

Grundtonsstrategier vid tonlösa segment

Filip von Kartaschew

fvk@stp.ling.uu.se

Magisterarbete i datorlingvistik
Språkteknologiprogrammet
Institutionen för lingvistik och filologi
Uppsala universitet

6 juni 2007

Handledare:
Bertil Lyberg, Linköpings Universitet

Sammandrag

Prosodimodeller som bl.a. kan användas i talsynteser grundar sig ofta på analyser av tal som består av enbart tonande segment. Framför tonlös konsonant saknar vokalsegments grundtonskurvor möjlig fortsättning och blir dessutom kortare. Detta brukar då justeras med hjälp av trunkering av grundtonskurvan. Tidigare studier har i korthet visat att skillnader, förutom trunkering, i vokalers grundtonskurva kan uppstå beroende på om efterföljande segment är tonande eller tonlöst. Med utgångspunkt från dessa studier undersöks i detta examensarbete grundtonskurvan i svenska satser. Även resultaten i denna studie visar att olika strategier i grundtonskurvan används, och att trunkering inte räcker för att förklara vad som sker med grundtonskurvan i dessa kontexter. Generellt visar resultaten på att det verkar viktigt för försökspersonerna att behålla den information som grundtonskurvan ger i form av max- och minimumvärde, och att fall och stigningar så långt det går bibehålls.

Innehåll

1	Inledning	3
1.1	Syfte och frågeställning	3
1.2	Disposition	3
2	Bakgrund	4
2.1	Grundton, ord- och satsaccent	4
2.2	Hypoteser	5
3	Experiment	6
3.1	Inspelnings- och analysverktyg	6
3.2	Försökspersonerna	6
3.3	Det språkliga materialet	6
3.4	Instruktioner till försökspersonerna	7
3.5	Inspelningarna	7
3.6	Datainsamling	7
3.6.1	Reliabilitet	8
4	Resultat	9
4.1	Orden med accent I	9
4.1.1	CV:CVC, icke-final position	10
4.1.2	CV:CVC, final position	11
4.1.3	CVC:VC, icke-final position	12
4.1.4	CVC:VC, final position	13
4.2	Orden med accent II	14
4.2.1	CV:CVC, icke-final position	14
4.2.2	CV:CVC, final position	15
4.2.3	CVC:VC, icke-final position	16
4.2.4	CVC:VC, final position	17
4.2.5	Ord med accent II och avslutande tonande/tonlös konsonant	18
4.3	Analys av första vokalen	18
4.3.1	Border adjustment	18
4.3.2	Compensation adjustment	19
4.3.3	Trunkering	19
4.3.4	Rate adjustment	20
4.3.5	Timing adjustment	20
4.4	Analys av andra vokalen	20
5	Diskussion	22
5.1	Duration	23
6	Sammanfattning	24
6.1	Vidare forskning	24

Förord

Jag vill rikta ett stort tack till dem som hjälpt mig färdigställa detta examensarbete. Bland dem märks särskilt min handledare Bertil Lyberg. Dessutom Krister, Anna, Per och Amelie som bidragit med sina röster, Johan Linder och Emma Forsgren som korrekturläst uppsatsen samt Erik Nordström som bidragit med teknisk support.

1 Inledning

Detta examensarbete består av en empirisk undersökning av grundtonskurvan i svenska satser. Tidigare studier (Eriksson & Alstermark 1972) och (Lyberg & Sangari 2000) har visat att det kan förekomma skillnader i samma vokalfonems grundtonskurvor beroende på om efterföljande segment är tonande eller tonlöst. Förutom att vokalen blir kortare före tonlösa segment (Elert 1965) så kan inte heller grundtonskurvans rörelser fortsätta in i det efterföljande tonlösa segmentet. Detta examensarbete försöker klargöra vad vår talapparat har för strategier för att kompensera grundtonskurvan i sådana kontexter. Både denna och de tidigare studier som presenteras här baseras på svenska så som man talar i Stockholm och Uppsala. Resultaten kan användas för att komplettera de prosodimodeller som bland annat används vid generering av grundtonskurvor för talsynteser.

1.1 Syfte och frågeställning

Syftet med detta examensarbete har varit att undersöka de prosodiska skillnaderna, i framför allt grundtonskurvan men även indirekt i duration, mellan vokaler i ord med tonande respektive tonlös efterföljande konsonant. Min hypotes är att det torde råda skillnader i grundtonskurvan i dessa kontexter. Dels beroende på den skillnad i duration som konsonanters egenskaper orsakar (Elert 1965), och dels att informationen som grundtonskurvan ger i betoningar glider över i efterföljande segment, så länge det segmentet är tonande. Vilket då torde leda till att denna kurva på något sätt måste kompensera för bortfallet av möjlig fortsättning. Frågeställningen är således; om och i så fall vad vi har för strategier för att behålla informationen i grundtonskurvan när efterföljande segment är tonlöst.

1.2 Disposition

Denna uppsats består av sex kapitel. Kapitel 2 ger en bakgrund och beskriver tidigare studier och teorier i det aktuella ämnet. Kapitel 3 beskriver metoden och hur experimentet har gått till. I kapitel 4 presenteras och analyseras insamlad data. I kapitel 5 diskuteras resultaten. I kapitel 6 sammanfattas arbetet och förslag på fortsatta studier ges.

2 Bakgrund

Vid generering av syntetiskt tal är det viktigt för naturligheten och även förståeligheten att göra så korrekta prosodimodeller som möjligt. Prosodimodellerna bör så gott det går överensstämma med naturligt tal. De prosodimodeller som finns och bland annat används för att generera grundtonskurvor i talsynteser utgår ofta från analyser av tal där man använder sig av enbart tonande segment genom hela satser, vilket således ger en oavbruten grundtonskurva (Bruce 1977). Vad som verkar vara föga studerat är hur grundtonskurvan beter sig när det förekommer tonlösa segment. Tidigare studier visar att segment påverkas i duration beroende på omgivande segments egenskaper (Elert 1965). Om ett segment påverkas i duration bör man kunna sluta sig till att även grundtonskurvan på något sätt påverkas. Detta brukar då avhjälpas med trunkering, se figur 1.

Resultaten i detta examensarbete kan alltså, förutom att vara intressant ur ett rent fonetiskt perspektiv, även vara ett steg till att göra prosodimodeller vid generering av syntetiskt tal mer precisa. Framtida system för tal-till-text kan även tänkas behöva göra närmare analyser av prosodi.

2.1 Grundton, ord- och satsaccent

I detta arbete har grundtonskurvan (även kallad F₀-kurvan) varit föremål för analys. Grundtonens frekvens är det som används för att beskriva tonhöjdförloppet (satsmelodin) i tal. Det är den lägsta av deltonerna i harmoniska ljud. Grundtonens frekvens anges i hertz, Hz. En Hz är en svängning per sekund. Vid 100 Hz vibrerar alltså stämbanden 100 gånger per sekund. (Ladefoged 1993).

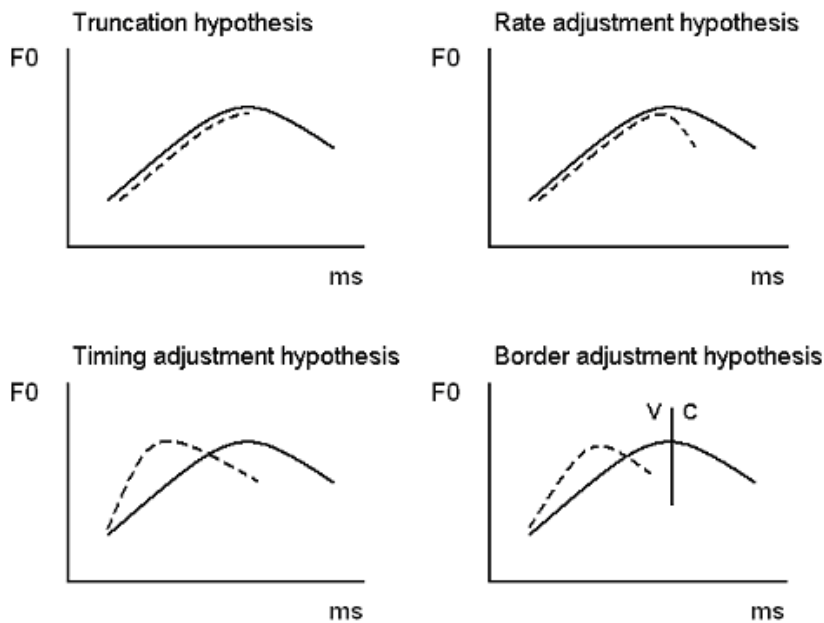
I svenskan skiljer man på två ordaccenter; akut accent (accent I) och grav accent (accent II). I ord med akut accent är det en stavelse som har starkare tryck i form av högre intensitet och tonhöjd. Andra eventuella stavelser har alla lägre tonhöjd. Ord med en stavelse har i svenskan alltid akut accent. I ord med grav accent är det två stavelser som har starkare tryck och högre tonhöjd. Exempel: stölen (akut accent) stölar (grav accent). I denna studie är alla testord tvåstaviga. Både ord med akut och grav accent finns representerade i testdatan.

Distribution av satsaccent är ofta semantisk betingat; ord med satsaccent markerar ofta något nytt eller att man vill lägga fokus något särskilt i satsen. Därav synonymen fokal accent. Satsaccent markeras genom variationer i tonhöjdförlopp och ljudstyrka. I denna studie har alla testord satsaccent. Enligt Bruce (1977) är det typiska tonhöjdförloppet för ord med satsaccent och akut accent, en stigning i första vokalen och ett fall i den andra. Grundtonskurvan i konsonanten mellan vokalerna är i det närmaste horisontellt rak. När ordet står i final position är stigningen något mer plan i första vokalen, medan fallet i andra vokalen faller till en lägre punkt i en brantare kurva. I konsonanten är det en brant stigning. I ord i icke-final position med grav accent och satsaccent är det typiska tonhöjdförloppet ett grundtonsfall i första vokalen och en stigning i den andra. Här händer det liksom i akut accent inte så mycket i konsonanten. När ordet står i final position liknar grundtonskurvan i första vokalen motsvarande segment i icke-final position. I efterföljande konsonant påbörjas en stigning som fortsätter in i den andra vokalen. Stigningen övergår dock till ett fall tidigt i vokalen. Observera att detta grundar sig på analyser av tal med endast tonande segment.

2.2 Hypoteser

Det finns sedan tidigare tre teorier (Lyberg & Sangari 2000) som försöker klargöra vad som händer med grundtonskurvan vid tonlösa segment. Dessa tre och en fjärde, *border adjustment* som förklaras närmare i kapitel 4.3.1, visualiseras i figur 1 nedan.

- Trunkering. Ingen avpassning sker, grundtonskurvan skärs av vid CV-gränsen och tappar fortsatt fall eller stigning.
- Rate adjustment. Lutningen på kurvan ändras så att den faller snabbare.
- Timing adjustment. Avpassning sker i form av tidigareläggning av kurvans extremer; högsta och lägsta punkter.
- Border adjustment. Grundtonskurvan faller till ett lägre minimum, som i motsvarande ord uppnås först i efterföljande tonande konsonant.



Figur 1: Exempel på hur dessa teorier kan se ut. Den streckade linjen representerar vokal med efterföljande tonlös konsonant och den heldragna linjen vokal med efterföljande tonande konsonant.

3 Experiment

Det som har varit föremål för undersökning är om och i så fall hur man i tal kompenserar grundtonskurvans stigningar och fall i vokaler med efterföljande tonlösa konsonanter jämfört med vokaler som efterföljs av tonande konsonanter. Datainsamlingen för detta ändamål grundar sig på inspelningar av fyra personers tal. Inspelningarna består av ett antal satser ur vilka grundtonskurvan vid de aktuella segmenten plockats ut.

3.1 Inspelnings- och analysverktyg

Inspelningarna har skett på en PC med integrerat ljudkort. Mikrofonen som användes var en Shure SM58. Mjukvaran som använts för både inspelningarna och datainsamlingen är Wavesurfer, versionerna 1.6.0-1.8.5. Wavesurfer är ett program för ljudvisualisering och -modifiering, och är utvecklat på Centrum för talteknologi (CTT) på KTH. Med Wavesurfer kan man bland annat spela in ljud och sedan visualisera det i form av spektrogram, grundtonskurvor mm.

3.2 Försökspersonerna

I undersökningen har fyra försökspersoner använts. Valet av dessa personer har gjorts genom att ta personer med liknande språklig och social bakgrund. De har svenska som modersmål och är eller har varit studerande vid Uppsala Universitet. Testgruppen bestod av två kvinnor; AL 24 år, härstammande från Stockholm och AP 27 år, härstammande från Uppsala. Samt två män; PM 25 år och KM 29 år, båda härstammande från Uppsala.

3.3 Det språkliga materialet

Det som har undersökts är vilka skillnader som kan uppstå i grundtonskurvan mellan vokaler som följs av tonande respektive tonlösa konsonanter. För att få fram jämförbara data skapades först satser med enbart tonande segment. I mitten av var och en av satserna finns ett huvudord, se tabell 1 nedan, som vid inspelningarna ges fokal accent. Huvudorden är tvåstaviga och representerar varianter med accent I och accent II samt fonologiskt lång och kort vokal (CV:CVC respektive CVC:VC) i alla, fyra, kombinationer. Det är framför allt första vokalen i orden som varit föremål för analys.

För dessa huvudord skapades även satser där konsonanterna i mitten av huvudorden ersattes med dess tonlösa motsvarighet. I dessa huvudord är det /g/ och /k/ som jämförts. Detta ger totalt åtta satser. Antalet satser dubblerades genom att för varje huvudord även ha en sats med huvudordet i final position. Satserna är på formen *Lägga X i ladan* respektive *Lägga ladan i X*, där X är huvudordet.

Senare spelades det in kompletterande satser för att se hur grundtonskurvan i den sekundära betoningen i accent II-ord, gravisfallet, beter sig i absolut final position med en avslutande tonlös konsonant. För detta ändamål skapades satserna *Bilen är packad* och *Lasset är packat*.

Huvudorden är lika varandra fonemiskt. Bl.a. används samma vokalfonem i de aktuella stavelserna i alla testord. Detta för att undvika inherent olikheter i till exempel duration och tonhöjd (Lehiste 1970).

tonande	tonlös
Accent I	
dågen	båken
dåggen	båcken
Accent II	
dågar	båkar
båggen	båckar

Tabell 1: Huvudorden

3.4 Instruktioner till försökspersonerna

Instruktionerna till försökspersonerna försöktes i början av inspelningarna hållas till ett minimum, för att på så sätt få mer naturliga data på grund av mindre styrning. De fick veta att de skulle betona huvudordet och de blev instruerade att tala som de normalt gör, avslappnat och naturligt. Det visade sig dock att försökspersonerna hade en tendens att tala onaturligt tydligt genom att bland annat överdriva betoning och artikulation. Det förekom dessutom kortare pauseringar efter huvudorden när de stod i icke-final position och före huvudorden när de stod i final position. Så testpersonerna fick till slut instruktioner om att försöka tala naturligt och att inte överdriva betoningen genom för starkt tryck och pauseringar. I och med detta låter meningarna mer som delar av en större diskurs i stället för att vara lösryckta satser som med största sannolikhet inte skulle låta naturligt i kontinuerligt tal.

3.5 Inspelningarna

Satserna spelades in med ett par sekunders mellanrum i omgångar om 16 satser. I varje omgång är ordningen på satserna slumpvist omkastad. Detta för att minimera risken för variation i bland annat tryck och duration på satserna som spelas in i början och i slutet i varje omgång. Varje enskild sats lästes in totalt åtta gånger. Satserna med huvudordet i mitten och i final position har spelats in för sig.

3.6 Datainsamling

Ur inspelningarna plockades sedan grundtonskurvan ut. De segment varifrån data har plockats ut är de två vokalerna och konsonanten i mitten av huvudorden, t.ex. /age/ i dagen. Viktigt för datainsamlingen har varit att göra så riktiga och likställda segmenteringar som möjligt. Segmenteringen har i detta arbete utgått ifrån den definition som Elert (1965) använder. Vokalsegmentens början och slut kan definieras som snabba stigningar respektive fall i intensitet. Dessa stigningar och fall sammanfaller också med spektrogrammet där segmentskiften blir desto tydligare visuellt, särskilt i dessa satser där vokalen följs av en klusil. Frånvaron av periodisk ljudkälla i de tonlösa segmenten är också en tydlig indikation på segmentskifte.

Efter segmenteringen har en uppskattning av medellängden på varje segment gjorts, detta för att kunna bestämma lämpliga plottningsintervall. Plottningsintervallen bestämmer de mätpunkter som sedan plockats ut från grundtonskurvan. Dessa intervall delar varje segment av samma fonem i samma ord i lika

många delar. Antalet delar beror på hur långt segmentet är. Längden på intervallen varierar mellan drygt 20 till 60 ms. I de flesta fall är antalet delar fem, men när segmenten är så korta att längden mellan delarna blir mindre än c:a 20 ms så har antalet delar minskats, i vissa fall ända ned till tre delar. Detta ger mellan fem respektive tre mätpunkter, i början, slutet och i de punkterna inuti segmenten. Ur de tonande konsonantsegmenten har endast två mätpunkter plottas. Medelvärdena för varje mätpunkt har sedan satts samman till kurvor genom att förbinda dem med linjer. Kurvorna för ordparen som skulle jämföras sattes sedan in i samma diagram. För att lättare se skillnader mellan kurvorna för varje segment flyttades kurvorna för de tonlösa ordens vokaler till samma CV-gränser i diagrammen som de helt tonande orden. Den faktiska durationen på de tonlösa segmenten framgår alltså inte i diagrammen.

I de tonlösa segmenten har jag ignorerat F0 dels eftersom de uppenbart ska sakna ton och dels för att den F0-kurva som ändå ibland skapats i Wavesurfer består av brus och klickljud vid aspiration och ocklusion och därför har varit väldigt spridd. Det verkar och ska inte ha rört sig om en periodisk ljudkälla.

Man kan inte räkna ut några medelkurvor mellan testpersonerna. Detta eftersom människor talar olika i form av hastighet, tonhöjd, intensitet mm.

3.6.1 Reliabilitet

För att mäta reliabiliteten på insamlad data har standardavvikelsen för varje mätpunkt räknats ut. Denna ligger i regel runt 5 Hz och överstiger sällan 10 Hz för de manliga försökspersonerna, och runt 9 Hz och överstiger sällan 15 Hz för de kvinnliga försökspersonerna. Värt att notera är att detta alltså gäller mätpunkterna i alla segment varifrån data har plottats, inklusive de tonande klusilerna, som förväntades ha en större spridning. Undantaget är de två kvinnliga försökspersonerna som i orden med accent I i final position inte verkar sträva efter att få en tydlig avslutande stavelse. Sista stavelsen liknar mer ett skorrande ljud.

4 Resultat

I detta kapitel presenteras insamlad data i form av diagram. Varje figur visar ett av de jämförda ordparen. Figurerna är i sin tur uppdelade i fyra delfigurer, en för varje försöksperson. Delfigurerna *(a)* och *(b)* visar de två kvinnliga försökspersonernas kurvor, AL respektive AP, och *(c)* och *(d)* visar de två manliga försökspersonernas kurvor, KM respektive PM. De segment varifrån grundtonskurvan plottats är understrukna i figurtexten.

I diagrammen har vokalsegmenten i orden med tonlös konsonant flyttats till samma CV-gränser som motsvarande ord med tonande konsonant. Dessa segmentskiften markeras av de vertikala linjerna. Durationen på de tonlösa konsonanterna framgår alltså inte. Den faktiska sammanlagda durationen på de tre segmenten är ungefär lika långa i de jämförda orden. Den totala durationen på de tre segment med tonlös konsonant är i regel mellan 0-20 ms kortare. Den andra vokalens grundtonskurva är med andra ord framskjuten något. Orden med tonlös konsonant representeras av de röda kurvorna. De kontinuerliga blå kurvorna representerar orden med enbart tonande segment.

I analysen av kurvorna i kapitel 4.3 ligger fokus på framför allt första vokalen i testorden. I kapitel 4.4 behandlas även andra vokalen kort. Analyserna sker med utgångspunkt från de i inledningen nämnda teorierna; trunkering, rate adjustment, timing adjustment och, den i denna undersökning nyttillkomna, border adjustment. En mer generell diskussion av resultaten ges i kapitel 5.

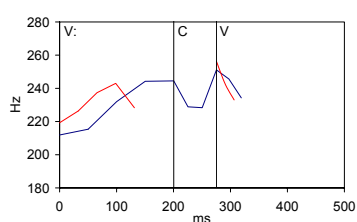
4.1 Orden med accent I

Här presenteras diagrammen för orden med akut accent, icke-final och final position. I figurerna 2 och 3 är första vokalen fonologiskt lång. I figurerna 4 och 5 är vokalen fonologiskt kort. Observera att grundtonskurvorna i figurerna 3(b), 5(a) och 5(b) sträcker sig över ett större frekvensområde. Y-axeln är därför komprimerad i dessa diagram.

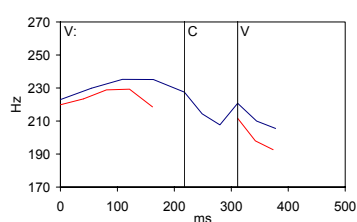
4.1.1 CV:CVC, icke-final position

I första vokalen tidigarelägger samtliga försökspersoner fallet. Alla utom PM tidigarelägger dessutom max. Alla utom PM kan alltså sägas ha timing adjustment. PM:s kurva påminner mer om rate adjustment.

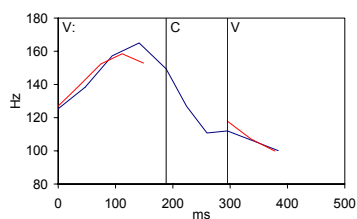
AL har till skillnad från de andra försökspersonerna ett högre max i andra vokalen. Efter tonlös konsonant är vokalsegmentet kortare och har en högre startpunkt. De justeras med rate adjustment och kurvan når ner till samma minimum. AP:s kurvor är likadana, förutom att kurvan efter tonlös konsonant är lägre. Skillnaden i frekvens är samma som skillnaden i sista mätpunkten i den första vokalen.



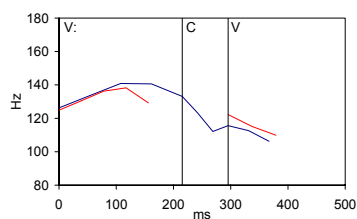
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



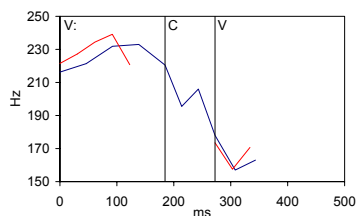
(d) Försöksperson PM

Figur 2: icke-final position, dägen/bäken

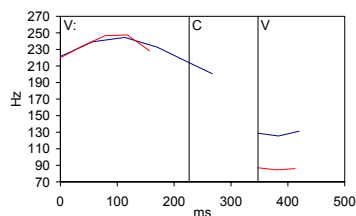
4.1.2 CV:CVC, final position

I AP:s och KM:s första vokal trunckeras praktiskt taget grundtonskurvan. I AL:s och PM:s kurvor nås max ungefär samtidigt, därefter faller kurvan snabbare. Dessa kurvor betar sig alltså som rate adjustment. AL:s kurva är dessutom högre, men har ett nästan dubbelt så långt fall och når därmed samma lägsta punkt.

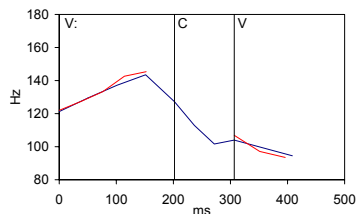
Eftersom orden här står i final position så är grundtonskurvan i andra vokalen lägre än i samma ord i icke-finalposition. PM:s och KM:s kurvor innehåller ett svagt fall. AL börjar med ett fall och har en stigning i slutet av vokalen. AP:s kurva är nästan helt rak. Efter tonlös konsonant är den dessutom lägre. Man kan även se att AP:s grundtonskurva försvinner i det tonande konsonantsegmentet. Detta händer bara i ord med akut accent som står i final position och visar sig även i några av de följande diagrammen, se 5(a), 5(b) och 7(b).



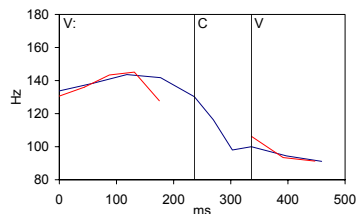
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



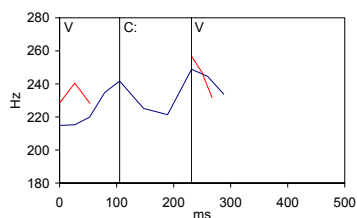
(d) Försöksperson PM

Figur 3: Final position, dágen/báken

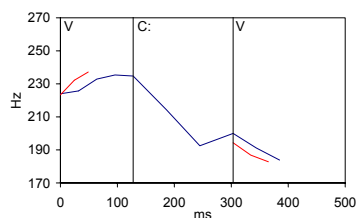
4.1.3 CVC:VC, icke-final position

Liksom i figur 2(a) lägger AL fallet i den tonande konsonanten, och flyttar därför fallet i ordet med tonlös konsonant in i föregående vokal. I AP:s kurva återstår endast stigningen. Transportsträckorna till och från stigningen trunkeas. KM:s kurva stiger snabbare till ett något högre max. PM försöker få med samma kurva på kortare tid, men lyckas inte nå samma max. Däremot infinner sig ett svagt fall.

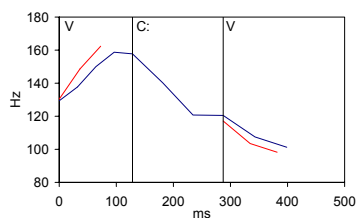
I båda kurvornas andra vokal är AL:s max återigen (se figur 2(a)) betydligt högre än för de andra försökspersonerna. Vokalen med föregående tonlös konsonant som är ännu något högre anpassas med rate adjustment för att nå ner till samma minimum. AP:s och KM:s kurvor är något lägre när grundtonen kommer tillbaka. Kurvorna har dock nästan identisk lutning. I PM:s kurva är det större skillnad i både lutning och frekvens.



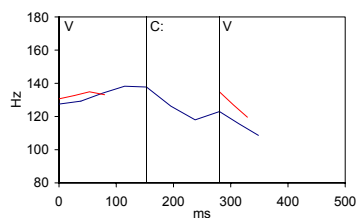
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



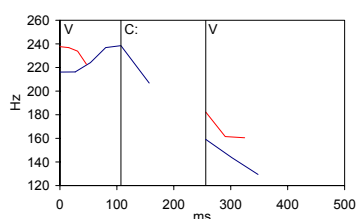
(d) Försöksperson PM

Figur 4: icke-final position, däggen/bäcken

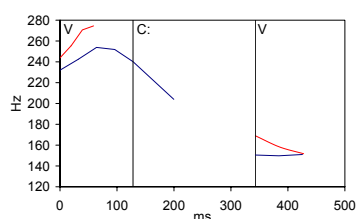
4.1.4 CVC:VC, final position

Återigen flyttar AL fallet som helt är i den tonande konsonanten in i första vokalen. Stigningen som helt försvinner verkar inte vara viktig att behålla här. AP kompenserar istället den korta durationen med en högre kurva. Maxet skiljer sig hela 20 Hz. KM gör ungefär likadant, men kurvan liknar mer vanlig trunker- ing förutom att den är något högre. PM komprimerar kurvan, liksom i samma ord i icke-finalposition. Här är dock maxet högre.

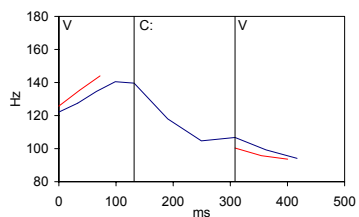
I AL:s, AP:s och PM:s andra vokal är det en högre startpunkt efter tonlös konsonant. Efter startpunkten ändras AP:s och PM:s kurvor och når samma frekvens i sista mätpunkten. AL:s kurva når aldrig ned till samma frekvens.



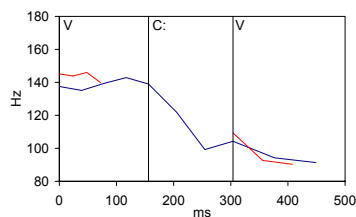
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



(d) Försöksperson PM

Figur 5: Final position, däggen/bäcken

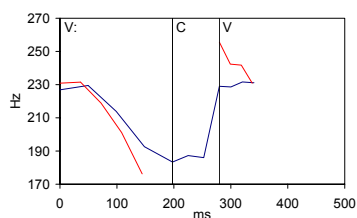
4.2 Orden med accent II

I detta kapitel presenteras diagrammen för orden med grav accent, icke-final position och finalposition. I figurerna 6 och 7 är första vokalen fonologiskt lång. I figurerna 8 och 9 är vokalen fonologiskt kort.

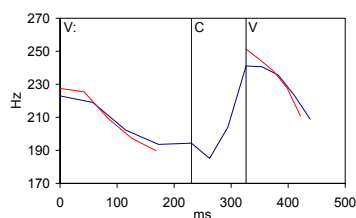
4.2.1 CV:CVC, icke-final position

Här är försökspersonernas strategier i första vokalen snarlika. Allas kurvor följs först åt, sedan sker en svag rate adjustment. I AP:s och KM:s kurvor är rate adjustment så svag att det närmast rör sig om trunkering. AP:s kurva har dock en högre startpunkt vilken gör att kurvan har ett längre fall.

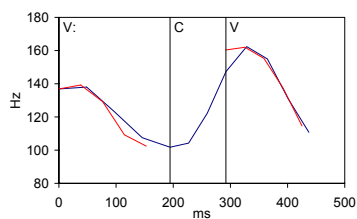
AL:s kurvor i andra vokalen är nästan varandras spegelbilder. De andra försökspersonerna har en högre startpunkt men kurvan ansluter sig sedan till motsvarande vokal med föregående tonande konsonant.



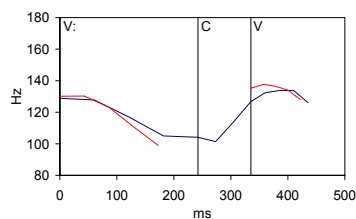
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



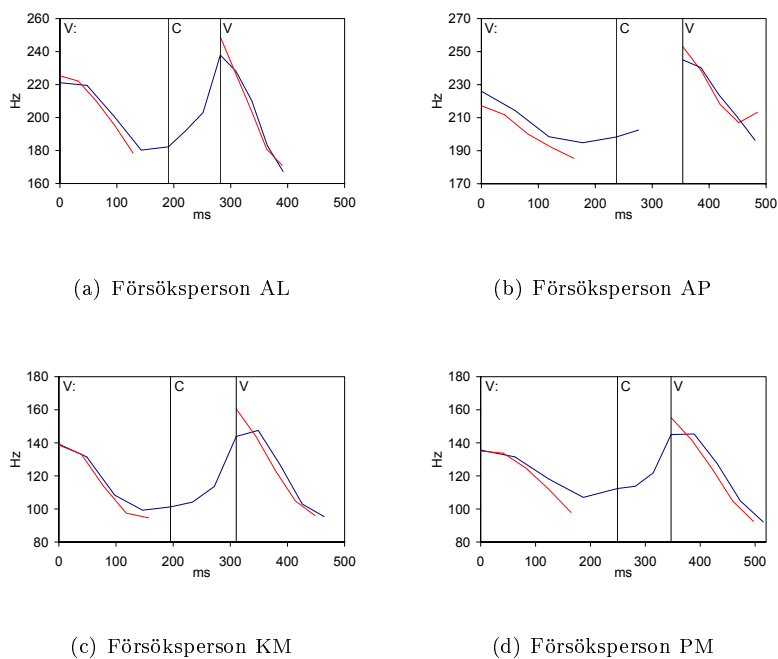
(d) Försöksperson PM

Figur 6: icke-final position, dågar/båkar

4.2.2 CV:CVC, final position

Även i final position växlar kurvorna i första vokalen mellan att vara svaga rate adjustment och trunkeringar. AP anpassar dessutom kurvan genom att låta den vara lägre. Här är durationen på vokalsegmentet så långt att försökspersonerna verkar hinna få med den rörelse i kurvan som behövs.

Andra vokalens kurvor är snarlika för alla försökspersoner, förutom den högre startpunkten.

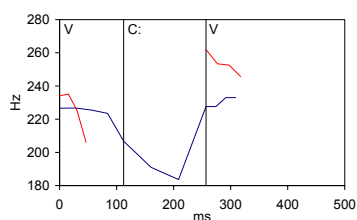


Figur 7: Final position, dägar/bäkar

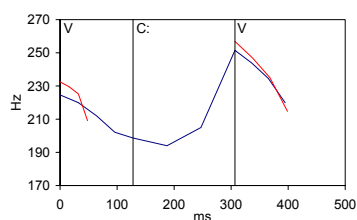
4.2.3 CVC:VC, icke-final position

I första vokalen liknar PM:s kurva trunkering, men liksom AP så kan man nog säga att det är rate adjustment. Bådas kurvor börjar med en högre frekvens. Lutningen är dock densamma fram till andra mätpunkten där kurvan börjar falla brantare. AL kompenserar den extremt korta durationen med ett högre max som följs av ett väldigt brant fall. KM:s kurva ser ut att vänstertrunkeras. Den lilla stigning som finns försvinner i vokalen med efterföljande tonlös konsonant.

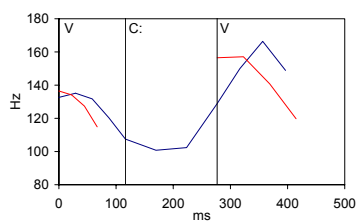
I andra vokalen är AL:s kurvor återigen varandras motsatser; en stigning efter tonande konsonant och ett fall efter tonlös konsonant. KM:s kurvor är också väldigt olika varandra. Det är endast ett fall efter tonlös konsonant och en kraftig stigning efter tonande konsonant, som följs av ett mindre kraftigt fall. Förutom ett något högre max efter tonlös konsonant är det ingen större skillnad i AP:s andra vokal.



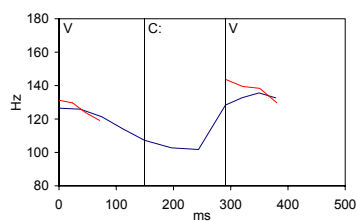
(a) Försöksperson AL



(b) Försöksperson AP



(c) Försöksperson KM



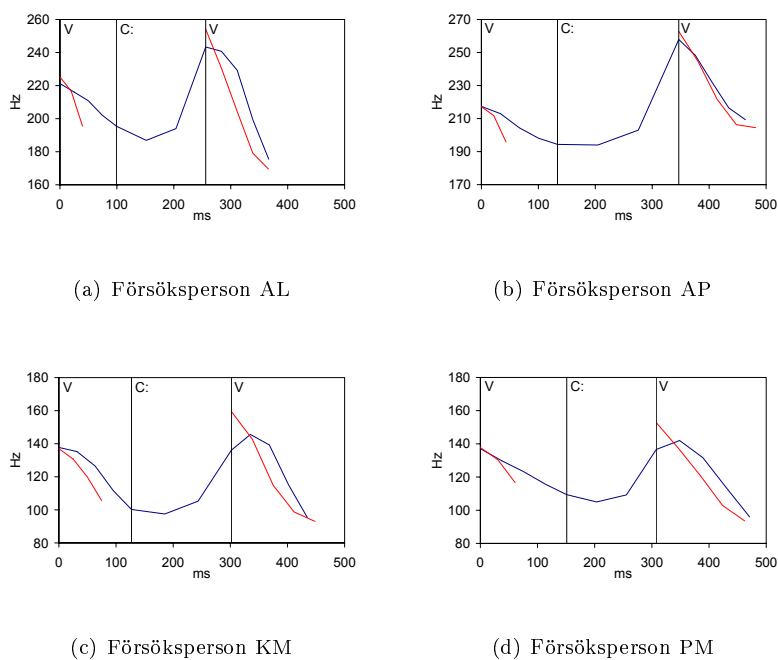
(d) Försöksperson PM

Figur 8: icke-final position, bäggen/bäckar

4.2.4 CVC:VC, final position

I första vokalen blir det tydligt rate adjustment för alla försökspersoner utom KM, som liksom i icke-final position, figur 8(c), vänstertrunkerar kurvan och tidigarelägger fallet. I både final och icke-final position har KM längre duration än de andra försökspersonerna på vokalen innan tonlös konsonant.

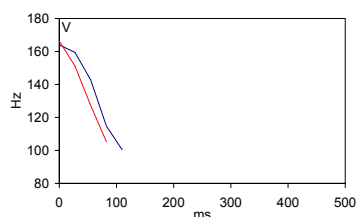
I andra vokalen efter tonlös konsonant har AL:s, KM:s och PM:s grundtonskurva en högre start som sedan faller i en rak brant kurva. KM:s och PM:s vokal med föregående tonande konsonant börjar med en stigning.



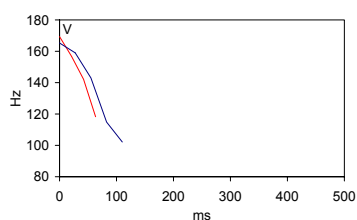
Figur 9: Final position, bäggen/bäckar

4.2.5 Ord med accent II och avslutande tonande/tonlös konsonant

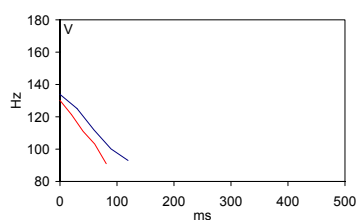
Denna inspelning gjordes bara med försökspersonerna AP, KM och PM. Här står huvudordet med accent II i final position med avslutande tonlös/tonande konsonant. Här har endast grundtonskurvan i sista vokalen plottats. Försökspersonerna har snarlika lutningar, men kurvorna i vokalen före det tonlösa segmentet vänstertrunkeras för AP och KM. PM löser det genom att ha en lägre startpunkt och en något brantare lutning på kurvan.



(b) Försöksperson AP



(c) Försöksperson KM



(d) Försöksperson PM

Figur 10: Final position, påckad/päckat

4.3 Analys av första vokalen

Man kan se skillnader i första vokalens grundtonskurva mellan orden som står i icke-final position jämfört med de som står i final position. Försökspersonerna har dock i stort sett samma strategier oberoende på om orden står i final eller ickefinal position.

4.3.1 Border adjustment

De redan etablerade teorierna har observerats i de insamlade data. Till dessa kan läggas till ännu en teori som jag har valt att kalla *border adjustment*. Fenomenet har upptäckts vid flera tillfällen och kan inte fullt ut beskrivas med någon av de nämnda teorierna. Det kan sägas vara en tilläggsstrategi eftersom den alltid förkommer i kombination med timing och/eller rate adjustment. Border adjustment innebär att grundtonskurvan i vokaler med efterföljande tonlöst segment

faller till en lägre punkt än vad de motsvarande vokalerna med efterföljande tonande konsonant gör. De når ett minimum som i de helt tonande orden nås först en bit inne i konsonanten. Det är alltså som en motsats till trunkering.

I de kurvor där border adjustment är tydligast har försökspersonen flyttat maxet i de helt tonande orden så långt fram att fallet helt hamnat inne i den tonande konsonanten. Denna företeelse har även observerats av (Eriksson & Alstermark 1972). Observationerna har i olika hög grad gjorts hos alla testpersoner men framförallt hos AL, se figurerna 2(a) och 4(a). Ett annat exempel på detta ses i figur 5(a). Där har både max och grundtonsfallet flyttats från VC-gränsen till CV-gränsen. Dessa kurvor är även exempel på timing adjustment eftersom maxet är tidigarelagt.

4.3.2 Compensation adjustment

Det förekommer även andra situationer då kurvorna skiljer sig åt så mycket och är av sådan karaktär att de nämnda teorierna inte kan beskriva skillnaderna. Compensation adjustment kommer härnäst att referera till dessa kurvor som ett samlingsbegrepp på övriga strategier. Compensation adjustment förekommer då vokalsegmentet blir så kort att speciella strategier måste tillämpas för att behålla den information som grundtonskurvan ska ge. Det rör sig om orden med fonologiskt kort vokal, då de är kortare än 60 ms. Det leder ofta till att kurvorna skiljer sig åt i max. Se till exempel figur 8(a), där fallet är närapå dubbelt så långt. Kurvan har där ett högre max men det branta fallet ser till att grundtonen når samma minimum.

När det blir en högre kurva som innehåller ett fall så har den ofta border adjustment-tendenser eftersom fallet blir längre. Om man då tänker sig att man flyttar kurvan så att startpunkterna överensstämmer så skulle det bli border adjustment. Detta sker ofta när durationen blir så korta att försökspersonerna verkar vilja kompensera den korta tid vokalen har på sig att ge grundtonskurvan samma information. Se även figur 5(b) där kurvan trunkeras, men den korta durationen kompenseras med en högre kurva.

4.3.3 Trunkering

Trunkering är i viss mån vanligt förekommande, men att hela kurvan bara är trunkerad är ovanligt. Ofta börjar kurvan med något slags avpassning och trunkeras på slutet. Några exempel på nästan total trunkering finns dock, se figurerna 3(c) och 6(b).

I några kurvor ser det ut som att början är trunkerad (vänstertrunkering). Om man flyttar fram kurvan till andra mätpunkten i motsvarande vokal så följer kurvorna varandra om inte exakt så åtminstone snarlikt. Det verkar som att det vid dessa tillfällen är viktigt att bibehålla själva rörelsen i kurvan. Det som då trunkeras är delar av segmentens kurvor som inte innehåller så mycket rörelse, se figurerna 4(b) och 9(c). Även 2(a) är ett exempel på en sådan kurva. Den börjar med en vänstertrunkering men avslutas med border adjustment. Informationen bibehålls medan de oviktiga transportsträckorna till och från fallen och stigningarna hoppas över.

En kombination av trunkering och compensation adjustment är en annan återkommande strategi då kurvorna har samma lutning men kompenseras genom att kurvan är högre, se figurerna 5(b) och 5(c).

4.3.4 Rate adjustment

Rate adjustment är en vanligt förekommande strategi. Särskilt i orden med grav accent. PM använder även strategin i orden med akut accent. Det typiska i orden med accent I är då att kurvorna följs åt fram till tre fjärdedelar eller hälften av vokalen med efterföljande tonlös konsonant. Där faller kurvan sedan snabbare och når ned till ungefär samma minimum, se figurerna 3(d) och 3(a). Ibland blir det även högre minimum (trunkering) eller lägre minimum (border adjustment). I orden med accent II är det typiska att kurvan börjar falla snabbare redan efter andra mätpunkten, se till exempel kurvorna 6(a) och 7(d).

4.3.5 Timing adjustment

Även timing adjustment förekommer frekvent, och då framför allt i orden med akut accent. Kurvorna 2(a), 4(a) och 5(a) är tydliga exempel på tidigareläggning av max. Dessa tre kurvor är samtidigt exempel på border adjustment, eftersom AL här i de helt tonande orden lägger fallet i den konsonanten. Det förekommer också kurvor som ser ut att vara vänstertrunkerade. Dessa bör dock också kunna kategoriseras som timing adjustment eftersom maxet blir tidigarelagt, se t.ex. figurerna 4(b) och 8(c).

4.4 Analys av andra vokalen

Till skillnad från första vokalen skiljer sig inte durationen nämnvärt i andra vokalen. I orden med akut accent är det generellt inte heller någon större skillnad på grundtonskurvorna i andra vokalen. AP låter kurvorna vara något lägre i orden med akut accent, men de är av samma karaktär. Undantaget är i orden med fonologiskt kort konsonant i final position, se 5(b). I och med compensation adjustment i första vokalen, där kurvan är 20 Hz högre, låter hon också grundtonskurvan i andra vokalens startpunkt vara 20 Hz högre. AM:s kurvor skiljer sig åt i lutning i kurvorna 2(a) och 4(a), men trots den något kortare durationen och den något högre startpunkten når kurvorna ned till samma minimum.

I orden med grav accent skiljer sig kurvorna åt desto mer. Där skiljer sig kurvorna nästan genomgående åt genom att stigningen i början av vokalen saknas. Vokalerna med föregående tonande konsonant fortsätter med stigningen som påbörjats i konsonanten till andra mätpunkten, och börjar sedan efter ett svagt fall att falla snabbt. I vokalerna efter tonlös konsonant är det istället bara ett snabbare fall med högre startpunkt. Den högre startpunkten beror på att det slutna stadiet av den tonlösa klusilen är längre. Därmed bildas ett högre lufttryck i de supraglottala kaviteterna. När detta tryck släpps i ocklusionen blir luftströmmen genom glottis starkare vilket leder till en maskering av max. Detta fenomen kallas för *perturbation* och verkar vara starkare ju lägre grundtonen är. (Lyberg 1984). Det visar även kurvorna i denna studie, tendensen är starkast för de två manliga försökspersonerna. Exempel på detta finns i flera av kurvorna. Se kurvorna i figurerna 6 till och med 9. Lägg särskilt märke till de manliga försökspersonernas kurvor. I AM:s kurvor i icke-final position är det drastisk skillnad i orden med grav accent. Kurvorna är varandras spegelbilder. Istället för en stigning är det endast ett fall i vokalerna med föregående tonlös konsonant.

I figur 10 visas de grundtonskurvor som plottats från sista vokalen där huvudordet med grav accent står i final position med avslutande tonlös/tonande

konsonant. De tre försökspersonerna använde sig alla i denna kontext av en liknande strategi. AP och KM har inte med det mjuka fallet som finns i vokalen med efterföljande tonande konsonant, utan faller direkt i en rak kurva. PM:s kurva börjar med ett något lägre max och fallet är något brantare i vokalen med efterföljande tonlös konsonant. Dessa strategier bör kategoriseras som rate adjustment. AP och PM lyckades dessutom komma ned till samma minimum medan KM:s grundtonskurva trunkeas.

5 Diskussion

Det mest uppenbara som framkommit i och med denna studie och som besvarar frågeställningen är att avpassning av grundtonskurvan, förutom trunkering, tydligt sker. Avpassningen sker ofta genom att på något sätt tidigarelägga fallet. Att enbart använda trunkering i en intonationsmodell vid dessa kontexter verkar alltså inte vara optimalt. Generellt verkar det vara viktigt för försökspersonerna att på något sätt se till att fallet i grundtonskurvan behålls för att inte förlora den prosodiska informationen som den innehåller. Försökspersonerna verkar också sträva efter att inte förlora max och minimum, även om minimum ibland inte hinns med (trunkering) och ibland till och med är lägre (border adjustment).

I de extrema förutsättningarna där fonologiskt korta vokaler blir kortare än 60 ms så verkar försökspersonerna vilja kompensera det genom att öka hela grundtonskurvans frekvens, eller låta startpunkten vara högre (compensation adjustment) för att sedan låta grundtonskurvan falla i en väldigt brant kurva.

Utifrån resultaten i denna studie går det inte att sluta sig till att en viss av de nämnda teorierna bör användas mer än någon annan. Många gånger är det svårt att kategorisera kurvorna till en specifik teori. Det är ofta en kombination av flera teorier. Vad resultaten däremot visar är att enbart trunkering tar bort mycket av grundtonskurvan som innehåller rörelse och därmed information. Särskilt när skillnaden i duration är stor. Således torde en mer precis prosodi-modell innehålla avpassning. Generellt nås alltid samma max i grundtonskurvan i vokalerna. Ibland nås till och med ett högre max.

Det verkar inte bara vara durationen som gör att försökspersonerna vill anpassa grundtonskurvan. Om stor del av fallet fortsätter in i konsonanten kan det vara viktigt att få med det fallet redan i vokalen. I de mest tydliga situationerna med border adjustment är skillnaden i duration inte nämnvärt stor. Ibland ter det sig som att ju mindre durationsskillnaderna är desto mindre avpassning. Detta märks tydligast i orden med grav accent där första vokalen är fonologiskt lång.

Försökspersonernas grundtonskurvor är olika men har sinsemellan liknande strategier. Undantaget är AL som avviker genom att i de helt tonande orden med akut accent i första betoningen ofta lägga fallet senare än de andra försökspersonerna. Så långt stämmer det överens med Bruce (1977). Men efter vokalen faller kurvan i den tonande konsonanten. Det verkar vara viktigt att fallet inte försvinner, eftersom hon i vokalen med efterföljande tonlös konsonant tidigarelägger fallet genom att helt flytta det till första vokalsegmentet.

För att återknyta till hur satsaccenten beskrivs enligt Bruce (1977), se kapitel 2.1, kan sägas att resultaten för grundtonskurvan i de helt tonande orden stämmer ganska väl överens, med några undantag. Försökspersonerna i denna studie låter, nästan genomgående, grundtonskurvan börja falla i slutet av första vokalen i *dågen*. Fallet fortsätter sedan in i den tonande konsonanten, och det blir en lägre grundtonskurva i andra vokalen. Återigen med undantag för AL; som i orden med akut accent i icke-final position har en högre grundton i andra vokalen. Enligt Bruce (1977) är det bara en stigning i första vokalen, och andra vokalen innehåller ett grundtonsfall som startar på samma frekvens som den första vokalen slutade.

I *dågen* är det bara en stigning i första vokalen. I konsonanten faller dock grundtonen. Det leder till att hela grundtonskurvan i andra vokalen i regel är lägre än vad som beskrivs i Bruce (1977). Även här är undantaget AL som har

en högre grundtonskurva i andra vokalen i icke-final position.

I orden med grav accent stämmer grundtonskurva i första vokalen och konsonanten väl överens med Bruce (1977). Den stigning som nämns i andra vokalen i icke-final position förekommer däremot bara hos AL och PM. Denna stigning uppkommer dock inte när vokalen föregås av tonlös konsonant, där återstår endast ett grundtonsfall (perturbation).

5.1 Duration

Liksom tidigare studier om skillnader i duration visat (se bl.a. (Eriksson & Alstermark 1972) och (Elert 1965)) skiljer sig samma vokalfonem i duration beroende på vad som omger dem. Denna studie visar också att det ofta blir en betydande skillnad i duration mellan vokalerna som jämförts. Detta är troligtvis fysiologiskt betingat eftersom samma fenomen, om än inte alltid i samma höga grad som i svenska, observerats i studier av andra språk, (Elert 1965). Detta leder till att den kortare vokalen inte hinner med samma vågformade kurva, utan blir så gott som rak och ter sig sträva efter att nå max och minimum så snabbt som möjligt, se figur 4(b). Samtidigt verkar max då också överdrivas i extremfallen, se figur 5(b).

AL, AP och PM har överlag knappt hälften så långa vokaler innan tonlös konsonant i orden med fonologiskt kort vokal. I orden med fonologiskt lång vokal så är de med efterföljande tonlös konsonant knappt tre fjärdedelar så långa. Ungefär samma förhållande råder för KM, som dock låter vokalerna vara drygt hälften respektive tre fjärdedelar så långa. I detta avseende är försökspersonerna mycket konsekventa.

Huruvida föregående konsonant är tonlös eller ej inverkar inte på durationen i andra vokalen, enligt resultaten i denna studie.

6 Sammanfattning

I detta examensarbete har en fonetisk studie av grundtonskurvan i svenska satser gjorts. Syftet har varit att undersöka om det förekommer skillnader i vokalers grundtonskurvor, beroende på om efterföljande konsonantsegment är tonande eller tonlöst.

Med hjälp av fyra försökspersoner har det spelats in ett antal satser där de aktuella vokalerna placerats i olika kontexter. De olika kontexterna består av variationer av ordaccent, kvantitet, ordets position i satsen och tonalitet på efterföljande konsonant. Ur dessa inspelningar har grundtonskurvan vid de aktuella segmenten plottats och medelvärddats. Sedan har grundtonskurvan i vokalerna i orden som endast skiljer sig åt i konsonantens tonalitet jämförts. Orden med de aktuella vokalerna har fokal accent.

Resultaten visar att försökspersonerna anpassar sina grundtonskurvor när tonlösa segment förekommer. Anpassning sker generellt genom tidigareläggning av händelser i grundtonskurvan. Tidigareläggning sker genom antingen ändring av hastigheten på det avslutande fallet i vokalen, tidigareläggning av maxvärdet eller en kombination av dessa. Anpassningen verkar ske för att man inte ska gå miste om grundtonsfallet. I regel nås samma maxvärde och grundtonen faller ner till samma frekvens i slutet av vokalen oavsett om efterföljande konsonant är tonande eller tonlös. Enligt resultaten i denna studie torde man alltså kunna sluta sig till att anpassning av vokalers grundtonskurva vid prosodimodellering bör ge ett naturligare resultat än om man bara trunkerar kurvan. I en prosodimodell kan man då skapa regler som gör att vokaler i ord med fokal accent innan tonlös konsonant tilldelas anpassade fall och stigningar, och behåller de max- och minimumvärden i grundtonskurvan som motsvarande vokal med efterföljande tonande konsonant har.

6.1 Vidare forskning

Ett rimligt antagande är att om man tar hänsyn till detta prosodiska fenomen vid modellering av grundtonskurvor för talsynteser, så kan man öka naturligheten och eventuellt också förståeligheten. För att kontrollera om detta verkligen har någon betydelse kan ett lyssnartest göras där resultaten i denna studie är implementerade i en talsyntes. I ett sådant test får lyssnare bedöma om det ökar förståeligheten och naturligheten i det syntetiskt generade talet. Det skulle också vara intressant att göra dessa analyser på befintliga talsynteser för att se vad som sker med grundtonskurvan i dessa kontexter.

Eftersom durationen verkar vara en viktig faktor vid dessa prosodiska företeelser skulle det vara intressant att studera skillnader i grundtonskurvan mellan öppna och stängda vokaler, då det visat sig i tidigare studier (Elert 1965) att ju öppnare en vokal är desto längre duration och vice versa.

Det finns många varianter på detta fenomen som skulle behöva studeras. Något som skulle vara intressant att studera mer ingående är hur kurvorna skiljer sig åt i andra betoningen i ord med grav accent beroende på om efterföljande konsonant är tonande eller inte. Här har bara klusilerna /k/, /g/, /t/ och /d/ analyserats, men min uppfattning är att det råder ett liknande förhållande mellan de andra konsonanterna som har tonande och tonlösa motsvarigheter. T.ex. kan det vara intressant att studera frikativor i dessa kontexter.

Referenser

- Bruce, G. (1977). *Swedish word accents in sentence perspective*, PhD thesis, Lunds universitet.
- Elert, C.-C. (1965). *Phonologic Studies of Quantity in Swedish*, Almqvist & Wiksell, Uppsala.
- Eriksson, Y. & Alstermark, M. (1972). Fundamental frequency correlates of the grave word accent in swedish: The effect of vowel duration, *STL-QPSR* .
- Ladefoged, P. (1993). *A Course in Phonetics*, Harcourt Brace College Publishers, Fort Worth.
- Lehiste, I. (1970). *Suprasegmentals*, the MIT-press, Cambridge.
- Lyberg, B. (1984). Some fundamental frequency perturbations in a sentence context, *Journal of Phonetics* .
- Lyberg, B. & Sangari, S. (2000). Some observations on different strategies for the timing on fundamental frequency events, *Proceedings of ICSLP* .