

Sammanfall mellan betonat kort /i/ och kort /e/ i Götaland

en fonetisk och metodologisk undersökning

Håkan Hansson

**Kandidatuppsats i fonetik
Språk- och litteraturcentrum, Lunds Universitet
HT 2016
Handledare: Mechtild Tronnier**

Sammanfattning

Sammanfallet mellan betonat kort /i/ och betonat kort /e/ i Götaland har studerats baserat på materialet i SweDia-korpusarna. Ett exempel på det sammanfallet är att *fisk* kan uttalas ”fes^k”. En dataanalys av korpusen för undersökning av fonem gjordes, liksom en analys av korpusen för spontant tal.

Resultatet av dataanalysen visar inte på att uttalet med /e/ skulle vara vanligare i Västsverige än i övriga Götaland, eftersom uttalet med /e/ nästan saknas i det eliciterade materialet. Däremot finns det uttalet i den förväntade omfattningen i spontantalskorpusen. Orsaken till skillnaden är att man inte har fått med det dialektala uttala uttalet vid eliciteringen.

Det är svårt att göra bra kvantitativa undersökningar av dialekter enbart från SweDia-materialet, eftersom man är beroende av det material som redan samlats in. Däremot ger spontantalskorpusen goda möjligheter att bedöma dialekterna, som en förberedelse för insamling av eget material.

I studien beräknades skillnaden mellan vokalerna i orden *sitt* och *lett* genom att först beräkna det akustiska avståndet för varje enskild informant och sedan använda detta för vidare analyser. Eftersom en stor spridning oavsett orsak ser ut som ett stort akustiskt avstånd tenderar den metoden att ge ett för högt värde på det akustiska avståndet.

Abstract

The coalescence of a short stressed /i/ and a short stressed /e/ in Götaland has been studied based on the SweDia databases. An example of the coalescence is that the *fisk*, the Swedish word for fish, is pronounced “fes^k”. An analysis of the database with elicited material was done, as well as an analysis of the database for spontaneous speech.

The result of the analysis does not indicate that the pronunciation with /e/ is more common in Western Sweden than in the rest of Götaland, as the pronunciation with /e/ is almost absent in the elicited database. On the other hand this pronunciation is found to the expected extent in the database for spontaneous speech. The reason for the difference between the databases is that the pronunciation with /e/ was not successfully elicited.

It is difficult to make good quantitative studies of dialects based on just the SweDia databases, as they are limited to material that has already been collected. But the database with spontaneous speech offers a good opportunity to listen to different dialects as a preparation for collecting a dedicated material.

In this study the acoustic distance between the vowels in *sitt* and *lett* was calculated by first calculating the distance for each informant and then using this data for further analysis. As a large straggle, independent of the reason, automatically gives a high value for the acoustic distance, the values of acoustic distance tend to be overestimated.

Innehåll

1. Inledning	4
2. Forskningsfrågor	5
3. Tidigare forskning.....	6
3.1 Tidigare forskning om korta betonade /i/ och /e/ i södra Sverige	6
3.2 SweDia-projektet.....	8
3.3 Metoder för att analysera vokaler	9
3.3.1 Källa-filer-perspektivet	9
3.3.2 Formantmetoden	11
3.3.3 Andra metoder än formantmetoden.....	13
4. Material och metod	15
4.1 Material	15
4.2 Metod	16
4.2.1 Val av metod i tidigare studier	16
4.2.2 Val av metod i den här studien.....	17
5. Resultat och analys.....	19
5.1 Mätningar av akustiskt avstånd mellan kort /i/ och kort /e/	19
5.1.1 Auditiv bedömning av uttalet av kort /i/ och kort /e/ i SweDia-materialet.....	19
5.1.2 Geografisk fördelning	19
5.1.3 Åldersfördelning	24
5.2 Felkällor knutna till insamlingen av materialet.....	24
5.2.1 Auditiv kontroll av punkter med litet eller stort akustiskt avstånd	24
5.2.3 Auditiv kontroll av en inspelningarna från en tre orter	25
5.3 Felkällor knutna till analysmetoden	27
5.3.1 Skillnad mellan formanterna F1, F2, och F3.....	27
5.3.2 Skillnad mellan upprepningar av orden och olika extraktioner av formanterna	28
5.4 Summering av resultaten	31
6. Diskussion och slutsatser	32
6.1 Diskussion	32
6.1.1 Sammanfall mellan betonade kort /i/ och kort /e/.....	32
6.1.2. SweDia-materialet.....	32
6.1.3 Akustiskt avstånd	33
6.2 Slutsatser	34
7. Referenser	35
Appendix 1	37
Appendix 2.....	38
Appendix 3.....	39

1. Inledning

Det är känt att språk förändras med tiden och det gäller också dialekter (Bruce 2010, s. 205-226). I en studie över dialektutjämning i Västsverige (Svahn & Nilsson 2014) visas hur olika dialektala drag försvagas eller försvinner från språket. Ett sådant dialektalt drag som är typiskt för Västsverige är sammanfallet mellan ett betonat kort /i/ och ett betonat kort /e/ (/i/ => /e/). Det är det som gör att ordet *fisk* kan uttalas som "fesök".

I den här studien undersöks om detta sammanfall är vanligare i Västsverige än i övriga delar av Götaland. Samtidigt borde man kunna hitta ett sammanfall åt andra hållet (/e/ => /i/) så att *hemma* uttalas "himma" i andra delar av Götaland.

Metoden går ut på att bestämma frekvensen för vokalernas formanter och sedan beräkna avståndet mellan frekvenserna för vokalerna i två ord. Detta kallas akustiskt avstånd och kan uttryckas i Hz eller i Bark. Bark är en anpassning av frekvensen till örats perception. De ord som i första hand studeras är *sitt*, som innehåller ett /i/, och *lett*, som innehåller ett /e/.

De flesta liknande studier undersöker det akustiska avståndet mellan det typiska uttalet av två vokaler (Wenner 2010, Ferrange & Pellegrino 2010a), men det räcker inte i det här fallet. Man kan anta att det bara är en mindre del av befolkningen i Västsverige som har uttalet med /e/, och då kan man inte nöja sig med att bara undersöka ett medelvärde eller medianvärde av uttalet. I stället måste man ta reda hur stor del av talarna som har sammanfallet.

Materialet som används i studien samlades in i SweDia-projektet omkring år 2000 för att undersöka svenska dialekter. Inspelningar gjordes då på drygt hundra orter i Sverige och svenskspråkiga Finland. En av de korpusar som ställdes samman är inriktad på fonem och en annan innehåller spontant tal. Den första av dem används i den här studien för en matematisk beräkning av det akustiska avståndet och spontantalet används för att lyssna till normalt tal.

Informanterna i SweDia-materialet är indelade i två åldersgrupper och då finns det också en möjlighet att se om någon av grupperna har en större andel talare med sammanfallet /i/ => /e/ än den andra. Eftersom dialekterna i Sverige har utjämnats under de senaste hundra åren kan man vänta sig att äldre talare har tydligare dialektala drag.

Samtidigt som sammanfallet /i/ => /e/ undersöks noteras också i vilken utsträckning SweDia-materialet kan användas för den här typen av undersökningar. Slutligen utvärderas hur lämplig den använda metoden är för att beräkna det akustiska avståndet.

2. Forskningsfrågor

Studien undersöker tre frågor:

1. Finns det någon påvisbar skillnad i det akustiska avståndet mellan ett betonat kort /i/ och ett betonat kort /e/ för dialekter i Götaland?
2. Vilka möjligheter och begränsningar finns att använda materialet från SweDia-projektet för att undersöka vokaler i svenska dialekter?
3. Är det lämpligt att använda metoden med akustisk distans för att undersöka dialektala skillnader?

Till de två sista frågorna finns ingen formulerad hypotes, men det är ett rimligt antagande att både materialet och metoden måste vara användbara på något sätt. Däremot finns det två hypoteser till den första frågan.

1. Hypotesen är att ett sammanfall mellan kort /i/ och kort /e/ är vanligare i Västsverige än i andra delar av Götaland.
2. Hypotesen är att det är vanligare att äldre än yngre informanter har ett sammanfall mellan kort /i/ och kort /e/.

Den första hypotesen stöder sig på en intuitiv uppfattning att /i/ => /e/ är vanligare i Västsverige än i andra delar av Götaland, och den andra hypotesen stöder sig på resultat från Svahn & Nilsson (2014, s. 79-80).

Den skillnaden man kan se borde vara att en viss andel av informanterna i material inte skiljer på korta /i/ och /e/. Men de som inte har det sammanfallet borde ha ett normalt akustiskt avstånd, även om man kan tänka sig att också det varierar över materialet.

3. Tidigare forskning

3.1 Tidigare forskning om korta betonade /i/ och /e/ i södra Sverige

Den mest kompakta beskrivningen av svenskans vokaler finns i Bruces ”Vår fonetiska geografi” (Bruce 2010, s. 101-140). Framställningen bygger på tidigare litteratur kompletterat med eget material, som i huvudsak kommer från SweDia-projektet. Svenskan är känt för att vara ett vokalrikt språk. Man brukar räkna med nio långa vokaler och nio korta, där vissa av dem kan sammanfalla beroende på dialekt. Vid undersökningarna har Bruce studerat vokalernas formantfrekvenser och kompletterat med att studera spektrogram och avlyssning av inspelat material.

I exemplen här och senare i texten markerar kursiv stil att det är ett ord i modern svenska med normal stavning. Ord inom citationstecken markerar en ortografisk återgivning av ett dialektalt ord. Ord eller tecken mellan snedstreck är en fonematisk transkription och mellan hakparenteser är det en fonetisk transkription.

För Götaland skiljer Bruce på sydsvenska mål, götamål och gotländska (Bruce 2010, s. 29). Detta ansluter till den uppdelning som Wessén tidigare gjort. (Wessén, 1960, s. 13-14). Till sydsvenska mål hör de dialekter som talas i Skåne, Blekinge och en bit upp i Halland och Småland. Det sammanfaller ungefär med de områden som fram till mitten av 1600-talet tillhörde Danmark. Götamål har sitt centrum i Västergötland, men talas också i Halland, Bohuslän, Dalsland, Småland, i södra Östergötland och på Öland. Gränsen mellan götamål och sveamål i Östergötland ligger ganska långt söderut, så större delen av Östergötland hör språkligt till sveamålen (Wessén 1960, s. 23). Gotländskan betraktas som en egen dialekt. Den här uppdelningen är naturligtvis grov och det finns sällan tydliga gränser mellan språkvarianterna.

Under senmedeltiden inträffade några genomgripande vokalförändringar i det svenska språket (Wessén 1968, s.76-77). Ett /i/ övergick till ett kort /e/ många fall, t ex ”skip”=> *skepp*. I några fall där övergången uteblev i sveamålen finns den i götamålen. Exempel på sådana ord är *fisk* => ”fesk”, *skrift* => ”skreft”, ord som *bitit* => ”betet” och *rivit* => ”revet” och där den följande stavelsen innehåller halv vokalen /j/ (*vittja* => ”vettja”). I många fall behölls ett kort /i/ också i götamålen. Det gäller ofta framför nasaler (*ingen* => ”ingen”).

Det är svårt att följa hur uttalet har förändrats i olika delar av landet under de senaste femhundra åren, men från slutet av 1800-talet och första halvan av 1900-talet finns det ett stort antal uppteckningar av folkmål. De är ofta mycket ambitiösa och gjorda av entusiastiska lingvister, men täcker ibland ett ganska litet geografiskt område.

En noggrann genomgång av /e/ och /i/ i västgötska görs av Landtmanson (1952, s. 27-31). Fornsvenska /ai/ har normalt utvecklats till ett /e/ (”stain” => *sten*), men ett kort /ai/ framför konsonantgrupper blir ofta ett /ɛ/ (*gäspa*). Framför konsonantgrupper kunde man också hitta övergångar till /i/ (*hemma* => ”hemma”, ”hämna” eller ”himma”). Ett långt /i/ i riksspråket brukar bevaras i västgötskan, men kan ibland bli ett /e/ (*gris* => ”gres”). Ett kort /i/ i riksspråket kan kvarstå som ett /i/ eller omvandlas till ett /e/ enligt exemplen från Wessén ovan.

Det finns också en separat studie över korta /i/ och /e/ i Blekinge (Benson 1981, s. 5-20). Han konstaterar att det ofta är svårt att skilja ett kort /i/ från ett kort /e/ i blekingskan. En genomgång av tidigare forskning som refereras visar att samma sak gäller för södra Halland och södra Öland,

och Benson antar att detta sammanfall är av gammalt datum och inte beror på påverkan i modern tid. I Blekinge förekommer både fallet att /i/ => /e/ och att /e/ => /i/.

I en avhandling om språket i östra Småland (Areskog 1936, s. 67-90) noteras att ett kort /e/ i småländskan kan uttalas antingen som en diftong eller som ett /i/ (*hem* => ”him”). Ett kort /i/ i standardsvenska motsvaras nästan alltid av ett kort /i/, men i undantagsfall kan det vara ett kort /e/. Långa /i/ har i allmänhet behållits om långa /i/.

I en senare beskrivning av språket i Kalmar (Magnusson 1978, s. 18-21) noteras att ett kort /e/ kan övergå i ett kort /i/ (*ledsen* => ”lissen”, *spets* => ”spitts”). Ett kort betonat /i/ blir ofta /e/ (*skicka* => ”skecka”, *visp* => ”vesp”), men kan också behållas som /i/. I en motsvarande studie av språket i Jönköping (Lönnerholm 1972, s. 20-23) noteras att /i/ ofta uttalas som /e/ på samma sätt som i andra Götalandsmål. Ett kort betonat /e/ får ibland en dragning mot /ε/, men det nämns inget om att det kan uttalas som /i/.

I en avhandling om skånska dialekter (Hansson 1969, s.76-79) konstateras att det i norra Skåne finns de tre korta fonemen /i/, /e/ och /ε/, men i södra Skåne finns bara /i/ och /e/. I en studie om språket i Lund (Ingers 1970, s. 29-37) visas att ett långt /e/ i standardspråket ofta blir ett /ε/ (*deras* => ”däras”) eller behålls som /e/ (”erhålla”). Kort /e/ blir ofta ett kort /i/ (*hem* => ”himm”, *sett* => ”sitt”). Ett standardspråkligt /i/ behålls ofta som ett /i/, men korta /i/ kan övergå i korta /e/ (*inte* => ”ännte”, ”ennte”).

I gotländskan behålls standarduttalet av kort betonat /i/, men vokalen kan förlängas (Wessén 1968, s. 48).

Det kan vara på plats att göra en snabb jämförelse med danskan också (Karman 2000, s. 65-66, 100). I danskan uttalas ett långt /e/ normalt som [e:] och ett kort betonat /e/ i allmänhet som [œ] eller som [æ]. I diftonger kan /e/ uttalas som [a] (*dejlig* => [dajli], *nej* => [naj]). Ett /i/ uttalas normalt som ett /i/, men korta /i/ uttalas ofta som [e] (*spil* => [spel]).

Det materialet om svenska dialekter som har refererats hittills bygger i allmänhet på information från personer födda på 1800-talet. Det finns få noggranna studier om svenska dialekter från mitten av 1900-talet och fram till idag. En av avsikterna med SweDia-projektet var att fylla den luckan.

Bruce (2010, s. 205-214) menar att de svenska dialekterna förändras med tiden och att en viss utjämning har skett. Orsaken antas vara att folk flyttar oftare och längre bort och på det viset träffar folk som talar på ett annat sätt än de själva, snarare än påverkan från massmedia. Det som ändrar sig långsammast i dialekterna är uttalet och ljudsystemet. Från SweDia-materialet drar han också slutsatsen att yngre personer har färre dialektala drag än äldre och att samma sak gäller för kvinnor jämfört med män.

För dialekterna i Västergötland och Bohuslän finns det dock en omfattande studie, ”Dialektutjämning i Västsverige” (Svahn & Nilsson 2014, s. 37-39, 79-80, 106), där förändringar i dialekten undersöks. Författarna har jämfört inspelat material med sex informanter från mitten av 1900-talet med inspelningar från sex informanter från 2000-talet och intervjuer med gymnasieelever i samband med att studien skrevs. Informanterna i första gruppen var födda under 1800-talet och i den andra gruppen på 1950- eller 1960-talen, medan gymnasieeleverna var födda på 1990-talen. Sammanfallet /i/ => /e/ finns hos alla informanterna i gruppen födda på 1800-talet och ingen av dem har standardspråksuttalet /i/. I gruppen med informanter födda i mitten av 1900-talet har alla både dialektuttalet och standarduttalet i olika omfattning, medan nio

av de knappt hundra gymnasieeleverna har dialektuttal i någon omfattning. Detta tyder på att dialektuttalet håller på att försvinna, men att det fortfarande finns kvar i den äldre generationen.

I Leinonens "An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects" (Leinonen 2010) undersöks variationen över landet i uttalet av varje vokal för sig. Avhandlingen bygger på samma material som används i den här studien men för hela landet. Där görs en bedömning av uttalet av vokaler /i/ och /e/, som undersöks i den här studien.

Ordet *lett* från verbet "leda" uttalades i fornnordiska uttalades som "lajt" och i SweDia-materialet finns detta uttal bevarat hos några informanter (Leinonen 2010, s. 92). Yngre talare har en tendens till ett öppnare uttal än det normala uttalet [e], och variationen i uttalet av den här vokalen varierar förhållandevis mycket. Ordet *leta* hade i fornnordiska diftongen /ai/, som finns kvar på Gotland. Den diftong som finns i sydsvenska mål skiljer sig från den ursprungliga diftongen (Leinonen 2010, s. 91).

Vokalen i ordet *disk* uppvisar liten variation över generationer och områden i landet (Leinonen 2010, s. 89). Samma sak borde gälla för vokalen i *sitta*, som inte behandlas separat. För vokalen i ordet *dis* tenderar yngre talare att ha ett mer öppet uttal än äldre (Leinonen 2010, s. 88-89).

Om man sammanfattar detta så borde man kunna hitta uttalet "sett" för *sitt* åtminstone i götamålsområdena, eventuellt också i områdena för sydsvenska mål. Däremot inte i norra delen av Östergötland eller på Gotland. Sammanfallet /i/ => /e/ behöver inte ske för samma ord i götamål och sydsvenska mål. För *lett* kan man förvänta sig uttal med både /e/ och /ɛ/, och i de sydsvenska områdena borde man troligen kunna hitta uttalet /i/.

3.2 SweDia-projektet

I SweDia-projektet samlades exempel på talad svenska in från drygt hundra orter i Sverige och svensktalande Finland under åren 1998 – 2003 (Schaeffler 2005, s. 57-61, Leinonen 2010, s. 45-53). Projektet var ett samarbete mellan universiteten i Lund, Stockholm och Umeå. Materialet har sedan använts för att studera variationer i svenskt tal.

På varje ort valdes tolv informanter ut. Det var tre äldre män, tre äldre kvinnor, tre yngre män och tre yngre kvinnor, även om det i vissa fall finns någon informant mer eller mindre i någon grupp. De flesta personer i den äldre gruppen var mellan 55 och 75 år och i den yngre gruppen var de mellan 20 och 35 år. Det motsvarar födelseår på 1925-45 för den äldre gruppen och 1965-80 för den yngre gruppen.

Det insamlade materialet är uppdelat i fyra korpusar. En av dem innehåller spontantal, där informanterna samtalar med eller berättar något för intervjuaren i minst 15 minuter. En del av detta material finns tillgängligt på <http://swedia.ling.gu.se/>. De tre andra korpusarna är avsedda för språkvetenskaplig forskning. De är inriktade på prosodi, kvantitet och fonem.

För de tre vetenskapliga korpusarna har intervjuaren försökt elicitera de sökta orden genom att ge informanten ledtrådar. Informanten upprepade sedan det rätta ordet fem gånger. Detta har spelats in och sedan bearbetats vidare.

Figur 1 visar vilka orter i södra Sverige som användes för insamling av materialet. Man har företrädesvis valt orter på landsbygden och undvikit större städer. Orterna finns förtecknade i Appendix 1. Malmö, Göteborg och Stockholm har inte använts för insamling av material, utan är med på kartan för att underlätta orienteringen.



Figur 1: Geografisk placering av orterna i SweDia-projektet i södra Sverige. Från Leinonen (2010, s. 51).

I den del av spontantalskorpusen som är tillgänglig för forskning, men inte i den del som är tillgänglig för allmänheten, finns uppgifter om informanternas namn och ålder. Ljudinspelningarna är benämnda som "floby_om_1" för "old man" nummer 1 i Floby, och i databasen med extraherade formantfrekvenser är datan lagrad under ett löpnummer för varje informant. Det finns sedan en lista som kopplar löpnumret till beteckningarna för ljudinspelningarna. I den här studien används löpnumret för identifiering av informanterna och i Tabell 5, 6 och 7 används också identifieringen från ljudinspelningarna men inte informanternas namn.

3.3 Metoder för att analysera vokaler

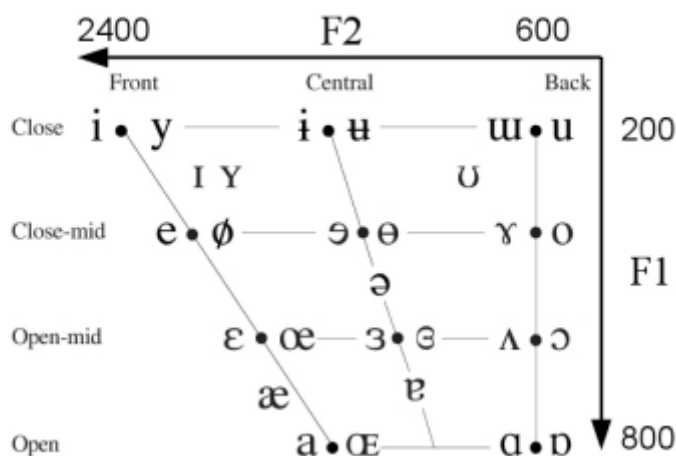
Detta avsnitt ger en bakgrund till hur vokaler kan beskrivas ur källa-filterperspektivet och redogör för två metoder att karakterisera vokaler.

3.3.1 Källa-filer-perspektivet

En enkel ansats för att studera språkljud fonetiskt är källa-filter-perspektivet. Det fungerar både för konsonanter och vokaler, men för konsonanterna är det lite mer komplicerat än för vokaler. Framställningen följer "Talets akustik och perception" (Lindblad 2011) där inget annat anges.

Vokaler är alltid tonande, och det innebär att det är stämläpparna som är källan för ljudet. Talröret från struphuvudet till läpparna betraktas som ett resonansrör. Man kommer då att få resonans för de frekvenser som har en udda multipel av en kvartsvåglängd som är lika lång som talröret. En vuxen person får då resonanser vid ungefär 500 Hz, 1500 Hz, 2500 Hz och så vidare. Dessa resonansstoppas kallas inom fonetiken för formanter. Genom att ändra läget på tungan och läpparna kan utformningen av talröret påverkas och på det sättet förskjuts resonansstopparna.

Det finns två enkla sätt att beskriva vokaler fonetiskt. Det ena är att karakterisera dem med tre parametrar: bakre eller främre, öppen eller sluten och med eller utan läpprundning. Det andra är att beskriva vokalerna med frekvenserna för formanterna. Detta åskådliggörs i Figur 2. Fonetiska tecken för vokaler är inplacerade i en romb som brukar kallas vokalfyrsidingen. Den påminner om en genomskärning av munhålan med läpparna åt vänster. Placeringen av varje vokal motsvarar då ungefär tungans högsta punkt i munnen. [i] kan beskrivas som en främre sluten vokal utan läpprundning, och [i] sitter då längst upp till vänster i vokalfyrsidingen, medan [a] är en främre öppen vokal utan läpprundning och finns då längst ned till vänster. Denna uppställning kan kombineras med frekvensskalorna för de två första formanterna F1 och F2 som i Figur 2.



Figur 2: Vokalfyrsidingen med inritade axlar för formanterna F1 och F2. Från Leinonen (2010, s. 25)

Trots att den här beskrivning inte återger läpprundningen eller frekvensen för F3, som är knuten till läpprundningen, har vokalfyrsidingen visat sig mycket användbar. Den återfinns både i läroböcker och artiklar.

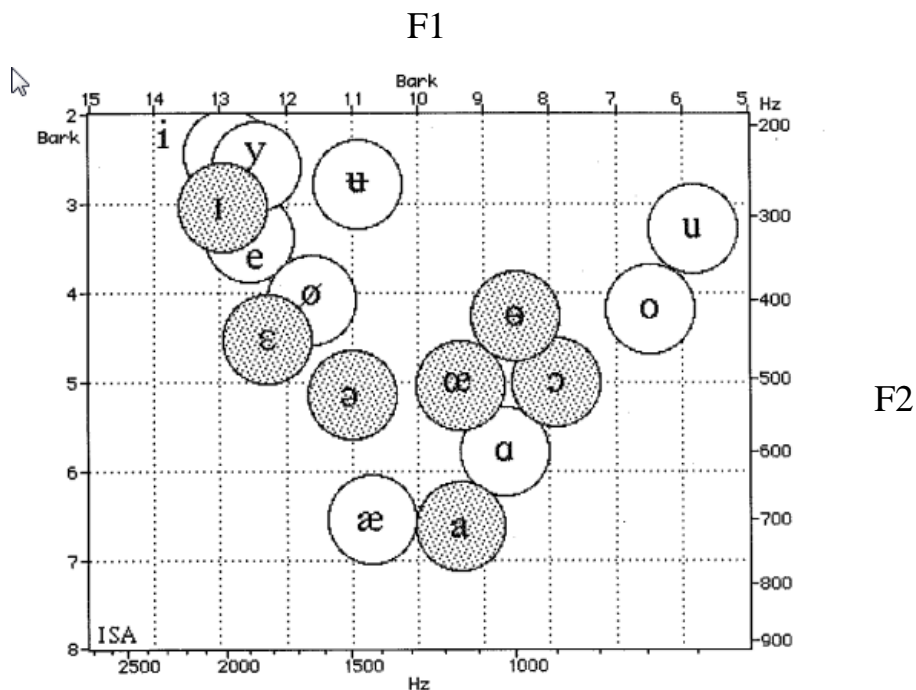
Det som avgör om en röst låter ljus eller mörk är frekvensen för stämtonen som bildas vid stämläpparna. Ju större struphuvudet är desto lägre frekvens får stämtonen. Motsvarande gäller för formanternas resonansfrekvenser. Ju längre avståndet från struphuvudet till läpparna desto lägre blir formantfrekvenserna. Detta kan uppfattas som att vokalfyrsidingen i Figur 2 förskjuts medan axlarna för F1 och F2 ligger fast. Effekten av detta är att om en barnröst och en vuxen persons röst uttalar en vokal med samma formantfrekvenser, så kan det uppfattas som ett [u] (vokalen i *bot*) i första fallet och ett [o] (vokalen i *båt*) i andra fallet.

Detta är ett besvärligt problem när man analyserar tal. Man vill ofta komma till rätta med det genom att skjuta tillbaka vokalfyrsidingarna så att de sammanfaller med varandra. Detta kallas normalisering. För att en normalisering ska fungera bra vill man ta bort faktorer som är knutna till personens röstläge men behålla dialektala drag och andra egenheter.

Det finns flera sätt att göra normaliseringen, och Lobanovs metod är en av dem som oftast ger bäst resultat (Adank mfl 2004). Då relaterar man formanterna i den vokal man vill undersöka till en sammanvägning av formanterna för alla vokalerna i språket, uttalade av samma person. Det värde man får då är svårare att tolka än en frekvens i Hz, men å andra sidan kan man på enkelt sätt jämföra olika talare med varandra.

Ett vanligt sätt att representera vokalljuden är Barkskalan, som är knuten till perceptionen av talet. Den är konstruerad så att vi uppfattar en skillnad mellan två vokaler om skillnaden i frekvens mellan dem är mer än 1 Bark. Upp till 500 Hz är 1 Bark ungefär 100 Hz, däröver blir

1 Bark ungefär 20% av frekvensen. Figur 3 visar båda skalorna i samma diagram. Där kan man notera att Barkskalan i diagrammet är linjär medan frekvensskalan i Hz inte är det. Därigenom blir alla ringarna i figuren lika stora. Man noterar också att några av ringarna ligger ovanpå varandra, och då ska man komma ihåg att man egentligen borde ha en tredimensionell bild som också omfattar F3, men den bilden blir svårare att tyda. Två ringar som ligger nästan ovanpå varandra i Figur 2 kan alltså vara skilda i F3.



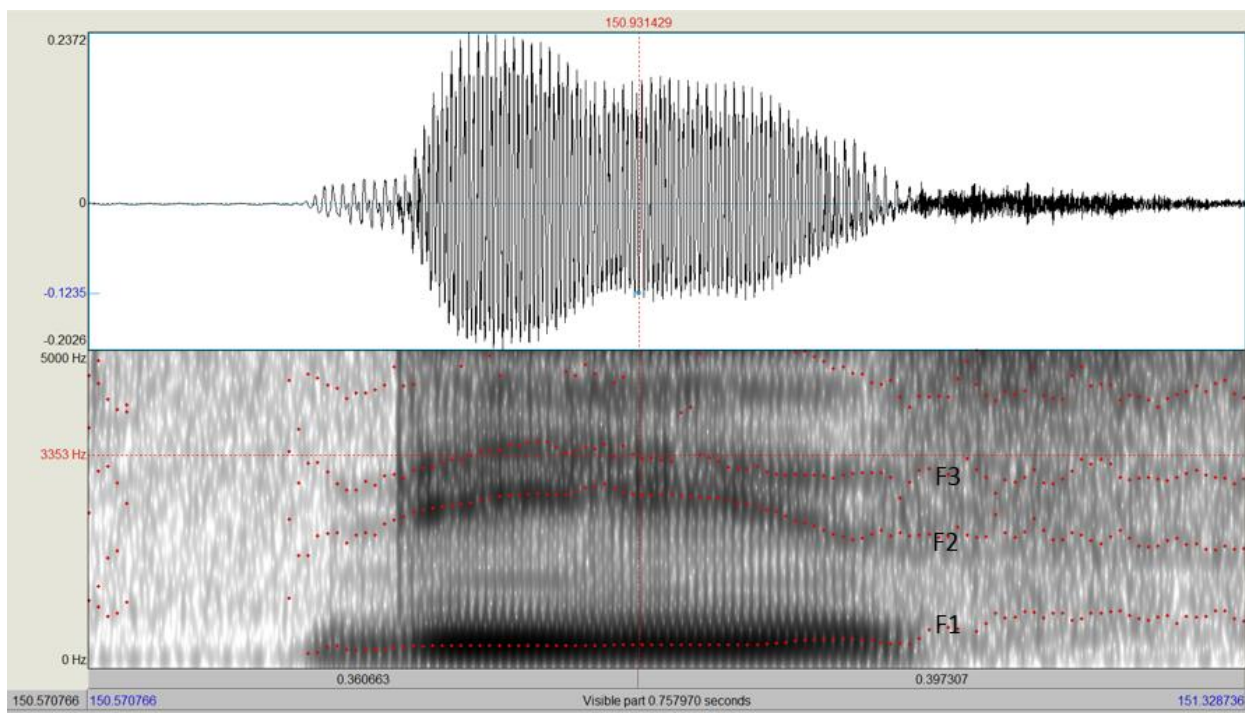
Figur 3: Samband mellan akustisk perception och formantfrekvenser. Mörka cirklar markerar de korta vokalerna och ljusa cirklar de långa. Från Iivonen (1994, s. 85)

Ringarna i Figur 3 markerar var de olika vokalerna typiskt ligger i svenskan, men det finns en stor individuell variation. Det är heller inte säkert att en och samma talare alltid ligger nära mitten av cirkeln för en vokal, utan man kan utnyttja hela ytan. Var i cirkeln en vokal ligger beror bland annat på vilka konsonanter som står intill vokalen och hur de påverkar uttalet. Om cirklarna för två vokaler närmar sig eller överlappar varandra för en persons uttal kommer de som lyssnar inte att kunna skilja vokalerna åt.

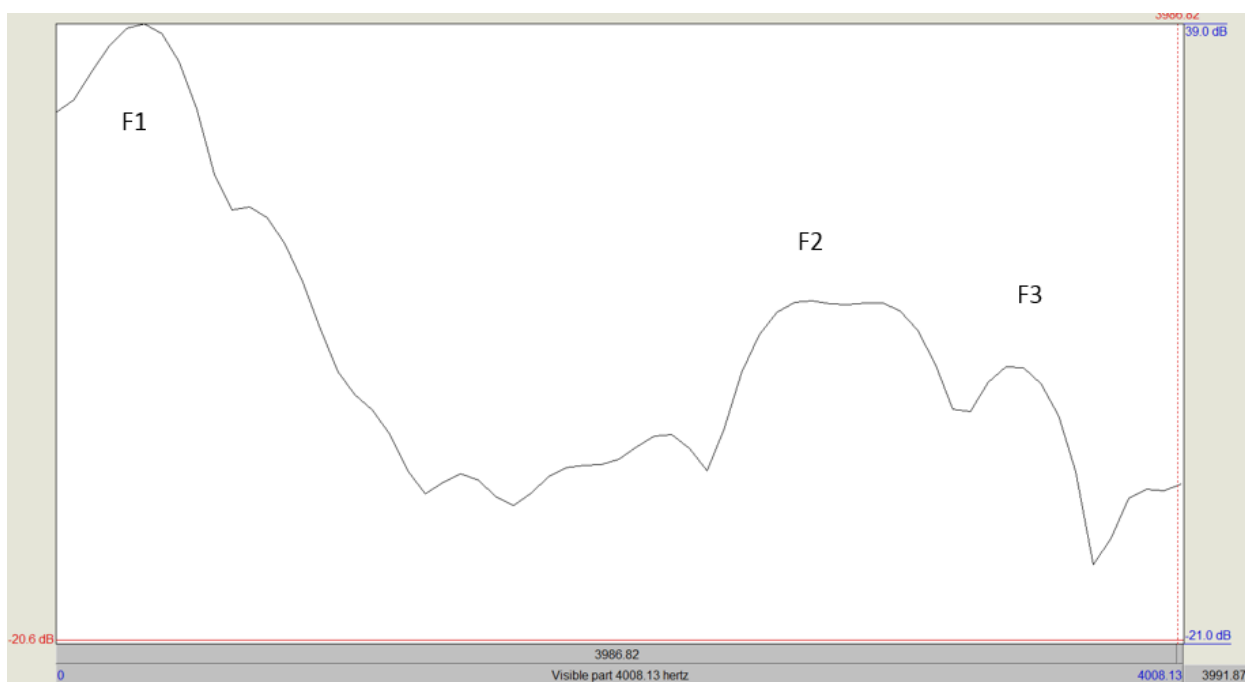
Två fenomen som påverkar analysen av vokaler är koartikulation och diftonger. Koartikulation beror på att käke, tunga och läppar inte hinner ställa om sig mellan fonemen, utan fonemen färgar av sig på varandra. Detta gäller i synnerhet korta vokaler, som är det som i första hand studeras i den här studien. Diftonger rör däremot främst långa vokaler, där en vokal kan börja som ett fonem och sluta som ett annat. I båda fallen är det svårt att göra automatiska analyser.

3.3.2 Formantmetoden

Den vanligaste metoden att analysera vokaler är att studera formanterna (Bruce 2010, s.106-110). Den kan illustreras med ett exempel på hur det har gjorts på det material som används som utgångspunkt för den här studien. Den övre kurvan i Figur 4 är ett vågdiagram som visar hur amplituden av ljudtrycket från inspelningen varierar över tiden. Den nedre kurvan är ett spektrogram, som visar frekvensinnehållet i den akustiska signalen över tiden. Svärtningen i spektrogrammet anger amplituden för en viss frekvens. Figur 5 visar hur amplituden varierar med frekvensen för en viss tidpunkt i Figur 4.



Figur 4: Vågdiagram (överst) och spektrogram (underst) för ordet "dis", intalat av informant 316 i SweDia-materialet. De tre formanterna F1, F2 och F3 är markerade.



Figur 5: Diagram över amplitud (y-axeln) som funktion av frekvens för ordet "dis", tagen vid det svaga lodräta strecket i Figur 4. De tre formanterna F1, F2 och F3 är markerade.

Bestämningen av formantfrekvenserna från en ljudinspelning innehåller följande steg:

1. Den akustiska signalen spelas in och lagras digitalt. För SweDia-materialet samplades den med 32 kHz vid inspelningen och i det lagrade material har man gått ned till 16 kHz. I varje samplingspunkt finns ett mått på signalens amplitud. Det är detta som syns i vågdiagrammet i Figur 4.

2. Sedan tar man ut ett tidsfönster ur vågdiagrammet och omvandlar det till en kurva av den typen som finns i Figur 5. Man kan tänka sig att vågdiagrammet är sammansatt av periodiska signaler med olika frekvens och amplitud. Det Figur 5 visar är hur amplituden varierar med frekvensen vid en enda sådan tidpunkt. Vid extraktionen av formantfrekvenserna som används i den här studien integrerades antagligen signalen inom ett intervall på 25 ms.
3. Genom att flytta det 25 ms långa tidsfönstret framåt i tiden får man en följd av ögonblicksbilder som fogas samman till spektrogrammet. Mera svärtning motsvarar här högre amplitud i Figur 5. Vid formantextraktionerna som används i den här studien var tidsintervallen 10 ms. I diagrammet i Figur 4 ligger de tätare.
4. När man sedan bestämmer formanterna för en vokal i spektrogram eller automatiskt väljer man ut vid vilken tidpunkt man vill göra extraktionen och tar ut en kurva av typen i Figur 5 från spektrogrammet. Där letar man upp de frekvenser där amplituden har lokala maxima. Detta är vokalens formanter.

Man noterar i Figur 5 att formanterna kan ha spetsiga eller flacka toppar. Det påverkar hur bra man kan bestämma frekvensen. Man ser också att amplituden för formanterna varierar, men efter att man väl har bestämt formantfrekvensen används inte informationen om amplituden.

Om man lagrar datan med 16 kHz samplingsfrekvens betyder det att man har 16 000 samplingspunkter per sekund. Man bör sampla med dubbla hastighet jämfört med den maximala frekvensen man vill kunna studera. Det betyder att vi borde kunna analysera frekvenser upp till 8 kHz i det här materialet. Vi vet också att frekvenserna för F1, F2 och F3 ligger under 5 kHz för vokalerna, så detta utgör ingen begränsning.

Tidsfönstret på 25 ms innehåller 400 mätpunkter och vi får in 40 tidsfönster på en sekund. Detta betyder att varje tidsfönster får in en hel period av en signal med 40 Hz. De lägsta frekvenserna för F1-formanten ligger inte under 100 Hz, så detta ger heller ingen begränsning för frekvenserna på den låga sidan.

När man väljer längden på tidsfönstret måste man ta hänsyn till tidskonstanterna för talgesterna. I svenska är medelvärdet för längden på en kort vokal drygt 100 ms och för en lång vokal ungefär 160 ms (Schaeffler 2005, s. 12). I SweDia-materialet som används i den här studien finns korta vokaler som är ned till 50 ms. För att få bra signalkvalitet vill man egentligen ha ett långt tidsfönster, men det får inte vara så långt att frekvensbilden ändrar sig. För korta vokaler kan det vara ett problem. På 50 ms får man in 5 extraktionsspunkter med 10 ms tidsskillnad, men med ett integrationsfönster på 25 ms kanske det bara bli ett par punkter som återger vokalen utan störningar från de intilliggande fonemen.

3.3.3 Andra metoder än formantmetoden

Ett problem med att extrahera vokalernas formantfrekvenser är att det kan bli mycket fel när extraktionen görs automatiskt (Ferrange & Pellegrino 2010a). De senaste decennierna har det blivit vanligare att använda andra metoder för att extrahera akustisk data. En sådan metod användes av Leinonen (2010, s.71-83), och den har två steg.

1. Man börjar med att filtrera ljudet med bandbredden 1 Bark. Det behöves ungefär 20 sådana filtreringar för att täcka upp det frekvensintervall där de tre första vokalformanterna ligger. För varje sådan filtrering får man ett amplitudvärde. Integrationen görs över 13 ms och fönstret flyttas framåt i tiden över hela vokalens utsträckning i tiden.

2. När man har gjort Barkfiltreringarna av hur en person uttalar de olika vokalerna i ett språk gör man en matematisk analys av detta för att se vad som varierar och vad som är konstant. Det som varierar beskrivs med ortogonala principalkomponenter för den personen.

Principalkomponenterna motsvarar på något sätt formantfrekvenserna. I Leinonens avhandling förklarar de två första principalkomponenterna ungefär 65% av variationen mellan vokalerna och tar man med den tredje blir det 83%. Principalkomponent 1 och 2 har bra korrelation med formantfrekvenserna F1 och F2. Tolkningen av vad en principalkomponent egentligen är är svårare än för en formantfrekvens. I Leinonens avhandling täcks vokalerna in med ± 2 enheter av de båda första principalkomponenterna.

En fördel med denna metoden jämfört med formantmetoden är att mer akustisk information tas med i analysen. Bland annat finns amplituden för den akustiska signalen med i analysen, vilket den inte gör i formantmetoden. En annan fördel är att det går lättare att automatisera analysen än med formantmetoden och att den inte kräver manuella korrigeringar i samma utsträckning (Leinonen 2010, s. 26-28).

4. Material och metod

4.1 Material

Materialet i den här studien utgår från SweDia-korpusen om fonem. Där har informanterna talat in varje ord fem gånger. Vokalerna för informanter i Götaland har segmenterats ut. I samband med det har vokalen bedömts auditivt. Därefter har vokalernas formanter extraherats. Intervallet mellan extraktionerna är 10 ms och vid extraktionerna har troligen ett integrationsfönster på 25 ms använts. Denna information är lagrad i tre filer, som beskrivs närmare i Appendix 3. Detta material var tillgängligt från början av den här studien. Utöver detta har korpusarna för spontantal och fonem använts för lyssning.

En uppfattning om materialets omfång ges i Tabell 1. Det är alltså 457 talare från 37 orter, vilket innebär ungefär 12 talare per ort. Det är i de flesta fall tre äldre män, tre äldre kvinnor, tre yngre män och tre yngre kvinnor. För varje informant har i genomsnitt 29 olika ord segmenterats och formantfrekvenserna extraherats. De tre ord som används i den här studien har talats in av åtminstone 448 talare. I Appendix 2 finns en förteckning över de vanligaste orden i materialet och exempel på hur de har eliciterats. Man ser också att antalet extraktioner per intalad vokal är 12, vilket betyder att en typisk vokal är $12 \times 10 \text{ ms} = 120 \text{ ms}$ lång. Den här studien rör korta vokaler och där vi kan ha ned till 5 extraktioner per vokal i enstaka fall.

Antal orter	37
Antal talare	457
Antal talare * antal olika vokaler	Cirka 13 100
Antal inlästa vokaler	65 547
Antal formantextraktioner	770 090
Antal talare per ort	12
Antal intalade ord per talare	29
Antal extraktioner per intalat ord	12

Tabell 1: Nyckeltal för mängden data för Götaland i den här studien.

De tre ord som har valts för studien finns i Tabell 2. Urvalet är begränsat till de knappt 30 ord som finns i Swedia-materialet, vilket betyder att de flesta av svenskans 18 vokalljud finns med en eller två gånger. Helst skulle man vilja ha minimala par av ord där konsonantkontexten är densamma och bara vokalen skiljer. Detta är inte möjligt i SweDia-materialet, även om konsonanterna i Swedia-materialet har valts så att de är alveolara. För att studera kort /i/ finns orden *sitt* och *disk*, och för kort /e/ finns bara *lett*. Som huvudord för kort /i/ valdes *sitt* eftersom det ordet har samma avslutande konsonant som *lett*, även om den inledande konsonanten skiljer. Motsvarande långa vokaler i SweDia-materialet finns bara i orden *leta* och *dis*.

Ord	Vokal	Kommentar
sitt	/i/	Av verbet <i>sitta</i> eller det possessiva pronominet <i>sitt</i>
lett	/e/	Av verbet <i>leda</i> eller verbet <i>le</i>
disk	/i/	Använda tallrikar och bestick

Tabell 2: Ord som studerats i den här uppsatsen.

Det är olyckligt att orden inte har eliciterats på samma sätt på alla orter. Det är inte självklart att *lett* från verben *leda* och *le* uttalas på samma sätt, och samma sak gäller för de två varianterna av *sitt*.

4.2 Metod

För att välja metod för undersökningen har i första hand tre avhandlingar och två artiklar studerats.

- Leinonens avhandling "An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects" (Leinonen 2010), där variationen i uttal av enskilda vokaler i SweDia-materialet studeras. Leinonen använder metoden med principalkomponenter.
- Weners avhandling "När lögnare blir lugnare" (Wenner 2010) studerar sammanfall mellan kort /ø/ och kort /œ/ i Uppland. Undersökningen bygger på eget intervju-material och formantmetoden användes.
- Schaefflers avhandling "Phonological Quantity in Swedish Dialects" (Schaeffler 2005) använder kvantitetskorpusen från SweDia-projektet. Där studeras kvantitet, dvs längd för vokaler och konsonanter, men också vokalernas formanter analyseras.
- Ferrange & Pellegrinos artiklar "Formant frequencies of vowels in 13 accents of the British Isles" (Ferrange & Pellegrino 2010a) och "Vowel systems and accent similarity in the British Isles" (Ferrange & Pellegrino 2010b). Båda behandlar samma utgångsmaterial med brittiska dialekter. I den första artikeln görs analysen med formantmetoden och i den andra en alternativ metod som är besläktad med den metod Leinonen använde.

4.2.1 Val av metod i tidigare studier

Wenner (2010, s. 55-62) använde minimala ordpar, som bara skiljer sig på om vokalen är /œ/ eller /ø/. På det viset kunde påverkan från de omgivande konsonanterna för uttalet av vokalen undersökas. I studien intervjuades totalt 78 personer, 60 olika ord användes och 3570 förekomster av vokalerna analyserades. Wenner använde automatisk bestämning av formantfrekvenser med två metoder, men på grund av risken för fel kontrollerades alla formantfrekvenserna i spektrogrammen. Eftersom båda vokalerna hon studerade är rundade har den väsentliga analysen begränsats till F1 och F2, men F3 har tagits med när akustiska avstånd har beräknats. Båda de undersökta vokalerna är korta, och för att minska påverkan från koartikulation har formantvärden nära mitten av vokalen extraherats. I undersökningen har ingen normalisering använts, förutom i vokaldiagrammen där Lobanovs metod har använts. Informanterna har upprepat varje ord minst tre gånger, och Wenner har normalt använt den första inspelningen av ordet för analysen.

Wenner (2010, s. 67-87) har först beräknat medelvärden för vokalernas formantfrekvenser och sedan använt dessa medelvärden för att bestämma det akustiska avståndet. Hon jämför 13 minimala ordpar. Ordparet med minst skillnad har ett akustiskt avstånd på strax under 1 Bark och ordparet med störst skillnad har strax över 2 Bark. Detta stämmer ungefär med Figur 3. En informant kan ha en spridning över 6 Bark för uttalet av en och samma vokal i olika konsonantkontexter.

Schaeffler (2005, s. 64, 97-99) studerade fonemens längd, men tittade också på hur detta hängde samman med vokalernas kvalitet. Han undersökte skillnaden i akustiskt avstånd mellan de långa och korta varianterna av vokalerna. Då använde han bara F1 och F2 och extraktioner nära vokalens mittpunkt. Därefter tog han medianvärdet av F1 och F2 för de fem inspelningarna av vokalen, och det var dessa värden han använde för att beräkna det akustiska avståndet för varje informant. Detta uttrycktes i Bark. Motiveringen för att välja medianvärdet i stället för medelvärdet var att minska påverkan från avvikande värden.

Vid en jämförelse mellan /a/ i *tak* och *tack* fick Schaeffler (2005, s.97-99) ett akustiskt avstånd på strax under 1 Bark för svensktalande Finland, som har liten skillnad i kvalitet. För stora delar Götaland var det ungefär 3 Bark. Detta är ett något större akustiskt avstånd än vad man skulle vänta sig från Figur 3.

Leinonen (2010, s. 55-83) använde sig av metoden med principalkomponenter för att beskriva vokaler. Anledningen är att detta är en säkrare metod som fungerar bättre vid automatisk extraktion av datan. I analysen använde hon de två första principalkomponenterna. För varje vokal gjorde hon nio extraktioner. Den första låg vid 25% av vokalens längd och den i sista vid 75% av längden. Alla värdena användes för beräkning av principalkomponenterna, men de diagram som blev resultatet av avhandlingen baseras på den första och den sista extraktionen. På det sättet kunde hon avgöra om en vokal var en diftong eller inte. Varje extraktion integrerades över 13 ms.

Ferrange & Pellegrino (2010a) använde en korpus sammanställd 2003 för att studera vokaler i 13 olika brittiska dialekter med formantmetoden. Materialet bestod av inläsningar av 19 olika vokaler i en /hVd/-kontext, där V markerar vokalen. Varje vokal talades in fem gånger. Materialet innehåller både män och kvinnor, men bara inspelningarna från de 145 männen användes, eftersom mäns och kvinnors röster skiljer så mycket att det är svårt att analysera dem tillsammans i samma datamängd även efter normalisering. En automatisk bestämning av formantfrekvenserna gjordes och om den inte matchade spektrogrammet uteslöts den observationen från studien. På det sättet sorterades 20% av materialet bort. För varje vokal gjordes sedan en ny extraktion av formanterna så att man fick 13 punkter för varje vokal. Detta följdes av en Barktransformering och en linearisering för att utesluta kraftigt avvikande punkter. Slutligen gjordes en Lobanovnormalisering. Eftersom detta kräver att varje informant har läst in alla vokaler förlorade man 70% av de kvarvarande informanterna i det här steget.

På grund av det stora bortfallet av material gjorde författarna (Ferrange & Pellegrino 2010b) om studien med en alternativ extraktionsmetod. Ljudsignalen filtrerades i 12 frekvensintervall med ett integrationsfönster på 20 ms och ett intervall mellan extraktionerna på 10 ms. Då kunde också material med både manliga och kvinnliga röster användas.

4.2.2 Val av metod i den här studien

Metoden att behandla datan valdes med de tidigare studierna som bakgrund.

- Formantmetoden för extraktion av formantfrekvenser användes. Detta motiveras med att denna data redan fanns tillgänglig för Götaland, även om det objektivt sett borde varit fördelaktigt att göra en ny extraktion med den metod som Leinonens använde. Men de flesta analyserna av vokaler har gjorts med formantmetoden, så också den metoden borde vara tillräckligt bra.
- Datan från den mittersta extraktionen och den mittersta inläsningen användes. Eftersom studien rör korta vokaler valde jag, precis som Wenner, att använda den mittersta extraktionen för att minska påverkan av koartikulation. Jag valde också att använda den mittersta inläsningen, eftersom jag tyckte att den var mest representativ. Det var attraktivt att använda ett medianvärde eller medelvärde av de fem upprepningarna för att minska spridningen i frekvens, men genom att alltid basera analysen på en bestämd extraktion kunde jag gå tillbaka och analysera den på ett sätt som man inte kan göra med medelvärden eller medianer.

- Både F1, F2 och F3 användes i analysen. Det är naturligt att använda alla de tre första formanterna vid beräkningarna, eftersom alla tre har betydelse för perceptionen av vokalerna. I diagram visas normalt bara F1 och F2 för att undvika tredimensionella bilder.
- Beräkning av akustiskt avstånd gjordes i Bark. Alternativet är att använda skillnaden i frekvens uttryckt i procent av den ena frekvensen. Detta är lättare att förstå tekniskt sett än Bark. Att Bark ändå valdes beror på att det återger perceptionen på ett bättre sätt. 1 Bark motsvarar ungefär det akustiska avståndet mellan två närliggande vokaler. Formeln som användes för omvandling från Hz (f i formeln) till Bark (z i formeln) är (Traummüller, 1990):

$$z = \frac{26.81 * f}{1960 + f} - 0.53 \quad (1)$$

Det akustiska avståndet AD mellan två vokaler räknades ut som roten ur summan av kvadraterna på skillnaden i frekvens för F1, F2 och F3 uttryckt i Bark (här betecknar x och y de två vokalerna som jämförs):

$$AD = \sqrt{(F1y - F1x)^2 + (F2y - F2x)^2 + (F3y - F3x)^2} \quad (2)$$

I Appendix 3 beskrivs också hur databehandlingen har gjorts i den här studien. I princip har de tre filerna från formantextraktion ur SweDia-materialet slagits ihop och sedan har den mittersta extraktionen av vokalen och den mittersta inspelningen selekterats ut. Därefter har bara de fem ord som innehåller /e/ eller /i/ behållits. Den filen innehåller då en rad per talare för vart och ett av de fem orden. Det är den filen som har använts vid all dataanalys utom när spridning mellan upprepningar eller extraktioner från en och samma vokal studerats.

Vid den vidare behandlingen av datan har formantextraktionerna från två ord lagts ihop i en fil. Sedan har skillnaden i de tre formanternas frekvenser och det akustiska avståndet i två och tre dimensioner beräknats. Denna data har sedan använts för att generera diagrammen i studien.

Eftersom det är känt att formantmetoden ofta ger felaktiga värden på de extraherade frekvenserna använde man en automatisk korrektion vid extraktionen av F1 och F2 i det råmaterial som har använts i den här studien. Rättningen genomfördes på F1 och F2, men inte på F3. 0.5% av extraktionerna i materialet hade korrekationer på F1 och 5% hade det på F2. Det som oftast händer när det blir fel är att F2 får värdet för F1 och vid rättningen ges då F2 den frekvensen som F3 hade från början. Då får F2 och F3 samma värde. De automaträttade extraktionerna togs av den anledningen bort innan analyserna i den här studien gjordes.

Sammanfallet kort /i/ => /e/ undersöks i studien genom att analysera det akustiska avståndet mellan vokalerna i *sitt* och *lett*. Sedan kontrolleras också samma sak mellan *sitt* och *disk* för att få grepp om variationen i uttalet av /i/. Vid behov kunde också jämförelser med de långa vokalerna i *dis* och *leta* göras.

De fyra tidigare studierna har alla i första hand studerat medelvärde eller median medan spridningen har haft mindre betydelse. I den här studien är det också intressant att undersöka fördelningarnas svansar långt från medianen, eftersom det är där man kan se sammanfallet mellan /e/ och /i/. Detta beror på att bara en mindre del av talarna borde ha det sammanfallet.

Dataanalysen har gjorts i R (version 3.3.1) i utvecklingsmiljön RStudio (version 0.99.903). Ljudfilerna har analyserats i Praat (version 6.0.23).

5. Resultat och analys

5.1 Mätningar av akustiskt avstånd mellan kort /i/ och kort /e/.

Idén med den här studien är att undersöka det akustiska avståndet mellan vokalerna i *sitt* och *lett* för att se om landskapen i Västsverige har fler talare med en tendens att uttala vokalen i *sitt* som /e/ i stället för /i/.

5.1.1 Auditiv bedömning av uttalet av kort /i/ och kort /e/ i SweDia-materialet

I samband med segmenteringen av SweDia-materialet gjordes en bedömning av uttalet av vokalerna. Om man går igenom de auditiva bedömningarna får man resultatet som visas i Tabell 3 för *sitt* och i Tabell 4 för *lett*. Vid sammanräkningen har olika allofoner slagits ihop till fonem.

	/i/	/e/	/y/
Antal observationer	441	10	1
Procent	98%	2%	0%

Tabell 3: Bedömningar av vokalen i *sitt* i SweDia-databasen.

Av de 452 talare i Götaland som talat in ordet *sitt* skulle det alltså bara vara 10 som haft uttalet med /e/ och dessutom en som haft uttalet /y/. Det är mycket lägre än väntat och antyder att /e/-uttalen knappast finns kvar i Götaland. Av de tio som har /e/-uttalet är åtta äldre män, en äldre kvinna och en yngre kvinna. Den geografiska fördelningen var fyra från Skåne, tre från Västergötland, två från Bohuslän och en från Småland. Detta stämmer med slutsatsen i avsnitt 3.1 om var man kan förvänta sig att hitta sammanfallet /i/ => /e/, men inte med den intuitiva uppfattningen att det är vanligast i Västsverige. Antalet observationer är dock för lågt för att man ska kunna dra tydliga slutsatser.

	/i/	/e/	/ɛ/	/a/
Antal observationer	13	326	98	7
Procent	3%	73%	22%	2%

Tabell 4: Bedömningar av vokalen i *lett* i SweDia-databasen.

444 talare i Götaland har talat in ordet "lett" i SweDia-databasen. Här är resultatet mera splittrat och uttalet har en större spridning. Om man går från [i] till [a] täcker man hela vänstersidan på vokalfrysningen i Figur 3. Uttalet verkar vara delat mellan /e/ och /ɛ/, och där är det i första hand den yngre gruppen som har uttal med /ɛ/. /e/-uttalet finns i hela Götaland. Av de 13 uttalen med /i/ finns 8 i Blekinge, och det är i första hand äldre personer som har det uttalet. Uttalen med /a/ finns i Småland, Skåne och Blekinge, och också här dominerar äldre talare.

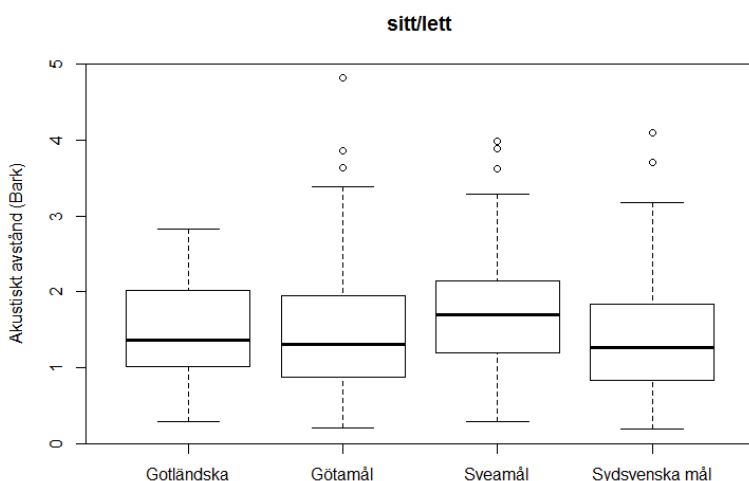
Vid beräkningen av det akustiska avståndet betyder detta att en variation i första hand beror på en spridning i uttalet av /e/. En jämförelse med Figur 3 antyder att akustiska avstånd upp till 2 Bark borde vara ganska vanliga, medan avstånd på över 3 borde vara ovanliga.

5.1.2 Geografisk fördelning

Boxplottar över det akustiska avståndet för vokalerna i ordparet *sitt/lett* i Götaland visas i Figur 6, Figur 7 och Figur 8 uppdelat på dialekt, landskap och ort. I Appendix 1 finns en tabell över i vilka landskap de olika orterna ligger och vilket dialektområde de tillhör.

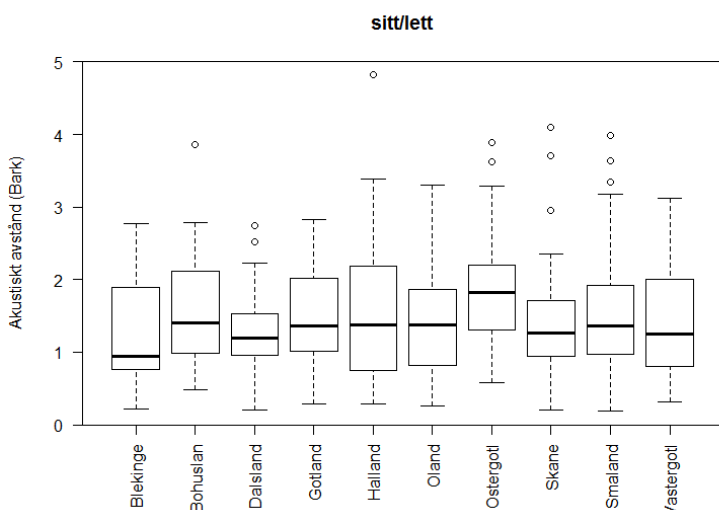
I boxplottarna i anger det tjocka vågräta strecket medianvärdet för det akustiska avståndet. De nedre och över gränserna för boxarna anger första och tredje kvartilen, dvs 25 % av värdena ligger utanför boxarna på var sida om boxen. De små vågräta strecken ovanför och under boxarna anger statistiska beräkningar för var de normala fördelningarna slutar. Om det finns små ringar i diagrammen så markerar det värden som avviker så mycket att de ligger utanför den normala fördelningen.

Figur 6 visat att medianvärden och spridning för de fyra dialekterna i Götaland ser ganska lika ut, förutom att det akustiska avståndet för sveamålet ligger lite högre. Detta gäller i första hand Östergötland och beror på att uttalet av /e/ är öppnare där, enligt diskussion i avsnitt 5.1.1. Medianen för avståndet ligger mellan 1 och 1.5 Bark för de tre andra dialekterna. Detta är lite högre än väntat men inte orimligt.



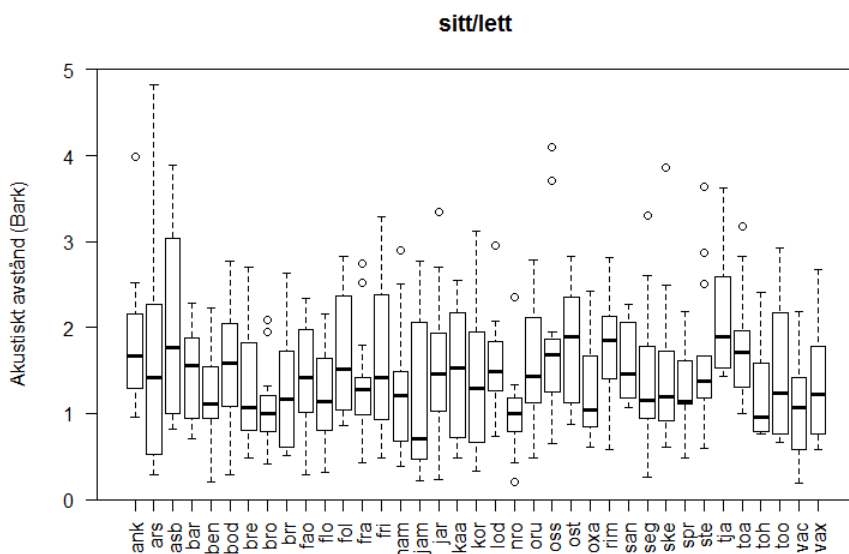
Figur 6: Akustiskt avstånd för vokalerna i sitt och lett uppdelat på dialekt. Y-axeln visar avståndet i Bark.

Figur 7 visar samma sak uppdelat på landskap. Medianvärdet ligger från 1 Bark upp till 1.5 Bark utom för Östergötland, som ligger lite högre.



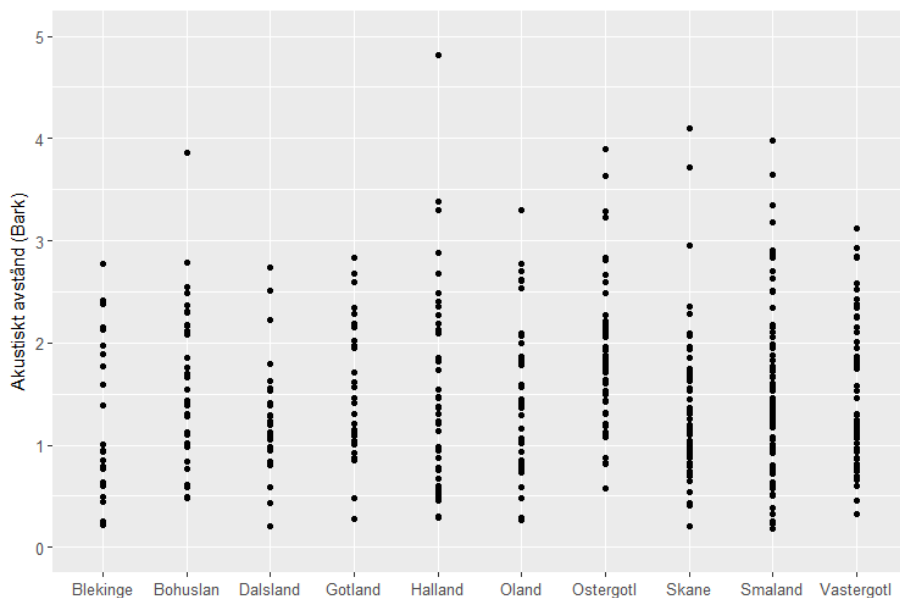
Figur 7: Akustiskt avstånd för vokalerna i sitt och lett uppdelat på landskap. Y-axeln visar avståndet i Bark.

Figur 8 är uppdelad per ort. Eftersom det är färre observationer på en ort än i ett landskap blir spridningen mellan orterna större än mellan landskapen. Medianvärden sprider från strax under 1 Bark upp till 2 Bark.



Figur 8: Akustiskt avstånd mellan *i* sitt och *lett* uppdelat på orter. Y-axeln visar avståndet i Bark. Trebokstavsförkortningarna för orterna återfinns i Appendix 1.

Variationen i akustiskt avstånd i boxplottarna har inte med sammanfallet */i/ => /e/* att göra utan visar att typiska talare på olika orter har en skillnad i uttalet av minst en av vokalerna. Eftersom man kan anta att bara en mindre del av informanterna har sammanfallet */e/ => /i/* eller */i/ => /e/* måste man studera diagram av den typen som finns i Figur 9 för att undersöka var man har de sammanfallen. I diagrammet finns en prick för varje informant.



Figur 9: Akustiskt avstånd för vokalerna i *sitt* och *lett*. Y-axeln är avståndet i Bark.

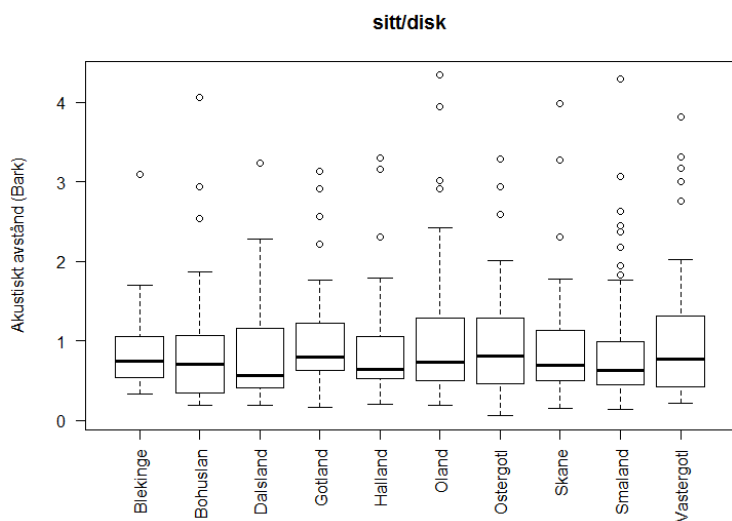
Vi förväntar oss att ett typiskt avstånd mellan */i/* och */e/* ska vara knappt 1 Bark enligt Figur 3. För ett sammanfall borde avståndet rimligen vara under 0.5 Bark. Det finns ett antal sådana

observationer, men de är inte särskilt många. De är ganska jämnt spridda mellan landskapen och andelen sådana observationer för de västsvenska landskapen Västergötland och Bohuslän är färre än för de flesta andra landskapen. Man noterar också att det finns åtskilliga observationer med ett avstånd på mer än 3 Bark.

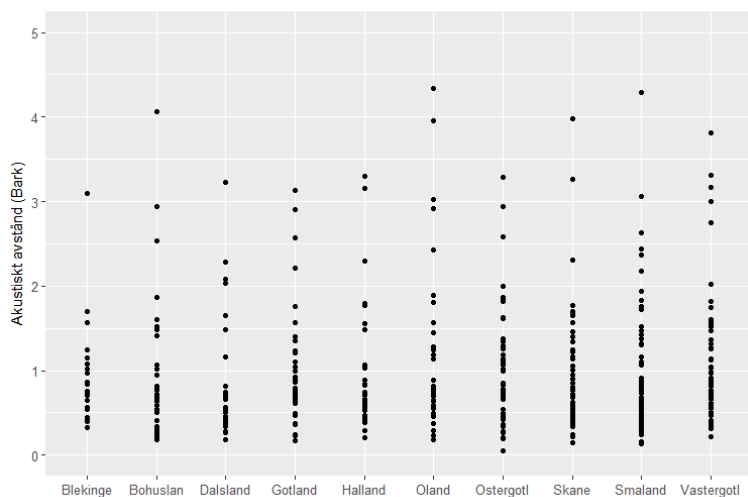
Om man lägger ihop de 13 informanter som hade sammanfallet /e/ => /i/ med de 10 som hade /i/ => /e/ får man totalt 23 personer med sammanfall mellan vokalerna. Det stämmer ungefär med antalet informanter som har ett avstånd under 0.5 Bark i Figur 6. En noggrannare kontroll visar att det inte är någon total överensstämmelse mellan de som har bedömts ha sammanfall och de som fått ett lågt avstånd vid beräkningen.

Så här långt finns det inget som tyder på att sammanfallet /i/ => /e/ skulle vara särskilt vanligt i Götaland. För att förstå detta bättre undersöks också det akustiska avståndet mellan vokalerna i *sitt* och *disk*. Det är i båda fallen ett kort /i/, och då förväntar man sig ett sammanfall för alla utom dem uttalar de båda vokalerna som olika fonem.

Figur 10 och Figur 11 visar diagrammen för *sitt* och *disk* uppdelat på landskap. Det är få observationer som ligger nära 0 Bark. Medianvärdet är omkring 0.7 Bark med liten spridning mellan landskapen. Detta antyder att det finns en märkbar skillnad i uttalet av vokalen i de båda orden. Det finns också många värden som ligger över 3 Bark, och detta är en orimligt stor skillnad.



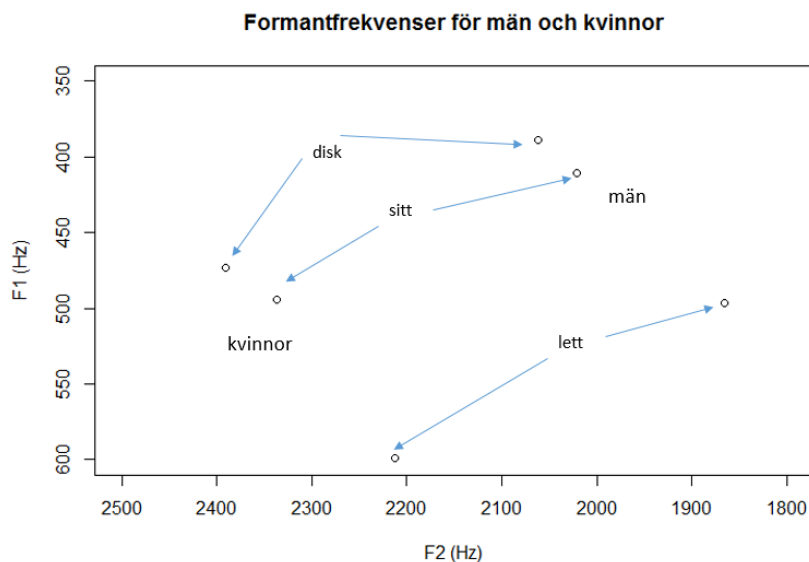
Figur 10: Akustisk avstånd för vokalerna i *sitt* och *disk*. Y-axeln visar avståndet i Bark.



Figur 11: Akustiskt avstånd för vokalerna i *sitt* och *disk*. Y-axeln är avståndet i Bark.

Figur 12 visar medelvärdena av frekvenserna för F1 och F2 för orden *sitt*, *disk* och *lett*. Datan är inte normaliserad och av den anledningen visas män och kvinnor separat. Det finns en stor spridning för varje punkt som inte syns i diagrammet. Vid beräkningen av medelvärdet finns ingen kvadratisk term med. Värdena för enskilda informanter ligger spridda runt medelvärdena.

Här förväntar vi oss att *sitt* och *disk* ska ligga nära varandra medan *lett* borde ligga på ett avstånd på drygt 1 Bark. Det finns en systematisk skillnad mellan vokalerna i *sitt* och *disk*, som kan bero på olika konsonantkontexter. Avståndet mellan vokalerna i *sitt* och *disk* är 0.25 Bark om man beräknar det för både F1, F2 och F3. Skillnaden är systematisk, så det finns en verklig skillnad i uttalet, men den är inte lika stor som medianvärdet i Figur 10 antyder. Förutom den systematiska skillnaden i avstånd måste det finnas ytterligare något bidrag till det man ser i Figur 10.



Figur 12: Frekvenserna för formanterna F1 och F2 för hela Götaland för vokalerna i *sitt*, *disk* och *lett*.

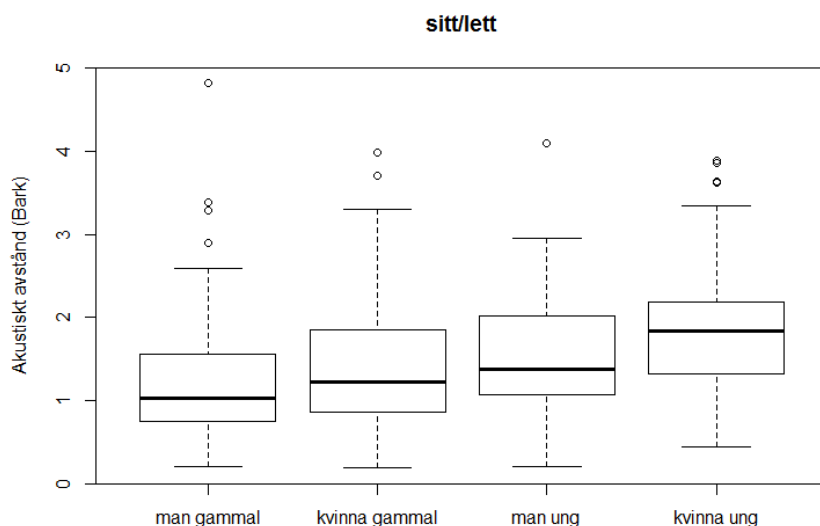
I Figur 9 kan man inte se att Västsverige skulle ha en högre andel sammanfall där /i/ => /e/ än andra delar av Götaland. Det finns däremot en viss spridning i medianvärdet av det akustiska avståndet mellan landskapen. De delar av Götaland där man talar sveamål har ett större avstånd

än andra områden. Dessutom verkar det finnas en tydlig skillnad i akustiskt avstånd mellan vokaler i *sitt* och *disk*. Detta tillsammans gör att man inte kan dra några tydliga slutsatser av resultaten av undersökningen, och jämförelserna med de långa vokaler blir inte meningsfulla. Däremot är det fortfarande intressant att undersöka vad det här kan bero på.

5.1.3 Åldersfördelning

Man kan använda samma material för att undersöka om det finns en skillnad i akustiskt avstånd mellan gamla och unga personer. SweDia-material ger en möjlighet att undersöka grupperna gammal mot ung och man mot kvinna. I flera andra studier har man påpekat att även om man använder normalisering helst bör analysera data för män och kvinnor separat (Wenner 2010, s. 67 och Leinonen 2010, s. 67-70, 151-152).

En sådan analys för hela materialet visas i Figur 13. Man ser att spridningen är ungefär samma men det akustiska avståndet är något lägre för äldre personer än för yngre. I båda åldersgrupperna har männen något lägre akustiskt avstånd än kvinnorna. Eftersom andelen informanter i hela materialet med sammanfall mellan kort /e/ och /i/ är lågt kan man inte vänta sig att se någon tydlig skillnad mellan yngre och äldre i det avseendet.



Figur 13: Akustiskt avstånd för hela materialet för ordparet *sitt/lett* delat på äldre män, äldre kvinnor, yngre män och yngre kvinnor.

5.2 Felkällor knutna till insamlingen av materialet

För att göra en bra bedömning av materialet måste man också lyssna på inspelningarna för att se om det ger någon ledtråd till de oväntade resultaten. En slumpvis avlyssning av det eliciterade materialet från olika landskap bekräftar bilden att uttalet med /e/ i *sitt* är ovanligt.

5.2.1 Auditiv kontroll av punkter med litet eller stort akustiskt avstånd

En auditiv kontroll gjordes av vokaler som hade låga och höga avstånd i Figur 9. Man kan lägga märke till att skillnaden i uttal mellan de fem upprepningarna av samma ord kan vara stor, och ofta är det den första som avviker mest. Man noterar också att sättet eliciteringen gjorts på

varierar. Olika meningar används vid eliciteringen, ibland med olika betydelse för de sökta orden. I några fall har intervjuaren läst upp orden för talaren, något som inte var meningen. Det finns också en skillnad på hur intervjuaren talar. Ibland är det på standardsvenska och ibland är det på ett sätt som ligger närmare informantens dialekt.

De lägsta värdena på akustiskt avstånd låg strax under 0.5 Bark. De flesta var sammanfall där *lett* uttalade som "litt" och de kom geografiskt från Småland och Blekinge.

De höga värdena på akustiskt avstånd låg på 4 till 5 Bark. Geografiskt låg tyngdpunkten i Halland och Västergötland. I många fall var det talare som uttalade *sitt* med /i/ och *lett* med /ε/. Jämför man detta med frekvensskalan överlagrad på vokalfyrstidningen i Figur 3 ser man att detta kan vara en rimlig förklaring till värden på akustiskt avstånd på åtminstone upp till 2 Bark. Det finns emellertid många punkter med avstånd på 3 Bark som vid lyssning inte verkar ha så stor skillnad i uttal.

Summan av detta är att många av de höga värdena verkar vara korrekta. De beror oftast på stor spridning i uttalet av *lett*, men det finns också punkter med stort uppmätt avstånd som inte motsvarar skillnaden man hör när man lyssnar på inspelningarna.

5.2.3 Auditiv kontroll av en inspelningarna från en tre orter

Tre orter valdes ut för en kontroll av det eliciterade materialet. Öxabäck ligger i södra Västergötland och är den ort i SweDia-materialet som ligger närmast därifrån jag själv kommer. Där borde /e/-uttalet av /i/ finnas, åtminstone i den äldre gruppen. Jämshög i Blekinge är orten med flest fall där ett kort /e/ uttalades som /i/. Ankarsrum i nordöstra Småland är en av de orter som har stort akustiskt avstånd mellan vokalerna i *sitt* och *lett* kombinerat med förhållandevis liten spridning.

I tabellerna nedan jämförs den ursprungliga bedömningen av uttalet av vokalen i *sitt* och *lett* i den eliciterade delen av SweDia-materialet med min egen bedömning. Dessutom har jag gjort en grov bedömning av informanternas uttal av samma vokaler i spontantalsdelen. Där är det svårare att göra en bra bedömning, eftersom varje ord bara uttalas en gång i taget och ofta inte lika tydligt som i den eliciterade delen. Man bör också försöka basera bedömningen på vokaler i ungefär samma konsonantkontext som *sitt* och *lett*. Slutligen anges det beräknade akustiska avståndet för det eliciterade materialet.

Tabell 5 visar resultatet för det inspelade materialet från Öxabäck. Där borde /e/-uttalet finnas, och jag borde själv kunna avgöra när något i inspelningen låter konstigt.

Både bedömningen när SweDia-materialet samlades in och min egen bedömning är att alla tolv informanterna har tydliga /e/ i *lett* och tydliga /i/ i *sitt* vid eliciteringen. Vi borde då förvänta oss att det akustiska avståndet skulle vara någonstans omkring 1 Bark, men det är stor spridning. Det finns några informanter som ligger lite lägre och några som ligger lite högre.

Detta jämfördes sedan med spontantalet. Man kan allmänt säga att särskilt de äldre, men också flera av de yngre, hade ett uttal med en tydlig dialektal prägel. För de äldre talarna övervägde /e/-uttalet av kort /i/ i ord där man kunde förvänta sig det uttalet. Det gäller uttal som "netton", "feck", "jeck", "metten", "mett" och "feske" för *nitton*, *fick*, *gick*, *mitten*, *mitt* och *fiske*. Vid eliciteringen har man uppenbarligen missat detta. Man får en känsla av att informanterna har lagt sitt språk tillrätta för att anpassa sig till intervjuarna, som hade ett mer standardspråkligt uttal. Detta gäller särskilt vid eliciteringen, men också vid spontantalet.

Informant	speaker_id	Bedömning SweDia <i>sitt/lett</i>	Bedömning idag <i>sitt/lett</i>	Bedömning spontantal <i>sitt/lett</i>	Akustiskt avstånd (Bark)
Öxabäck äldre man	313	i e	i e	i e	0.7
Öxabäck äldre man	314	i e	i e	e e	1.5
Öxabäck äldre man	315	i e	i e	e/i e	1.0
Öxabäck äldre kvinna	316	i e	i e	e/i e	1.0
Öxabäck äldre kvinna	317	i e	i e	-	2.4
Öxabäck äldre kvinna	318	i e	i e	-	1.8
Öxabäck yngre man	319	i e	i e	i e	1.0
Öxabäck yngre man	320	i e	i e	i/e e	0.6
Öxabäck yngre man	321	i e	i e	i e	1.4
Öxabäck yngre kvinna	322	i e	i e	i e	2.4
Öxabäck yngre kvinna	323	i e	i e	i/e e	0.9
Öxabäck yngre kvinna	324	i e	i e	i/e e	0.7

Tabell 5: Bedömning av uttalet av vokalerna i *sitt* och *lett* i materialet från Öxabäck. *i/e* betyder att */i/* överväger och *e/i* att */e/* överväger. Avståndet är för det eliciterade materialet.

Tabell 6 för Jämshög visar en ganska bra överensstämmelse mellan bedömningen i SweDia-materialet och mig. Det finns ett tydligt */i/* i *sitt*, och för *lett* finns det en spridning från */i/* över */e/* till */ε/*. Uttalen med */i/* är vanligast i den äldre generationen och då särskilt för män. Uttalet av båda vokalerna var suddigare i Jämshög än i Öxabäck och av den anledning svårare att bedöma. Man får också en känsla av att informanterna i Jämshög är mer komfortabla med sin dialekt och inte rättar till den vid intervjutillfället.

Vid spontantalet kan man notera en lite svagare tendens mot att uttala kort */e/* som */i/*, men det är en subjektiv bedömning. Det akustiska avståndet ligger för ungefär hälften av informanterna i närheten av 0.5 Bark, men det finns också värden på över 2 Bark.

Informant	speaker_id	Bedömning SweDia <i>sitt/lett</i>	Bedömning idag <i>sitt/lett</i>	Bedömning spontantal <i>sitt/lett</i>	Akustiskt avstånd (Bark)
Jämshög äldre man	64	i i	i i	i i	0.2
Jämshög äldre man	65	i i	i e	i i/e	0.3
Jämshög äldre man	66	i i	i i	i i	1.8
Jämshög äldre kvinna	67	i i	i i	i i/e	0.5
Jämshög äldre kvinna	68	i e	i ε	i i	0.6
Jämshög äldre kvinna	69	i e	i e	i/e e	2.4
Jämshög yngre man	70	i i	i i	i i/e	0.8
Jämshög yngre man	71	i i	i i	i i	0.6
Jämshög yngre man	72	i ε	i ε	i e	2.1
Jämshög yngre kvinna	73	i e	i ε	i e	2.8
Jämshög yngre kvinna	74	i e	i ε	i e	2.0
Jämshög yngre kvinna	75	i i	i e	i e	0.4

Tabell 6: Bedömning av uttalet av vokalerna i *sitt* och *lett* i materialet från Jämshög. *i/e* betyder att */i/* överväger. Avståndet är för det eliciterade materialet.

Tabell 7 för Ankarsrum visar att *sitt* uttalas med ett tydligt */i/* av alla informanterna, men *lett* varierar mellan */e/* och */ε/*. I spontantalet fanns det någon informant som tenderade att uttala kort */i/* som */e/*, men överensstämmelsen med det eliciterade talet var annars bra.

Informant	speaker_id	Bedömning SweDia <i>sitt/lett</i>	Bedömning idag <i>sitt/lett</i>	Bedömning spontantal <i>sitt/lett</i>	Akustiskt avstånd (Bark)
Ankarsrum äldre man	263	i e	i ε	i e	1.7
Ankarsrum äldre man	264	i e	i e	i e	1.3
Ankarsrum äldre man	265	i e	i ε	i e	1.5
Ankarsrum äldre kvinna	266	i e	i e	i e	1.6
Ankarsrum äldre kvinna	267	i e	i e	i e	4.0
Ankarsrum äldre kvinna	268	i e	i ε	i e/ε	1.0
Ankarsrum äldre kvinna	269	i e	i e	i e	2.5
Ankarsrum yngre man	270	i ε	i ε	i e	2.3
Ankarsrum yngre man	271	i ε	i ε	i e	2.1
Ankarsrum yngre man	272	i e	i ε	i/e i	1.2
Ankarsrum yngre kvinna	273	i e	i ε	i ε	1.3
Ankarsrum yngre kvinna	274	i e	i ε	i e	2.1
Ankarsrum yngre kvinna	275	i e	i e	i e	1.2

Tabell 7: Bedömning av uttalet av vokalerna i *sitt* och *lett* i materialet från Ankarsrum. *i/e* betyder att /i/ överväger och *e/ε* att /e/ överväger. Avståndet är för det eliciterade materialet.

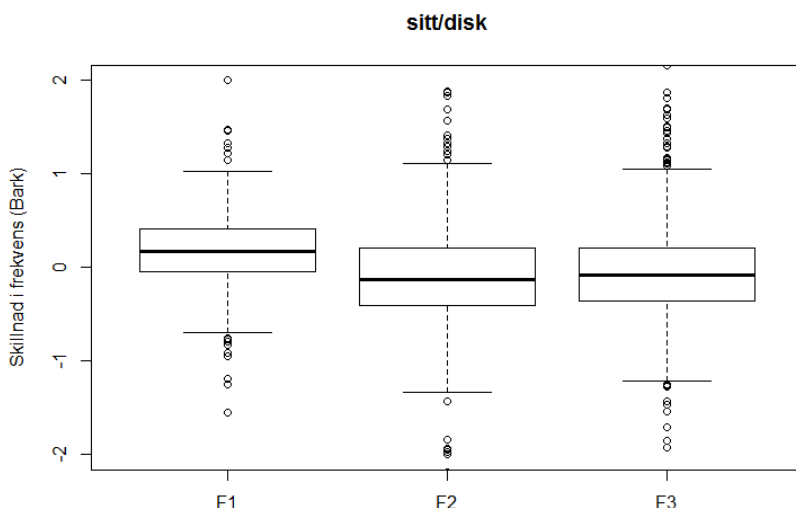
5.3 Felkällor knutna till analysmetoden

Det man noterar först är att värdet av det akustiska avståndet enligt Ekvation (2) aldrig kan bli negativt. Om man tänker sig att man jämför två formantextraktioner som ger exakt samma värden på alla de tre formantfrekvenserna så blir avståndet 0. Om det finns en skillnad mellan extraktionerna för någon formant kommer man alltid att få ett högre värde. Om skillnaden i var och en av de tre formanterna var 0.3 Bark så skulle avståndet vara roten ur 0.27, dvs ungefär 0.5 Bark. Detta stämmer ungefär med den nedre gränsen för fördelningen i Figur 9. Dessutom tittar vi inte bara på extraktionen av en formantfrekvens utan på skillnaden mellan två extraktioner, där båda extraktioner kan ha ett fel.

För att undersöka detta närmare kan man studera om det är så att någon av de tre formanterna sprider mer än de båda andra eller om vi har stor spridning mellan olika inläsningar eller extraktioner.

5.3.1 Skillnad mellan formanterna F1, F2, och F3

Man kan misstänka att någon av formanterna bidrar mer till spridningen än de andra. Detta kan studeras i en boxplot där man tittar på spridningen i frekvens för var och en av formanterna. I Figur 14 visas detta för avståndet mellan vokalerna i *sitt* och *disk*. Eftersom värdet inte är kvadrerat kan skillnaden vara både positivt och negativt. Här förväntar vi oss att alla formanterna ska vara centrerade runt 0 Bark, helst med liten spridning. Alla tre formanterna ser ungefär likadana ut, men man kan notera att F1 har ett medianvärde som är lite positivt och F2 ett som är lite negativt. Detta stämmer med skillnaden mellan *sitt* och *disk* i Figur 12.



Figur 14: Spridning för formanterna F1, F2 och F3 för avståndet mellan *sitt* och *disk*.

Det är fortfarande så att spridningen för varje formant är oväntat stor. Enligt resonemanget ovan ger en stor spridning ett högt värde på det akustiska avståndet. Vi borde heller inte ha några värden på 1 Bark eller mer, utom för informanter som uttalar vokalerna i *sitt* och *disk* som olika fonem. Ringarna i Figur 13 markerar enskilda värden. I materialet ingår mer än 400 värden, så det är ändå en liten andel av dem som ligger över 1 Bark.

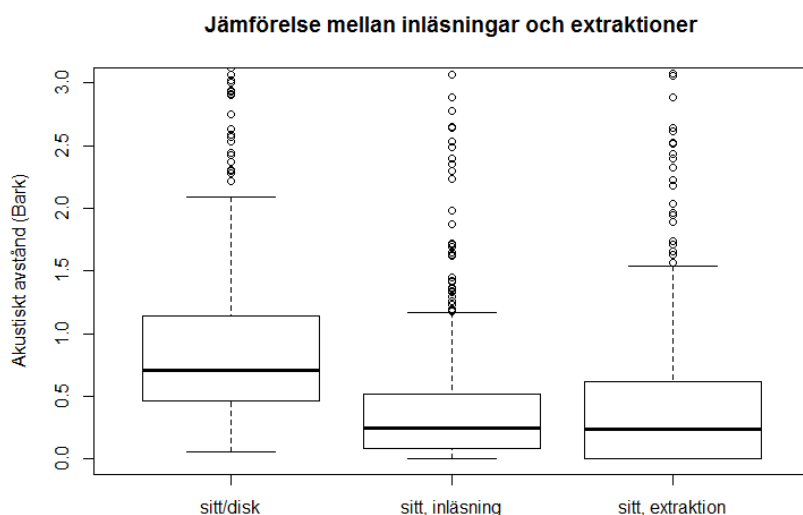
Slutsatsen av det här resonemanget är att orsaken till den stora spridningen och de höga värdena på det akustiska avståndet i Figur 9 är att vi har stor spridning vid extraktionen av formantfrekvenserna. Detta kan antingen bero på att varje informant har en stor naturlig variation i sitt tal eller att metoden för att extrahera frekvenserna inte är tillräckligt bra.

5.3.2 Skillnad mellan upprepningar av orden och olika extraktioner av formanterna

För att förstå varifrån den stora spridningen i avstånd kommer kan man studera upprepade inspelningar av samma ord. Den möjligheten har man i SweDia-materialet, eftersom informanterna upprepade varje ord fem gånger.

Varje kort vokal har också åtminstone fem extraktioner av formanterna med 10 ms tidsskillnad. På samma sätt som för de fem upprepningarna av varje vokal kan man titta på två efterföljande extraktioner för samma vokal för att se hur mycket de varierar.

Figur 15 jämför det akustiska avståndet mellan vokalerna i *sitt* och *disk* med vad man får om man jämför två intilliggande upprepningar av *sitt* av samma informant och två efterföljande formantextraktioner med 10 ms tidsmellanrum. Att medianvärdet är högre i boxplotten längst till vänster beror på att finns en verklig skillnad i uttal av vokalerna i de två orden förutom ett bidrag från spridningen, så som Figur 11 och Figur 12 antyder. Men det är anmärkningsvärt att spridning är nästan lika i stor i de båda boxplottarna till höger.



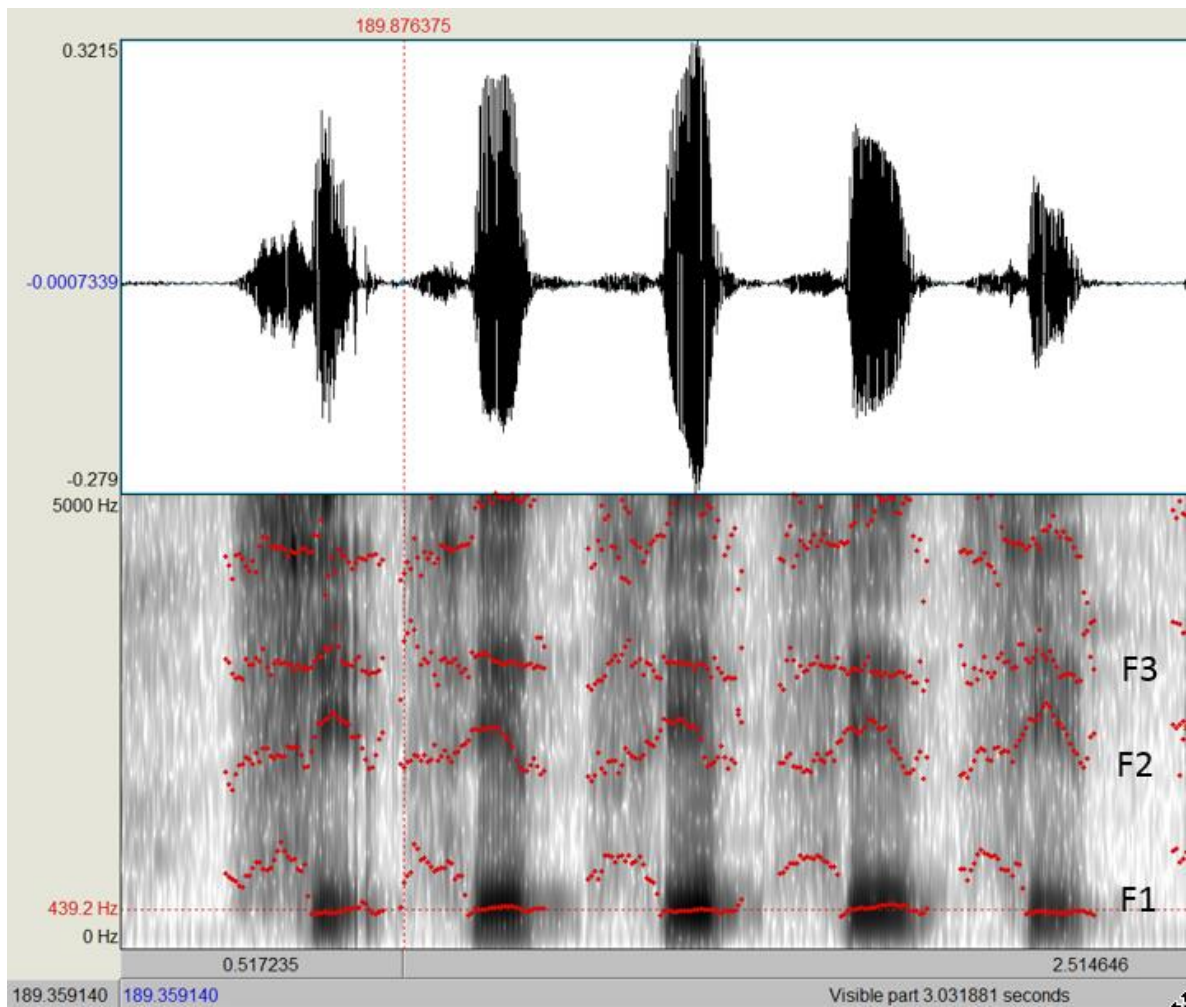
Figur 15: Jämförelse av akustiskt avstånd för hela Götaland mellan *sitt* och *disk*, två påföljande upprepningar av *sitt* och två påföljande extraktioner från en inspelning av *sitt*.

Den stora spridningen mellan olika upprepningar och mellan olika extraktioner bidrar säkert till att det uppmätta avståndet mellan *sitt* och *disk* blir större än vad borde vara. Det kan vara intressant att se hur det ser ut ett spektrogram.

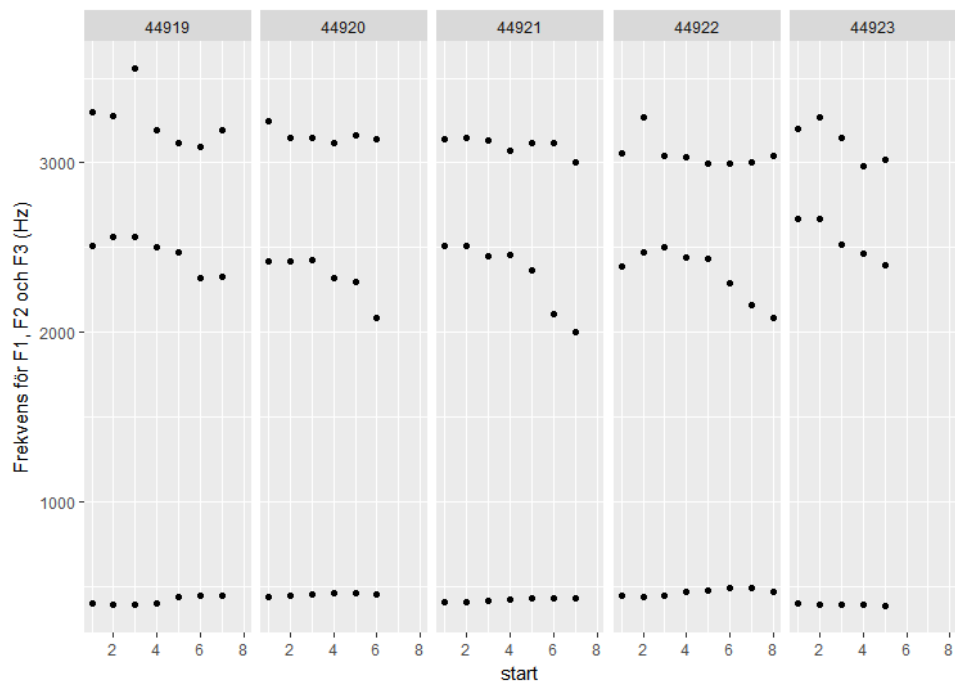
Figur 16 visar vågdiagrammen och spektrogrammen för de fem upprepningarna av *sitt* av en informant. Vågdiagrammen antyder att det är stor skillnad mellan de fem inläsningarna. I spektrogrammen är det de mörka partierna som är vokalen, och särskilt för F2 verkar frekvensen variera kraftigt. För de två sista upprepningarna har man en stark gradient också mitt i vokalen. Detta leder till att man misstänker att det finns en stor naturlig variation i formantfrekvenserna för korta vokaler mellan upprepningar och mellan extraktioner. Figur 4 visar ett spektrogram av *dis* intalat av samma informant. Där är vokalen lång och variationen är mindre men inte försumbar.

Figur 17 visar extraktionerna av vokalsens formantfrekvenser när en informant talar in *sitt* fem gånger. 1 Bark motsvarar ungefär 20% av frekvensen. För F3 finns två punkter som avviker ungefär 0.5 Bark från intilliggande extraktioner. Detta fel hänger antagligen samman med inspelningen eller extraktionen. Man ser också att nivån på F1 skiljer närmare 100 Hz mellan de två sista inläsningarna. Detta återspeglar troligen en verklig skillnad i uttalet av vokalen. F2 varierar kraftigt under sista delen av vokalsens längd, och vilket värde man får vid extraktion beror på vilken av extraktionspunkterna man råkar välja.

Det är tydligt att spridningen i formantfrekvens kan orsakas både av en variation i informantens uttal och av en variation som hänger samman med extraktionen. Inom ramen för den här studien går det inte att avgöra vilken av de två faktorerna som har störst vikt. Vid en kontroll har några orimligt höga akustiska avstånd kunnat förklaras med de spridningar som syns i Figur 16.



Figur 16: Fem inläsningar av ordet sitt av informant 316 i SweDia-materialet. Mellanrummet mellan inläsningarna är bortklippt.



Figur 17: Extraktionerna av formanterna för vokalen i sitt för informant 316.

5.4 Summering av resultaten

Fråga 1: Skillnader i sammanfall mellan kort /i/ och /e/ beroende på region eller ålder.

Analysen av det eliciterade materialet gav inte belägg för att sammanfallet mellan kort /i/ och kort /e/ skulle vara vanligare i Västersverige än i andra delar av Götaland. Det var helt enkelt för liten andel av informanterna som hade det sammanfallet för att man ska kunna dra den slutsatsen. Samma sak gäller skillnaden mellan äldre och yngre talare. Däremot går det att hitta något tiotal informanter som har sammanfallet /i/ => /e/ eller /e/ => /i/.

Det finns också en variation i medianvärdet av det akustiska avståndet mellan kort /i/ och /e/ och i områden där sveamål talas är avståndet större. Variationen är knuten till uttalet av /e/, som varierar mer än uttalet av /i/.

En stickprovsanalys av materialet med spontantal antyder att den förväntade skillnaden verkligen finns, både för Västsverige i förhållande till andra delar i Götaland och mellan äldre och yngre talare i Västsverige. Detta har man missat i det eliciterade materialet.

Om man bara baserar sig på det eliciterade materialet ska alltså båda hypoteserna förkastas. Men en stickprovskontroll av spontantalet pekar ändå på att de kan vara riktiga.

Fråga 2: Möjligheter och begränsningar med SweDia-materialet

SweDia-materialet kan mycket väl användas för analyser av svenska dialekter, men det finns inte bara fördelar utan också nackdelar med det.

Fördelar:

- Materialet innehåller dialektprover från hela landet och det är tillgängligt.
- Korpusen med spontantal gör att man kan undersöka fenomen som inte hade planerats vid insamlingen av det eliciterade materialet.

Nackdelar:

- För de eliciterade korpusarna är man begränsad till det material som har samlats in.
- Som den här studien visar kan man misstänka att informanter i flera fall har lagt sitt uttal till rätta får att komma närmare standardspråket.

Fråga 3: Metoden med akustiskt avstånd baserat på extraherade formantfrekvenser.

Den naturliga spridningen i uttalet eller osäkerheten i extraktionen av formantfrekvenserna gav en för stor spridning, som ledde till att värden på det akustiska avståndet blev större än det borde vara. I den här studien var det inte möjligt att avgöra om det var den naturliga spridningen i uttal eller extraktionsmetoden som bidrog mest till felet.

6. Diskussion och slutsatser

6.1 Diskussion

6.1.1 Sammanfall mellan betonade kort /i/ och kort /e/

Avsikten var från början att använda det eliciterade materialet för att studera uttalet av /e/ och /i/. Men eftersom sammanfallet /i/ => /e/ nästan saknades där måste analysen av detta basera sig på ett stickprov av spontant tal. Detta gav ett för litet underlag för att dra några säkra slutsatser, men vissa tendenser kan man ändå se.

Materialet från korpusen för spontantal bekräftar Svahn & Nilssons slutsats att personer från Västsverige som är födda i mitten av 1900-talen har både uttalet med kort /e/ och med kort /i/ av ord som har kort /i/ i standardspråket. Samma sak gäller slutsatsen att äldre personer tenderar att ha det dialektala uttalet i större utsträckning än yngre personer. Det är också intressant att uttalet av ett kort /e/ som /i/ fanns kvar i Blekinge med omnejd, som man kunde förvänta sig.

Analysen av det eliciterade materialet visar att yngre personer generellt har ett större akustiskt avstånd mellan kort /i/ och kort /e/. Detta stämmer väl med Leinonens slutsats att yngre personer tenderar att ha ett mer öppet uttal av /e/ än äldre samtidigt som variationen i uttalet av /i/ är liten.

Däremot gav en stickprovskontroll inget belägg för att uttalet med /e/ skulle finnas i hela Götaland, som man kunde förvänta sig från genomgången av tidigare forskning i avsnitt 3.1. Man kan misstänka att det uttalet har försvunnit sedan uppteckningarna gjordes. Dessutom väljs ofta äldre informanter när man gör dialektstudier för att få en så genuin dialekt som möjligt. Det kan alltså hända att uttalet med /e/ var på väg att försvinna i stora delar av Götaland redan när uppteckningarna gjordes under första halvan av 1900-talet.

6.1.2. SweDia-materialet

Den mest intressanta iakttagelsen om SweDia-materialet var att man inte hade fått med uttalet av kort /i/ => /e/ vid eliciteringen i samma omfattning som det finns i spontantalet. Detta är inte ett okänt problem. Leinonen (2010, s. 46-47, 154) kommenterar att intervjuaren kan påverka informanten, och noterar att intervjuarna i SweDia-materialet många gånger uppmanar informanten att använda det dialektala uttalet i stället för standardspråksuttal. Detta har jag också observerat.

Också Svahn & Nilsson (s. 219-251) tangerar problemet med hur intervjuaren påverkar informanten i sin undersökning. Deras material innehåller inte någon eliciterad del, men förutom en inspelad intervju fick några informanter göra en egeninspelning utan intervjuare. Det konstaterades att yngre informanter talade ungefär likadant i båda fallen, men flera äldre informanter hade ett mer dialektalt drag vid egeninspelningen.

Jag själv är något yngre än den äldre åldersgruppen i SweDia-materialet och jag har både /e/-uttal och /i/-uttal av ord med /i/ som standarduttal. Jag har en tendens att välja /e/-uttalet när jag pratar med personer som har samma dialekt som jag och /i/-uttalet när jag pratar med andra personer. Det speciella med just uttalet av betonat kort /i/ i Västsverige är att det finns två varianter och talarna kan växla fritt mellan dem. Det är alltså inte frågan om en gradvis övergång utan ett dialektalt drag som en person kan välja att använda eller att inte använda i sitt tal.

När man jämför det eliciterade talet med spontantalet lägger man märke till att informanterna uttalar de eliciterade orden högt och tydligt och på ett medvetet sätt, men att spontantalet är snabbare och mindre genomtänkt. Detta kan också vara en orsak till att spontantalskorpusen har starkare dialektala drag.

Vid eliciteringen vill man att informanterna ska använda sitt normala uttal, men det kan vara svårt att uppnå detta. I den här studien fick man inte fram det naturliga uttalet av /i/ => /e/ vid eliciteringen även om informanterna hade det uttalet vid spontantalet. Däremot verkar glidningen mellan /e/ och /ɛ/ för vokalen i *lett* ha fångats på ett riktigt sätt. Det är oklart hur man ska göra för att få ett så naturligt tal som möjligt vid eliciteringen, men en jämförelse med spontantal kan ge en antydning om ifall man har fått med de karakteristiska drag man är intresserad av.

6.1.3 Akustiskt avstånd

Metoden att ta ut formantfrekvenser för två vokaler, sedan beräkna det akustiska avståndet individuellt och därefter slå samman datan för grupper av informanter fungerade som avsett, men den stora spridningen i bestämningen av formantfrekvenserna gjorde att de beräknade avstånden tenderade att bli för stora. Det kan antingen bero på att formantmetoden inte är tillräckligt robust eller på att den individuella variationen i talet är för stor. I den här studien kan man inte avgöra vilken av de två förklaringarna som har störst betydelse.

I den här studien var det inte i första hand avsikten att studera en hel grupps medelvärde eller medianvärden av formantfrekvenser så som både Leinonen, Wenner och Schaeffler gjort. I stället gällde det att se om det fanns en mindre grupp med avvikande uttal i olika regioner. Då måste man kunna identifiera hur stor andel av informanterna som har ett avvikande uttal. Med metoden som användes var det möjligt att identifiera personer som hade sammanfallet mellan kort /i/ och /e/, men tydliga regionala skillnader kunde man inte se. En orsak till detta var att andelen informanter sammanfall mellan kort /e/ och kort /i/ var oväntat låg i det eliciterade materialet.

Det är ett systematiskt avstånd mellan medelvärdena för vokaler i *sitt* och *disk* på 0.25 Bark och ett avstånd på 0.7 Bark om man i stället börjar med att beräkna avståndet för varje informant och sedan tar medianvärdet av detta. Skillnaden beror på en spridning i bestämningen av formantfrekvenserna som antingen kan vara en naturlig spridning i talet eller vara knuten till extraktionen av frekvenserna. Den naturliga spridningen vill man ha kvar, men den delen som är knuten till extraktionen skulle man vilja minska.

Det finns vissa möjligheter att reducera spridningarna vid extraktionen med formantmetoden som användes i den här studien. Man kan göra en anpassning av de extraherade punkterna till en linje för varje formant och välja en formantfrekvens som ligger på linjen. På det sättet undviker man risken av att välja punkter som avviker mycket från sin omgivning. På samma sätt kan man minska spridningen genom att ta medianvärdet eller medelvärdet på fem upprepningarna av samma ord. Det är ändå tveksamt om detta är tillräckligt. De 5% av extraktionerna som automatkorrigeras tyder på att det finns ett mera fundamentalt problem.

Diskussionerna om behovet av manuell kontroll vid automatisk extraktion med formantmetoden hos Leinonen (2010, s 24-26), Wenner (2010, s. 60-63) och Ferrange & Pellegrino (2010a, s. 4-7) tillsammans med resultatet av den här studien visar att man bör välja någon av de robustare extraktionsmetoder som Leinonen (2010) eller Ferrange & Pellegrino (2010b) använde. Det borde ge mindre spridning knuten till extraktionsmetoden, men den naturliga spridningen i talet blir naturligtvis kvar.

6.2 Slutsatser

Undersökningen av sammanfallet mellan kort /i/ och /e/ i Västsverige gav inte något tydligt resultat. Det är uppenbart att SweDia-materialet inte har någon stor andel informanter som har detta sammanfall i den eliciterade delen, men enstaka informanter som har sammanfallet kunde hittas. Däremot antyder en analys av spontantalsdelen att /e/-uttalet är mer utbrett än vad som framgår av den eliciterade delen och motsvarade den förväntade omfattningen i Västsverige. Det är också tydligt att de är de äldre informanterna som har sammanfallet i större utsträckning än yngre.

Om man vill göra en större studie av dialekter i Sverige kan man knappast göra den enbart från materialet i SweDia eftersom man då är begränsad till de ord som eliciterats. Den naturliga metoden är då att samla in eget material. Däremot är det lämpligt att inleda med en analys av SweDia-materialet innan man samlar in eget material. Då är spontantalsdelen en bra utgångspunkt som kan ge indikationer om just det som man är intresserad av.

Den valda metoden att studera akustiskt avstånd tenderade att ge för höga värden på det akustiska avståndet. Spridningen i bestämningen av formantfrekvenserna var för stor. Det är inte klart om det största bidraget till detta beror på spridning i extraktionen med formantmetoden eller om det beror på att talarna har en stor spridning i sitt normala tal.

7. Referenser

- Adank, P. & Smits, R. & van Hout, R. (2004): A comparison of vowel normalization procedures for language variation research., *Journal of the Acoustic Society of America*, 116 (5), 3099-3107.
- Areskog, G. (1936). *Östra Smålands Folksmål*. Avhandling, Stockholms Högskola. Lund: Carl Bloms Boktryckeri.
- Benson, S. (1981). *Blekingska dialektstudier*. Lund: Gleerupska universitetsbokhandeln.
- Bruce, G. (2010). *Vår fonetiska geografi*. Lund: Studentlitteratur AB
- Ferrange, E. & Pellegrino, F. (2010a). Formant frequencies of vowels in 13 accents of the British Isles., *Journal of the International Phonetic Association*. 40 (1), 1-34.
- Ferrange, E. & Pellegrino, F. (2010b). Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics. *Journal of Phonetics*. 38, 526-539.
- Hansson, Å. (1969). *Fonematiska studier i skånska dialekter*. Lund: Studentlitteratur.
- Iivonen, A. (1994). A psychoacoustical explanation for the number of major IPA vowels. *Journal of the International Phonetic Association*, 24 (2), 73-90.
- Ingers, I. (1970). *Språket i Lund: dialekt och riksspråk i Lund och omnejd. 2:a upplagan*. Stockholm: Norstedts
- Karman, H. (2000). *Tio lektioner i danska*. Studentlitteratur.
- Landtmanson, S. (1952). *Västgötamålet*. Uppsala: Appelbergs Boktryckeri.
- Leinonen, T. (2010) *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects*. Groningen
- Lindblad, P. (2011) *Talets Akustik och Perception*. Lund: Lunds Universitet
- Lönnerholm, E. (1972). *Språket i Jönköping*. Stockholm: Läromedelsförlagen Svenska Bokförlaget.
- Magnusson, B (1978). *Språket i Kalmar*. Stockholm: Esselte Studium,
- Praat: (<http://www.fon.hum.uva.nl/praat/>)
- R: (<https://cran.r-project.org/>)
- RStudio: (<https://www.rstudio.com/>)
- Schaeffler, F. (2005) *Phonological Quantity in Swedish Dialects*. Umeå: Umeå Universitet
- Svahn, M. & Nilsson, J. (2014) *Dialektutjämning i Västsverige*. Göteborg: Institutet för språk och folkminnen

SweDia-materialet: <https://corpora.humlab.lu.se/ds/asv/?0> (allt inspelat material) och <http://swedia.ling.gu.se/> (delar av spontantalskorpuserna tillgängliga för allmänheten).

Traunmüller, H. (1990): Analytical Expressions for the tonoptic sensory scale, *J Acoust. Soc. Am.*, 88 (1), 97-100.

Wenner, L. (2010) *När lögnare blir lugnare*. Uppsala: Uppsala Universitet.

Wessén, E. (1960). *Våra folksmål. 6:e upplagan*. Lund:Fritzes.

Wessén, E. (1968). *Svensk språkhistoria. 8:e upplagan*. Stockholm: Almqvist&Wiksell

Appendix 1

Orter i Götaland som ingår i SweDia-projektet.

Dialekten för orterna ansluter sig till Wesséns geografiska indelning (Wessén 1960, s. 13-15, 23, 30).

Nummer	Ortsnamn	Kortform	Landskap	Dialekt
1	Ankarsrum	ank	Småland	Sveamål
2	Årstad	ars	Halland	Götamål
3	Asby	asb	Östergötland	Sveamål
4	Bara	bar	Skåne	Sydsvenska mål
5	Bengtsfors	ben	Dalsland	Götamål
6	Böda	bod	Öland	Sveamål
7	Bredsättra	bre	Öland	Sveamål
8	Broby	bro	Skåne	Sydsvenska mål
9	Burseryd	brr	Småland	Götamål
10	Fårö	fao	Gotland	Gotländska
11	Floby	flo	Västergötland	Götamål
12	Sproge	spr	Gotland	Gotländska
13	Stenberga	ste	Småland	Götamål
14	Tjällmo	tja	Östergötland	Sveamål
15	Torsås	toa	Småland	Sydsvenska mål
16	Torhamn	toh	Blekinge	Sydsvenska mål
17	Torsö	too	Västergötland	Götamål
18	Väckelsång	vac	Småland	Sydsvenska mål
19	Våxtorp	vax	Halland	Sydsvenska mål
20	Fole	fol	Gotland	Gotländska
21	Frändefors	fra	Dalsland	Götamål
22	Frillesås	fri	Halland	Götamål
23	Hamneda	ham	Småland	Götamål
24	Jämshög	jam	Blekinge	Sydsvenska mål
25	Järnsås	jar	Småland	Götamål
26	Kärna	kaa	Bohuslän	Götamål
27	Korsberga	kor	Västergötland	Götamål
28	Löderup	lod	Skåne	Sydsvenska mål
29	Norra Rörum	nro	Skåne	Sydsvenska mål
30	Orust	oru	Bohuslän	Götamål
31	Össjö	oss	Skåne	Sydsvenska mål
32	Östad	ost	Västergötland	Götamål
33	Öxabäck	oxa	Västergötland	Götamål
34	Rimforsa	rim	Östergötland	Sveamål
35	S:a Anna	san	Östergötland	Sveamål
36	Segestad	seg	Öland	Götamål
37	Skee	Ske	Bohuslän	Götamål

Appendix 2

Ord som eliciterades i Götaland i SweDia-projektet

De ord som finns i tabellen nedan är de till största delen enstaviga ord som eliciterades från nästan alla informanterna. Det är 28 ord. Exempelen på elicitering är från inspelningen av informant nummer 322. De ord som använts i den här uppsatsen har markerats med fetstil. Bedömning av vokalen baserar sig på Leinonen (2010, s. 48-50) och Lindblad (2011, s. 88)

Eliciteringen gjordes på olika sätt av olika intervjuare. Flertydiga ord kunde eliciteras med olika betydelser. Ytterligare flerstaviga ord eliciterades, men för dem är inte vokalförmanterna extraherade i det material som har använts i den här studien.

Ord	Vokal	Exempel på elicitering
lär	[æ:]	Prästen bör leva som han ...
blöt	/ø:/	Vad blir man när det regnar?
läs	/ε:/	När man inte ser vad det står på ett säger man till någon: Var god och ...
sot	/u:/	Vad tar man bort när man rensar skorstenen?
söt	/ø:/	En citron är sur, en sockerbit är ...
lus	/u:/	Flera löss, en enda ...
typ	/y:/	Ett annat ord för sort eller slag.
låt	/o:/	Ett annat ord för melodi och visa
leta	/e:/	Ett annat ord för söka
dör	/ø:/	Om man inte ger blommorna vatten de vissna. Det kan gå så långt att de ..
dis	/i:/	Tunn dimma.
lös	/ø:/	Min hund är fastbunden, din hund är ...
lat	/ɑ:/	Den som inte vill arbeta är ...
nät	/ε:/	En spindel spinner sitt ...
sitt	/ɪ/	Om man har en hund och vill att den ska sätta sig så säger man...
lott	/ɔ/	Om man är på en marknad och vill vinna något så kan man ta en ...
flytta	/ʏ/	Om man inte trivs där man bor så kan man alltid ...
lass	/a/	Glass i stora ...
läsk	/ε/	Socketdricka och Fanta är en typ av dryck som man kallar...
särk	[æ]	Ett äldre ord för nattskjorta.
disk	/ɪ/	Om man har ätit får man smutsigt porslin. Det kan man kalla...
ludd	/ø/	Om man har en torktumlare brukar man vara tvungen att ta bort ...
säll	/ɔ/	Någon som inte kan hålla en hemlighet läcker som ett ...
tysk	/ʏ/	Berlin är inte en engelsk stad, den är ...
dörr	[œ]	För att komma in i ett hus måste man öppna en ...
blött	/ø/	När det har regnat så är gräset ...
sytt	/ʏ/	Om man har tillverkat en klänning då har man ...
lett	/e/	I går var My glad så då log hon, men idag har hon inte ...

Appendix 3

Beskrivning av indata

Extraktionerna av vokalerna formantfrekvenser för Götaland finns lagrade i tre filer: t_village, t_speaker och t_formant. De kan kopplas ihop med fälten speaker_id och vowel_id. Fälten i de tre filerna beskrivs nedan.

t_village

village	char	trebokstavsbeteckning för orten enl. App 1
age	char	o = old, y=young
gender	char	m= man, w=woman
speaker_id	int	1 - 470, ett nummer för varje informant

t_speaker

speaker_id	int	1 - 470, ett nummer för varje informant
word	char	ordet som spelats in utskrivet
vowel	char	auditiv bedömning av vokalen i ordet
start	int	identifiering av startpunkt för inspelningen
end	int	identifiering av slutpunkt för inspelningen
vowel_id	int	1 – 68493 unik identifiering av varje inläsning
disabled	int	0=normalt, 1=rad bör tas bort vid analys

t_formant

vowel_id	int	1 – 68493 unik identifiering av varje inläsning
f1	int	extraherad frekvens för f1
f2	int	extraherad frekvens för f2
f3	int	extraherad frekvens för f3
f4	int	extraherad frekvens för f4
f5	int	extraherad frekvens för f5
pos	char	position inom varje segment; f=first, l=last, 0=annan position
tracked_f1	int	automatträttat värde av f1
tracked_f2	int	automatträttat värde av f2

Beskrivning av databehandlingen

Data som använts i studien har tagits fram så här:

- t_village, t_speaker och t_formant slås ihop till en fil med hjälp av fälten speaker_id och vowel_id.
- Rader med disabled=1 eller med automatkorrigering av F1 och F2 raderas från filen.
- Rader med village=ref och rader med village=yst raderas.
- Bara rader med orden *sitt*, *disk*, *lett*, *dis* och *leta* behålls.
- En kolumn med landskapsnamn läggs till och varje ort associeras med ett landskap.
- Raden med den mittersta extraktion för varje vokal selekteras.
- Raden med den mittersta av de fem upprepningarna selekteras.

Filen innehåller nu för varje informant en rad per intalad vokal av de fem orden som undersöks i den här studien. Den här filen har varit startpunkt för vidare databehandling utom när olika extraktioner eller olika upprepningar av samma vokal har studerats. Då har i stället en motsvarande fil genererats, där ”mittersta extraktionen -1” eller ”mittersta upprepningen-1” selekterats.

Data som använts för att studera akustiskt avstånd har tagits fram så här:

- Raderna för de två ord som ska jämföras har slagits ihop till en fil med hjälp av fältet speaker_id och kolumner för förändring i de tre formanterna och akustiskt avstånd för två och tre formanter lagts till.
- Värdena på F1, F2 och F3 omvandlas från Hz till Bark med ekvation (1) och avstånden mellan F1, F2 och F3 mellan de två orden liksom det akustiska avståndet för två eller tre formanter beräknas enligt Ekvation (2).

Denna fil används sedan för beräkningar och för att generera diagram.