiska skillnaden av alla undersökta talargrupper, medan konsonantlängd bara används av en del talare. Detta stämmer väl överens med Schaefflers (2005, 125) analys av kvantitet i svenska dialekter.

Enligt Schaefflers (2005) typologi (se avsnitt 2) kan det finlandssvenska standardduttalet sägas ha ett 3-vägssystem för kvantitet (dvs. VC-, V:C- och VC:-sekvenser i betonade stavelser), vilket skulle förutsäga tydlig kontrast i vokallängd (högt V:/V-värde), lång C:-längd och ingen kontrast för konsonantlängd. Precis detta kan man iaktta i Helsingfors, Mariehamn och bland i synnerhet äldre i Åbo: vokallängd används för att upprätthålla kontrasten V:C-VC:, medan konsonantlängd inte används för att signalera kontrasten.

Kontrasten mellan VC- och VC:-sekvenser studerades inte närmare i materialet, vilket gör att inga konklusioner kan dras om på vilket sätt kortstavigt uttal (VC) hålls isär från VC:. Reuters (1982, 173–175) mätningar antyder att V:C och VC: i medeltal är mer än dubbelt så

långa som VC, vilket kan tolkas som att totallängden fungerar som fonetisk signal.

Vasa har enligt Schaefflers teori ett överspecificerat system (Schaeffler 2005, 126) eftersom både vokal- och konsonantlängd skiljs väl åt trots att systemet saknar överlånga stavelser. En förklaring till detta kan vara att Vasa omringas av dialekter som har eller intill nyligen haft 4-vägssystem. Den fonetiska anpassningen till 3-vägssystemet kan enligt Schaeffler ta flera generationer av talare i anspråk.

Materialet från Åland för denna studie var mycket litet, i synnerhet vad gäller den äldre generationen. Från Åland finns däremot tidigare observationer som rimmar väl med den typologiska förklaringsmodellen. Sundberg (1993, 131) rapporterar uttal av typen [ro:ppa] *ropa* från Brändö. De åländska dialekterna saknar överlånga stavelser (Ivars 1988, 67), vilket enligt Schaeffler leder till konsonantförlängning i V:C-sekvenser.

Bland de äldre informanterna från Helsingfors och Åbo kan man skönja en tendens till de sociolingvistiska mönster som beskrivits av tidigare forskare. V:/C-värdet är lägre för kvinnor än för män och det verkar finnas ett visst samband mellan utbildningsgrad och uttal, så att hög utbildningsnivå innebär lågt V:/C-värde. De yngre informanterna från de båda städerna följer däremot inte detta mönster. I Åbo har i synnerhet de unga männen mycket höga V:/C-värden som motsvarar värdena från Vasa. I Helsingfors har de unga männen lägre V:/C-värden än kvinnorna.

Resultaten från Åbo är typologiskt sett de mest oväntade om man följer Schaefflers (2005) resonemang. Om man tolkar variationen mellan äldre och yngre informanter från Åbo som en språkförändring i skenbar tid, kan man säga att förändringen går från ett mer ekonomiskt system (där endast vokallängd används för att signalera den fonologiska kontrasten) till ett överspecificerat system. Förklaringen till språkförändringen ska i så fall kanske sökas på det sociolingvistiska snarare än på det typologiska planet. En tänkbar förklaring kunde vara inflyttning från Österbotten och den sydfinländska skärgården och landsbygden till Åbo. Ett annat alternativ kunde vara att det uttal som tidigare förknippats med den s.k. bildade Helsingfors- och Åbosvenskan har kommit att få en allt starkare association till bara Helsingforsregionen och kommit att stigmatiseras på övriga håll i landet. Detta understöds

av en perceptionsstudie utförd av Haapamäki & Wikner (u.u.) som visar att lyssnare som identifierar förlängda konsonanter i tal utgår från att talaren kommer från huvudstadsregionen. Haapamäki & Wikner konstaterar att detta är ett drag som kan betraktas som ett platsindex för Helsingfors(regionen) och som i vissa fall genererar negativa attityder.

Analysen av taltempots inverkan på V:/C-värdet visar att några av informanterna varierar konsonantlängden i V:C-sekvenser mer än vokallängden när taltempot ändras. Detta går stick i stäv med de resultat som Heldner & Strangert (2001) och Pind (1999) redovisat för svenska respektive isländska, där det är det fonologiskt långa elementet som förlängs mest när taltempot minskar, vilket förstärker kontrasten såväl mellan långa och korta segment som mellan V:C- och VC:-sekvenser. Om Helsingforssvenska V:C-sekvenser har avvikit från centralsvenska V:C-sekvenser i fråga om fonetisk kvantitet är det möjligen på detta plan skillnaderna står att finna; dvs. vid normalt taltempo och normal prominens är skillnaderna mellan Helsingforssvenska och centralsvenska knappast stora, men i starkt framhävd ställning där taltempot är långsamt kan konsonanterna i Helsingforssvenskan vara markant långa. Denna typ av konsonantförlängning finns i det undersökta materialet främst hos en del av de äldre informanterna, medan de yngre i större utsträckning använder ett uttal där både vokalen och konsonanten förlängs vid långsammare taltempo.

Språkförändring drivs i moderna västerländska samhällen oftast av unga kvinnor snarare än av unga män (Chambers & Trudgill 1998, 62). Om man antar att det är de unga kvinnorna som visar vägen för språkförändringen kunde man kanske dra slutsatsen att det uttal som har högst prestige i början av 2000-talet är ett relativt ”neutralt” uttal. De unga kvinnorna i de olika städerna uppvisar inte lika extrema V:/C-värden som andra informanter inom samma stad. Det undersökta samplet är naturligtvis mycket litet för långtgående slutsatser av detta slag, men teorin om att ett relativt neutralt uttal håller på att breda ut sig understöds av utdraget i avsnitt 2 samt av Stenberg-Sirén & Östman (2012) och Stenius (2012) som har noterat att användningen av de extremt låga V:/C-värdena som typiskt associeras med Helsingfors har minskat i radiospråket.

Som nämnades i introduktionen har kvantitet setts som en av de tydligaste sociala markörerna i svenskan i Finland. Höckerstedt (2010, 61)

menar att kontakten med kvantitetspråket finska stödjer finlandssvenskarnas förmåga att höra olika grader av segmentlängd. Man kunde också säga att förlusten av överlånga stavelser har skapat utrymme för att använda konsonantlängden i V:C-sekvenser som social markör eftersom den inte längre behövdes för den fonologiska kontrasten. De starka sociala associationer som segmentlängd haft och har i kombination med talarnas perceptuella sensitivitet för kvantitetsskillnader kan eventuellt vara de faktorer som lett till att ett mer neutralt uttal kanske håller på att bli vanligt. Om man varken vill låta "snobbig" eller "bonnig" väljer man helt enkelt ett uttal mitt emellan de två ytterligheterna. Huruvida detta faktiskt är den riktning språkutvecklingen håller på att ta och hur finlandssvenskan i så fall kommer att positionera sig i förhållande till sverigesvenska i fråga om kvantitet återstår att se. Flera attitydstudier och lyssnarstudier skulle säkerligen kunna kasta mera ljus över hur de sociala associationer variabeln har påverkar språkförändringen.

Del IV

ORDKLASSTAGGNING AV
FINLANDSSVENSKT TALSPRÅK

1 Introduktion

En korpus är en samling texter som representerar en eller flera genrer vid en eller flera tidpunkter. Genom att använda korpusar kan språkforskare studera hur språk används i sin naturliga kontext. Stora korpusar började byggas redan under 1960-talet, men från 1980-talet och framåt tog korpuslingvistikens fart som forskningsområde allteftersom datorerna möjliggjorde snabb bearbetning av stora mängder elektroniska texter. Korpusar förenklar arbetet för språkforskare eftersom det går mycket snabbt att hitta exempel på användningen av ord och fraser. Man kan också enkelt få fram uppgifter om hur vanligt ett ord eller uttryck är och göra komplicerade analyser där många olika språkliga faktorer beaktas på en och samma gång.

Korpusar över svenskt skriftspråk finns samlade hos Språkbanken i Göteborg i konkordansverket Korp, som för tillfället omfattar 1,7 miljarder token¹ av ett flertal genrer från tidningstext och romaner till blogg- och Twittertexter. Finlandssvenska texter bildar en egen delkorpus² i Korp om mer än 70 miljoner token.

De första stora korpusarna var skriftspråkskorpusar, men efterhand har även stora talspråkskorpusar byggts upp. Utmaningarna vid uppbyggandet av talspråkskorpusar är betydligt större, eftersom talet först måste omvandlas till någon slags skriven, maskinläsbar form för att kunna bearbetas och göras sökbart. I stora språkområden har man i regel haft tillgång till stora resurser för att bygga upp forskningsinfrastruktur, inklusive talspråkskorpusar. Samtidigt vet vi att språk som talas av ett stort antal personer typologiskt ofta saknar drag som man framför allt hittar i små språkgemenskaper (Trudgill 2011). Stora kulturbärande språk har dessutom gått igenom långa processer av språklig standardisering. Om den korpusbaserade språkforskningen koncentreras till

1. Ett token är den minsta enheten som används vid annoteringen av en korpus. Token omges i regel av blanksteg eller av blanksteg och skiljetecken. Ett grafiskt ord eller ett tal utgör ett token. Men även skiljetecken kan räknas som token, vilket innebär att antalet token i en text ofta är större än antalet ord. I Talko utgör t.ex. även paustecknen token.
2. Denna delkorpus har skapats inom projektet *Svenskan i Finland – i dag och i går* som ett samarbete mellan Svenska litteratursällskapet i Finland, Institutet för de inhemska språken och Göteborgs universitet.

stora språk, riskerar man alltså att få en förenklad bild av den språkliga mångfalden. Därför är det mycket viktigt att korpusar, och i synnerhet talspråkskorpusar, görs också över mindre språk och dialekter.

Vid Svenska litteratursällskapet i Finland pågår arbetet med att bygga upp den finlandssvenska talspråkskorpusen Talko. Materialet för korpusen utgörs i första hand av inspelningar från *Spara talet*-projektet. Målet är att talspråkskorpusen inte enbart ska bestå av sökbara transkriptioner kopplade till ljudfiler, utan att dessa dessutom ska vara lingvistiskt annoterade med ordklasstaggar och viss morfologisk information samt information om grundform.

Lingvistisk taggning ökar korpusens värde för forskningen genom att det utöver sökning på bokstavliga ordformer blir möjligt att begränsa sökningar till en viss ordklass (t.ex. endast substantivet *blomma* och inte verbet *blomma*) eller att göra bredare sökningar som inkluderar en hel ordklass. När arbetet med delarna I–III i denna volym utfördes hade den lingvistiska annoteringen av *Spara talet*-materialet ännu inte inletts. Det innebar att systematiska maskinella sökningar i materialet kunde utföras utifrån uttalsform och ortografisk form men inte utifrån t.ex. ordklass. Om en validerad ordklasstagning hade varit tillgänglig hade bl.a. excerperingen av infinitivmärket gått mycket snabbare, eftersom man hade kunnat söka direkt på taggen för infinitivmärke i inspelningarna i stället för att söka efter orden *att* och *till* och manuellt utesluta användning som subjunktion, preposition eller adverb. Även analysen av kontextens betydelse för formvalet, t.ex. valet av infinitivmärke (I.3.4.1) eller infinitivens form (I.3.4.2), hade kunnat göras både effektivare och mer tillförlitligt om det varit möjligt att göra en analys av kontexten utifrån ordklass i stället för utifrån ortografisk form.

En korpus utgör alltså en forskningsinfrastruktur som olika forskare kan utnyttja på olika sätt. Enskilda forskares arbete underlättas och effektiveras eftersom varje forskare inte behöver börja från början med att samla in och bearbeta språkmaterial, och på så vis får man snabbare fram forskningsresultat. Genom sökningar i korpusen Talko kan man snabbt få svar på frågor som *Vilka uttal av ordet mycket förekommer på olika orter i Svenskfinland?* eller *Vilka former av infinitivmärket används på olika håll?* Det är ändå självklart att en korpus inte kan vara uppbyggd för att fullt ut fylla alla forskares behov, utan den utgör en basresurs som forskare kan utnyttja och arbeta vidare med på olika sätt.

De olika delarna av korpusbygget beskrivs översiktligt av Södergård & Leinonen (u.u.). I den här artikeln beskrivs den lingvistiska annoteringen av Talko, som är påbörjad men fortfarande kräver vidare utveckling. Vid annoteringen av finlandssvenskt talspråksmaterial går det att utnyttja resurser som finns för svenskt skriftspråk. Här beskrivs hur en taggare baserad på svenskt skriftspråk kan tillämpas på finlandssvenskt talspråk och hur taggningsresultatet gradvis förbättras så att man så småningom kan uppnå ett resultat som är tillfredsställande för analysen av talspråk.

2 Automatisk ordklasstagning

Lingvistisk annotering kan ske automatiskt med olika typer av statistiska eller regelbaserade taggare. En taggare är en typ av programvara som läser in en text och förser varje token med lingvistisk information. Regelbaserade taggare kräver betydligt mer arbete än statistiska taggare, eftersom reglerna måste formuleras för hand. I en talspråkskorpus med många olika dialekter skulle antagligen olika regler dessutom krävas för olika dialekter. För *Spara talet*-korpusen lämpar sig därför statistisk taggning bäst.

Statistisk taggning kräver två moment: träning och taggning. I träningskedet lär sig taggaren sannolikheter och frekvenser för det språk som ska taggas. Taggaren behöver alltså ett korrekt taggat träningsmaterial, och för att resultatet ska bli bra måste sannolikheterna i träningsmaterialet motsvaras av sannolikheterna i det material som ska taggas automatiskt. Helst ska träningsmaterialet därför bestå av samma språkliga varietet och samma genre som det material som ska taggas automatiskt. Vid den automatiska taggningen använder sig taggaren sedan av de sannolikheter den lärt sig av träningsmaterialet för att välja ordklasstagning för varje ord i materialet.

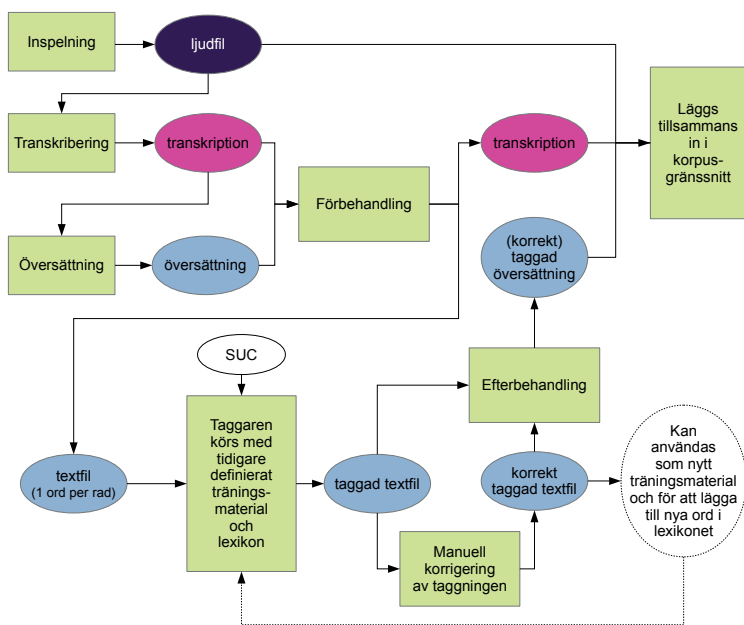
Automatisk taggning ger alltid en viss felprocent. Ju mer manuell taggat träningsmaterial som finns, desto lägre blir i allmänhet felprocenten vid automatisk taggning. Hur mycket handkorrigerat träningsmaterial som krävs för att man ska uppnå en hög korrekthet beror bl.a. på tagguppsättningens storlek. Vanligen rapporteras en uppnådd taggningsnoggrannhet på ca 95–97 procent (Manning & Schütze 1999, 371).

Flera olika fritt tillgängliga statistiska ordklasstagare finns till förfo-

gande. För taggningen av den finlandssvenska talspråkskorpusen valdes TreeTagger (Schmid 1994, 1995), som bl.a. också använts för taggning av norskt talspråk efter en evaluering av olika statistiska taggare (Søfteland & Nøklestad 2008).

3 Arbetsmoment och hjälpmedel

Vägen från dialektinspelningar till en ordklasstaggad talspråkskorpus har flera olika arbetsmoment. Figur IV.1 ger en schematisk bild över de olika arbetsmomenten. Första steget vid uppbyggandet av talspråkskorpusen har varit att transkribera dialektinspelningarna med en ljudenlig transkription. Detta har gjorts i programmet Transcriber (Barras et al. 2000). Transkriptionen har därefter ”översatts” halvautomatiskt ord för



Figur IV.1. Arbetsflöde. Den ljudenliga transkriptionen (rosa) och den ortografiska transkriptionen (ljusblå) följer var sin väg genom diagrammet och återförenas på slutet när de tillsammans med ljudfilen läggs in i ett gränssnitt. De gröna lådorna i figuren står för olika arbetsmoment. Eftersom det är den ortografiska översättningen som taggas har denna en längre väg genom figuren än vad den ljudenliga transkriptionen har.

ord till standardsvensk ortografi med hjälp av en dialekttranslitterator utvecklad vid Tekstlaboratoriet vid Universitetet i Oslo. Arbetet har skötts vid Språkarkivet vid SLS och riktlinjer har utvecklats för att transkriptionerna och översättningarna ska vara så konsekvent utformade som möjligt i korpusen. För korpusen har hittills sammanlagt ca 100 timmar tal transkriberats med en grov ljudskrift.

Översättningen till standardiserad ortografi är en förutsättning för ordklasstagningen. Ljudenliga transkriptioner innehåller för mycket variation för att statistisk taggning ska kunna tillämpas. Huvudprincipen vid översättningen har varit att det handlar om en ortografisk normalisering. En ord-för-ord-princip följs och det är inte meningen att betydelsen ska översättas utan det är bara stavningen (och vissa böjningsformer) som normaliseras.

Nedan beskrivs ordklasstagningen närmare. När materialet är taggat kan ljudfilen tillsammans med den ljudenliga transkriptionen och översättningen inklusive ordklasstaggar läggas in i ett korpusgränssnitt som möjliggör sökning i korpusen (sista steget i figur IV.1). Ljudfilen är genom tidsstämpling kopplad till transkriptionerna.

3.1 För- och efterbehandling av materialet

Före taggningen förbehandlades transkriptionerna för att få ett format som den statistiska taggaren kan läsa. Bl.a. krävs att varje token står på en egen rad i ett textdokument i inputfilen. I samband med förbehandlingen gjordes också flera olika automatiska kontroller av materialet, bl.a. kontroll av konsekvens vid standardiseringen av vissa dialektord.

Efter taggningen bearbetas materialet så att det taggade materialet återfår det format som används i korpusen. För- och efterbehandlingen gjordes med skript i programmeringsspråket Perl.

3.2 SUC

Stockholm-Umeå-korpusen (SUC; Gustafson-Capková & Hartmann 2006) är en manuellt taggad korpus som används som standard vid träning och evaluering av ordklasstagare på svenska. Korpusen består av ca 1 miljon ord svenskt skriftspråk av olika genrer från 1990-talet. Bl.a. det ordklasstagade material som ingår i Korp, inklusive det finlands-

svenska skriftspråksmaterialet, är taggat med SUC som träningsmaterial. Även den taggupsättning som används i SUC (Ejerhed et al. 1992) har blivit en standard för ordklasstagning på svenska.

Taggningen av *Spara talet*-materialet inleddes med att TreeTagger tränades på SUC. Den version som användes var SUC 3³. Av SUC 3 byggdes även en lexikonfil med det format som TreeTagger kräver.

När det gäller att tagga finlandssvenskt talspråk finns det två avsevärda skillnader i förhållande till SUC, som endast består av skriftspråk och av material från Sverige. I talspråk är meningar och satser inte alltid hela som i skriftspråket, utan olika typer av avbrott och upprepningar är vanliga och tydliga satsgränser i form av interpunktion saknas. Dessutom förekommer det ett stort antal diskurspartiklar, t.ex. uppbackningar och ansatsljud, som en taggare tränad på skriftspråk inte känner till (jfr Lindström 2008). En taggare tränad på skriftspråk kan ändå användas som ett första steg vid taggning av talspråk. Søfteland & Nøklestad (2008) rapporterar att en taggare tränad på norskt skriftspråk användes vid taggning av talspråk från Oslo. Taggningen handkorrigerades och användes därefter som nytt träningsmaterial, vilket gjorde att man till slut kunde uppnå en korrekthet på 96,9 procent vid automatisk taggning med TreeTagger. Vid taggningen av *Spara talet*-korpussen används samma tillvägagångssätt.

Vid en provtaggning av *Spara talet*-material med TreeTagger tränad på SUC kunde genast flera svagheter identifieras. Dels handlar det om ord som används på annat sätt i tal än i skrift, dels om skillnader mellan språkbruket i Sverige och Finland. Ordet *före* finns t.ex. endast som preposition i SUC-lexikonet men används ofta som subjunktion i *Spara talet*-materialet, medan *till* används som infinitivmärke i de finlandssvenska dialekterna vilket inte förekommer i skriftspråk. Ytterligare ett exempel är *nå*, som är en frekvent diskurspartikel i finlandssvenskt talspråk, men som tolkas som ett verb av taggaren tränad på skriftspråk. På grund av denna typ av skillnader måste annoteringen av en del av den finlandssvenska korpussen handkorrigeras och därefter användas som träningsmaterial vid taggning av den resterande korpussen. Endast så kan man uppnå ett tillfredsställande annoteringsresultat som motsvarar forskningens behov.

3. Fritt tillgänglig på Språkbankens webbplats i XML-format.

3.3 Tagguppsättningen

SUC-tagguppsättningen finns väl dokumenterad i en manual (Ejerhed et al. 1992). Taggarna består av ordklass och viss morfologisk information. Vid substantiv, adjektiv och pronomen anges genus, numerus, species och kasus, för adjektivens del dessutom komparationsgrad. Verben taggas för tempus och aktiv/passiv. Taggarna består av förkortningar som skrivs med kapitaler. En ordform som t.ex. *springer* får taggen VB.PRS.AKT, som står för verb–presens–aktiv.

Två tillägg gjordes till SUC-taggarna för Talko: DI och PAUS. Bägge markerar sådant som saknas i skriftspråk. Taggen PAUS används för paustecken av olika längd, medan DI-taggen behövs för olika typer av diskurspartiklar och tvekljud (jfr Nivre & Grönqvist 2001 och Rehbein & Schalowski 2013).

3.4 Automatisk taggning och handkorrigering

Taggning med TreeTagger innebär som nämnts två moment: träning och taggning. I inledningsskedet tränades TreeTagger med enbart SUC som träningsmaterial. Därefter taggades en av de finlandssvenska transkriptionerna och taggningsresultatet korrigerades för hand. Som hjälpmedel vid korrigeringen av den automatiska taggningen hade vi tillgång till en webbapplikation utvecklad vid Tekstlaboratoriet i Oslo.

Efter att en transkription handkorrigerats kan denna läggas till i träningsmaterialet för att förbättra taggningsresultatet för följande fil som ska taggas. Det nya materialet läggs till både i träningsmaterialet och i lexikonet. Proceduren upprepas för varje ny fil som har taggats och på så vis förbättras taggningsresultatet steg för steg.

4 Taggningsresultat

Den fråga som inställer sig vid automatisk taggning är hur bra resultatet blir. Noggrannheten brukar anges som den procentuella andelen korrekt taggade token. I token inkluderas som nämnt förutom ord också t.ex. interpunktion, eller i vårt fall bl.a. paustecken.

Hittills har transkriptionerna av åtta inspelningar ur korpussen handkorrigerats. Vid varje transkription handlar det om ett ca 20 minuter

långt klipp ur intervjun (som totalt varade ca 60 minuter), och antalet token per transkription är i medeltal drygt 3 200.

En utgångskorrektet fås genom att se hur hög korrektet är när det finlandssvenska materialet taggas med TreeTagger tränad enbart på SUC. Genom att jämföra den automatiska taggningen med den korrekta taggningen (dvs. den handkorrigerade) får man fram korrektet som i detta fall är 78,8 procent. Mer än en femtedel av alla token blir alltså fel taggade när man kör en automatisk taggning och taggaren är tränad endast på SUC.

Vid utvärderingen av taggningsresultat brukar man skilja mellan kända ord och okända ord, dvs. ord (eller token) som finns i det lexikon taggaren har tillgång till genom träningsmaterialet och ord som inte finns i träningsmaterialet. Andelen okända token i *Spara talet*-filerna är relativt hög (13,7 %) när taggaren har tränats enbart på SUC. Detta beror främst på ett stort antal paustecken, tvekljud, uppbackningar osv. i talspråksmaterialet som helt saknas i SUC. Bland okända ord finns också bl.a. ortnamn i Svenskfinland, dialektord och finska ord, men också ”vanliga svenska ord” som bara inte råkar finnas belagda i SUC-korpusen.

Korrektet för okända token är så låg som 20,4 procent, vilket har att göra med att taggarna PAUS och DI inte alls används i SUC. Taggaren har därmed inte någon som helst träning i hur dessa taggar används, och alla pauser, tvekljud etc. blir inledningsvis konsekvent feltaggade. Taggningsresultatet förväntas bli betydligt bättre när taggaren utöver SUC kan tränas också på *Spara talet*-material. Korrektet för kända token är 88,1 procent när taggaren har tränats på SUC. Även här finns alltså stort utrymme för förbättring.

4.1 Korsvalidering

Evaluering av ordklasstagare sker ofta genom s.k. tiodelad korsvalidering. Detta innebär att det handkorrigerade materialet delas upp i tio delar och taggaren turvis tränas på nio delar av materialet och testas på den tionde, vilket upprepas tio gånger. Medelkorrektet av de tio körningarna visar taggarens noggrannhet.

Eftersom hittills åtta transkriptioner har handkorrigerats gjordes i stället för en klassisk tiodelad korsvalidering en åttadelad korsvalide-

ring. Åtta olika körningar gjordes och tränade taggaren turvis på SUC + 7 *Spara talet*-filer, och korrektheten testades på den åttonde filen. Den genomsnittliga korrektheten för de åtta körningarna är 92,8 procent, dvs. avsevärt högre än de 78,8 procent som uppnås när taggaren är tränad enbart på SUC.

Ordklasserna där den största förbättringen sker när *Spara talet*-material läggs till i träningsmaterialet är förutom DI och PAUS infinitivmärke, utländska ord, egennamn, subjunktioner och pronomen.

Skillnaden i korrekthet mellan enskilda *Spara talet*-filer visade sig vara stor. Tabell IV.1 visar korrektheten för varje transkription när taggaren är tränad dels enbart på SUC (sista kolumnen i tabellen), dels på SUC plus de sju övriga *Spara talet*-filerna. Tabellen är ordnad efter fallande korrekthet enligt den åttadelade korsvalideringen. De österbottniska filerna har den lägsta korrektheten med Närpes på sista plats.

Även när man ser på resultaten när taggaren tränats enbart på SUC har de österbottniska filerna sämre resultat än de flesta sydfinländska. Sannolikt beror det sämre resultatet för de österbottniska filerna på att avståndet mellan österbottniskt talspråk och skriven standardsvenska är större än mellan sydfinländskt talspråk och skriven standardsvenska.

Tabell IV.1. Korrekthet per fil. Filerna är ordnade enligt fallande korrekthet i den åttadelade valideringen. Resultatet för de österbottniska filerna är klart sämre än för de sydfinländska.

FIL	TAGGAREN TRÄNAD PÅ: SUC + 7 FILER				BARA SUC
	KORREKTHET	ANTAL TOKEN			KORREKTHET
	(%)	RÄTT	FEL	TOTALT	(%)
abo_ym13	94,7	2 756	154	2 910	80,4
houtskar_yw01	94,2	3 137	194	3 331	82,1
abo_om04	93,6	2 429	167	2 596	74,5
korpo_ow01	93,1	3 174	234	3 408	83,1
helsingfors_om01- helsingfors_ow03	93,0	3 516	265	3 781	80,1
vasa_ym06	92,7	3 273	256	3 529	77,4
esse_om01	92,6	2 789	222	3 011	76,8
narpes_ym03	88,1	2 829	382	3 211	74,9

Den enda sydfinländska filen som får sämre resultat än de österbottniska när taggaren tränats enbart på SUC är abo_om04 (74,5 %). Detta beror antagligen på att det handlar om en inspelning med extra mycket paustecken och tvekljud. Resultatet för denna fil förbättras märkbart när *Spara talet*-material ingår i träningsmaterialet (93,6 %).

Resultaten antyder att det för de österbottniska filernas del krävs ännu mera handkorrigerat träningsmaterial (i form av österbottniska *Spara talet*-transkriptioner) för att uppnå en hög korrekthet än vad fallet är för de sydfinländska filerna.

4.2 Inlärningskurvan

För att testa hur mycket taggningsresultatet förbättras i takt med att mer handkorrigerat material läggs till i träningsmaterialet, gjordes en analys där andelen finlandssvenskt talspråksmaterial i träningsmaterialet ökades stegvis med ungefär 1 000 token i taget.

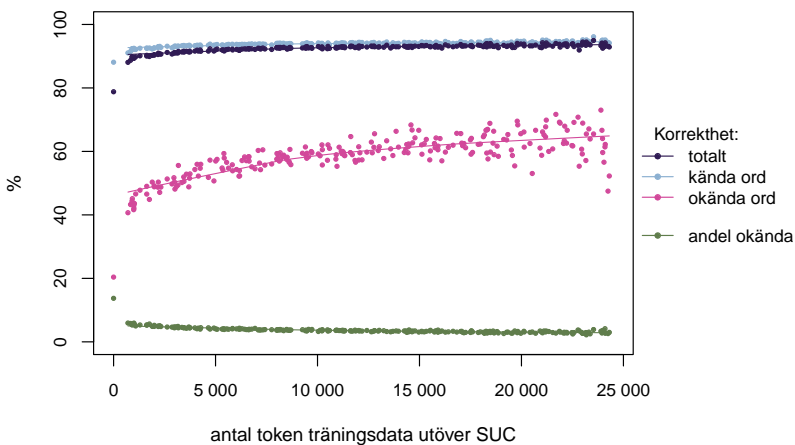
Eftersom en ordklasstagare använder sig av kontext för att tilldela varje ord i en korpus taggar, går det inte att slumpvis välja ut exakt 1 000 ord som träningsmaterial för evalueringen. I stället användes ett skript som följde följande procedur:

1. Räknar ut hur många token som i medeltal ingår i en samtalstur i det handkorrigerade materialet.
2. Räknar ut hur många turer som i medeltal utgör 1 000 token → ”faktor x ”.
3. Drar ett slumpmässigt sampel på x antal turer ur materialet, vilket leder till ungefär 1 000 token. Men antalet token kan variera ganska mycket eftersom en tur kan bestå av allt från ett enda token (t.ex. en uppbackning) till ibland mer än tvåhundra token om en informant får tala på utan att intervjuaren lägger sig i.
4. Taggaren tränas på SUC plus det sampel av turer som valdes ut slumpmässigt i steg 3. Efter detta taggas det övriga materialet (dvs. det som inte valdes ut i steg 3), och taggningsresultatet jämförs med den handkorrigerade taggningen.
5. Steg 3 till 4 upprepas 10 gånger, dvs. totalt dras 10 olika slumpsamplel att träna på. Varje sampel har ungefär 1 000 token, men vissa kommer att ha mindre än 1 000 och andra mer än 1 000.

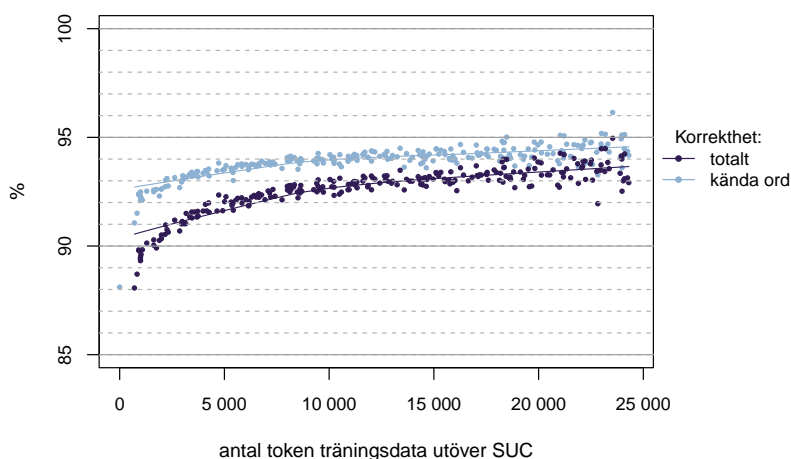
6. Steg 3 till 5 upprepas så att antalet turer i tränings sampel är $2x$, $3x$... nx , där n begränsas av det totala antalet turer i materialet minus x (så att det fortfarande finns minst cirka 1 000 ord kvar att testa på i den sista körningen). Det görs alltså totalt 10 körningar med tränings sampel på cirka 1 000 token, 10 körningar med tränings sampel på cirka 2 000 token, 10 körningar med tränings sampel på cirka 3 000 token osv. så länge materialet räcker.

Resultatet ses i figur IV.2. Figuren visar den totala korrektheten (mörkblått) samt korrektheten för kända (ljusblått) och okända token (rosa) separat. Dessutom visas den procentuella andelen okända token (grönt). Varje körning i den ovan beskrivna proceduren representeras av en punkt av varje färg i figuren. Vid 0 antal token redovisas resultatet när taggaren har tränats på endast SUC. De övriga punkterna är i princip centrerade kring 1 000, 2 000, 3 000 etc. token *Spara talet*-material utöver SUC, men i praktiken bildar de ett kontinuerligt moln. Linjerna som är dragna genom molnen av punkter visar utvecklingen.

Figuren visar att resultatet förbättras så fort tränings materialet förutom SUC inkluderar en del *Spara talet*-material. Korrektheten stiger och andelen okända token sjunker markant redan vid 1 000 token eget tränings material och taggningsresultatet fortsätter att förbättras ju mer *Spara talet*-material som ingår i tränings materialet.



Figur IV.2. Inlärningskurvan för taggaren. Andelen okända ord minskar och korrektheten ökar ju mer *Spara talet*-material som ingår i tränings materialet.



Figur IV.3. Inlärningskurvan för taggaren inzoomad på den översta delen av procentskalan.

För en tydligare bild av den totala korrektheten visar figur IV.3 bara den översta delen av procentskalan. Här ser man att den totala korrektheten förbättras till över 93 procent när ungefär 15 000 token *Spara talet*-material ingår i träningsmaterialet. Orsaken till att korrektheten här verkar högre än vid korsevalueringen (92,8 %) är att turer från en och samma intervju kan ingå i både träningsmaterial och evalueringsmaterial. Ofta handlar intervjuerna om ett specifikt ämne (t.ex. barbershopmusik, lokalhistoria); ett visst ordförråd och vissa ortnamn återfinns bara i en enda intervju och saknas alltså i träningsmaterialet om man utelämnar en fil som helhet. Andelen okända ord blir därmed högre. Vid korsevalueringen saknas i flera fall också den dialekt eller varietet som ska taggas helt i träningsmaterialet, vilket i sig kan försämrta taggningsresultatet. Korsevalueringen ger alltså en mindre fördelaktig bild av taggningsresultatet, och kanske också den bild som är mer relevant med tanke på taggningen av mera material.

I figur IV.3 ser man att taggningsresultatet förbättras snabbt till vänster i figuren, men att kurvan efterhand blir mindre brant, vilket är typiskt för ordklasstagare. Taggaren lär sig snabbt vissa saker, medan de fel som återstår blir allt svårare att korrigera. Till slut kommer taggaren att uppnå en nivå där korrektheten inte längre ökar trots att man tillför mera träningsmaterial. Kurvan planar alltså ut och de fel som

återstår klarar taggaren helt enkelt inte av att lära sig. Denna nivå har ännu inte uppnåtts i figur IV.3 utan kurvan pekar uppåt även längst till höger, vilket innebär att man kan räkna med att taggningsresultatet fortfarande kommer att förbättras i takt med att mera material handkorrigeras och läggs till i träningsmaterialet.

4.3 Svårigheter med tagguppsättningen

En genomgång av vilken typ av ord som ofta blir feltaggade i den automatiska taggningen visar att några notoriska problem med SUC-tagguppsättningen återkommer också i Talko (jfr Megyesi 2002, 57–61 och Östling 2013, 12–14). Ett av dessa är ambiguiteten mellan adverb och adjektiv. I SAG klassificeras neutrumformer av adjektiv som adjektiv även när de står som adverbial i satsen. Vid taggningen av SUC har man däremot valt att tagga dessa adjektiv som adverb (Ejerhed et al. 1992, 6). Resultaten tyder på att en statistisk taggare inte klarar av att göra denna distinktion vilket leder till många feltaggingar, dvs. en stor del adjektiv taggas som adverb och tvärtom. Den enklaste lösningen i detta fall skulle vara att följa SAG i stället för SUC. Då skulle ord som *svårt* och *bättre* alltid entydigt taggas som adjektiv. Att markera adverbial i satsen skulle sedan vara en uppgift för en eventuell senare syntaktisk parsning (automatisk satsdelsanalys) snarare än för ordklasstaggingen.

En liknande svårighet som leder till många homografer och därmed till många feltaggingar är distinktionen mellan självständigt och adjektiviskt använda pronomen. När ord som *annan*, *sådan* och *många* används adjektiviskt taggas de som adjektiv, men när de står självständigt taggas de som pronomen.

Ett ytterligare problem utgörs av taggen för partikel (PL). Enligt SUC-rapporterna (1992 och 2006) har denna först avlägsnats ur SUC, för att sedan återinföras. I SAG tolkas partikeladverbial som en satsdel snarare än en ordklass. Enligt denna tolkning skulle taggen strängt taget inte höra hemma i en ordklasstagging. Ett kriterium för partikeladverbial är enligt SAG att partikeln uppstår huvudbetoningen – ett kriterium som en taggare som utgår från skriftbilden inte kan använda sig av. Gränsen mellan partikeladverbial och övriga adverbial är enligt SAG överlag oskarp. En gränsdragning som är oskarp för mänskliga bedömare leder med högsta sannolikhet till inkonsekvenser vid manuell

taggning, vilket gör att man inte heller kan förvänta sig att en statistisk taggare ska lära sig distinktionen. PL-taggen leder till stor ambiguitet då det finns flera ord som kan tolkas som antingen adverb eller partikel (*upp, ut*) eller som adverb, preposition eller partikel (*runt, över*).

5 Slutsatser

Det material som hittills har handkorrigerats är mycket blandat i fråga om dialektområde. Detta var också avsikten med den inledande testningen av ordklasstaggningen, för att få en bred bild av taggningsresultatet. Frågan som infinner sig är hur hög korrekthet man vill uppnå. Frågan är subjektiv. En korrekthet på 92,8 procent är redan mycket bättre än de 78,8 procent som uppnås när taggaren är tränad enbart på SUC. Det vore ändå önskvärt att nå upp till åtminstone 95 procent och ännu hellre till 96 eller 97 procent (jfr avsnitt 2). Schmid (1995) rapporterar korrektheter kring 97 procent med TreeTagger vid taggning av tyska och engelska. Trots att det handlar om förbättringar om endast någon eller några procentenheter gäller det ett stort antal ord i transkriptionerna (se antal rätt och fel taggade token i tabell IV.1).

Hur mycket mer material som behöver handkorrigeras för att nå upp till den önskade korrektheten går inte att förutspå eftersom förbättringen går långsammare ju bättre resultat man uppnått. För vissa sydfinländska filer är korrektheten dessutom redan med det träningsmaterial som nu finns tillgängligt nära 95 procent (se tabell IV.1), medan det för österbottniska filer kan behövas betydligt mer handkorrigerat material för att uppnå samma resultat. Taggningsresultatet kan eventuellt förbättras mer effektivt genom att man delar upp materialet i olika dialektområden än genom att bara öka den totala mängden träningsmaterial. Detta behöver testas.

På sikt kunde man också överväga att helt avstå från SUC som träningsmaterial och bara använda *Spara talet*-material när en tillräckligt stor mängd material har handkorrigerats. Det är ändå tveksamt om det går att nå upp till så stora mängder handkorrigerat material att det helt kan ersätta SUC. SUC har omkring en miljon ord, vilket är ungefär lika mycket som hela den planerade finlandssvenska talspråkskorpusen sammanlagt.

Ytterligare en möjlighet till att förbättra taggningsresultatet skulle vara att inte bygga upp lexikonfilen endast utifrån träningsmaterialet utan att lägga till ord även från andra ordlistor med morfologisk information, vilket skulle minska andelen okända ord vid taggningen (jfr Östling 2013).

Målet är att förbättra taggningsresultatet ytterligare. För att kunna göra detta behövs framför allt ännu mer handkorrigerat material och flera utvärderingar och felanalyser. Men redan i sin nuvarande form kan korpusen och annoteringen vara till stor nytta för forskare, och annoteringen kanske till och med har en högre korrekthet än vissa delkorpor i Korp som är taggade med bara SUC som träningsmaterial.

Figur- och tabellförteckning

Figurer

1	Karta som visar de finlandssvenska regionerna med regionhuvudstäderna utmärkta	12
2	Stapeldiagram över <i>di</i> ~ <i>dom</i>	16
3	Stapeldiagram över andelen <i>i</i> -uttal i orden <i>se</i> och <i>tre</i>	18
4	Stapeldiagram över infinitivmärket	20
I.1	Karta över Åboland	27
I.2	Antal invånare i de åboländska kommunerna åren 1945–2005	30
I.3	Skärmdump ur Praat	40
I.4	Lådagram över andelen dialektvarianter av variabeln ' <i>ei</i> -difftong'	55
I.5	Lådagram över andelen dialektvarianter (<i>o</i>) av variabeln ' <i>kort u</i> '	66
I.6	Lådagram över andelen dialektvarianter (<i>ngg</i>) av <i>ng</i> -ljudet	69
I.7	Mosaikdiagram över variationen <i>ngg</i> ~ <i>ng</i>	70
I.8	Mosaikdiagram över variationen <i>ngg</i> ~ <i>ng</i> utan belägg på <i>inga</i>	71
I.9	Lådagram över kortstavighetsindex	84
I.10	Lådagram över andelen dialektvarianter av den maskulina slutartikeln	91
I.11	Lådagram över andelen dialektvarianter av neutralt pronomen	99
I.12	Stapeldiagram över antalet belägg på varianterna <i>ti</i> , <i>att</i> och <i>å</i> av infinitivmärket	106
I.13	Diagram över informanternas användning av de tre formerna av infinitivmärket	108
I.14	Lådagram över andelen dialektvarianter av infinitiver	117
I.15	Lådagram över andelen dialektvarianter i preteritum av 2:a och 3:e konjugationens verb	121
I.16	Lådagram över andelen dialektvarianter (<i>inga</i>) av negationen	124
I.17	Lådagram över variabler vilkas dialektvariant tidigare funnits i alla åboländska dialekter	129
I.18	Lådagram över lokala dialektvarianter i Åboland i varianternas ursprungliga spridningsområde	129

I.19	Lådagram över lokala dialektvarianter i Åboland utanför varianternas ursprungliga spridningsområde	130
I.20	Den genomsnittliga andelen lokala varianter på individnivå	132
I.21	Diagram över resultatet av principalkomponentanalysen	134
II.1	Distributionen av olika former av pronomenet <i>de</i> i de svenska dialekterna i Finland	149
II.2	Stapeldiagram över <i>de</i>	157
II.3	Stapeldiagram över <i>dem</i>	157
II.4	Förändringen från <i>de/dem</i> till <i>dom</i> över tid efter grammatisk kategori	158
II.5	Stapeldiagram över ordföljdens betydelse för variationen /di/ ~ /dåm/ hos personligt pronomen som subjekt	160
II.6	Stapeldiagram över andelen /dåm/ bland intervjuarna	161
III.1	Lådagram över V:/C-värde, V:-längd och C-längd efter stad och åldersgrupp	179
III.2	Punktdiagram över V:-längd och C-längd för varje informant	181
III.3	Punktdiagram över spridningen av vokallängd och konsonantlängd i V:C- resp. VC:-sekvenser	187
IV.1	Arbetsflöde	200
IV.2	Inlärningskurvan för taggaren	207
IV.3	Inlärningskurvan för taggaren, inzoomad	208

Tabeller

1	Antal informanter per stad, åldersgrupp och kön	15
2	Åldersfördelningen bland informanterna	15
I.1	Befolkning och näringsstruktur i Åboland år 2005	29
I.2	Antal informanter per ort, åldersgrupp och kön	41
I.3	Belägg på suffixet <i>-og</i> i materialet	96
I.4	Variabler i den kvantitativa analysen	126–127
II.1	Fördelningen på olika grammatiska kategorier efter stad	155
II.2	Andel <i>de här/där</i> av det totala antalet ord i materialet per stad	155
III.1	Koefficienter för fixa effekter i regressionsmodellen med $\log(V:/C)$ som beroende variabel	177
III.2	Slumpintercept för regressionsmodellen i tabell III.1	183
III.3	Slumpintercept för en regressionsmodell utan stad och ålder som fixa effekter	184–185
III.4	Somers D_{xy} -värden för fem olika regressionsmodeller körda separat för sju informantgrupper	188
IV.1	Korrekthet per fil	205

Referenser

Förkortningar

SNF	Studier i nordisk filologi
SSLS	Skrifter utgivna av Svenska litteratursällskapet i Finland

Arkivsamlingar

Förteckning över dialektsamlingar i Svenska litteratursällskapets arkiv i Helsingfors som hänvisas till i texten.

Samling	Innehåll, arkivbildare	År
SLS 663	Säljakten i Hitis (K. Zilliacus)	1954
673	Ord rörande fiske, båtar och kvinnosysslor upptecknade i Iniö (K. Zilliacus)	1955
701	Ordförrådet i Houtskär (K. Zilliacus)	1957
702	Utskrifter från band av språkprov (K. Zilliacus)	1958
1746	Folkkultursarkivets pristävling nr 32: Vad talar du? Dialekt – när, var, hur? (G. Harling-Kranck, K. Westman)	1992
1954	Språkarkivets specialfrågebrev nr 47: Dialekt och talspråk idag (Å. Mitts)	1998

Litteratur

- Ahlbäck, Olav. 1945. *Apokopen i de finlandssvenska folkmålen*. (SSLS 299. SNF 31–32:7.) Helsingfors.
- . 1946. *Studier över substantivböjningen i Finlands svenska folkmål*. (SSLS 308. SNF 33–34.) Helsingfors.
- . 1956. *Svenskan i Finland*. (Skrifter utgivna av Nämnden för svensk språkvård 15.) Stockholm: Svenska bokförlaget.
- Aho, Eija. 2010. Om tonala drag i dialekterna i Åboland. *Folkmålsstudier* 48. S. 13–35.

- Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19. S. 716–723.
- Akselberg, Gunnstein. 2003. Utviklinga av vossamålet ved tusenårsskiftet. I: *Nordisk dialektologi*, red. av G. Akselberg, A. M. Bødal & H. Sandøy. Oslo: Novus. S. 197–226.
- Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. I: *Perspectives on variation. Sociolinguistic, historical, comparative*, red. av N. Delbecque, J. van der Auwera & D. Geeraerts. Berlin: De Gruyter. S. 7–42.
- . 2011. Dialect vs. standard: a typology of scenarios in Europe. I: *The languages and linguistics of Europe. A comprehensive guide*, red. av B. Kortmann & J. van der Auwera. Berlin: De Gruyter. S. 485–500.
- Auer, Peter & Frans Hinskens. 1996. The convergence and divergence of dialects in Europe. *Sociolinguistica* 10. S. 1–30.
- Baayen, R. Harald. 2008. *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barras, Claude, Edouard Geoffrois, Zhibiao Wu & Mark Liberman. 2000. Transcriber: Development and use of a tool for assisting speech corpora production. *Speech Communication* 33. S. 5–22.
- Bates, Douglas, Martin Maechler & Ben Bolker. 2011. *lme4: Linear mixed-effects models using Eigen and Eigen++*. R package version 0.999375-41. <http://CRAN.R-project.org/package=lme4>.
- Behne, Dawn, Peter Czigler & Kirk Sullivan. 1996. Acoustic characteristics of perceived quantity and quality in Swedish vowels. I: *SST 1996 Proceedings*, red. av P. McCormack & A. Russell. Canberra: Australasian Speech Science and Technology Association. S. 49–54.
- Bergman, Gösta. 1942. Sydliga provinsialismer i det bildade svenska talspråket. *Arkiv för nordisk filologi* 56. S. 125–250.
- Bergroth, Hugo. 1928. *Finlandssvenska.Handledning till undvikande av provinsialismer i tal och skrift*. Helsingfors: Schildts. [Faksimiluppl. 1992 av 2 rev. och tillök. uppl.]
- Bergström, Susanne. 1984. *Mattkärr-dialekten. En fonematisk, fonotaktisk och morfologisk beskrivning*. Opublicerad pro gradu-avhandling, Åbo Akademi.
- Boersma, Paul & David Weenink (2013). *Praat: doing phonetics by computer*. <http://www.praat.org>.
- Bruce, Gösta. 1994. Prosodisk strukturering i dialog. I: *Svenskans beskrivning* 20. Lund: Lund University Press. S. 9–23.
- . 2010. *Vår fonetiska geografi. Om svenskans accenter, melodi och uttal*. Lund: Studentlitteratur.
- . 2012. *Allmän och svensk prosodi*. Lund: Studentlitteratur.
- Chambers, J.K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press, 2 uppl.

- Davis, Stuart. 2011. Quantity. I: *The handbook of phonological theory*, red. av J. Goldsmith, J. Riggle & A. Yu. Somerset, NJ: Wiley-Blackwell, 2 uppl. S. 103–140.
- Drager, Katie & Jennifer Hay. 2012. Exploiting random intercepts: Two case studies. *Language variation and change* 24. S. 59–78.
- Ejerhed, Eva, Gunnel Källgren, Ola Wennstedt & Magnus Åström. 1992. *The linguistic annotation system of the Stockholm-Umeå corpus project*. (Report 33.) Department of General Linguistics, University of Umeå.
- Elert, Claes-Christian. 1964. *Phonologic studies of quantity in Swedish. Based on material from Stockholm speakers*. Uppsala: Almqvist & Wiksell.
- Engstrand, Olle & Diana Krull. 1994. Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica* 50. S. 80–91.
- Eriksson, Anders. 2004. Swedia-projektet: dialektforskning i ett jämförande perspektiv. *Folkmålsstudier* 43. S. 11–32.
- Ewerth, Sten. 2008. Dom skriver fortfarande dem. *Språktidningen* 12/2008. S. 66–67.
- Fagerlund, L. W. 1878. *Anteckningar om Korpo och Houtskärs socknar*. (Bidrag till kännedom av Finlands natur och folk 28.) Helsingfors: Finska Vetenskaps-Societeten.
- FO = *Ordbok över Finlands svenska folkmål*. 1976–. Helsingfors: Forskningscentralen för de inhemska språken/Svenska litteratursällskapet i Finland.
- FOreg = *Ordregistret till Ordbok över Finlands svenska folkmål*. Helsingfors: Institutet för de inhemska språken.
- FOS = *Finlands officiella statistik*. Helsingfors: Statistikcentralen. E-publikation, hänvisat 6/1 2012. <http://www.stat.fi>.
- Garlén, Claes. 2003. *Svenska språknämndens uttalsordbok*. Stockholm: Svenska språknämnden/Norstedt.
- Giles, Howard, Nikolas Coupland & Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. I: *Contexts of accommodation*, red. av H. Giles, J. Coupland & N. Coupland. Cambridge: Cambridge University Press. S. 1–68.
- Grass, Eva. 1970. *Växlingen di/dåmm i Eskilstuna stadsspråk. Undersökning av ett recessivt språkdrag*. (FUMS Rapport 16.) Uppsala universitet.
- Gullmets-Wik, Marie-Charlotte. 2004. *Backa och backe, gläder och glad: Formväxling i talspråket i Jakobstad, Kristinestad, Ekenäs och Lovisa*. (SSLS 674. SNF 82.) Helsingfors.
- Gustafson-Capková, Sofia & Britt Hartmann. 2006. *Manual of the Stockholm Umeå Corpus version 2.0*. <http://spraakbanken.gu.se/parole/Docs/SUC2.0-manual.pdf>.
- Haapamäki, Saara & Sarah Kvarnström. 2014. Svenskan i Helsingfors och Åbo ur ett folkdialektologiskt perspektiv. I: *Svenskans beskrivning* 33, red. av J. Lindström et al. (Nordica Helsingiensia 37.) Helsingfors universitet. S. 109–118.

- Haapamäki, Saara & Sarah Wikner. u.u. Att undersöka lyssnares perception av regional variation. *Folkmålsstudier* 53.
- Hadding-Koch, Kerstin & Arthur Abramson. 1964. Duration versus spectrum in Swedish vowels: Some perceptual experiments. *Studia Linguistica* 18. S. 94–107.
- Hagren, Kristina. 2008. *Hur märks infinitiven? Infinitivkonstruktioner i svenska dialekter med fokus på infinitivmärket*. (Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 75.)
- Hakulinen, Lauri. 1979. *Suomen kielen rakenne ja kehitys*. Helsinki: Otava, 4 uppl.
- Harling-Kranck, Gunilla. 1981. De finlandssvenska dialekternas regionala indelning: förslag till gränsdragningar. I: *De finlandssvenska dialekterna i forskning och funktion*, red. av B. Loman. (Meddelanden från Stiftelsens för Åbo Akademi forskningsinstitut 64.) Åbo Akademi. S. 65–80.
- . 1998. *Från Pyttis till Nedervetil. Tjugonio prov på dialekter i Nyland, Åboland, Åland och Österbotten*. (SLS 610.) Helsingfors.
- Heldner, Mattias & Eva Strangert. 2001. Temporal effects of focus in Swedish. *Journal of Phonetics* 29. S. 329–361.
- Helgason, Pétur, Catherine Ringen & Kari Suomi. 2010. Quantity in Central Standard Swedish and Fenno-Swedish. I: *Proceedings from Fonetik 2010*. (Working Papers 54.) Centre for Languages and Literature, Lund University. S. 47–50.
- . 2013. Swedish quantity: Central Standard Swedish and Fenno-Swedish. I: *Journal of Phonetics* 41. S. 534–545.
- Helsingin seudun aluesarjat*. E-publikation, hänvisat 15/4 2015. Tilastokanta > Pääkaupunkiseutu alueittain > Väestö > Väestön koulutusrakenne. Källa: Statistikcentralen och den statistiska verksamheten vid Helsingfors stad. <http://www.aluesarjat.fi>.
- Hinskens, Frans. 2011. Koineization in the present-day Dutch dialect landscape: postvocalic /r/ and more. *Taal & Tongval* 63. S. 99–126.
- Hinskens, Frans, Peter Auer & Paul Kerswill. 2005. The study of dialect convergence and divergence: conceptual and methodological considerations. I: *Dialect change. Convergence and divergence in European languages*, red. av P. Auer, F. Hinskens & P. Kerswill. Cambridge: Cambridge University Press. S. 1–48.
- Hoppenbrouwers, Cor. 1983. Het genus in een Brabants regiolect. *TABU, Bulletin voor Nederlandse Taalkunde* 13. S. 1–25.
- Hultman, O. F. 1894. De östsvenska dialekterna. I: *Finländska bidrag till svensk språk- och folklifsforskning*. Helsingfors: Svenska landsmålsföreningen. S. 113–303.
- af Hällström-Reijonen, Charlotta. 2010. Ledare. *Språkbruk* 3/2010. S. 3.
- Höckerstedt, Leif. 2010. *Svenskan på plats! Helsingforssvenskan visar vägen: födelse, liv och död*. Helsingfors: Söderströms, epub-uppl.
- Itkonen, Terho. 1965. Über einige Sandhi-Erscheinungen im Finnmarklap-

- pischen, Finnischen und Finnlandsschwedischen. *Finnisch-Ugrische Forschungen* 35. S. 218–263.
- Ivars, Ann-Marie. 1986. *Från Österbotten till Sörmland. En undersökning av emigration och språklig anpassning.* (SSLS 532. SNF 66.) Helsingfors.
- . 1988. *Närpesdialekten på 1980-talet.* (SSLS 552. SNF 70.) Helsingfors.
- . 1996. *Stad och bygd. Finlandssvenska stadsmål i ett regionalt och socialt perspektiv.* (Folkmålsstudier 37.) Helsingfors: Föreningen för nordisk filologi.
- . 2003. Lokalt och regionalt i svenskan i Finland: Tendenser i språkutvecklingen i norr och söder. I: *Nordisk dialektologi*, red. av G. Akselberg, A. M. Bødal & H. Sandøy. Oslo: Novus. S. 51–81.
- . 2005. Town and country: when dialect meets standard in urban environments: the case of Finland Swedish. *Linguistics* 43. S. 1049–1071.
- . 2006. Dubbelsupinum i Finland. *Svenska landsmål och svenskt folkliv* 2006. S. 17–30.
- . 2009. Sydösterbottnisk dubbling. I: *En färd i språke. Festskrift till Marketta Sundman*, red. av K. M. Berg et al. Åbo universitet, Nordisk filologi. S. 53–65.
- . 2010. *Sydösterbottnisk syntax.* (SSLS 743. SNF 84.) Helsingfors.
- . 2012. Vad händer med språket i Kristinestad? *Folkmålsstudier* 50. S. 101–128.
- . 2015. *Dialekter och småstadsspråk.* Svenskan i Finland – i dag och i går vol. I:1. (SSLS 798.) Helsingfors.
- Ivars, Ann-Marie & Lisa Södergård. 2007. Spara det finlandssvenska talet. I: *Nordisk dialektologi og sociolingvistik*, red. av T. Arboe. Århus: Peter Skautrup Centret for Jysk Dialektforskning. S. 202–206.
- . 2009. *Spara det finlandssvenska talet. Slutrapport.* SLS elektroniska publikationsarkiv på biblioteksportalen Doria. <http://urn.fi/URN:NBN:fi-fe201004191691>.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Áfarli & Øystein Alexander Vangnes. 2009. The Nordic Dialect Corpus – an advanced research tool. I: *NODALIDA 2009 Conference Proceedings*, red. av K. Jokinen & E. Bick. (NEALT Proceedings Series 4.) S. 73–80.
- Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3. S. 359–383.
- Johnson, Keith. 2008. *Quantitative methods in linguistics.* Malden, Mass. etc.: Blackwell.
- Karlsson, Anne-Maj. 2006. *"Från fader till son". Dialektutjämningen i Nagu, en fallstudie i tre generationer i en familj.* Opublicerad pro gradu-avhandling, Åbo universitet.
- Kerswill, Paul & Peter Trudgill. 2005. The birth of new dialects. I: *Dialect change. Convergence and divergence in European languages*, red. av P. Auer, F. Hinskens & P. Kerswill. Cambridge: Cambridge University Press. S. 196–220.
- Kerswill, Paul & Ann Williams. 2002. "Saliency" as an explanatory factor in language change: evidence from dialect levelling in urban England. I: *Lan-*

- guage change. *The interplay of internal, external and extra-linguistic factors*, red. av M. C. Jones & E. Esch. Berlin: De Gruyter. S. 81–110.
- Kim, Yuni. 2006. Variation and Finnish influence in Finland Swedish dialect intonation. *Lund Working Papers in Linguistics* 52. S. 77–80.
- Kiparsky, Paul. 2008. Fenno-Swedish quantity: Contrast in stratal OT. I: *Rules, Constraints, and Phonological Phenomena*, red. av B. Vaux & A. Nevins. Oxford: Oxford University Press. S. 185–220.
- Kleiweg, Peter. 2013. *RuG/Lo4: software for dialectometrics and cartography*. <http://www.let.rug.nl/~kleiweg/Lo4/>.
- Kristiansen, Tore. 2003. Sproglig regionalisering i Danmark? I: *Nordisk dialektologi*, red. av G. Akselberg, A. M. Bødal & H. Sandøy. Oslo: Novus. S. 115–149.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Larsen, Amund B. 1917. Nabooopposition – knot. I: *Maal og minne 1916–1917*. S. 34–46.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, Massachusetts: MIT Press.
- Lehti-Eklund, Hanna (red.). 1998. *Samtalsstudier*. (Meddelanden från Institutionen för nordiska språk och nordisk litteratur vid Helsingfors universitet 19.) Helsingfors universitet.
- Lehtonen, Jaakko. 1970. *Aspects of quantity in standard Finnish*. (Studia philologica Jyväskyläensia 6.) University of Jyväskylä.
- Leinonen, Therese. 2010. *An acoustic analysis of vowel pronunciation in Swedish dialects*. (Groningen dissertations in linguistics 83.) University of Groningen.
- . 2012. Indelning av finlandssvenska dialekter med särskild hänsyn till Åboland. *Folkmålsstudier* 50. S. 155–173.
- . 2013. Phonetic quantity as a social marker in urban Finland Swedish. I: *Phonetics in Europe. Perception and Production*, red. av C. Gooskens & R. van Bezooijen. Frankfurt a. M.: Peter Lang. S. 181–204.
- . 2015. Dialektgeografi och dialektometri. I: *Sociolinguistik i praktiken*, red. av S. Boyd & S. Ericsson. Lund: Studentlitteratur.
- Lenz, Alexandra. 2003. *Struktur und Dynamik des Substandards. Eine Studie zum Westmitteldeutschen (Wittlich/Eifel)*. (Zeitschrift für Dialektologie und Linguistik. Beihefte 125.) Stuttgart: Steiner.
- Lindström, Jan. 2008. *Tur och ordning. Introduktion till svensk samtalsgrammatik*. Stockholm: Norstedt.
- Lundström, Gudrun. 1939. *Studier i nyländsk syntax*. Stockholm: Norstedt.
- Manning, Christopher D. & Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Megyesi, Beáta. 2002. *Data-driven syntactic analysis. Methods and applications for Swedish*. Stockholm: Kungl. Tekniska högskolan.
- Naert, Aino. 1995. O. F. Hultman och de finlandssvenska dialekternas indelning. *Svenska landsmål och svenskt folkliv 1994*. S. 7–20.
- Nerbonne, John, Rinke Colen, Charlotte Gooskens, Peter Kleiweg & Therese

- Leinonen. 2011. Gabmap: a web application for dialectology. *Dialectologia* Special Issue II. S. 65–89.
- Nivre, Joakim & Leif Grönqvist. 2001. Tagging a corpus of spoken Swedish. *International Journal of Corpus Linguistics* 6. S. 47–78.
- Nordberg, Bengt. 2012. Intraindividual conversation-internal variation: interactionally or linguistically motivated? I: *Sociolinguistic Symposium 19: Abstracts*, red. av M. Hüning & U. Reich. Berlin: Freie Universität Berlin. S. 164.
- Nyholm, Leif. 1976. *Formväxling i Helsingforsvenskan. En studie i intervjuSpråk.* (FUMS Rapport 44.) Uppsala universitet.
- . 1984. *Svenskt stadsmål i Helsingfors.* (Meddelanden från Institutionen för nordiska språk och nordisk litteratur vid Helsingfors universitet B:8.) Helsingfors universitet.
- Paolillo, John C. 2001. *Analyzing linguistic variation. Statistical models and methods.* Stanford, California: Center for the Study of Language and Information.
- Pind, Jörgen. 1999. Speech segment durations and quantity in Icelandic. *Journal of the Acoustical Society of America* 106. S. 1045–1053.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Wien. <http://www.R-project.org/>.
- Rabb, Viveca. 2007. *Genuskongruens på reträtt. Variation i nominalfrasen i Kvevlaxdialekten.* Åbo Akademi.
- . 2014. Dialektvariation i Vasa. I: *Svenskans beskrivning 33*, red. av J. Lindström et al. (Nordica Helsingiensia 37.) Helsingfors universitet. S. 372–382.
- . 2015. Standard eller dialekt? Några nedslag i modern Vasasvenska. I: *Vasa som språkgemenskap. Perspektiv på den tvåspråkiga stadens flerspråkighet*, red. av H. Lönnroth & L. Laukkanen. (Vaasan yliopiston julkaisuja. Tutkimuksia 304, Kielitiede 49.) Vasa universitet. S. 24–40.
- Rehbein, Ines & Sören Schalowski. 2013. STTS goes Kiez: Experiments on annotating and tagging urban youth language. *Journal for Language Technology and Computational Linguistics* 28. S. 199–227.
- Reuter, Mikael. 1971. Vokalerna i finlandssvenskan: en instrumentell analys och ett försök till systematisering enligt särdrag. I: *Festskrift till Olav Ahlbäck.* (SSLS 446. SNF 58.) Helsingfors. S. 240–249.
- . 1980. Kortstavighet i Helsingforsvenskan. I: *Festskrift till Carl-Eric Thors.* (SSLS 490. Studier i nordisk filologi 62.) Helsingfors. S. 222–233.
- . 1982. *Kvantitet i Helsingforsvenskan.* Opublicerad licentiatavhandling, Helsingfors universitet.
- . 1992. Swedish as a pluricentric language. I: *Pluricentric languages. Differing norms in different nations*, red. av M. G. Clyne. (Contributions to the sociology of language 62.) Berlin: De Gruyter. S. 101–116.
- Riad, Tomas. 1992. *Structures in Germanic prosody. A diachronic study with special reference to the Nordic languages.* Department of Scandinavian languages, Stockholm University.

- . 2014. *The Phonology of Swedish*. (The phonology of the world's languages.) Oxford: Oxford University Press.
- Rietveld, A. C. M. & V. J. van Heuven. 2009. *Algemene fonetiek*. Bussum: Coutinho, 3 uppl.
- Saari, Mirja. 1975. *Talsvenska. En sociolingvistisk studie över syntaktiska drag i intervjuvar*. (SSLS 468. SNF 60.) Helsingfors.
- . 2000. *Den sir läcker ut den där salladen*. Om användningen av adnominala *den där*-fraser i svenskan. I: *Denna – den här – den där. Om demonstrativer i tvärspråklig belysning. En minnesskrift till Elsie Wijk-Andersson*. (Asla information 26:2.) S. 103–138.
- SAG = Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens grammatik* vol. 1–4. Stockholm: Svenska Akademien/Norstedt.
- Sandström, Caroline. 2010. *Genus i östra Nyland. Från dialektutjämning till dialektmarkör*. (Nordica Helsingensia 20.) Helsingfors universitet.
- Sankoff, David. 1988. Variable rules. I: *Sociolinguistics. An international handbook of the science of language and society* vol. 2, red. av U. Ammon, N. Dittmar & K. J. Mattheier. Berlin: De Gruyter. S. 984–997.
- Sarkar, Deepayan. 2008. *Lattice. Multivariate data visualization with R*. New York: Springer.
- Schaeffler, Felix. 2005. *Phonological quantity in Swedish dialects. Typological aspects, phonetic variation and diachronic change*. (Phonum 10.) Umeå university.
- Schmid, Helmut. 1994. Probabilistic part-of-speech tagging using decision trees. I: *Proceedings of International Conference on New Methods in Language Processing*. Manchester.
- . 1995. Improvements in part-of-speech tagging with an application to German. *Proceedings of the ACL SIGDAT-Workshop*. Dublin. S. 47–50.
- Schwarz, Christian, Helmut Spiekermann & Tobias Streck. 2011. Primäre und sekundäre Dialektmerkmale: Empirische Befunde aus Dialekten und Standardvarietäten. I: *Akten der 10. Arbeitstagung für bayerisch-österreichische Dialektologie*, red. av H.-D. Pohl. (Klagenfurter Beiträge zur Sprachwissenschaft 34–36.) Wien: Praesens. S. 355–377.
- Selenius, Ebba. 1978. Studies in the development of the 2-accent system in Finland-Swedish. I: *Nordic Prosody. Papers from a symposium*, red. av E. Gårding, G. Bruce & R. Bannert. (Travaux de l'Institut de linguistique de Lund 13.) Lund University. S. 229–236.
- SO = *Svensk ordbok utgiven av Svenska Akademien*. 2009. Stockholm: Norstedt.
- Språkriktighetsboken*. 2005. (Skrifter utgivna av Svenska språknämnden 93.) Stockholm: Norstedts akademiska förlag.
- StatFin = *Statistikdatabasen StatFin*. Helsingfors: Statistikcentralen. E-publication, hänvisat 6/1 2012. www.stat.fi/tup/statfin.
- Statistisk årsbok för Finland*. 1879–. Helsingfors: Statistikcentralen.
- Stenberg-Sirén, Jenny & Jan-Ola Östman. 2012. Processes of standardization through the media in minority languages: Swedish in Finland. I: *Sociolin-*

- guistic Symposium 19. Abstracts, red. av M. Hüning & U. Reich. Berlin: Freie Universität Berlin. S. 183.
- Stenius, Yrsa. 2012. Språk och klass. I: *Åbo underrättelser*, 5 juni 2012.
- Strangert, Eva & Mattias Heldner. 1995. Labelling of boundaries and prominences by phonetically experienced and nonexperienced transcribers. *Phonum* 3. S. 85–109.
- Sundberg, Eva. 1993. *Dialekten i Ålands nordöstra skärgård*. Mariehamn: Ålands högskola.
- Sundman, Marketta. 2000. Svenska och finska demonstrativer som definithetsmarkörer i muntliga narrativer hos tvåspråkiga finländare. I: *Denna – den här – den där. Om demonstrativer i tvåspråklig belysning. En minnesskrift till Elsie Wijk-Andersson*. (Asla information 26:2.) S. 139–165.
- Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2008. *Finnish sound structure. Phonetics, phonology, phonotactics and prosody*. (Studia humaniora ouluensia 9.) University of Oulu.
- Szmrecsanyi, Benedikt. 2006. *Morphosyntactic persistence in spoken English. A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. (Trends in Linguistics. Studies and Monographs 177.) Berlin: De Gruyter Mouton.
- Södergård, Lisa. 2009. Det sparade talet: erfarenheter från ett insamlingsprojekt. I: *Svenskan i Finland 11*, red. av L. Collin & S. Haapamäki. (Skrifter från svenska institutionen vid Åbo Akademi 7.) S. 205–210.
- Södergård, Lisa & Therese Leinonen. u.u. Från dialektinspelning till tal-språkskorpus: beskrivning av ett korpusbygge. I: *Ideologi, identitet, intervention. Tionde nordiska dialektologkonferensen*, red. av J.-O. Östman et al. (Nordica Helsingiensia.) Helsingfors universitet.
- Søfteland, Åshild & Anders Nøklestad. 2008. Manuell morfologisk tagging av NoTa-materialet med støtte fra en statistisk tagger. I: *Språk i Oslo. Ny forskning omkring talespråk*, red. av J. B. Johannessen & K. Hagen. Oslo: Novus. S. 226–234.
- Tabachnik, Barbara G. & Linda S. Fidell. 2007. *Using multivariate statistics*. Boston: Pearson, 5 uppl.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language variation and change* 24. S. 135–178.
- Tandefelt, Marika. 1986. The mechanisms of Swedish-Finnish language shift. *Nordic Journal of Linguistics* 9. S. 83–96.
- . 1996. *På vinst och förlust. Om tvåspråkighet och språkförlust i Helsingforsregionen*. (Forskningsrapporter från Svenska handelshögskolan 35.) Helsingfors.
- (red.) 2015. *Gruppspråk, samspråk, två språk. Svenskan i Finland – i dag och i går* vol. I:2. (SSLS 799.) Helsingfors.
- Thelander, Mats. 1979. *Språkliga variationsmodeller tillämpade på nutida Burträsktal. Del 2. Teoretiska utgångspunkter och analys på språkartsnivå*.

- (*Studia philologiae Scandinavicae Upsaliensia* 14:2.) Uppsala universitet.
- . 1980. *De-dialectalisation in Sweden*. (FUMS Rapport 86.) Institutionen för nordiska språk vid Uppsala universitet.
- . 1982. A qualitative approach to the quantitative data of speech variation. I: *Sociolinguistic variation in speech communities*, red. av S. Romaine. London: Arnold. S. 65–83.
- Thurman, Janne. 1898. *Pargasmålet. Ljud- och formlära*. (Bidrag till kännedom om de svenska landsmålen ock svenskt folklied XV:4.) Stockholm: Landsmålsföreningarna i Uppsala, Helsingfors och Lund.
- . 1929. *Nagumålet*. (SSLS 205. SNF 19:3.) Helsingfors.
- Trudgill, Peter. 1986. *Dialects in contact*. (Language in society 10.) Oxford: Blackwell.
- . 2011. *Sociolinguistic typology. Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- UVF = *Uppslagsverket Finland – webbutgåva*. Helsingfors: Schildts. E-publication, hänvisat 6/1 2012. <http://www.uppslagsverket.fi>.
- Wessén, Elias. 1969. *Våra folkmål*. Stockholm: Fritzes, 9 uppl.
- Wessman, V.E. V. 1936. *De finlandssvenska dialekternas labiala vokaler I. Kort o, u, å i huvudtonig stavelse framför geminata och konsonantgrupper. Ljudhistorisk och dialektgeografisk undersökning*. (Folkmålsstudier 4.) Helsingfors: Föreningen för nordisk filologi.
- . 1938. *De finlandssvenska dialekternas labiala vokaler II. Kort o, u, å i huvudtonig stavelse framför kort konsonant. Ljudhistorisk och dialektgeografisk undersökning. Folkmålsstudier 5*. S. 21–134.
- Wide, Camilla. 2011a. Pronomenen *den här* och *den där* som planeringspartiklar i finlandssvenska dialekter. I: *Interaktionell dialektologi*, red. av G. Bockgård & J. Nilsson. Uppsala: Institutet för språk och folkminnen. S. 251–306.
- . 2011b. Proximala demonstrativa pronomen i finlandssvenska dialekter. I: *Studier i dialektologi och sociolingvistik*, red. av L.-E. Edlund, L. Elmevik & M. Reinhammar. Uppsala: Kungl. Gustav Adolfs Akademien för svensk folkkultur. S. 331–339.
- Zilliacus, Kurt. 1990. *"Finska skären"*. *Studier i åboländsk kulturhistoria*. Helsingfors: Konstsamfundet.
- . 1992. *Åboländska*. Helsingfors: Konstsamfundet.
- Östling, Robert. 2013. Stagger: an open-source part of speech tagger for Swedish. *Northern European Journal of Language Technology* 3. S. 1–18.
- Östman, Jan-Ola & Leila Mattfolk. 2011. Ideologies of standardisation: Finland Swedish and Swedish-language Finland. I: *Standard languages and language standards in a changing Europe*, red. av T. Kristiansen & N. Coupland. Oslo: Novus. S. 75–82.

Bilaga A

Regressionsanalyser, åboländska

1 Ljudlära

1.1 *ei*-diftongen

Logistisk regression med blandade effekter över *ei*-diftongen. Dialektvariant är *ei/äi* och standardvariant är *e*: $AIC = 634,9$, Somers $D_{xy} = 0,794$.

Tabell A.1. Stokastiska effekter, *ei*-diftongen, $n = 962$; grupper: informanter = 43, ord = 4.

FAKTOR	STOK. EFFEKT	STDAVV.
informant	intercept	0,980
ord	intercept	0,515

Tabell A.2. Fixa effekter, *ei*-diftongen.

	ESTIMAT	MEDELFEJ	Z-VÄRDE	P
Intercept	-1,735	0,957	-1,814	0,070
ort = <i>Hi</i>	-0,606	1,741	-0,348	0,728
ort = <i>Ho</i>	0,592	1,092	0,542	0,588
ort = <i>Iö</i>	-0,485	1,294	-0,375	0,708
ort = <i>Ki</i>	-1,006	1,285	-0,783	0,434
ort = <i>Ko</i>	-1,697	1,234	-1,376	0,169
ort = <i>Na</i>	-0,983	1,194	-0,824	0,410
ort = <i>Pa</i>	-1,222	1,159	-1,054	0,292
ort = <i>Vf</i>	1,874	1,387	1,351	0,177
ålder = <i>yngre</i>	-0,909	1,111	-0,818	0,413
<i>Hi:yngre</i>	-0,603	2,262	-0,266	0,790
<i>Ho:yngre</i>	3,301	1,476	2,236	0,025
<i>Iö:yngre</i>	2,993	1,783	1,678	0,093
<i>Ki:yngre</i>	-0,011	1,770	-0,006	0,995
<i>Ko:yngre</i>	2,325	1,777	1,308	0,191
<i>Na:yngre</i>	-0,769	1,638	-0,469	0,639
<i>Pa:yngre</i>	-0,774	1,809	-0,428	0,669
<i>Vf:yngre</i>	-15,830	751,382	-0,021	0,983

1.2 Kort *u*

Logistisk regression med blandade effekter över variabeln kort *u*.
 Varianter är *o* och *u*. $AIC = 549,5$, Somers $D_{xy} = 0,859$.

Tabell A.3. Stokastiska effekter, kort *u*, $n = 637$; grupper: informanter = 43.

FAKTOR	STOK. EFFEKT	STDAVV.
informant	intercept	1,717

Tabell A.4. Fixa effekter, kort *u*.

	ESTIMAT	MEDELFEJL	Z-VÄRDE	P
Intercept	3,094	0,921	3,360	0,001
ort = <i>Hi</i>	-1,243	1,408	-0,882	0,378
ort = <i>Ho</i>	-2,076	1,137	-1,826	0,068
ort = <i>Iö</i>	-3,509	1,385	-2,535	0,011
ort = <i>Ki</i>	-1,470	1,264	-1,162	0,245
ort = <i>Ko</i>	-7,364	1,915	-3,845	< 0,001
ort = <i>Na</i>	-4,251	1,144	-3,717	< 0,001
ort = <i>Pa</i>	-2,001	1,092	-1,833	0,067
ort = <i>Vf</i>	0,624	1,614	0,387	0,699
kön = <i>kvinna</i>	-1,650	0,639	-2,580	0,010

1.3 *ng*-ljudet

Logistisk regression med blandade effekter över variabeln $ng(g)$. Varianter är dialektvarianten *ngg* och den standardsvenska varianten *ng*. $AIC = 739,7$, Somers $D_{xy} = 0,882$.

Tabell A.5. Stokastiska effekter, *ng*-ljudet, $n = 1992$; grupper: informanter = 43, ord = 2.

FAKTOR	STOK. EFFEKT	STDAVV.
informant	intercept	2,873
ord (<i>inga</i> vs annat ord)	intercept	0,539

Tabell A.6. Fixa effekter, *ng*-ljudet.

	ESTIMAT	MEDELFEJL	Z-VÄRDE	P
Intercept	-3,191	0,771	-4,141	< 0,001
ålder = <i>yngre</i>	-4,273	1,389	-3,076	0,002
ställning = <i>slutljud</i>	-1,210	0,308	-3,931	< 0,001
stavelse = <i>obetonad</i>	-1,111	0,335	-3,319	< 0,001

2 Formlära och formord

2.1 Infinitivmärket

$ti \sim att/\ddot{a}$

Logistisk regression med blandade effekter över variabeln infinitivmärket. Varianter är ti och att/\ddot{a} . $AIC = 379,0$, Somers $D_{xy} = 0,993$. Trots att inget av p -värdena för ort är signifikanta i tabell A.8 har ort bibehållits som variabel i modellen eftersom likelihood-kvottestet visar att en modell med ort som faktor är signifikant bättre ($p = 0,0085$) än en modell utan ort (jfr Tabachnik & Fidell 2007, 459).

Tabell A.7. Stokastiska effekter, infinitivmärket ($ti \sim att/\ddot{a}$), $n = 520$; grupper: föregående ord = 204, informanter = 43.

FAKTOR	STOK. EFFEKT	STDAVV.
föregående ord	intercept	2,388
informant	intercept	4,454

Tabell A.8. Fixa effekter, infinitivmärket ($ti \sim att/\ddot{a}$).

	ESTIMAT	MEDELFEJL	Z-VÄRDE	P
Intercept	2,158	2,244	0,962	0,336
ort = Hi	1,246	3,478	0,358	0,720
ort = Ho	4,327	2,937	1,474	0,141
ort = $I\ddot{o}$	7,039	4,870	1,445	0,148
ort = Ki	-1,473	3,113	-0,473	0,636
ort = Ko	-0,509	3,278	-0,155	0,877
ort = Na	-2,025	2,690	-0,753	0,452
ort = Pa	-5,014	3,098	-1,619	0,106
ort = Vf	4,273	3,781	1,130	0,258
ålder = $yngre$	-4,688	1,750	-2,679	0,007

$att \sim \hat{a}$

Logistisk regression med blandade effekter över variabeln infinitivmärket. Varianter är *att* och *å*. $AIC = 264$, Somers $D_{xy} = 0,920$.

Tabell A.9. Stokastiska effekter, infinitivmärket ($att \sim \hat{a}$), $n = 233$; grupper: föregående ord = 120, informanter = 30.

FAKTOR	STOK. EFFEKT	STDAVV.
föregående ord	intercept	1,173
informant	intercept	1,606

Tabell A.10. Fixa effekter, infinitivmärket ($att \sim \hat{a}$).

	ESTIMAT	MEDELFEJL	Z-VÄRDE	P
Intercept	0,716	0,938	0,764	0,445
ort = <i>Hi</i>	-1,088	1,545	-0,704	0,481
ort = <i>Ho</i>	1,120	1,713	0,654	0,513
ort = <i>Iö</i>	-1,591	2,356	-0,675	0,500
ort = <i>Ki</i>	-2,129	1,438	-1,481	0,139
ort = <i>Na</i>	-2,398	1,246	-1,925	0,054
ort = <i>Pa</i>	-1,611	1,192	-1,351	0,177

2.2 Infinitiv

Logistisk regression med blandade effekter över variabeln infinitiv. Varianter är apokoperad infinitiv och fullform. $AIC = 421,9$, Somers $D_{xy} = 0,842$.

Tabell A.11. Stokastiska effekter, infinitiv, $n = 481$; grupper: informanter = 43.

FAKTOR	STOK. EFFEKT	STDAVV.
informant	intercept	1,631

Tabell A.12. Fixa effekter, infinitiv.

	ESTIMAT	MEDELFEJ	Z-VÄRDE	P
Intercept	-4,018	0,907	-4,430	< 0,001
kontext = <i>upptakt</i>	2,421	0,391	6,197	< 0,001
ort = <i>Hi</i>	0,292	1,430	0,205	0,838
ort = <i>Ho</i>	4,047	1,178	3,436	< 0,001
ort = <i>Iö</i>	5,407	1,443	3,747	< 0,001
ort = <i>Ki</i>	0,403	1,285	0,313	0,754
ort = <i>Ko</i>	2,932	1,314	2,231	0,026
ort = <i>Na</i>	0,642	1,116	0,575	0,565
ort = <i>Pa</i>	-0,034	1,088	-0,031	0,975
ort = <i>Vf</i>	1,802	1,493	1,207	0,227

2.3 Preteritum

Logistisk regression med blandade effekter över variabeln preteritum-apokope. Varianter är den dialektala apokoperade preteritumformen och den standardspråkliga fullformen. $AIC = 992,2$, Somers $D_{xy} = 0,796$.

Tabell A.13. Stokastiska effekter, preteritum, $n = 1600$; grupper: ord = 96, informanter = 43.

FAKTOR	STOK. EFFEKT	STDAVV.
ord	intercept	0,605
informant	intercept	1,475

Tabell A.14. Fixa effekter, preteritum.

	ESTIMAT	MEDELFEJ	Z-VÄRDE	P
Intercept	2,680	0,890	3,013	0,003
ort = <i>Hi</i>	-0,370	1,306	-0,284	0,777
ort = <i>Ho</i>	0,413	1,149	0,359	0,720
ort = <i>Iö</i>	0,803	1,501	0,535	0,593
ort = <i>Ki</i>	-1,246	1,153	-1,081	0,280
ort = <i>Ko</i>	-0,445	1,238	-0,360	0,719
ort = <i>Na</i>	-0,556	1,041	-0,535	0,593
ort = <i>Pa</i>	-2,126	1,004	-2,118	0,034
ort = <i>Vf</i>	-0,150	1,573	-0,096	0,924
ålder = <i>yngre</i>	1,097	0,583	1,881	0,060