

Effekten av talarformanten och F0's styrka på otränade röstbedömare

Helena Bowin

Institutionen för lingvistik
Kandidatexamensarbete 15 hp
Fonetik
Vårterminen 2013
Handledare: Ulla Sundberg
Examinator: Henrik Liljegen
English title: The effect of speaker formant and the intensity of F0 on untrained voice assessors



Stockholms
universitet

Effekten av talarformanten och F0's styrka på otränade röstbedömare

Helena Bowin

Sammanfattning

Tal och röst styrs av komplicerade processer, där vitt skilda funktioner används i dess produktion och perception. Hur tal och röst uppfattas skiljer sig på flera sätt från den objektiva mätbara akustiska talsignalen. Grundtonens, F0's, styrka och det man kallar talarformanten är två parametrar som påverkar hur rösten låter. Studien hade som syfte att undersöka om grundtonens styrka och med eller utan talarformant påverkar vilka stimuli otränade röstbedömare tycker bäst om. För att undersöka detta skapades ett syntetiskt /a/ efter formantfrekvenser för svenska vokaler (Fant, 1959), som varierades med utan eller med talarformant och med sju olika grundtonsstyrkor, totalt fjorton stimuli. Tolv kvinnor och åtta män, fick lyssna på /a/:na och skatta dem från det /a/ som de gillade mest till det /a/ som de gillade minst. Studiens frågeställningar bekräftades, variationer av grundtonens styrka och med eller utan talarformant påverkade vilka stimuli otränade röstbedömare föredrog, och män och kvinnor föredrog olika stimuli. De två /a/:na med svagast och det /a/ med starkast grundton, gillades minst av hela gruppen. De fyra /a/:na däremellan tyckte informanterna om ungefär lika mycket. Studerade man var för sig vad män och kvinnor skattade blev resultatet annorlunda. Kvinnor föredrog stimuli med talarformant mer än de utan talarformant, medan män föredrog stimuli med stark grundton. Ytterligare studier kunde vara intressanta att göra, för att undersöka om stimulus egenskaper med avseende på talarformant och grundtonsstyrka har någon effekt på hur tal uppfattas och tolkas.

Nyckelord/Keywords

Röstperception, röstkvalitet, F0-styrka, talarformant

The effect of speaker formant and the intensity of F0 on untrained voice assessors

Abstract

Speech and voice is governed by complicated processes, where a variety of different functions are used in its production and perception. Speech and voice differs in many ways from the objective measurable acoustic speech signal. The intensity of the fundamental frequency, F0, and what is called the speaker formant are two parameters that affect how voice is perceived. The purpose of the study was to investigate if the varied intensity of F0 and with or without speaker formant affect what stimuli untrained voice assessors like the most. To do so, a synthetic /a/ was created, formed by formant frequencies for Swedish vowels (Fant, 1959), varied with with or without speaker formant, and seven different intensities of the fundamental frequency, creating a total of 14 stimuli. Twelve women and eight men were asked to listen to and grade the stimuli, from the /a/ that was liked the most to the one that was liked the least. The expectations of the study were confirmed, the variation of the intensity of F0 and with or without speaker formant, did affect which stimuli the listeners preferred, and also confirmed that men and women preferred different stimuli. The two /a/'s with the weakest and the /a/ with the strongest intensity of F0, were least liked. The four /a/s with varying intensity of F0 in between, were liked approximately the same. If men and women ratings were studied separately from one another, the result differed. Women preferred stimuli with speaker formant more than the ones without, whereas men more appreciated stimuli with strong intensity of F0. If the study was to be retested and the results would be confirmed, further tests would be of interest to do, to investigate if the differences of preferences affect how speech is perceived and interpreted.

Keywords

Voice perception, voice quality, F0 intensity, speaker formant

Innehållsförteckning

Inledning	1
Bakgrund	1
Talproduktionen.....	1
Andningsapparaten.....	2
Fonation.....	2
Röstkällan.....	2
Röstkällansegenskaper.....	2
Flödesglottogrammet.....	2
Fonationssätt.....	3
Artikulationen.....	3
Talarformanten.....	4
Röstens och talets perception.....	5
Syfte och frågeställningar	5
Material och metod	6
Val av metod och försökspersoner.....	6
Stimuli.....	6
Metod och genomförande av studie.....	7
'Visor'.....	7
Testinstruktioner, samt rutiner för för och efterarbete.....	7
Frågeformulär.....	7
Genomförande av test.....	7
Analys.....	8
Resultat	8
Lyssnartesten.....	8
Spridningsdiagram.....	11
Individuella kvalitetsvärden	11
Frågeformuläret.....	11
Diskussion	12
Metoddiskussion.....	12
Resultatdiskussion.....	13
Allmän diskussion.....	14
Slutsatser	16
Referenser	17
Bilagor	18
Bilaga 1. Testinstruktioner samt rutiner för för- och efterarbete i samband med test	
Bilaga 2. Frågeformulär i samband med test	

Bilaga 3. Spridningsdiagram sid 1-4

Bilaga 4. Individuella kvalitetsvärden för skattningar sid 1-3

Inledning

Med röstens hjälp klarar vi komplexa kommunikationsprocesser. Vi använder rösten när vi talar och sjunger, men vi kan också använda den till att till exempel gråta, skratta och hosta. Vi omges av röster med olika kvaliteter, vissa har stor klang, andra låter tunna och svaga. Det finns många åsikter om vad en bra röst är, vilket har lett till funderingar kring om vad det kan vara vi gemensamt uppskattar i en röst. Röst får här främst stå för de ljud som alstras vid talproduktion, till vilken det behövs andningsapparat, stämband, och ansatsrör (dvs. svalg, mun- och näshåla). Tal- och röstproduktionen kan kortfattat beskrivas så att andningsapparaten producerar en luftström, som sätter stämbanden i vibration varvid ett ljud bildas. När detta ljud, eller vågrörelse, passerar ansatsröret omvandlas det genom artikulationen till olika språkljud. Röstens tonhöjd, styrka och klangfärg, kan varieras så att varje röst klingar unikt för varje individ. Men människors röster har också många likheter, som gör att lyssnare som regel lätt kan uppfatta om en röst tillhör en man, kvinna eller ett barn. Så på samma gång som det unika i rösten har stor betydelse, har röstens många gemensamma nämnare stor betydelse för att vi kan använda rösten till den fantastiska kommunikationsapparat som den är. För att kanske få svar på ytterligare någon fråga i detta komplicerade system undersöktes i den här studien om /a/ med olika grundtonstyrka och med eller utan talarformant hade någon effekt på vad otränade röstbedömare tyckte om vokalklangen. Eftersom testgruppen består av ungefär lika många män som kvinnor, jämförs också män och kvinnors resultat med varandra.

Bakgrund

Tal och röst är komplicerade processer, där vitt skilda funktioner används i deras produktion, perception och tolkning. Talapparaten består av en synkronisering av flera organ, som lungor, in- och utandningsmuskler, struphuvud, svalg, munhåla, tunga, läppar. Den akustiska signal som skapats med hjälp av talapparaten når sedan örat via öronmusslan, hörselgången och mellanörat, som på olika sätt förstärker och omvandlar den inkommande signalen. Inne i innerörat uppfattas frekvenser olika då frekvensernas representation utefter basilarmembranet gör att vi är känsligare för frekvenskillnader i det lägre frekvensområdet jämfört med i det högre, samtidigt som de högre frekvenserna "klustras" till enheter lättare. Hela denna process gör att den slutliga tolkningen av talsignalen skiljer sig på flera sätt från den objektiva mätbara akustiska talsignalen.

Talproduktionen

För att producera tal behöver vi andningsapparaten, stämbanden och ansatsröret, som består av svalget, munhålan och näshålan. Stämbanden producerar med hjälp av luftflödet ljud, som när det förs vidare genom ansatsröret, omvandlas till olika talljud.

Andningsapparaten

Andningsapparaten, dvs. lungorna, mellangärdet, bukväggens muskulatur och de yttre- och inre interkostalmusklerna som sitter mellan revbenen, uppgift är att upprätthålla och kontrollera ett övertryck i lungorna, det subglottiska trycket.

Fonation

När luften pressas genom stämbandsspringan, kallad glottis, sätts stämbanden i vibration, varvid de alstrar ett ljud. Att alstra ett ljud med hjälp av stämbandsvibrationer kallas fonation. Genom att stämbanden initialt förs samman, adduseras, byggs ett tryck upp under stämbanden, som när det blir tillräckligt högt separerar stämbanden. I och med det uppbyggda trycket får luften extra fart när stämbanden går isär, varvid de sugs ihop igen (delvis beroende på Bernouille effekten, turbulent brus bildat av accelererande luftpartiklar), vilket underlättas av stämbandets elasticitet och uppbyggnad. Genom det pulserande luftflödet genom glottis, som sätter stämbanden i vibration, alstras ett ljud, röstkällan.

Röstkällan

Röstkällan består av många samtidigt ljudande toner, ett ackord av olika samtidigt ljudande toner med olika frekvenser. Dessa samtidigt ljudande toners styrka och frekvens kan ses representerade i ett spektrum. Tonerna i ett spektrum kallas för deltoner. I ett spektrum anges varje deltons frekvens och ljudnivå i dB. Den första deltonen, den frekvens som normalt uppfattas som tonhöjden, kallas grundtonen och de övriga övertoner eller deltoner bidrar till röstklangen. Röstkällans spektrum är harmoniskt, dvs. deltonernas frekvenser är heltalsmultiplar av grundtonsfrekvensen, med jämnt avtagande amplitud om ca 6 dB/oktav. Det betyder att om grundtonens frekvens är 100 Hz, blir nästa delton 200 Hz, därefter 300 Hz osv.

Röstkällans egenskaper

Man kan beskriva röstkällans akustiska egenskaper med dess grundtonsfrekvens, amplitud och spektrum. Grundtonsfrekvensen, också kallad F₀, upplever vi som tonhöjd, amplituden som röststyrka och spektrum som röstens klangfärg. Grundtonsfrekvensen ändras genom att muskler i larynx förlänger eller förkortar stämbanden och genom en förstävning av stämbanden som åstadkoms av en sammandragning av vocalismusklerna inne i stämbanden. Röstkällans amplitud styrs av det subglottala trycket, på så vis att om fonationsstyrkan ska höjas höjs det subglottala trycket (Cleveland & Sundberg, 1983). Röstkällans klang, återspeglar dess spektrum, och är en effekt av hur stämbanden vibrerar. Det i sin tur beror sammantaget på subglottiska trycket, larynxmuskulaturen och stämbandets mekaniska egenskaper. I verkligheten hör vi aldrig röstkällan utan anstatsrörets filtrering, så för att studera röstkällan, behöver röstkällan separeras från röstklangen. Det kan man göra med en inversfiltrering. En inversfiltrering fungerar som en "omvänd filtrering", eftersom man vill undersöka källsignalen utan ansatsrörets resonansverkan, använder man filter som tar bort främst de starkaste formanterna, dvs. första och andra.

Flödesglottogram

Röstkällan kan beskrivas med ett flödesglottogram. I flödesglottogrammet kan man utläsa hur mycket luft som passerar glottis per tidsenhet, det transglottala luftflödet, hur länge glottis är stängt och hur länge det är öppet. Detta upprepade mönster visar ljudets periodiska luftflödesvariationer, ljudets vågform. Vågformen har en öppenfas och en slutenfas. Under öppenfasen ökar luftflödet successivt,

och under slutenfaser minskar det. Flödesglottogrammet toppamplitud, som mäts från luftflödets högsta värde, dess topp, till dess dal, när luftflödet är som minst, kallas för pulsamplitud. I början av öppningsfasen passerar luftströmmen långsammare genom glottis för att mot slutet av slutenfaser komma upp i en högre hastighet, vilket gör att öppningsfasen är längre än slutenfaser (Rothenberg, 1981). Ett viktigt samband mellan röstkällans vågform och spektrum är ju större pulsamplitud i flödesglottogrammet, desto starkare grundtonsamplitud i källspektrum. Detta fann Sundberg och Gauffin (1979) genom att studera röstkällan i några sångröster. Fant (1979) fann ett annat samband mellan vågform och spektrum genom uträkningar som visade att övertonernas styrka har med slutningshastigheten att göra. Detta styrktes senare av Gauffin och Sundberg (1980), det vill säga, ju snabbare slutningshastighet, desto starkare övertoner jämfört med de lägre deltonerna. I boken "Röstlära" (s. 101) skriver Sundberg (2001) så här:

"Flödesglottogrammet pulsamplitud bestämmer amplituden på röstkällans grundton, medan dess slutningshastighet bestämmer övertonernas styrka, och en ökning av slutningshastigheten ger större utdelning i de högre än de lägre deltonerna."

Fonationssätt

Flödesglottogrammet pulsamplitud kan varieras beroende på fonationssättet. Man brukar prata om pressad, normal, flödig och läckande fonation. En pressad fonation kännetecknas av högt subglottiskt tryck kombinerat med stor adduktionskraft, vilket ger en liten pulsamplitud och därmed även liten grundtonsamplitud. En pressad röst kännetecknas alltså av en svag grundton. En flödig fonation, vilken skapas genom ett lågt subglottiskt tryck i kombination med svag adduktionskraft (den kraft med vilken stämbanden förs samman) ger en starkare grundtonsstyrka. Det vill säga, ju flödigare fonation, desto starkare blir grundtonshalten i källspektrum (Sundberg & Gauffin, 1979). Hur flödesglottogrammet pulsamplitud och slutningshastighet hänger samman med ljudproduktionen, kan sammanfattas så här:

- Genom att variera fonationssättet bestäms pulsamplituden.
- Pulsamplituden bestämmer röstkällans grundtonsamplitud.
- Fonationsstyrkan, som styrs av subglottiska trycket bestämmer flödesglottogrammet slutningshastighet
- Flödesglottogrammet slutningshastighet bestämmer övertonernas styrka i röstkällan.

När vi varierar det subglottiska trycket och adduktionskraften kan alltså olika fonationstyper skapas.

Artikulationen

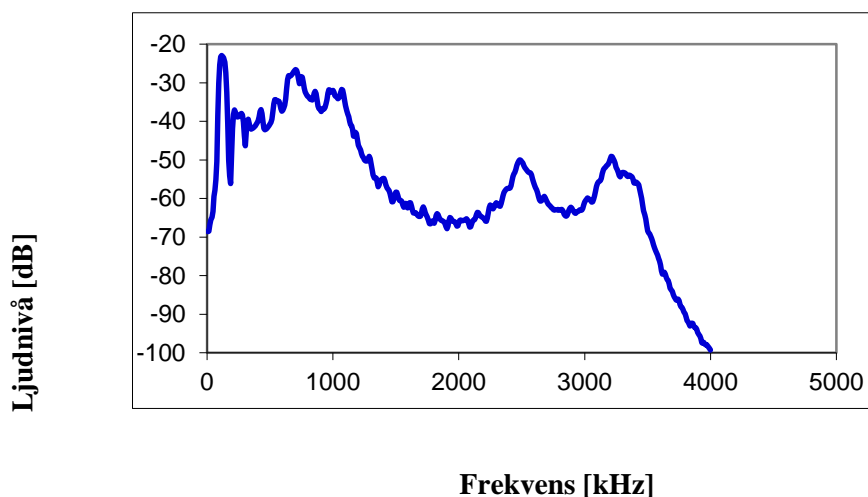
När röstkällan passerar ansatsröret, förstärks vissa frekvenser medan andra dämpas, beroende på ansatsrörets längd och form (Fant, 1960). Ansatsröret fungerar som en resonator som gynnar särskilda frekvenser. De frekvenser som ansatsröret förstärker bäst kallas formantfrekvenser. De deltoner i röstkällan som ligger närmast formantfrekvenserna förstärks mest. Om man vet rörets längd från glottis till läpparna, går det att beräkna resonansfrekvenserna i ansatsröret (Fant, 1960). Ju längre ansatsrör desto lägre formantfrekvenser och ju kortare ansatsrör, desto högre formantfrekvenser.

Ansatsrörets längd och form är viktig för klangfärgen. Med hjälp av artikulatorerna, läppar, tunga, underkäke, velum, larynx och svalgets konstriktormuskler, kan man forma ansatsröret på olika sätt, och beroende formen på ansatsröret, artikulationsprofilen, resonerar olika frekvenser, varvid det bildas olika vokalkvaliteter. Varje vokal får genom att artikulatorerna ställs in olika, olika profiler av ansatsröret (Fant, 1960). Utifrån de olika profilernas genomskärningsyta och ansatsrörets avstånd från

glottis till läpparna, kan man beskriva ansatsrörets form med en kurva, areafunktionen. Förändringen av ansatsrörets form, påverkar formantfrekvensernas läge. Första formantens läge bestäms mestadels av käköppningens storlek, medan andra formanten påverkas mest av tungans form. Man kan bilda alla vokaler genom att variera första och andra formanten. Totalt har ansatsröret upp till sex formanter som anses vara av betydelse. De två första formanterna bestämmer vokalkvaliteten och de högre formanterna är viktiga för den personliga röstkvaliteten.

Talarformanten

Man har länge känt till att den manliga operarösten har en förhöjning i röstspektrum kring 2800 Hz, kallad sångformanten (Bartholomew, 1934; Sundberg, 1974). Denna förhöjning i röstspektrum bildas genom att 5:e formanten, F5, dras ned mot 3:e formanten, F3 och har betydelse för att manliga opera- och konsertsångare kan höras över en orkester utan mikrofonförstärkning. Genom att mäta LTAS (Long-time average spectrum) hos manliga skådespelares röster av bra kvalitet har man funnit att det också finns en topp i talröstens spektrum vid 3 500 Hz (Leino, 1994). Leino lät olika yrkeskategorier inom röstområdet få lyssna på inspelningar av professionella skådespelare och därefter göra en subjektiv bedömning av deras röstkvaliteter på en 6 gradig skala. Leino fann en samstämmighet mellan rösternas kvalitet och en förhöjning vid 3 500 Hz i röstspektrum, vilket antogs vara ett resultat av att F4 och F5 närmade sig varandra. Den spektrala lutningen konstaterades också vara brantare ju sämre röstkvaliteten bedömdes vara, även om vad som bedömdes vara en bra eller dålig röst berodde mindre på den spektrala lutningen än på förhöjningen vid 3 500 Hz. Frekvensförhöjningen i röstspektrum som fanns i de röster som bedömdes ha god kvalitet fick namnet talarformant. Talarformanten ger rösten ökad bärighet och gör att den uppfattas som mer klangfull, även om Leino också fann att några röster med stark talarformant hade bedömts som dåliga och att några röster med svag talarformant hade bedömts som goda. Jämfört med den förhöjning i röstspektrum kring 2 800 Hz, som kallas sångformanten, har talarformanten, se figur 1, alltså sin peak vid den högre frekvensen 3 500 Hz och dess amplitud är jämförelsevis svagare. Observera att sångformanten inte finns representerad i figur 1.



Figur 1. Bilden visar talarformantens frekvensförhöjning vid ca 3 500 Hz. Spektrum som visas är av stimuli 7 (se tabell 1) från studien i detta arbete. Grundtonen är förstärkt med +7,5 dB, se topp vid ca 125 Hz och har tillagd talarformant, se topp vid 3 200-3 400 Hz. F1, F2 och F3 syns också som frekvenshöjningar vid respektive ca 700, 1 000 och 2 400 Hz.

Röstens och talets perception och tolkning

Röstkällans klang är ett resultat av hur stämbanden vibrerar. På vilket sätt de vibrerar beror på subglottiska trycket, larynxmuskulaturen och hur stämbanden ser ut och är uppbyggda. När röstkällan sedan passerar ansatsröret förändras den akustiska signalen utifrån hur vi artikulerar, omformar denna signal. Det ljud, den akustiskt mätbara röstsignalen, fångas sedan upp av vårt auditiva system med dess specifika förutsättningar. Hur vi faktiskt uppfattar signalen beror på många faktorer.

Vid röststörningar används ofta olika auditiva metoder för röstbedömning. Hammarberg et al. (1980) undersökte olika perceptoriska kategorier som användes vid auditiva röstbedömningar, för att skapa en ökad samstämmighet mellan bedömare. Detta ledde fram till fem grundläggande perceptoriska bipolära dimensioner, instabil- stabil, läckande-alltför tät, hyper-hypokinetisk, grov-ljus och falsett-bröstregister.

Att göra perceptuella bedömningar av röststörningar som har god reliabilitet och validitet har visat sig svårt (Kreiman & Gerratt, 1993), trots de ansträngningar som gjorts för att skapa akustiska analysmetoder för att på ett standardiserat sätt kunna mäta olika röstdimensioner, exempelvis av B. Hammarberg et al. (1980). Skillnader på resultat mellan olika bedömare, bedömningstillfällen etc har bl.a. varit några av svårigheterna. De analytiska mätmetoderna används vid diagnosställande och utvärderingar av terapiers effekt vid olika röstproblem på logopedkliniker. Det är därför av vikt att söka mer objektiva parametrar vid bedömning av röster och då inte bara för röster med störningar.

Rogerson och Dodd (2005) har visat att om barn lyssnar på en talare med en dysfunktionell röst får de sämre resultat på frågor på innehållet, jämfört med barn som fick lyssna på en röst av god kvalitet. Detta betyder att det inte bara är ett problem för talaren om rösten inte fungerar bra, utan det påverkar hur mottagaren tar till sig vad som sägs. Barn har dessutom en mindre flexibilitet än vuxna i sina perceptionsstrategier och har svårare än vuxna att vänja sig vid tal med olika störningar (Hazan & Barrett, 2000).

Syfte och frågeställningar

Syftet med studien var att försöka hitta mätbara parametrar i röstspektrum som har betydelse för vad lyssnare uppfattar som en positiv röstklang. Talarformanten och stark grundtonshalt kan kopplas till viktiga delar av röstens spektrum och är betydelsefulla för en god röstkvalitet. Talarformanten, som uppmätts i skådespelarröster kan kopplas till en positiv röstklang (Leino, 1994; Master et al., 2008). En röst med flödig fonation, vilket innebär stark grundtonshalt, ger rösten en bärig kvalitet (Sundberg & Gauffin, 1979). I den studie som har genomförts valdes därför dessa parametrar att varieras vid skapandet av stimuli. Utifrån studiens upplägg och syfte har följande tre frågeställningar formulerats:

- Har talarformanten någon betydelse för vilket /a/ otränade röstbedömarna föredrar?
- Spelar F0's styrka någon roll för vilket /a/ som otränade röstbedömarna föredrar?
- Är det några skillnader mellan vilka /a/:n män och kvinnor föredrar?

Metod och material

Val av metod och försökspersoner

Som metod valdes att göra ett lyssnartest utarbetat för perceptuella skattningar, kallat VSR, Visual Sort and Rate (Granqvist, 2003). VSR är en vidareutveckling av Visual Analog Scale, VAS. Testet genomförs i programmet Visor, också det skapat av Svante Granqvist. I 'Visor' underlättas jämförelser av stimuli som liknar varandra, vilket också gör rankningen av stimuli lättare. I det för denna studie skapade lyssnartestet får röstbedömaren till uppgift att sortera stimuli, 14 olika /a/:n, utifrån sina preferenser. De olika /a/:na varierar med utan eller med talarformant och olika styrka på grundtonen.

Tjugo otränade röstbedömare deltog i studien, 12 kvinnor och 8 män. Av dem var 15 stycken mellan 39-60 år, fyra stycken mellan 69-77 år och en deltagare var 18 år. Testpersonerna är grannar och tillfälliga arbetskollegor till uppsatsskrivaren, utan någon professionell bakgrund inom vare sig tal, röst eller sång och musik, alltså otränade röstbedömare. Alla tillfrågade testpersoner ställde upp och genomförde testet.

Stimuli

För att ha så stor kontroll som möjligt över vad som kan tänkas påverka testpersonens val, valdes att framställa en syntetiskt framställd vokal. Ett manligt /a/ skapades efter formantfrekvenser för svenska vokaler (Fant, 1959) i röstsyntesprogrammet Madde (Granqvist).¹

Först skapades 2 olika /a/:n, ett med talarformant och ett utan talarformant. Talarformanten, vilket beskrivits ovan, kännetecknas av en topp i röstspektrum kring 3 500 Hz, vilken har uppmätts hos talare med röster av god röstkvalitet (Leino, 1994). Därefter varierades dessa /a/:ns grundtonsstyrka med steg av 2,5 dB, från lägsta värdet på -7,5 dB till högsta värdet på +7,5 dB. På så sätt fick man 14 olika stimuli, sju stycken /a/:n med talarformant och varierad styrka på grundtonen om steg med 2,5 dB, och sju /a/:n utan talarformant, också med samma variation på grundtonsstyrkan i steg om 2,5 dB. Dessa spektrumvariationer gjordes med hjälp av programmet Addsynt (Granqvist, 1996). Se tabell 1.

Tabell 1. 14 stimuli med varierande FO's styrka i dB samt med eller utan talarformant.

Stimuli nr		FO's styrka	Stimuli nr		FO's styrka
1	med talarformant	-7,5 dB	8	utan talarformant	-7,5 dB
2	med talarformant	-5 dB	9	utan talarformant	-5 dB
3	med talarformant	-2,5 dB	10	utan talarformant	-2,5 dB
4	med talarformant	0 dB	11	utan talarformant	0 dB
5	med talarformant	2,5 dB	12	utan talarformant	2,5 dB
6	med talarformant	5 dB	13	utan talarformant	5 dB
7	med talarformant	7,5 dB	14	utan talarformant	7,5 dB

¹ Stimuli framställdes i samarbete med professor Johan Sundberg och lektor Svante Granqvist, båda från KTH.

Metod och genomförande av studie

'Visor'

För att kunna genomföra lyssnartestet installerades programmet Visor på en MacBook Pro. För att lyssna på stimuli i testet användes hörlurar av märket Koss, modell KTXPro 1. Visor, ett test enligt VSR metoden som nämnts ovan, fungerar så att till höger på datorskärmen finns ett fält med små paket som innehåller stimuli, i det här fallet 14 st paket. Dessa slängs av programmet slumpvis ut på skärmen inför varje test. Under varje paket döljer sig stimuli, i detta test de olika varianterna av /a/. Genom att klicka på paketen, kan testpersonen lyssna på de olika /a/:na. Uppgiften för testpersonen är att lyssna på stimuli, dra dem från höger fält, till vänster fält, och där rangordna dem. Det /a/ som testpersonen tycker bäst om placeras högst upp på skärmen och därefter i tur och ordning nedåt till det /a/ som testpersonen tycker minst om längs ned på skärmen. Det går att skatta flera /a/:n som likvärdiga, dvs. att ranka flera /a/:n lika högt eller lågt. Testpersonen kan lyssna och flytta på stimuli obegränsat antal gånger, under obegränsad tid. På så sätt kan testpersonerna fritt och enligt sina egna preferenser rangordna de 14 olika /a/:na. Eftersom testpersonerna fritt kan jämföra stimuli med varandra under testet, kan varje stimuli användas som ankarstimuli för varandra, vilket minskar graden av lyssnarens interna referenser (Gerratt et al., 1993).

Testinstruktioner, samt rutiner för för och efterarbete

Inför lyssnartestetets genomförande skapades ett testprocessdokument, "Instruktioner av försöksperson, samt för- och efterarbete i samband med Vokaltest". Se bilaga 1. Instruktionerna lästes upp för testpersonerna inför alla testtillfällen av samma testledare (författaren). Testledaren befann sig i bakgrunden vid alla testtillfällen. Detta för att finnas till hands om något oförutsett och oväntat problem skulle dyka upp.

Frågeformulär

Efter testet fick alla försökspersoner svara på frågorna i "Frågeformulär i samband med Vokaltest". Se bilaga 2. Genom frågeformuläret gavs besked om testpersonernas ålder, kön, utbildning, yrke, sång- eller musikkbakgrund, eventuell hörselskada, flerspråkighet, samt modersmål. Testpersonerna fick också skatta på en skala mellan 1-5, hur lätt/svårt det var att uppfatta skillnad på stimuli, samt hur lätt/svårt det var att rangordna desamma. Se bilaga 2.

Testpersonerna fick innan testet veta att resultaten skulle användas anonymt i en uppsats och har godkänt detta skriftligt.

Genomförande av test

Testmiljön var vid alla tillfällen tyst och utan störande moment och vid inget tillfälle behövdes testet avbrytas. Testet inleddes med att testpersonen skrev under ett godkännande av att testresultaten och svaren i frågeformulären används anonymiserade i detta arbete. Därefter instruerades testpersonen om hur testet skulle gå till genom att de skriftliga instruktionerna lästes upp. Se bilaga 1. När testet var genomfört fick försökspersonen svara på frågorna i "Frågeformulär i samband med Vokaltest". Se bilaga 2.

Analys

Varje /a/ fick genom sin placering i Visor-testet ett kvalitetsvärde (Granqvist, 2003). Kvalitetsvärdet är det värde, på en dold skala mellan 0-1 000, som testpersonens rankning av stimuli ger respektive /a/. Resultaten lades över i text-filer och därefter i excel-filer, en för varje testperson. Där kan utläsas:

- Kvalitetsvärdet, med ett värde mellan 0-1 000, för varje /a/
- Datum och tidpunkt för när de olika testen genomfördes
- Hur länge varje test varade
- Hur många gånger varje /a/ lyssnades på

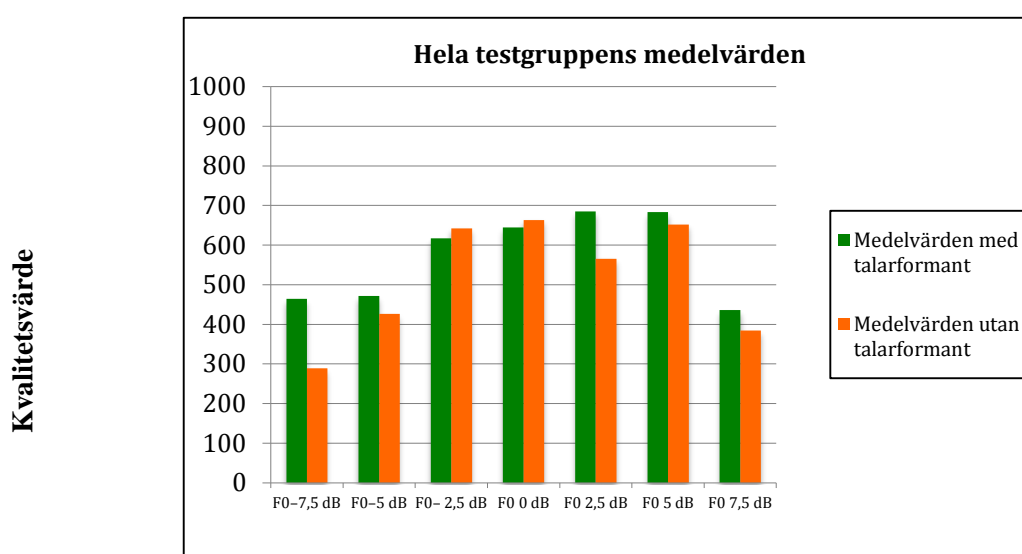
Därefter lades varje /a/:s kvalitetsvärde från de tjugo försökspersonerna ihop och ett medelvärde för varje /a/ räknades fram. Samma sak gjordes för män och kvinnor separat.

Resultat

Lyssnartesten

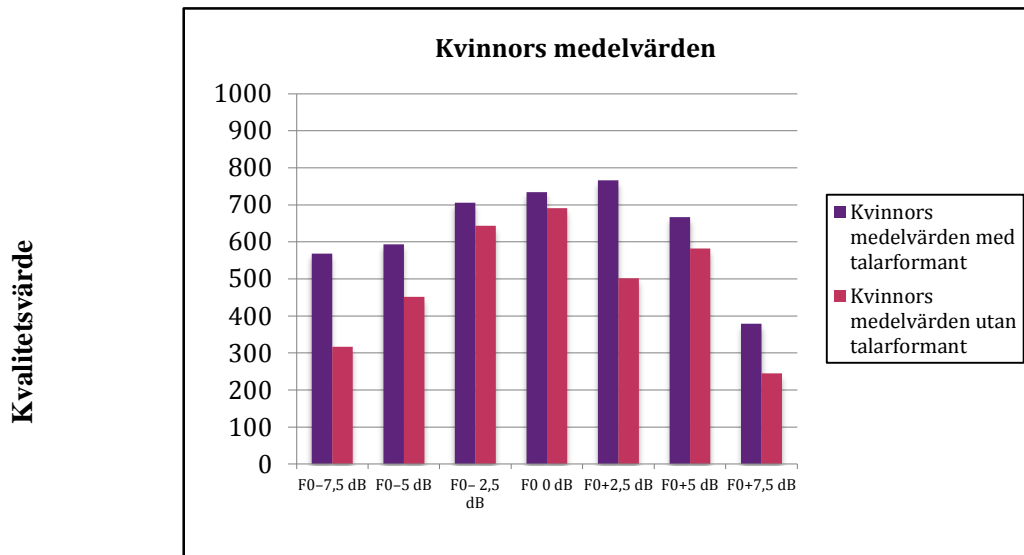
Medelvärdena på testresultaten presenteras nedan i fem olika stapeldiagram. Varje /a/:s medelvärde på kvalitetsvärdet visas i en stapel. De två /a/:nas staplar, med samma F0 styrka mätt i dB, men med närvaro av respektive frånvaro av talarformanten, ligger bredvid varandra.

Studerar man figur 2, kan man konstatera att de /a/:n som fått högst medelvärde är de som har en F0-styrka från -2,5 dB till +5 dB. De vokaler som fått lägst medelvärden, är de med +7,5 dB i F0's styrka och de där F0 har -7,5 och -5 dB i styrka. Däremot verkar närvaro av respektive frånvaro av talarformant spela mindre roll, förutom när F0 styrka krupit ned till -7,5 dB, då får /a/ med talarformant högre medelvärde, jämfört med /a/ utan talarformant. De fyra /a/:na däremellan skattas ungefär lika högt.



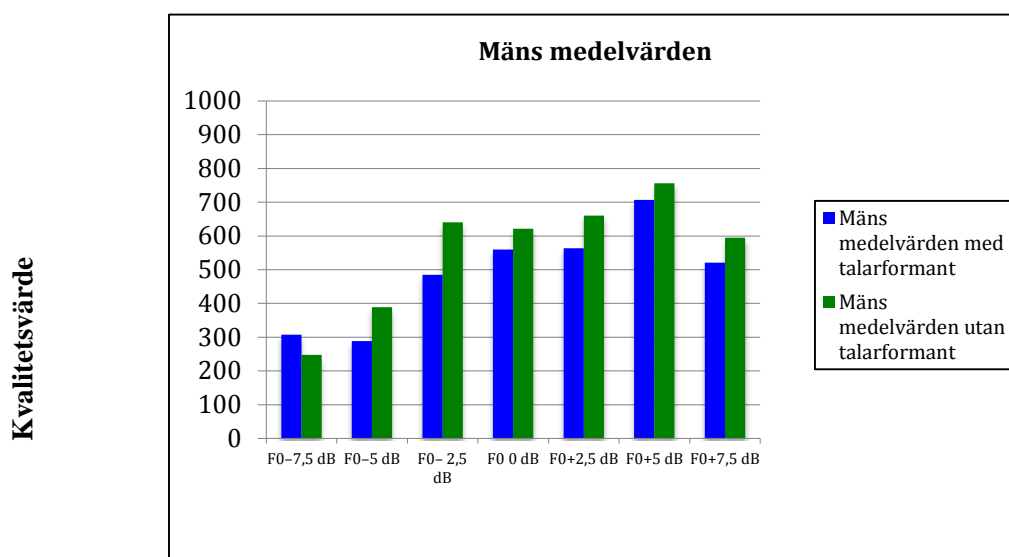
Figur 2. Alla testpersoners medelvärden för de 14 /a/:na med varierande styrka på F0 och med eller utan talarformant.

Om män och kvinnors skattningar separeras från varandra, blir resultatet något annorlunda. Kvinnors skattning skiljer sig inte vad gäller vilka /a/:n som får högst medelvärde, jämfört med hela testgruppens medelvärden, dvs. stimuli med grundtonsstyrkorna -2,5 dB; 0 dB; +2,5 dB och +5 dB får högst medelvärde. Däremot skattar kvinnor /a/:n med talarformant högre än /a/:n utan talarformant, jämfört med hela testgruppen, tydligast på de två /a/:n som har svagast F0, samt /a/ med +2,5 dB. Se figur 3.



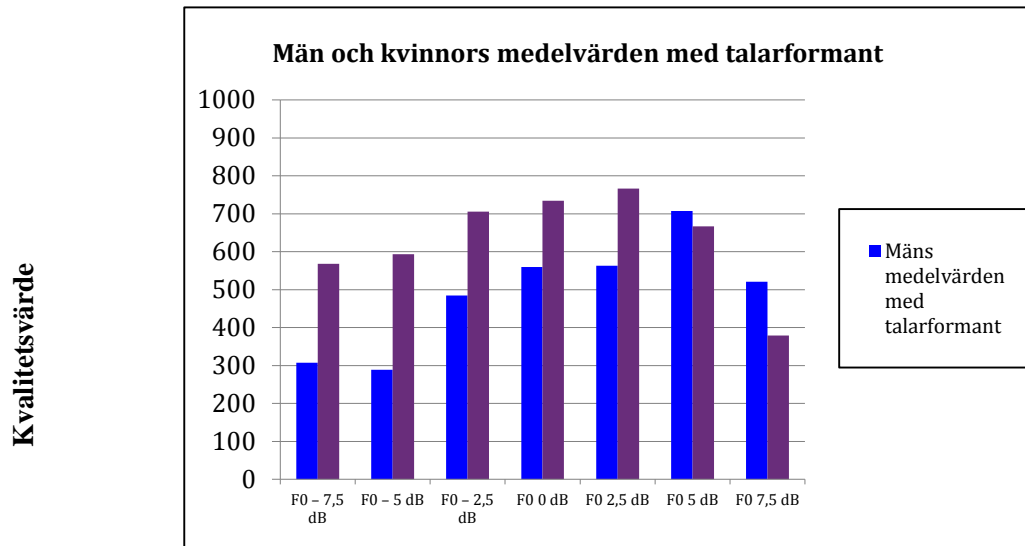
Figur 3. Kvinnors medelvärden på /a/ med varierande styrka på F0 och med eller utan talarformant.

Män skattar tvärtemot de /a/:n utan talarformant något högre, jämfört med de /a/:n med talarformant, förutom det /a/ med svagast grundtonstyrka, -7,5 dB. Skillnaden är dock inte så stor mellan tillagd talarformant eller utan talarformant. Det som verkar spela större roll för män, är F0's styrka. Värdet stegras uppåt mot F0-styrkan 5 dB, för att falla något på 7,5 dB. Se figur 4.



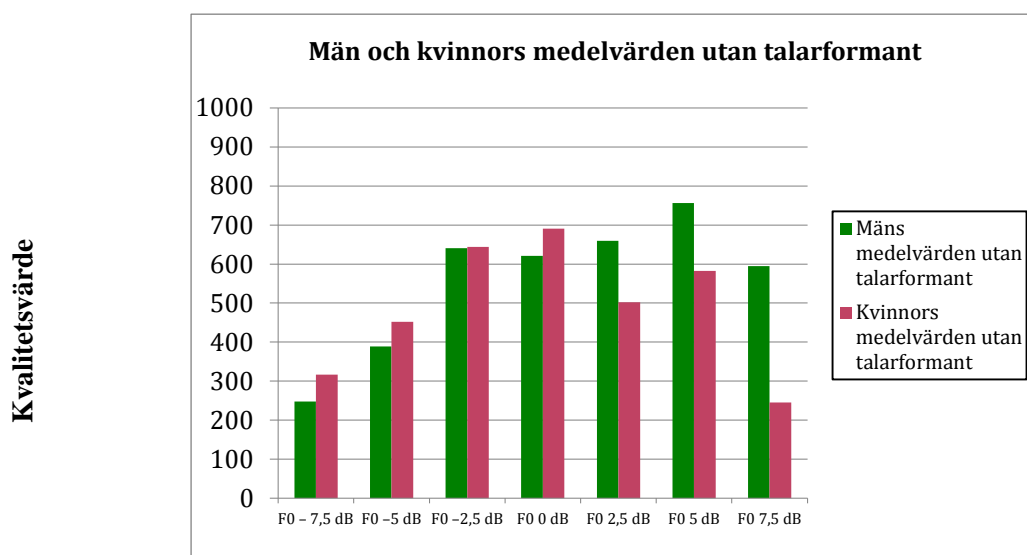
Figur 4. Mäns medelvärden på /a/ med varierande styrka på F0 och med eller utan talarformant.

Jämför man män och kvinnors skattningar med tillagd talarformant ser man att kvinnor skattar /a/ med talarformant högre än män för alla stimuli, utom de med starkast F0, +5 dB och +7,5 dB. Se figur 5.



Figur 5. Jämförelse mellan män och kvinnors medelvärden på /a/ med talarformant.

I figur 6, där mäns och kvinnors skattningar utan talarformant jämförs, skattar män och kvinnor ganska lika på de /a/:n med en grundtonsstyrka mellan -7,5 dB och 0 dB. För stimuli där F0's styrka är +2,5 dB; +5 dB och +7,5 dB skiljer sig skattningarna däremot åt, kvinnor skattar dessa lägre jämfört med hur män skattar.



Figur 6 Jämförelse mellan män och kvinnors medelvärden på /a/ utan talarformant.

Spridningsdiagram

Kvinnor och mäns resultat kan studeras för sig i två olika spridningsdiagram var, ett med tillagd talarformant och ett utan talarformant, där utfallet för varje stimulus kvalitetsvärde kan ses för sig, se bilaga 3. Även om de inte är så genomträngliga visar de överskådligt spridningen på resultaten. En samstämmighet kan konstateras mellan medelvärdet och de individuella värdena, (presenteras nedan) dvs. värdenas tyngdpunkt verkar vara samlade kring medelvärdets tyngdpunkt. I

”Spridningsdiagrammet kvinnor med talarformant” har alla stimuli fått två eller flera kvalitetsvärden på över 800, de fyra med högst medelvärden, medan F0 med minus 2,5 dB; 0 dB; 2,5 dB och 5 dB har inga kvalitetsvärde under 275. I ”Spridningsdiagram kvinnor utan talarformant”, kan en mer divergerad skattning ses vilket också bekräftas av de individuella värdena som är mer spretiga för stimuli utan talarformant. I ”Spridningsdiagram män med talarformant” ses en viss avvikelse från medelvärdets tyngdpunkt för F0 minus 7,5 dB och -2,5 dB, resterande värden är mer samlade kring medelvärdet. ligger stimuli med F0-styrka I ”Spridningsdiagram män utan talarformant” stöds bilden av ett stegrande favoriserande av stark grundton i jämförelse med motsvarande stapeldiagram. Samstämmighet kring medelvärdet råder, förutom vid totalt 4 värden, med ökande kvalitetsvärden mot stimuli med en F0-styrka på 5 dB.

Individuella kvalitetsvärden

I bilaga 4, visas 20 stapeldiagram, ett för varje testpersons kvalitetsvärden för respektive stimuli. Man kan se att det framträder vissa profiler av stapeldiagrammens form. Exempelvis en där testpersoner tydligt föredrar stimuli med tillagd talarformant, t.ex. ”Kvinna 57 år, ekonom”, en annan profil för de som mer gillar stimuli utan talarformant, t.ex. ”Man 18 år, student” och en tredje för testpersoner som mer lyssnar efter grundtonens styrka, oavsett tillagd talarformant eller inte, t.ex. ”Man 69 år, creative director. Sedan finns det några testresultat med profiler där resultaten spretar, t.ex. ”Kvinna 75 år”. Bland kvinnor finns flera som tydligt föredrar /a/ med talarformant, till exempel ”Kvinna 52 år, sjuksköterska” och ”Kvinna 57 år, ekonom”. ”Kvinna 42 år, bildpedagog” verkar lyssna mer efter vilken styrka F0 har och gilla /a/:n med stark grundtonsstyrka. Bland männen resultat är det flera som föredrar /a/ utan talarformant, till exempel ”Man 45 år, konsult”, även om också några män föredrar /a/ med tillagd talarformant, exempelvis ”Man 49 år, läkare”. Man kan alltså konstatera att de skillnader i värden som framkommer när män och kvinnor jämförts på gruppnivå också finns här på individnivå, även om det som är typisk skattningsprofil för kvinnor även kan ses hos en man och tvärt om.

Frågeformuläret

De svar som gavs på frågorna i ”Frågeformulär i samband med Vokaltest” var följande. Frågorna kan ses i bilaga 2. En testperson var 18 år, 8 stycken mellan 39-49 år, 7 stycken mellan 50-60 år och 4 stycken mellan 69-77 år. Totalt deltog 20 personer i studien, varav 8 män och 12 kvinnor. Av dessa hade 15 personer högskoleutbildning, 4 gymnasieutbildning och en hade genomgått grundskola. På frågan om yrke angav 1 sjuksköterska, 2 ekonom, 2 förskolelärare, 1 civilingenjör, 1 bildpedagog, 2 konsult, 2 läkare, 1 egenföretagare, 1 hemmafru, 1 antikhandlare, 1 student, 1 creative director, 1 kapitalförvaltare, 1 scenograf och 2 inget yrke. Två stycken uppgav att de hade hörselnedsättning, men ingen av dessa använde hörapparat under testet. Ingen testperson har utbildning inom musik eller sång. 10 personer utövar eller har utövat musik aktivt, men endast på amatörnivå. (Frågan om på vilken nivå som musiken utövades ställdes ej skriftligt, utan gjordes som ett muntligt tillägg av författaren.) 9 personer uppgav att de var flerspråkiga. 19 personer hade svenska som modersmål.

I frågeformuläret ställdes ytterligare två frågor. Fråga nummer 12 gällde hur lätt/svårt det var att uppfatta skillnaden mellan de olika stimuli och fråga nummer 13 gällde hur lätt/svårt det var att rangordna stimuli. Svaren gjordes på en skala mellan 1 och 5, där 1 var lätt och 5 var svårt. Svaren redovisas nedan i tabell 2 och 3.

Tabell 2. Redovisning av svaren på fråga 12 i frågeformuläret: "Hur gick det att uppfatta skillnaderna på de olika /a/:na?"

Lätt				Svårt
1	2	3	4	5
1 person	4 personer	9 personer	5 personer	1 person

Tabell 3. Redovisning av svaren på fråga 13: "Hur gick det att rangordna de olika /a/:na?"

Lätt				Svårt
1	2	3	4	5
0 personer	3 personer	8 personer	9 personer	0 personer

På frågan "Hur lätt gick det att uppfatta skillnaderna på de olika /a/:na?", svarade tre personer 2, åtta personer 3 och nio personer 4. På frågan "Hur gick det att rangordna de olika /a/:na?" svarade en testperson 1, fyra stycken 2, nio stycken 3, fem stycken 4 och en testperson 5. Sammanfattningsvis kan sägas att testpersonerna upplevde att det var något lättare att uppfatta skillnaderna mellan stimuli jämfört med att rangordna dem.

Diskussion

Metoddiskussion

Syftet med studien var att försöka hitta mätbara parametrar i röstspektrum som har betydelse för vad som uppfattas positivt av lyssnare. Som metod valdes att göra ett lyssnartest där otränade röstbedömare fick till uppgift att rangordna stimuli.

Testpersonerna tillfrågades i sin egenskap av otränade röstbedömare. Att de hittades bland uppsatsskrivaren grannar och tillfälliga arbetskollegor bör inte ha påverkat resultaten. Testmiljön var vid alla tillfällen tyst och utan störande moment, vilket eftersträvades eftersom skillnaderna i stimuli är små och det därför var extra viktigt att skapa en lugn miljö. Byte av testplats tros därför inte ha påverkat resultaten.

Testförfarandet i Visor med VSR möjliggör lättare jämförelser mellan olika snarlika stimuli genom möjligheten att själv flytta och placera de ljudstimuli som upplevs lika nära varandra på skärmen. När stimuli sedan rangordnas vertikalt på skärmen bildar varje stimulus en extern referens för resterande stimuli som liknar fördelarna i test med ankarstimuli (Granqvist, 2003). Syntetiska ankarstimuli

visades i en studie av Gerrat et al. (1993), ge en förbättrad reliabilitet vid lyssningstest. Chan & Yiu (2002), har också kunnat visa att reliabiliteten vid perceptuella röstbedömningar kan öka med hjälp av ankarstimuli. Genom att använda syntetiskt stimuli söktes skapa ett test som så mycket som möjligt säkrade att variationen av F0's styrka och talarformantens betydelse påverkade testpersonernas skattningar. Ett test med ett syntetiskt stimuli skapar också en mer objektiv utgångspunkt för eventuella vidare studier. Visor-testet gav varje stimulus ett kvalitetsvärde från alla testpersoner vilka har varit grundmaterialet vid studerandet av studiens resultat. Koncentration, olika erfarenhet bland lyssnare, liksom olika känslighet för det studerade kan vara sådant som påverkat resultatet (Kreiman et al, 1990).

Testinstruktionerna fungerade väl och att alla testpersoner klarade att genomföra testet utan svårigheter. I frågeformuläret är svaren på frågan om flerspråkighet oklara, eftersom det framkom under testperioden att det rådde delade meningar om vad som är flerspråkighet. Även om de otränade röstbedömarna upplevde det som ganska svårt att rangordna stimuli, pekar resultatens konsekvens mot att otränade röstbedömare klarar att särskilja och gradera de i studien förhållandevis små variationerna av stimuli.

I en studie av Kreiman & Gerrat (2000) fann man att lyssnare var mer överens vid syntetiskt stimuli jämfört med naturliga röster, vilket man tror hänger samman med svårigheten i att uppfatta en enstaka perceptuella dimension i komplexa stimuli, jämfört med syntetiskt mer renodlat stimuli. Det pekar mot att om denna studie skulle upprepas med naturliga röster, skulle det bli svårare för röstbedömarna att uppfatta skillnaderna mellan stimuli, vilket skulle kunna testas i en framtida studie. Hur testets längd påverkar resultaten och om antalet gånger stimuli lyssnas på skulle också kunna vara intressant att undersöka i framtida studier.

Resultatdiskussion

Tre frågeställningar formulerades inför studien som syftade till att hitta mätbara parametrar i röstspektrum som har betydelse för vad lyssnare uppfattar som en positiv röstklang.

- Har talarformanten någon betydelse för vilket /a/ otränade röstbedömarna föredrar?
- Spelar F0's styrka någon roll för vilket /a/ som otränade röstbedömarna föredrar?
- Är det några skillnader mellan vilka /a/:n män och kvinnor föredrar?

Talarformantens betydelse för vilket stimuli som föredrogs. För gruppen som helhet skattades stimuli med tillagd talarformant något högre än de utan talarformant. För kvinnor skattades stimuli med tillagd talarformant klart högre än för stimuli utan talarformant. På individnivå fanns mycket tydliga preferenser för talarformant, framför allt hos kvinnor.

F0's styrkas betydelse för vilket stimuli som föredrogs. Det verkar finnas en gräns för hur svag eller stark styrkan på F0 är. Blir den för svag, faller rankingen, blir den för stark faller den också. Detta mönster syns för hela gruppen, för kvinnor och för män. Männen värden ger dock staplarnas en något annan profil. De stimuli som har svag F0-styrka tappar ännu mer i värde, medan det med högsta styrkan på F0 inte faller lika lågt som i hela gruppen och som hos kvinnor. Man ser tydligt att män uppskattar F0 med stark styrka, däremot stimuli med svag F0-styrka gillas betydligt mindre. I de individuella resultaten framträder skillnaderna även här tydligare, flera av testpersonerna ger F0 med svag styrka riktiga bottenplaceringar, särskilt i gruppen män.

Skillnader i svaren beroende på kön. Svaren visade på ganska tydliga könsskillnader vilket redan nämnts i de två tidigare frågeställningarna. Man kan se både skillnad i val mellan vilken av talarformant och utan talarformant, där kvinnor föredrar stimuli med talarformant och män de utan talarformant. För styrkan på F0 syns också skillnader mellan könen. Män skattar stimuli med tillagd talarformant **och** stark F0-styrka högt, medan kvinnors skattningar dras åt det andra hållet – stimuli med talarformant ges ett lägre värde med den starkaste F0-styrkan medan de med svag F0-styrka ges ett förhållandevis högt värde.

Svarens skillnader väcker tankar kring att lyssnare vid sina val skulle kunna styras av olika preferenser som framträder som profiler i stapeldiagrammen. Dessa profiler kan ses exempelvis där individuella testpersoner tydligt föredrar stimuli med tillagd talarformant, eller för de som mer gillar stimuli utan talarformant, och en tredje för de testpersoner som mer lyssnar efter grundtonens styrka, oavsett tillagd talarformant eller inte. Man kan alltså konstatera att de skillnader i värden som framkommer när män och kvinnor jämförts på gruppnivå också finns här på individnivå, även om det som är typisk skattningsprofil för kvinnor även kan ses hos en man och tvärt om. Talarformanten (Bartholomew, 1934; Sundberg, 1974) och stark grundtonsstyrka (fås genom flödig fonation, (Sundberg & Gauffin, 1979)) påverkar båda klangen på rösten och kan vara orsaken till de resultat som fås.

Frågeformuläret. Svaren visar att vissa frågor visade sig ha betydelse för resultatet, medan andra just i den här studien inte gav något utslag, men skulle kunna vara intressanta att undersöka vidare. Testpersonernas ålder, 15 mellan 39-60, fyra mellan 69-77 och en 18-åring verkade inte ha någon effekt på resultaten. Detta kan dock vara en fråga för vidare undersökning, då testpersonernas ålder kunde vara representerade i alla åldersgrupper och studeras ur just den specifika aspekten. Man kan inte utesluta att preferensen av stimuli ändras med stigande ålder då en naturlig nedsättning av hörseln sker, som påverkar de högre frekvenserna. Män och kvinnors resultat skilde sig åt, som vi sett ovan, men inget stöd för detta har hittats i litteraturen. En majoritet av testpersonerna, 15 stycken, var högskoleutbildade, och man kan därför inte säga något om att det skulle påverka resultaten i studien. Två personer hade hörselnedsättning. Ingen kunskap om hur stor hörselnedsättningen det gällde finns, men det skulle kunna vara intressant att studera vidare kring detta, jämför ovan vid frågan kring åldrar. Ingen av testpersonerna var utbildade inom musik eller sång, hälften av personerna hade eller höll på med musik aktivt på en amatörnivå. Frågan berör en infallsvinkel som kan vara intressant att undersöka, den om musikalitet påverkar testpersonernas svar. Det går inte på grundval av denna studie dra några slutsatser om det. Frågan om flerspråkighet kan förbättras, eftersom osäkerhet kring vad flerspråkighet är fanns. Alla utom en testperson hade svenska som modersmål. Testpersonens modersmål skulle kunna ha en betydelse för vilka stimuli som föredras och i detta fall byggde stimuli på formantfrekvenserna för svenska vokalen /a/ (Fant, 1959).

Allmän diskussion

Att stimulus F0-styrka påverkade resultaten är inte så överraskande då grundtonens styrka har koppling till röstens flödighet. En flödig röstkvalitet upplevs av lyssnare som fri och kärnfull och kännetecknas av stark F0. Talarformanten ger å sin sida rösten starkare övertoner, vilket ofta gör att rösten uppfattas som mer klangfull, inom ett frekvensområde som är betydelsefullt för taluppfattbarhet, vilket kan vara en anledning till att många testpersoner föredrog stimuli med tillagd talarformant jämfört med stimuli utan talarformant. En röst med talarformant bedöms också av röstexperter vara av god kvalitet (Leino, 1994; Bartholomew, 1934).

Det subjektiva i vår perception skapar en inbyggd svårighet när vi samtidigt behöver kunna mäta hur vi uppfattar ljud och dessutom vara överens om vad vi menar när vi beskriver ljud i människans

talkommunikationsprocess. Som nämnts ovan har metoder för röstbedömningar gjorts av dessa skäl (Hammarberg et al., 1980), men som har problem med att bedömningarna inte ger stabila resultat. Kreiman et al. (1990); Kreiman et al. (1992) visade att resultat av lyssningstest påverkas av lyssnarens tidigare erfarenheter vid röstbedömningar och deras bedömningars varierade i resultat. Kreiman et al. (2004), önskar att större ansträngningar görs för att hitta mer objektiva mätmetoder inom röst, istället för att som hittills fokus mestadels legat på att identifiera och definiera beskrivande etiketter eller särdrag för rösten. Detta sätt att mäta röstkvalitet bygger på en lång tradition snarare än på teorier, som utvecklar metoder som kan utvärdera interaktionen mellan lyssnare och signal, istället för att behandla röstkvalitet endast som ett resultat av själva talsignalen. Kreiman et al. (2004) försvarar genom detta ANSI-definitionens (American National Standards Institute, upprättar akustiska terminologier) formella beskrivning av ett ljud, som uttrycks när en lyssnare kan särskilja två ljud med samma styrka och tonhöjd från varandra (ANSI, 1960). Denna beskrivning har fått kritik för att definiera vad en kvalitet inte är, snarare än vad den är.

Rogerson & Dodd (2005) visade att barn lär sig sämre om läraren har röstproblem, se ovan. Ingen liknande studie har gjorts med vuxna än, vad uppsatsskrivaren känner till, men skulle resultaten vara liknande som för barn, blir betydelsen av en god röstkvalitet ännu mer stärkt, då vårt samhälle i stor utsträckning bygger på vår förmåga att kunna kommunicera med tal. Om det skulle finnas olika lyssnarpreferenser, eller kanske strategier, dyker frågor upp om vad i perceptionen som i så fall avgör vilken strategi man väljer, och om det går att ändra på, genom träning exempelvis. En förvärvad hörselskada skulle kunna förändra en sådan eventuell hörselstrategi. På vilket eller vilka sätt det sker, och om en hörselnedsättning ger liknande preferenser hos olika personer, kan vara möjliga forskningsfrågor. Det har varit svårt att finna någon forskning som stöder eller kan förklara orsaken till att resultaten visar skillnad beroende på kön. Att män och kvinnor tycker olika är förstås inte ovanligt i många andra sammanhang. Det finns därför anledning att söka vidare för att finna ett svar på denna fråga, att undersöka om resultaten går att upprepa, liksom att vidare undersöka om det som uppfattas som en "positiv röstklang" också har ett samband med hur tal uppfattas, tolkas och förstås. Studiens resultat skulle kunna få implikationer inom områden som röstbehandling och talsyntes.

Slutsatser

Frågeställningarna i studien har bekräftats, F0's styrka och talarformanten verkar ha betydelse för vilka stimuli otränade röstbedömare föredrar, och män och kvinnor föredrar olika stimuli. Kvinnor gav stimuli med tillagd talarformant högre kvalitetsvärden än män och män föredrog stimuli med starkare styrka på F0 mer än kvinnor gjorde. Resultaten pekar alltså mot en skillnad i preferens som kopplat till kön. Viktigt är att komma ihåg att på individnivå finns stora variationer, det som är typisk skattningsprofil för kvinnor kan ses hos en man och tvärt om.

Att ha en röst som fungerar väl och låter bra spelar stor roll i många yrken. En lärare, skådespelare, politiker, försäljare, föredragshållare med flera är inte bara beroende av att rösten fungerar men också att rösten bär fram budskapet till lyssnaren. Hur en röst klingar har också konsekvenser i nästan alla våra sociala interaktioner, som den exempelvis mellan förälder och barn, elev och lärare, terapeut och patient. Om studien upprepas och resultaten bekräftas kan det vara intressant att undersöka om resultaten påverkar hur vi lyssnar och tar in information och om det har någon betydelse i något sammanhang.

Referenser

- ANSI (1960). Acoustical terminology, *American National Standards Institute, New York*.
- Bartholomew T. (1934). A physical definition of 'good voice quality' in the male voice, *J. Acoust Soc Amer.* 6, 25-33.
- Chan, K. & Yiu, E. (2002). The Effect of Anchors and training on the Reliability of Perceptual Voice Evaluation. *Journal of Speech, Language & Hearing Research*, 45, 111-126.
- Cleveland T. F. & Sundberg J. (1983). Acoustic analysis of three male voices of different quality. *Proc of Stockholm Music Acoustics Conference*.
- Fant G. (1959). An acoustic analysis and synthesis of speech with applications to Swedish. *Ericsson Technics*, 1-108.
- Fant G. (1960). Acoustic Theory of Speech Production. *The Hague, Mouton*.
- Fant G. (1979). Glottal source and excitation analysis. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 1, 85-107.
- Gauffin J. & Sundberg J. (1980). Data on the glottal voice source behavior in vowel production. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 2-3, 61-70.
- Gerratt B. R., Kreiman J. Antonanzas-Barroso, Berke G. S. (1993). Comparing Internal and External Standards in Voice Quality Judgements. *Journal of Speech and Hearing*, 36(2), 14-20.
- Granqvist S. (1996). Addsynt, an additive voice synthesiser for PC. *TMH_QPSR; KTH*, 4, 57-60.
- Granqvist S. (2003). The Visual Sort and Rate method for perceptual evaluation in listening tests. *Computer methods for voice analysis. Royal Institute of Technology Speech, Music and Hearing*.
- Hammarberg, B., Fritzell J., Gauffin J., Sundberg J., & Wedin L. (1980). Perceptual and Acoustic Correlates of Abnormal Voice Qualities. *Acta Otolaryngol*, 90, 441-451.
- Hazan, V. and Barrett, S. (2000) The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics*, 28, 377-396.
- Kreiman J. & Gerratt B.R. (2000). Disagreement in voice quality assessment. *J. Acoustical Society of America*, 108, 1867-1876.
- Kreiman J., Gerratt, B.R. & Precoda K. (1990). Listeners experience and perception of voice quality, *Journal of Speech and Hearing Research*, 33, 103-115.
- Kreiman, J., Gerratt, B.R., Precoda, K. & Berke, G.S. (1992). Individual differences in voice quality perception, *Journal of Speech and Hearing Research*, 35/3, 512-520.
- Kreiman Jody, Vanlancker-Sidtis D., Gerratt B. (2004). Defining and Measuring Voice Quality. *Sound to Sense*, June 11- June 13, MIT
- Leino T. (1994). Long-term-average spectrum study on speaking voice quality in male actors. *Proceedings of the Stockholm Music Acoustics Conference 1193 (SMAC 93), Kgl Musikaliska Akademiens skriftserie 79, Stockholm*, 206-210.
- Master S., De Biase N., Chiari B.M. & Laukkanen A.-M. (2008). Acoustic and Perceptual Analyses of Brazilian Male Actors' and Nonactors' Voices: Long-term Average Spectrum and the "Actor's Formant" *Journal of Voice*, Vol 22/2, 146-154.
- Rogerson J., Dodd B. (2005). Is there an Effect of Dysphonic Teachers' Voices on Children's Processing of Spoken Language? *Journal of Voice*, Vol 19/1, 47-60.
- Rothenberg M. (1981). The voice source in singing. *Research Aspects of Singing, KGL Musikaliska Akademiens Skriftserie*, 33, 15-33.
- Sundberg J. (1974). Articulatory interpretation of the 'singing formant'. *J Acoust Soc Amer*, 55, 834-844.
- Sundberg J. & Gauffin J. (1979). Waveform and spectrum of the glottal voice source. B Lindblom & S Öhman, *Frontiers of Speech Communication Research, Festschrift for Gunnar Fant, Academic Press, London*, 301-320.
- Sundberg J. (2001). *Röstlära. Johan Sundberg och Proprius förlag, Stockholm*.

Bilagor

Instruktioner av försöksperson, samt för- och efterarbete i samband med "Vokaltest".

Inled varje nytt test med att ta fram en ny testsida i programmet Visor.

Namnge testet.

Instruera därefter försökspersonen enligt följande:

1. Testet går ut på att du ska rangordna 14 olika /a/ vokaler, utifrån vilket du föredrar mest till det du föredrar minst.
2. Under testet kommer du att använda hörlurar.
3. Under de fjorton paketen till höger på skärmen döljer sig fjorton olika /a/.
4. När du dubbelklickar på ett paket hörs ett /a/ i hörlurarna.
5. För att rangordna drar du paketen till rutan till vänster.
6. Det /a/ du tycker om mest lägger du högst upp i rutan och det du tycker om minst längst ned. De övriga placeras däremellan.
7. Det är möjligt att lägga vokaler på samma nivå.
8. Det går bra att ändra rangordningen obegränsat antal gånger under testet.
9. Du får lyssna på vokalerna hur många gånger du vill genom att dubbelklicka på dem.
10. Det finns ingen tidsgräns på testet.

Klicka på ramen ovanför rutan till höger när försökspersonen är klar med testet för att avsluta.

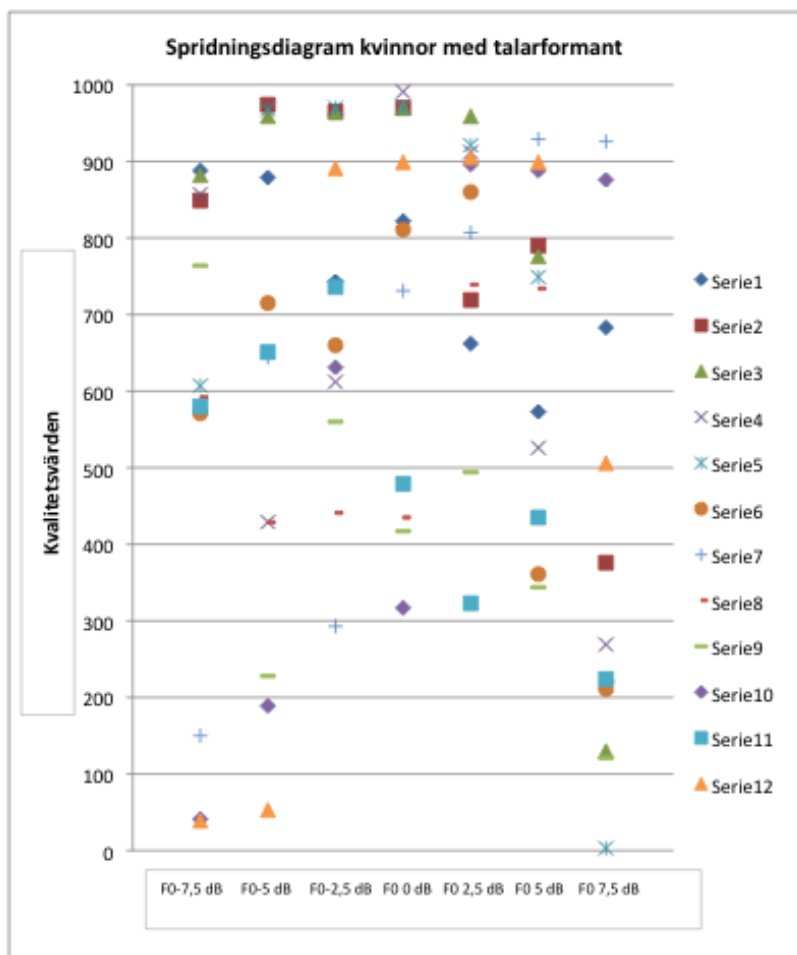
Be försökspersonen svara på frågorna i "Frågeformulär i samband med Vokaltest".

Frågeformulär i samband med "Vokaltest"

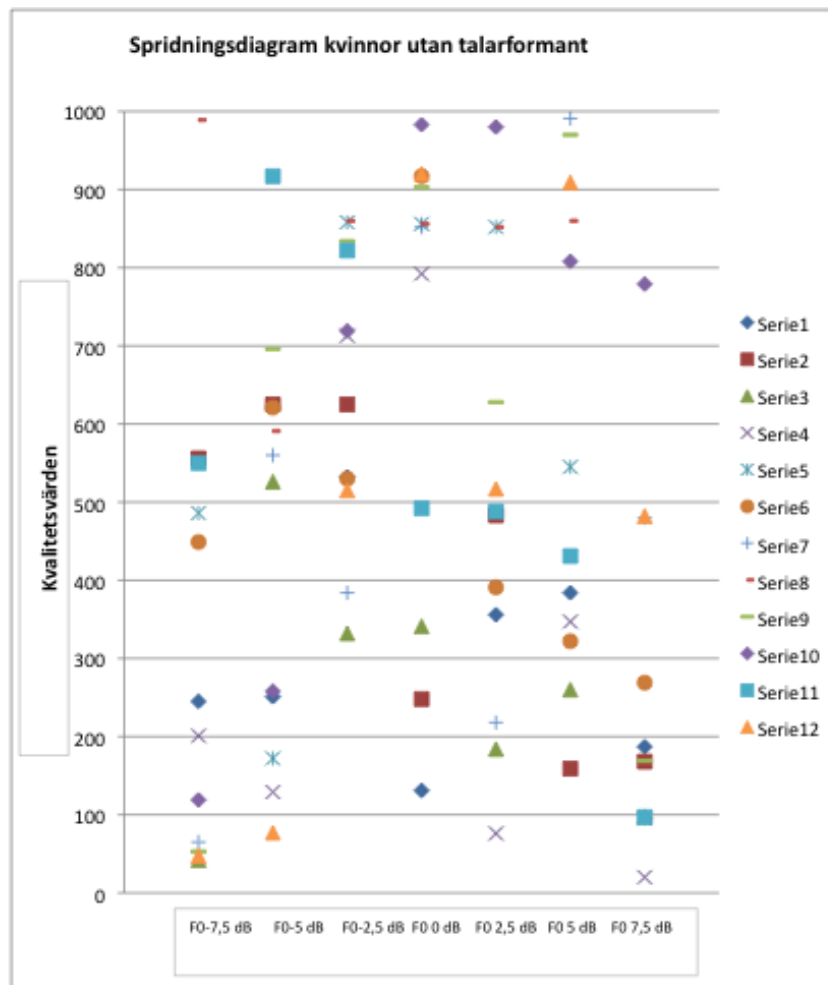
- 1 Vilken ålder har du? år
- 2 Vilket är ditt kön? Man
Kvinna
- 3 Har du gymnasieutbildning? Ja
Nej
- 4 Har du högskoleutbildning? Ja
Nej
- 5 Vad har du för yrke?
- 6 Har du hörselnedsättning? Nej
Ja
- 7 Om ja, använde du hörapparat under testet? Ja
Nej
- 8 Har du musik eller sångutbildning? Ja
Nej
- 9 Utövar du, eller har du utövat musik aktivt? Ja
Nej
- 10 Är du flerspråkig? Ja
Nej
- 11 Är svenska ditt modersmål? Ja
Nej
- 12 Hur gick det att uppfatta skillnaderna Lätt Svårt
på de olika /a/ vokalerna? 1 2 3 4 5
- 13 Hur gick det att rangordna /a/ vokalerna? Lätt Svårt
1 2 3 4 5

Försökspersonens namn i testet:

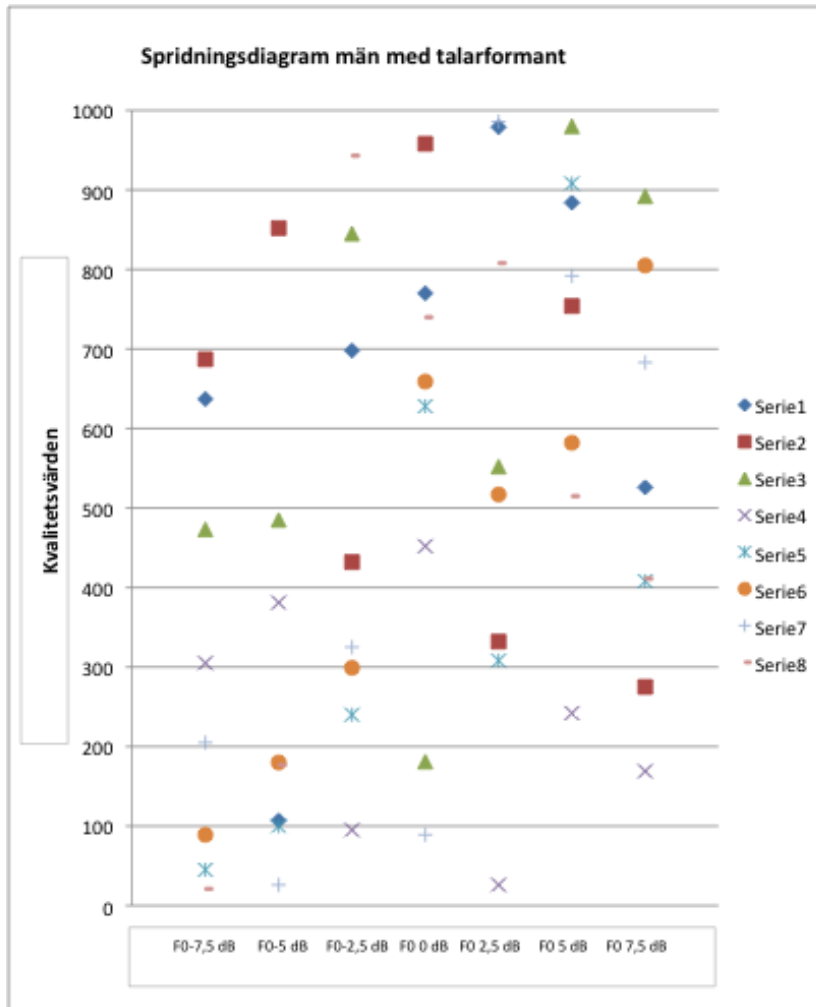
Testtid:



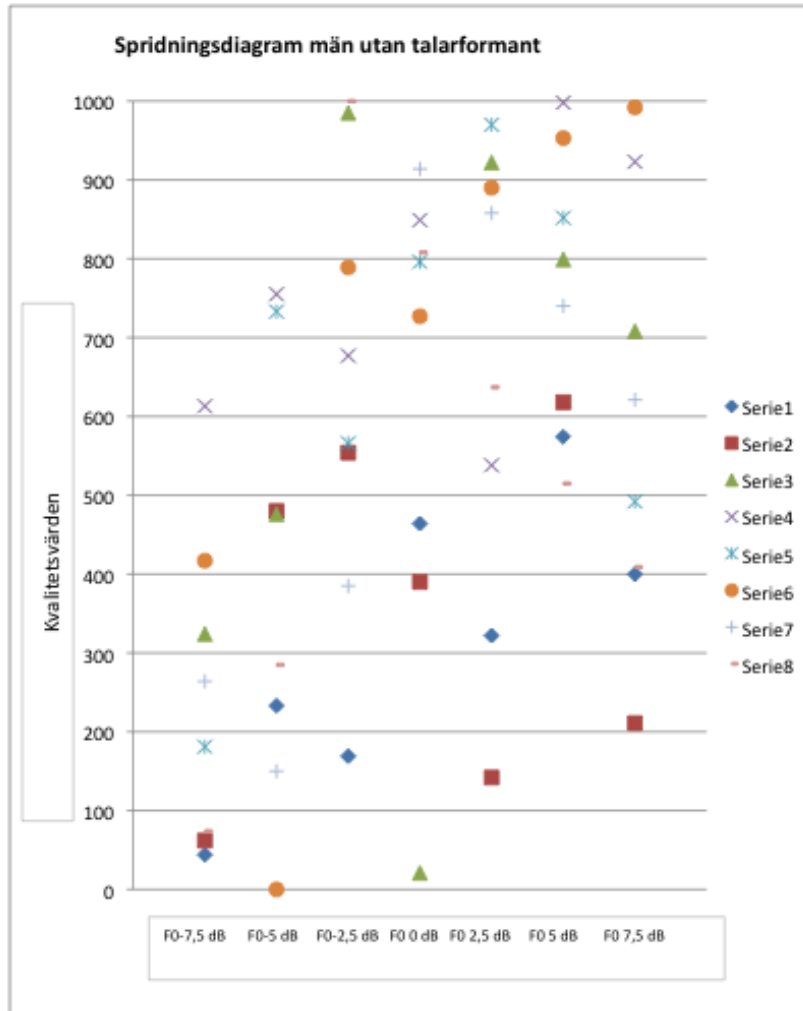
Kvinnors individuella ranking av stimuli med talarformant. Serie 1-12 representerar de olika kvinnorna. X-axeln visar stimuli med talarformant.



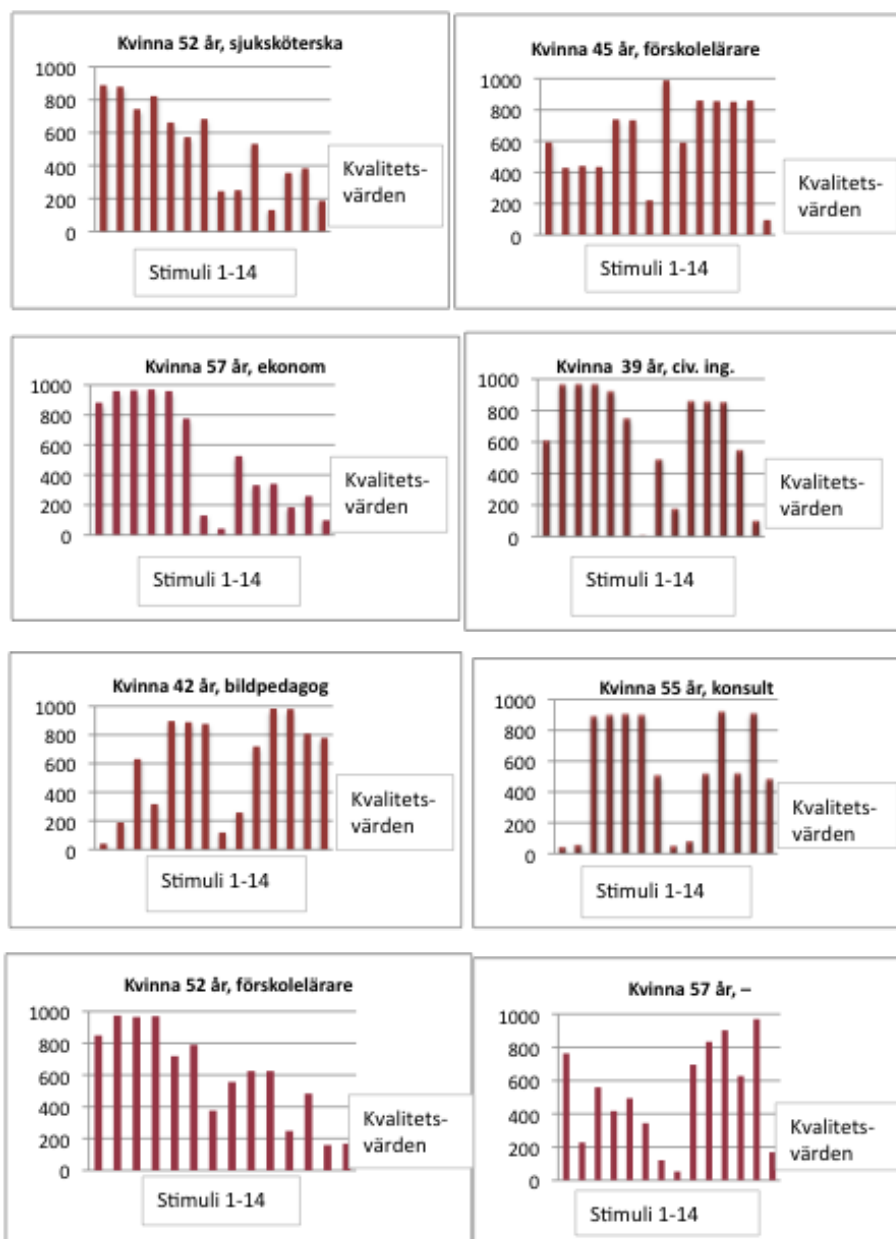
Kvinnors individuella rankning av stimuli utan talarformant. Serie 1-12 representerar de olika kvinnorna. X-axeln visar stimuli utan talarformant.



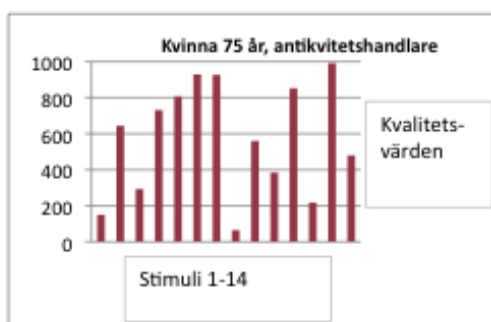
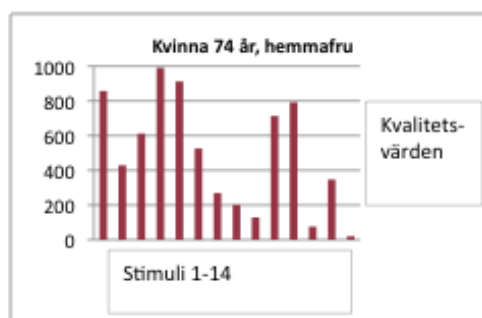
Mäns individuella ranking av stimuli med talarformant. Serie 1-8 representerar de olika männen. X-axeln visar stimuli med talarformant.



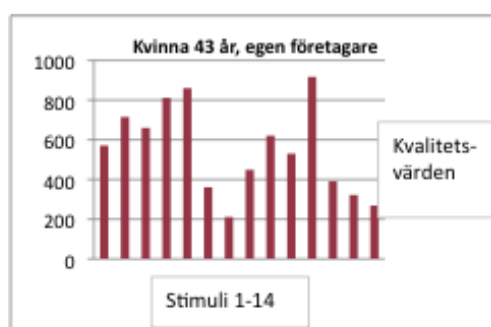
Mäns individuella ranking av stimuli utan talarformant. Serie 1-8 representerar de olika männen. X-axeln visar stimuli utan talarformant.

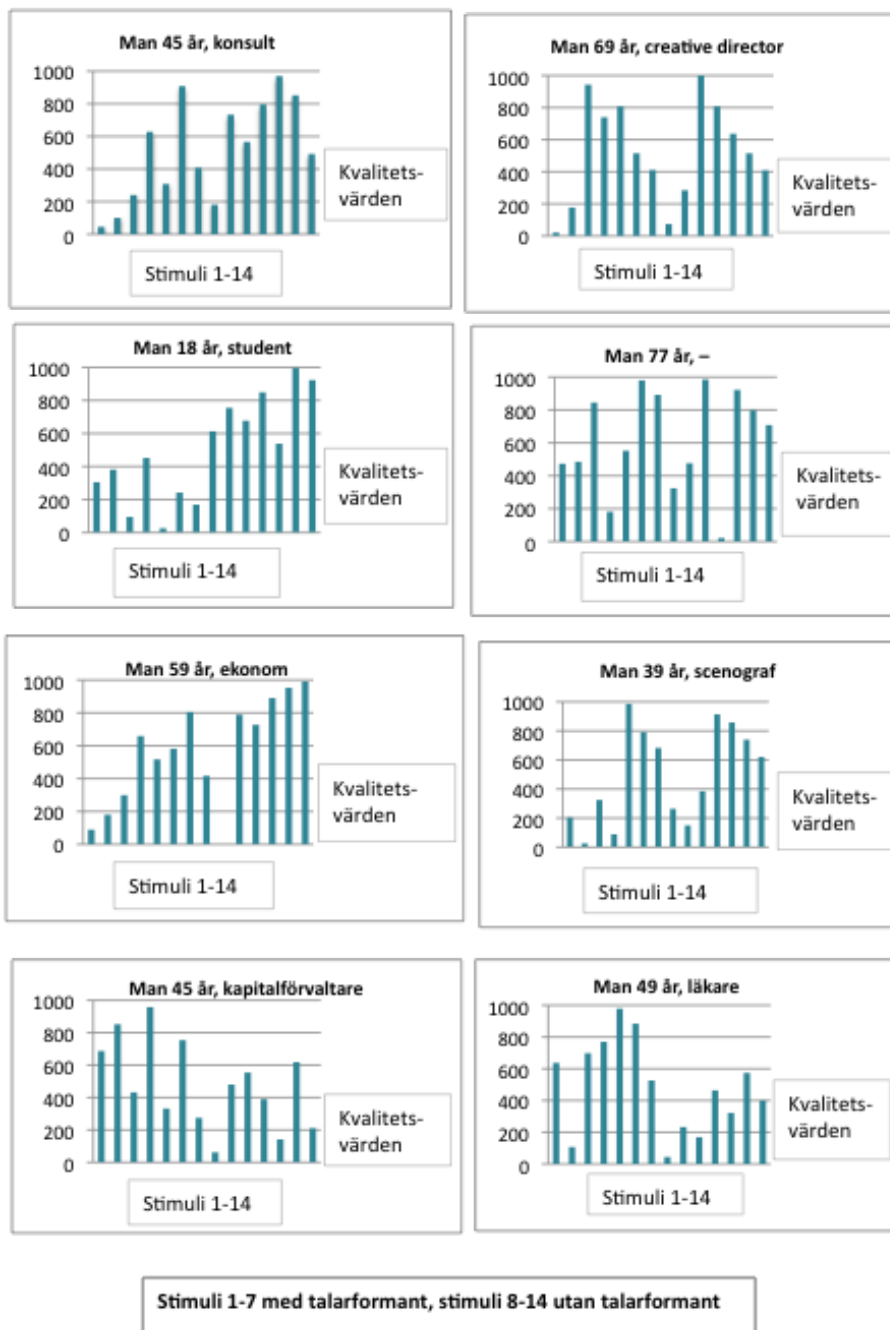


Stimuli 1-7 med talarformant, stimuli 8-14 utan talarformant



Stimuli 1-7 med talarformant, stimuli 8-14 utan talarformant





Stockholms universitet/Stockholm University
SE-106 91 Stockholm
Telefon/Phone: 08 – 16 20 00
www.su.se



Stockholms
universitet