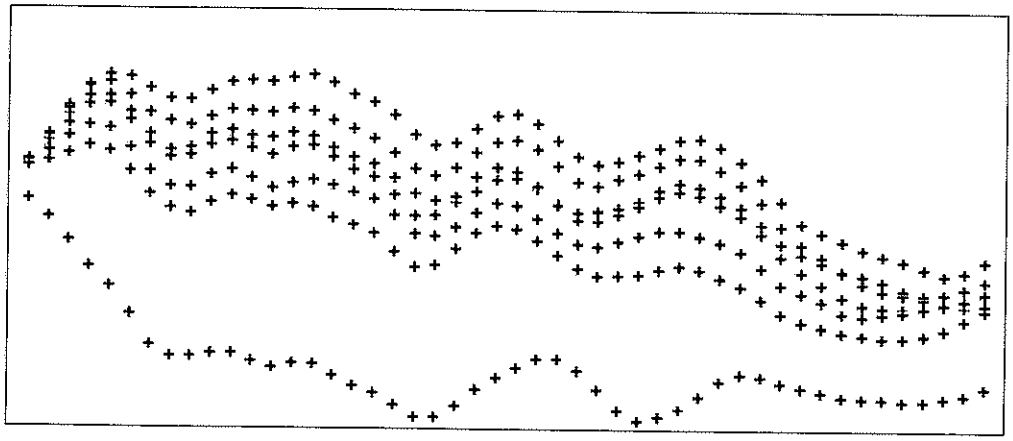
MEDICINSKA FAKULTETEN
Lunds universitet
Institutionen för logopedi, foniatri och audiologi

LTAS' förändring vid variation av röststyrkan?

Maria Nordenberg



Logopedutbildningen, 2003
Vetenskapligt arbete, 20 poäng
Handledare: Johan Sundberg

LITEN ORDLISTA

LTAS Ett långtidsmedelvärdesspektrum (eng. Long Term Average Spectrum, LTAS) visar den genomsnittliga spektrala energifördelningen för en talsignal och representerar egenskaper hos både röstkällan och ansatsrörets överföringsfunktion. Energin anges i form av genomsnittlig ljudnivå (dB).

Sound Pressure Level (SPL) Definieras som logaritmen av förhållandet mellan två tryck (ljudtryck) enligt formeln: $dB_{SPL} = 20 \times \log_{10} (P_1/P_r)$.

Leq Tidsmedelvärdet av ljudnivån beräknad på basis av ljudeffekten. Anges i dB.

K_i Nivåökningsfaktorn är specifik för ett frekvensområde och anger graden ökning av genomsnittlig ljudnivå vid en ökning av röstintensiteten (Leq).

SAMMANFATTNING

Långtidsspektrum (LTAS) är ett effektivt sätt att dokumentera röststatus. Utseendet hos LTAS påverkas dock av hur starkt man talar, varför syftet med denna studie är att bidra med normativa data för hur LTAS konfiguration förändras vid variation av röststyrkan, hos manliga och kvinnliga otränade röster. Kliniskt skulle dessa data kunna användas för att ge en hänvisning om i vilken utsträckning en förändring av rösten beror på en förändrad röststyrka. Eventuellt skulle också framkomna resultat kunna användas till att kvantifiera röstavvikelser. 31 försökspersoner, varav 16 kvinnor och 15 män, gavs uppgiften att läsa en standardtext sju gånger, med maximal variation av röstintensiteten uppmätt i form av genomsnittlig ljudnivå (Leq). Inspelning gjordes på DAT-bandspelare och överfördes därefter till dator för akustisk analys. Sambandet mellan genomsnittlig ljudnivå, vid ett givet frekvensområde, och röstintensiteten (Leq) kan beskrivas av ekvationens lutningskonstant, den s.k. nivåökningsfaktorn (K_i). Beräknade korrelationsvärden visade på ett mycket starkt linjärt samband. K_i är frekvensberoende och användes i studien för att beskriva hur LTAS' konfiguration förändras vid variation av röstintensiteten (Leq). Medelvärde för denna variation beräknades för gruppen män respektive kvinnor och visade sig ha en låg standardavvikelse i båda grupperna. Framkomna medelvärden bekräftar resultat från tidigare forskning, som säger att energinivån vid en ökad röstintensitet (Leq), ökar i större utsträckning vid höga frekvenser ($\geq$ 750 Hz) än vid låga frekvenser (< 750 Hz). Könsskillnader fanns vad gäller K_i's variation med frekvensen. Studiens resultat och metodik kan användas i det kliniska arbetet.

INNEHÅLLSFÖRTECKNING

Sid

1. INLEDNING 1
   1.1 Röstens Intensitet 1
      1.1.1 Skillnader mellan män och kvinnor 3
      1.1.2 Talarformanten 3
      1.1.3 Röstens mognad och åldrande 4
   1.2 Långtidsmedelvärdesspektrum – LTAS 4
      1.2.1 Användning av LTAS 5
      1.2.2 Variation mellan individer 6

2. SYFTE 7

3. METOD 8
   3.1 Försökspersoner 8
   3.2 Procedur 8
   3.3 Inspelning 9
   3.4 Analys 10
      3.4.1 Behandling av data 13

4. FRÅGESTÄLLNING OCH HYPOTESER 14
   4.1 Frågeställningar 14
   4.2 Hypoteser 14

5. RESULTAT 15
   5.1 Extrema röststyrkor och korrelationen mellan dessa 15
      5.1.1 Medelvärden för extrema röststyrkor 15
      5.1.2 Korrelation vid extrema röststyrkor 15
   5.2 Nivåökningsfaktor, K_i 17
      5.2.1 K_i hos män 18
      5.2.2 K_i hos kvinnor 20
   5.3 Differensen: beräknad–uppmätt 23
      5.3.1 Differensvärde för män 23
      5.3.2 Differensvärde för kvinnor 25
   5.4 Beräknad LTAS 27

6. DISKUSSION 29
   6.1 K_i's variation 29
      6.1.1 K_i's variation hos män 29
      6.1.2 K_i's variation hos kvinnor 30
      6.1.3 Skillnader mellan män och kvinnor 31
   6.2 Klinisk tillämpning 31

7. SLUTSATSER 33

SLUTORD 34

REFERENSER 35

BILAGOR

1. INLEDNING

1.1 Röstens intensitet

Röstens intensitet är den akustiska motsvarigheten till den perceptoriska egenskapen röststyrka (Lindblad, 1992), som egentligen är närmare korrelerad med det subglottala trycket (Ladefoged, 1961). Enligt Titze (1994) anges röstintensiteten vanligen som Sound Pressure Level (SPL) i enheten decibel (dB), och regleras i huvudsak med hjälp av tre olika mekanismer. Ett sätt att reglera intensiteten är genom ett förändrat lungtryck, dvs. subglottala trycket. En sådan ökning ökar stämläpparnas slutningshastighet vilket alltså resulterar i en ökad intensitet hos rösten. Beroende på om hastigheten ökar eller minskar erhåller övertonerna i källspektrum mer eller mindre energi, vilket innebär större eller mindre amplitud hos desamma. Likaså mängden aerodynamisk energi som omvandlas till akustisk energi i larynx reglerar röstintensiteten. Hur effektiv överföringen från en energiform till en annan är, beror i huvudsak på storleken hos den glottala öppningen i stämläpparnas svängningsförlopp. Slutligen kan intensiteten regleras till följd av en omformning av ansatsröret. Ansatsröret fungerar som en resonator och försvagar således vissa frekvenser emedan det förstärker andra. Förstärkta frekvenser kallas formanter. Ansatsrörets förmåga att överföra ljud från röstkällan till omvärlden beror bl.a. på röstens grundtonsfrekvens och frekvensavståndet mellan formanterna. Med röstkällan avses här det ljud som produceras i det vibrerande glottis. Enligt Sundberg (2001) bestäms ljudnivån hos röstkällan till större delen av slutningshastigheten hos stämläpparna.

Som nämndes ovan kan röstens intensitet anges i form av SPL i enheten dB. SPL definieras som logaritmen av förhållandet mellan två tryck (ljudtryck) enligt formeln:

$dB_{SPL} = 20 \times \log_{10} (P_1/P_r)$ (Baken, 1987)

I den aktuella studien mäts röststyrkan för en sammanhängande läsning i form av genomsnittlig ljudnivå, och anges i den s.k. ekvivalensnivån (Leq). Energimängden hos signaler, inklusive talsignaler, anges i Watt. Den genomsnittliga energimängden för en signal av en viss tidslängd kan räknas ut och anges i Watt. Omvandlar man detta genomsnittliga värde till decibel har man fått fram nivån för den genomsnittliga energin, den ekvivalenta energinivån (Leq), för signalen. SPL uttrycker energinivån som förhållandet mellan en given energinivå och en referensenerginivå, båda angivna i decibel. För en talsignal kan man räkna ut den genomsnittliga energinivån, eller genomsnittlig SPL. Leq är skilt från genomsnittlig SPL (AB Nyvalla DSP, 1996).

Ett flödesglottogram visar hur mycket luft som passerar glottis per tidsenhet. Vid fonation liknar glottogrammet en vågform där en svängningsperiod består av en stängd fas, adducerade stämläppar, och en öppen fas, abducerade stämläppar. Enligt en studie av Sundberg & Gauffin (1979) bestämmer flödesglottogrammets pulsamplitud grundtonens amplitud så att en stor pulsamplitud ger en stor grundtonsamplitud i källspektrum. I samma studie framkom det att fonationssätt har betydelse för pulsamplitudens storlek. En pressad fonation ger en liten pulsamplitud och därmed liten grundtonsamplitud. En flödig fonation däremot resulterar i en stor pulsamplitud och en stor grundtonsamplitud.

Flödesderivatan visar hur mycket flödeshastigheten förändras i varje skede av perioden i ett flödesglottogram. Den största förändringen uppstår när stämläpparna sluts. Ju större den negativa flödesderivatan är desto mer energi förläggs till övertonerna i källspektrum. Ökad

energi hos övertoner i källspektrum innebär en större amplitud. I en studie av Gauffin och Sundberg (1980) fann man ett nära samband mellan SPL-nivå och flödesderivatans negativa maxamplitud. Dvs. en ökning av röstens SPL-nivå innebar en ökning av den negativa flödesderivatan, vid givna frekvensvärden för grundton och formanter. I ett flertal studier har det framkommit att amplituden hos högfrekventa övertoner tilltar i högre utsträckning än amplituden hos lågfrekventa övertoner, vid en intensitetsökning av rösten (Fant, 1973; Gauffin & Sundberg, 1980; Sundberg, Ternström, Perkins och Gramming, 1988; Gramming, 1991). Sammantaget betyder det att en ökad slutningshastiget hos stämläpparna ger oss en större negativ flödesderivata. Det sistnämnda innebär en ökning av röstens SPL-nivå samt mer energi åt högfrekventa övertoner. Vid den ökade röstintensiteten minskar alltså källspektrums lutning, förutsatt att grundton och fonationssätt hålls konstanta.

Fant (1973) och Ternström (1989) bekräftade de just omnämnda tendenserna i källspektrum, med hjälp av långtidsmedelvärdesspektrum (LTAS). Resultaten från dessa två studier ger också något mer frekvensspecifik information om energifördelningen. Fant (1973) fann att frekvensområdet runt F0 har en hög energinivå i förhållande till högre frekvenser, vid låg genomsnittlig ljudnivå. Däremot vid hög genomsnittlig ljudnivå hos rösten har frekvenser upp till och med 700 Hz en lika hög eller högre energinivå än frekvensområdet vid F0. Analysen i Fants studie gjordes av en mycket kort talsekvens vid tre olika intensitetsnivåer (genomsnittlig ljudnivå). Ternström (1989) studerade LTAS' förändring till följd av intensitetsvariation vid körsång hos gosskör, ungdomskör samt vuxenkör. Förändringen beskrevs med hjälp av en konstant som motsvarar K_i (se ordlista). Hos två av de tre körerna var K_i=1 vid 700 Hz, vilket delvis bekräftar de fynd Fant (1973) gjorde.

Gramming (1991) undersökte sambandet mellan röstens genomsnittliga ljudnivå (SPL) och källspektrums utseende hos män och kvinnor. Vid låg SPL dominerar oftast en delton, vanligen grundtonen, den genomsnittliga SPL–nivån. Vid hög röstintensitet däremot bidrar ett flertal deltoner i källspektrum till den genomsnittliga SPL–nivån. I det senare fallet är den starkaste deltonen en överton.

Tidigare nämndes att Ternström (1989) undersökte hur lutningen i LTAS förändrades vid variation av intensitetsnivån vid sång av gosskör, ungdomskör samt vuxenkör. En linjär funktion visade sig kunna beskriva förhållandet mellan förändringen av intensitetsnivå (Leq) och resulterande förändring av genomsnittlig ljudnivå, i ett givet frekvensområde. Konstanten, "gain factor", i denna funktion är frekvensberoende och enhetslös. Den största spektrumförändringen sågs vid 3 kHz, där en ökad intensitetsnivå (Leq) av storleken 10 dB resulterade i en ökad genomsnittlig ljudnivå på ca 20 dB. Vet man den totala intensitetsförändringen hos rösten kan man med hjälp av den linjära funktionen få fram förväntad förändring hos spektrums energinivå för en given frekvens.

Slutningshastigheten hos stämläpparna är en av de faktorer som bestämmer vilken energinivå övertonerna har innan de förstärks eller försvagas i ansatsröret. Enligt Titze och Sundberg (1991) förstärks det ljud som produceras i röstkällan med 10-15 dB i ansatsröret. Som redan nämnts avgör grundtonsfrekvensens värde formanternas frekvensvärden och frekvens-avståndet dem emellan, vilka övertoner som förstärks respektive försvagas. Ju mindre frekvensavståndet mellan två formanter är desto mer energi kommer att förläggas vid sidan av och mellan dessa formanter. Dvs. överföringsfunktionen ökar vid detta frekvensområde samtidigt som den minskar vid övriga formanter. Exempel på ett sådant fenomen är sångformanten. Sångformanten uppstår, i huvudsak, till följd av att formanterna, F3, F4 och F5 närmar sig varandra vid sång (Sundberg, 2001). Sammanslutningen av dessa formanter ses

som en kraftig spektrumtopp i frekvensområdet 2-3 kHz och leder till en förbättrad ljudöverföringsförmåga hos ansatsröret.

SPL-nivån bestäms oftast av amplituden hos den överton som befinner sig närmast den första formanten (F1). Enligt Gramming (1991) kan första formanten sammanfalla med den starkaste övertonen om formantfrekvenserna hålls konstanta samtidigt som F0 ökas. Titze (1994) menar att ett sådan sammanslutning av en formant och den starkaste deltonen i källspektrum ger en avsevärd ökning av SPL-nivån hos rösten. Detta utnyttjas bland annat vid sång.

1.1.1 Skillnader mellan män och kvinnor

Män har längre ansatsrör än kvinnor. Eftersom ansatsrörets längd har betydelse för vilka frekvensvärden formanterna antar skiljer sig dessa åt vid en jämförelse mellan könen. Kvinnornas formanter har alltså högre frekvensvärden än motsvarande formanter hos män (Huber m.fl., 1999). Enligt Sundberg (2001) kan dock inte skillnader i formanternas frekvensvärden enbart tillskrivas skillnader i ansatsrörets dimensioner.

Skillnader i röstens SPL–nivå mellan könen är svåra att mäta och beräkna eftersom det är flera variabler som bestämmer SPL-nivån. Utifrån tidigare fysiologiska observationer av larynx, ansatsrör samt stämläpparnas vibrationsmönster hos män och kvinnor kom Titze (1988) fram till att en faktor på 1.4-1.6 beskrev storleksförhållandet mellan män och kvinnor för röstens SPL-nivå.

I en studie av Coleman m.fl. (1977) undersöktes vilka röstintensitetens extremvärden är hos män respektive kvinnor, vid mikrofonavståndet 15 cm. I studien fann man att medelvärdet för den maximala SPL-nivån var 113 och 117 dB och för den minimala SPL-nivån 55 och 58 dB, för män respektive kvinnor. I en studie av Gelfer och Young (1997) fann man att intensitetsnivåns medelvärde vid vanligt samtal var 70,42 dB för män, och 68,15 dB för kvinnor vid 30 cm avstånd mellan mun och mikrofon. Baken (1987) menar att SPL-nivån vid normal konversation är ca 70 dB.

1.1.2 Talarformanten

I studier där man försökt hitta akustiska korrelat till det som perceptuellt bedöms vara en god röstkvalitet (Leino, 1993; Nawka m.fl.,1997), har en topp hittats i LTAS vid ca 3,5 kHz. Nawka m.fl. (1997) ansåg att denna framträdande spektrumtopp kunde befinna sig någonstans i frekvensområdet 3150–3700 Hz. Spektrumtoppen har kommit att kallas talarformanten och anses vara en spektrumegenskap karakteristisk för den manliga talrösten, som perceptuellt bedöms ha en god kvalitet. I studien av Leino (1993) analyserades 48 manliga skådespelarröster med hjälp av LTAS. Han fann att spektrumtoppen vid 3,5 kHz var mycket högre hos de personer som bedömdes ha en god röstkvalitet. Däremot de som bedömdes ha en dålig röstkvalitet hade oftast en mycket liten spektrumtopp vid 3,5 kHz. Nawka m.fl. (1997) jämförde i sin studie LTAS från tre röstgrupper; skådespelar-, normala och dysfoniska röster, vid tre olika fonationsstyrkor. Vid en jämförelse mellan normala röster och skådespelarröster var talarformanten 10 dB högre hos de senare vid en fonationsstyrka på 60 dB. Försökspersonerna med dysfoniska röster saknade talarformant. Skådespelarna ombads även att läsa "standardmeningen" starkt samt att skrika. Vid ökning av röststyrkan från normal

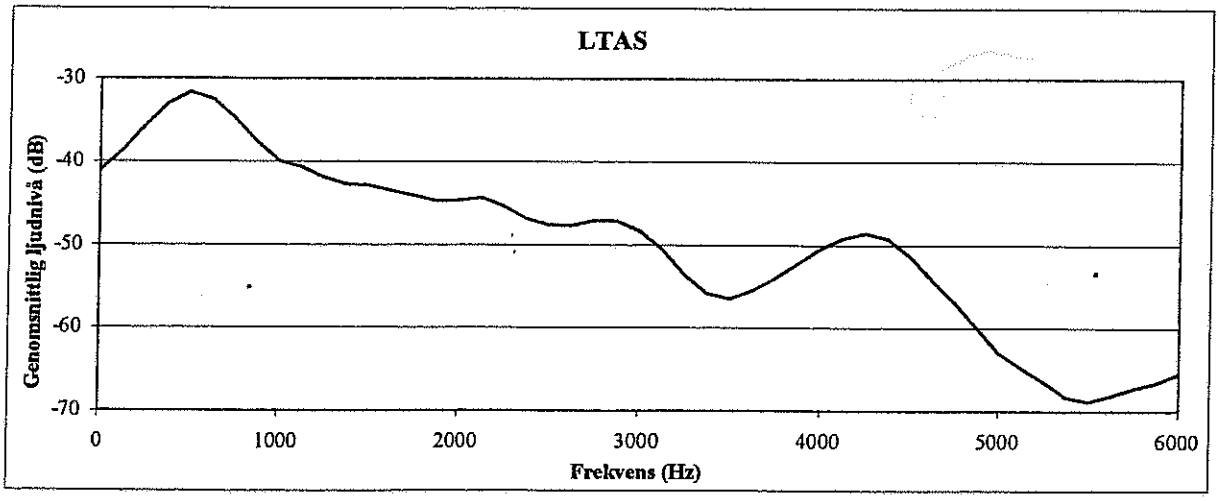
samtalsnivå till starkt tal ökade energinivån hos talarformanten. Däremot vid övergången från starkt tal (80 dB) till skrikande (100 dB) ökade energinivån hos frekvensband som angränsade till talarformanten. Talarformantens energinivå befanns dock vara oförändrad vid denna övergång. Cleveland m.fl. (2001) undersökte LTAS hos countrysångare vid sång och tal. De fann att talarformanten fanns både vid sång och tal. Ingen vet ännu med säkerhet hur talarformanten uppstår.

1.1.3 Röstens mognad och åldrande

Till en följd av struphuvudets tillväxt fortsätter röstens grundtonsfrekvens att förändras upp till en ålder av 20 år (Titze, 1994). Ansatsrörets tillväxt avstannar någon gång under tonåren, och som en följd av detta stabiliseras formantfrekvenserna under denna tid (Huber m.fl., 1999). Enligt Titze (1994) brukar röstkvaliteten, röststyrkans och grundtonfrekvensens stabilitet och omfång hålla sig relativt stabil till och med 60 års ålder. Därefter inträder ofta åldersrelaterade förändringar av vävnader och muskler i larynx, vilket medför en mer instabil och kvalitetsmässigt sett förändrad röst.

1.2 Långtidsmedelvärdesspektrum - LTAS

Ett långtidsmedelvärdesspektrum (eng. Long Term Average Spectrum, LTAS) visar den genomsnittliga spektrala energifördelningen för en talsignal och representerar egenskaper hos både röstkällan och ansatsrörets överföringsfunktion. Energin anges i form av genomsnittlig ljudnivå (dB). Eftersom LTAS visar den genomsnittliga energifördelningen för en hel talsignal får man information om de röstegenskaper som är bestående över tiden. Detta innebär att formanttopparna i LTAS utgör "medelvärdesformanter", både vad gäller frekvens och energinivå, för alla de fonem som ingår i den analyserade talsignalen. Likaså utgör lutningen hos LTAS ett medelvärde för lutningen hos alla röstkällspektrum i hela talsignalen.



Figur 1. *Grafen visar röstens genomsnittliga ljudnivå vid uppmätta frekvensområden hos en kvinnlig försöksperson.*

För att undvika att ett eller flera fonem dominerar i det resulterande LTAS, bör talsignalen vara av en viss längd. Enligt Fritzell m.fl. (1974) är en talsignal på 40 sek. tillräckligt lång.

Enligt Löfqvist och Mandersson (1987) ska man av samma skäl använda en standardtext vid inspelning av rösten. Fant (1973) menar att de röstegenskaper som är bestående över tiden även dominerar över språkspecifika egenskaper i LTAS.

Trots att LTAS används i stor utsträckning för att analysera och beskriva rösten, finns det inget vedertaget sätt för hur man ska extrahera data ur detta spektrum. Förslagen om kvantitativa mått som beskriver LTAS konfiguration och ger akustiska korrelat till perceptuella termer är många (Sundberg, 1986). Osäkerhet råder över hur extraherade data ska tolkas, vilket bl.a. grundar sig i att LTAS innehåller mycket information om både röstkälla och ansatsrörets överföringsfunktion.

Gauffin och Sundberg (1977) valde att dela upp LTAS i tre spektrumband, som avgränsades av frekvensvärdena 0, 2, 5 och 8 kHz. Maxvärdet inom 0-2 kHz bestämdes, liksom minsta skillnad i energinivå mellan de tre frekvensbanden. I en studie av Kitzing (1986) försökte man hitta akustiska korrelat till utförda perceptuella röstbedömningar i LTAS. Ett flertal sätt att extrahera data ur LTAS prövades. Tre sätt att mäta indikerade, med statistisk signifikans, de uppfattade röstkvaliteterna "klang" och "press". De tre sätten var: förhållandet mellan energi upp till1 kHz och energin över 1 kHz (< 1/ >1 kHz) , frekvensområdet i första formanten som mått på spektrumlutning, samt kvoten mellan grundtonsfrekvensens och första formantens amplitud (L_0/L_1). I sin studie av uppfattade egenskaper och dessas akustiska korrelat hos den professionella talarrösten, valde Bele (2002) att räkna ut differensen mellan första formantens energinivå och grundtonfrekvensens energinivå (L1-L0), och energinivåskillnaden mellan första formanten och frekvensbandet 2-3 kHz, respektive 3-4 kHz. Tidigare nämndes att Ternström (1989) i sin undersökning använde sig av en frekvensberoende konstant, "gain factor", för att beskriva intensitetsvariationens utseende hos LTAS. Det sistnämnda sättet används också i denna studie, där författaren valt att kalla konstanten nivåökningsfaktor (K_i).

Studier visar att interbedömarreliabiliteten vid lyssnarbedömning av röster ofta är god (Hammarberg m.fl., 1980; Hammarberg m.fl., 1984) Men faktum kvarstår att den perceptuella bedömningen är subjektiv, och är beroende av känslor samt graden av erfarenhet hos den person som utför bedömningen. Därför finns behov av en objektiv beskrivning av rösten som ett komplement och en bekräftelse av vad en kliniker hör. Behovet gäller framför allt vid funktionella röststörningar där optiska mätinstrument inte visar någon avvikelse i larynx, men där man auditivt kan fastställa att en röststörning föreligger. En akustisk analys erbjuder en sådan objektiv beskrivning och är relativt enkel att utföra i det kliniska arbetet. Den är icke invasiv och ställer låga krav på patienten vad gäller samarbete och deltagande. LTAS är ett röstanalysinstrument som, till skillnad från många andra sådana, erbjuder möjligheten att analysera vanligt tal (Kitzing & Åkerlund, 1993). Flertalet röstpatienter har problem med talrösten. Läsandet av en text ger en mer sanningsenlig bild av talröstbeteendet, än vad exempelvis en uthållen vokal gör. Det senare innebär för många patienter ett onaturligt och obekant röstbeteende.

1.2.1 Användning av LTAS

LTAS har tidigare använts för att försöka ta reda på vilka akustiska egenskaper som medför att man uppfattar en röstkvalitet som god eller dålig (Leino, 1993; Nawka m.fl., 1997). I en studie av Leino och Kärkkäinen (1995) undersöktes akustiska skillnader före och efter träning av talrösten, hos manliga skådespelarstudenter. LTAS efter träning visade en minskad spektrumlutning och en mer framträdande talarformant. Wedin m.fl. (1978) undersökte också

effekten av intensiv röstterapi hos normala röster med hjälp av LTAS. Könsbundna skillnader i röstkvalitet har undersökts med hjälp av LTAS (Mendoza m.fl., 1996). I sin forskning om barnrösten har White (2001) använt LTAS för att undersöka huruvida den perceptuella röstbedömningen har akustiska korrelat vad gäller könstillhörighet.

I ett flertal studier har man funnit att LTAS är ett användbart verktyg för att finna akustiska korrelat till uppfattad röstkvalitet (Gauffin & Sundberg, 1977; Hammarberg m.fl., 1980; Hammarberg m.fl., 1984; Hammarberg m.fl., 1986; Kitzing, 1986). I studien av Hammarberg m.fl. (1980) undersöktes både organiska och funktionella röststörningar. Man fann bl.a. att de båda beskrivande termerna läckage och hypofunktionell röst korrelerade med en kraftig spektrumlutning hos LTAS, medan hyperfunktionell röst korrelerade med en liten lutning. I samma studie påpekades vikten av att utgå ifrån en perceptuell röstbedömning vid sökandet efter akustiska korrelat, och inte diagnosen. En och samma diagnos hos två personer kan vid en kvalitativ jämförelse skilja sig mycket åt.

1.2.2 Variation mellan individer

Kitzing och Åkerlund (1993) undersökte kvalitativ förändring efter röstträning hos personer med funktionell dysfoni. Syftet med studien var att med hjälp av LTAS kunna påvisa relativt små röstförändringar efter utförd röstbehandling. Medelvärdes–LTAS räknades ut för män respektive kvinnor. Den förbättrade röstkvaliteten visade sig som en generell ökning av energinivån i LTAS, där första formanten dominerade något i energinivå. Den generella ökningen var statistiskt signifikant och större för de röster som av lyssnarpanelen bedömdes ha förbättrats i högre grad än övriga röster. Någon signifikant skillnad i konfigurationen av LTAS sågs inte, men kunde däremot observeras hos enskilda individer.

Problemet med den interindividuella variationen i LTAS omnämns i flera studier (Kitzing, 1986; Löfqvist, 1986; Kitzing & Åkerlund, 1993) och har lett till förslaget att LTAS enbart bör användas som akustiskt mått på röstförändringar hos samma individ. Enligt Kitzing (1986), och Kitzing och Åkerlund (1993) beror den interindividuella variationen till stor del på variation av intensitetsnivån. Som tidigare nämnts bestäms flera akustiska röstparametrar, tex. F0 och intensitet, av egenskaper i både röstkälla och ansatsrör. Detta har visat sig medföra svårigheter att bestämma exakt vilken egenskap som ligger bakom en förändring i LTAS. Ett sätt att lösa problemet är att försöka få patienten att kontrollera sin genomsnittliga röstintensiteten i den kliniska situationen (Kitzing, 1986). Denna uppgift, som är både onaturlig och svår för flertalet patienter, ökar emellertid risken för ett onaturligt och spänt röstbeteende i inspelningssituationen. Det finns ett behov av att undersöka LTAS som röstanalysmetod (Kitzing & Åkerlund, 1993), vilket för oss in på syftet med denna studie.

2. SYFTE

Syftet med studien är i första hand att undersöka hur LTAS förändras i olika frekvensband vid variation av röststyrkan, hos otränade vuxna män respektive kvinnor. Vidare bidrar studien med normativa data i ämnet. Förhoppningen är att framkomna resultat ska bidra med information om LTAS som röstanalysmetod samt underlätta användningen och tolkningen av LTAS i det kliniska arbetet. Röststyrkenivåerna mäts som genomsnittlig ljudnivå i form av den s.k. ekvivalentnivån (Leq), definierad som tidsmedelvärdet av ljudnivån beräknad på basis av ljudeffekten. Förändringarna hos LTAS beskrivs med hjälp av en frekvensberoende konstant, nivåökningsfaktor, definierad som lutningen hos den bästa linjära approximationen av hur mycket ljudnivån i ett givet frekvensområde ökar med Leq.

3. METOD

Innan testningen av försökspersonerna påbörjades genomfördes två provinspelningar med efterföljande analys. Dessa två personer deltog inte i studien och resultatet användes endast i syfte att upptäcka nödvändiga justeringar av det originalschema som upprättats för inspelning och analys.

3.1 Försökspersoner

Efter utfört bekvämlighetsurval spelades 32 personer in. 16 kvinnor och 16 män. Efter ett bortfall på en man ingick totalt 31 personer i studien. Orsaken till bortfallet var att DAT bandspelarens inspelningsnivå var felinställd. Försökspersonen gick inte att nå för ytterligare en inspelning. Medverkan var frivillig och de deltagande garanterades anonymitet. Försökspersonerna utgjordes av studerande och anställda på institutionen för Tal, Musik och Hörsel, vänner till författaren och ibland vännernas vänner. Inget arvode utgick för medverkan i studien. Ett representativt normalmaterial eftersträvades och följande inklusionskriterier användes i studien:

- Ålder 20-50 år
- Ingen pågående övre luftvägsinfektion vid testtillfället
- Inga uppenbara röstbesvär/röststörning.

Åldersgränserna valdes med tanke på röstens utveckling och åldrande. För att få bakgrundsfakta om försökspersonernas ålder, röstvanor, rösthygien och eventuella subjektiva röstbesvär användes en enkät av Fahlstedt och Morell, 1999 (bilaga 1). För de män och kvinnor som ingick i studien var genomsnittsåldern 29 respektive 28 år. Extremvärdena var för männen 23 och 35 år, och för kvinnorna 21 och 40 år. Med undantag av en kvinna (K1) hade de försökspersoner som ingick i studien, otränade talröster.

3.2 Procedur

Vid testning läste varje försöksperson en standardtext. Valet av text gjordes med tanke på att variation av läshastighet kan förekomma, och för att det skulle kännas naturligt att höja röststyrkan och hålla den på en jämn nivå. Standardtexten (bilaga 2) lästes sammanlagt 7 gånger, med ca 10s paus mellan varje läsning. För att få en spridning på den genomsnittliga röststyrkan fick försökspersonen lyssna till bakgrundsbuller i form av "partybabbel" (Widex ApS. Real-life environment sound examples), i hörlurar och ombads överrösta detta. För att minska graden av osäkerhet under inspelningen fick försökspersonen först läsa igenom texten tyst, och sedan pröva att läsa de första meningarna under verkliga testförhållanden.
Inför den 7:e och sista läsningen togs hörlurarna bort och försökspersonen läste texten med så svag röststyrka som möjligt utan att viska. Varje sekvens av "partybabbel" varade i 45s. Beroende på varierande läshastighet förekom en viss variation av talsignalens längd. Den understeg dock aldrig 41s. I talsignalens längd inräknas både tonande och tonlösa segment, samt pauser i talet.

3.3 Inspelning

Talsignalen spelades in med en DAT-bandspelare, via en huvudburen mikrofon på konstant avstånd från munnen. Inför varje inspelning mättes avståndet från mun till mikrofon som varierade mellan 29-30cm, utom i ett fall då avståndet var 32cm. Mikrofonen kalibrerades inför varje inspelning. Detta skedde genom att några kalibreringstoner av olika, konstanta ljudnivåer spelades in med mikrofonen intill en ljudnivåmätare. De avlästa ljudnivåerna angavs på inspelningen.

Med undantag av de åtta första försökspersonerna (bandspelare 1) gjordes inspelningen på två kanaler, med olika förstärkning. Syftet med detta var att få ett större utslag på läsningar med låg intensitet, så att dessa inte maskerades av inspelningsbruset. För att få en inspelning på två kanaler användes en mixer. Efter inspelning av tre försökspersoner ersattes denna, pga. tekniska problem, med en DAT-bandspelare. För de åtta första försökspersonerna kunde samma effekt ,vad gäller utslag, uppnås genom en överföring av ljudsignalen från bandspelare till dator via två kalibrerade inspelningskanaler.

"Partybabblet" spelades upp i fem realistiskt förekommande ljudnivåer (I) i ordningen enligt tabell 1 nedan. Dessa nivåer uppmättes i inspelningsstudion, genom att hörlurarna sattes intill ljudnivåmätaren under de 45s som partybabblet varade. Med hjälp av ljudnivåmätaren kunde sedan den genomsnittliga intensitetsnivån(Leq) mätas för aktuell sekvens. Samma procedur genomfördes för alla sex sekvenser med partybabbel.

Tabell 1. Turordning och värde på realistiskt förekommande intensitetsnivåer hos partybabblet.

Läsning	I (dB)
1	85
2	95
3	85
4	81
5	88
6	77

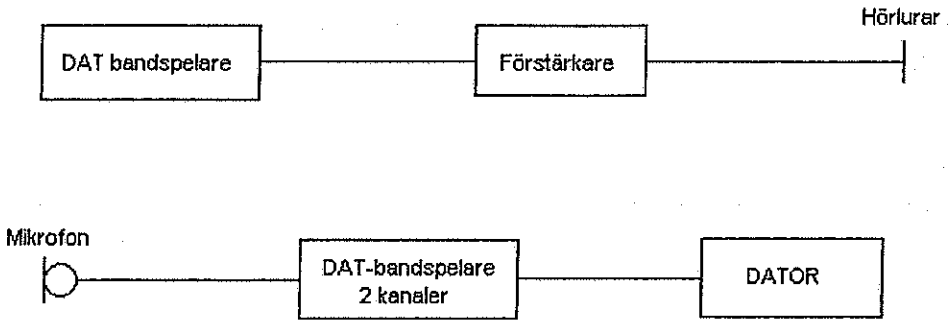
Vid användande av hörlurar försvinner en del av den "medhörning" man normalt har på sin egen röst. En ocklusionseffekt uppstår och för att minska denna valde författaren att använda hörlurar som inte helt tog bort "medhörningen". Detta medförde att partybabblet till viss del spelades in. Innan testningen med försökspersonerna startade kontrollerades om detta eventuellt gav missvisande resultat. Det visade sig gå bra också vid de svagaste läsningarna med babbel. Vid jämförelse av LTAS för den svagaste läsningen och LTAS för det starkaste partybabblet var minsta skillnaden dem emellan 8 dB. Samtliga försökspersoner genomförde inspelningen sittande i en ljudisolerad studio, med texten framför sig i lämplig höjd. För sex försökspersoner var en andra inspelning nödvändig pga. tekniska problem vid inspelningstillfället. Som nämnts var inte försöksperson M8 tillgänglig, vilket resulterade i ett bortfall. I samtliga fem fall genomfördes den andra inspelningen vid ett annat tillfälle(dag).

Utrustning som användes vid inspelningen var:

- Huvudburen mikrofon: AV-JEFE,TCM 110
- Bandspelare 1(inspelning): DAT Walkman, Digital recording, TCD-D100, SONY
- Bandspelare 2(inspelning): SONY TCD-D10, DIGITAL AUDIOTAPE-CORDER

- Bandspelare ("partybabbel"): SONY TCD-D10 PRO, DIGITAL AUDIOTAPE-CORDER
- Förstärkare ("partybabbel"): Fostex, personal monitor, Model 6301B
- Hörlurar: MONITOR, K240(600ohms), AKG
- Mixer: Behringer, Eurorack, MX602A, Ultra Low Noise Design
- Decibelmätare 1: LA-210, SOUND LEVEL METER, ONO SOKKI
- Decibelmätare 2: BRUEL & KJAER TYPE 2215, PRECISION SOUND LEVEL METER
- Decibelmätare 3: RION, NL-05, Integrating Sound Level Meter

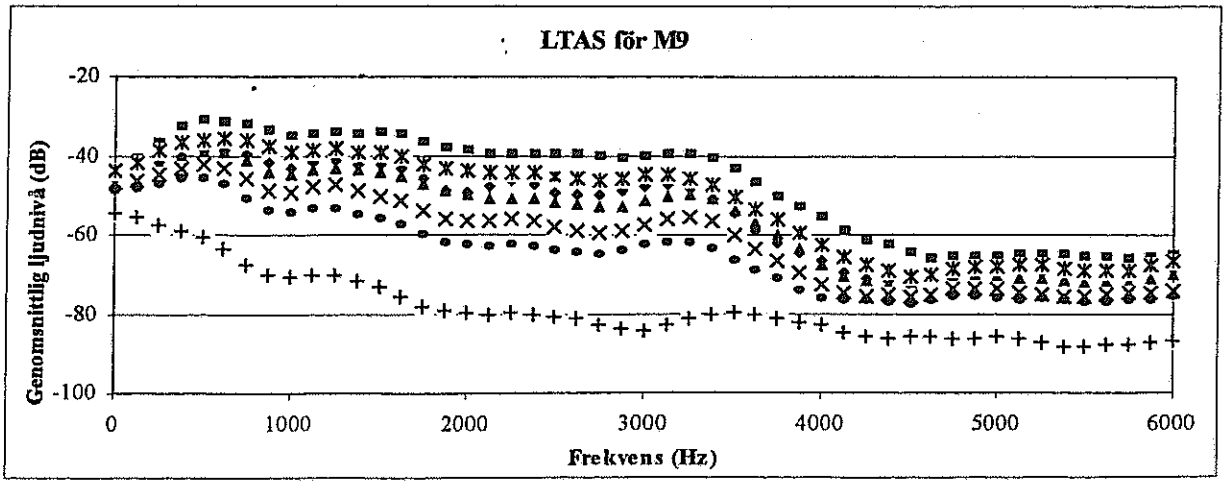
Alla tre decibelmätare som användes var likvärdiga och det faktum att tre olika mätare användes påverkar ej resultaten.



Figur 2. Kopplingsschema vid inspelning och efterföljande analys.

3.4 Analys

De inspelade talsignalerna överfördes från DAT-bandspelare till filer i signalanalysprogrammet Soundswell Signal Workstation 4.0, Soundfile Editor (AB Nyvalla, DSP,1996), med samplingsfrekvensen 16000 Hz. LTAS erhölls därmed för varje lässekvens. Alla sju LTAS, från varje försöksperson, samlades på en och samma ljudfil.



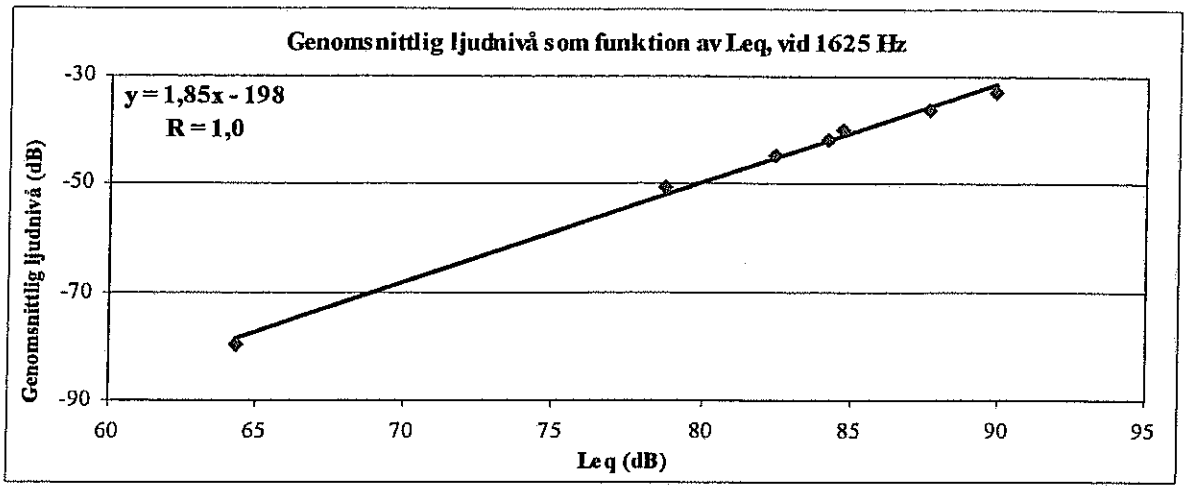
Figur 3. Graf som visar LTAS av de sju lässekvenserna för försöksperson M9.

Ur grafen i figur 3 kan man utläsa att den genomsnittliga ljudnivån ökar över hela frekvensområdet vid en ökning av röstintensiteten (Leq). Man kan man även se att den största ökningen för M9 befinner sig mellan 1000 och 3500 kHz.

Samtliga lässekvenser avlyssnades, och det visade sig i några fall förekomma överstyrning. Denna överstyrning omfattade en liten del av den totala tiden men för att vara säker på att felaktiga utslag på högre frekvenser inte förekom eliminerades dessa avsnitt. Diskontinuiteter som ibland uppstod till följd av denna process restaurerades, så att en kontinuerlig talsignal erhölls. Eftersom dessa inslag utgjorde en mycket liten del av totala talsignalen påverkades inte slutgiltiga resultat av eliminering och efterföljande restaurering.

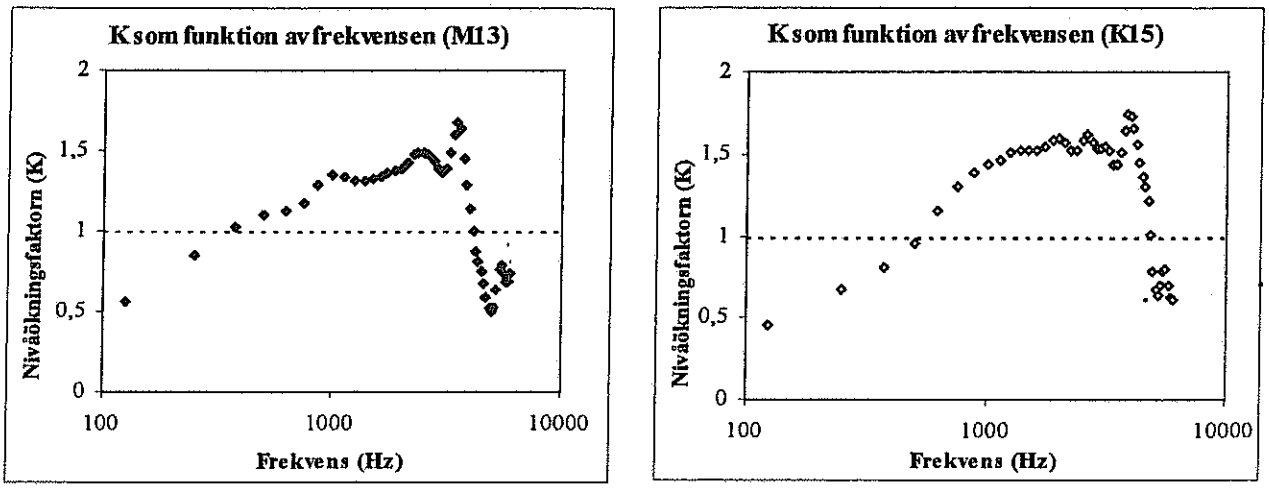
Såväl tonande som tonlösa ljud inkluderades i analysen av talsignalen. Enligt Gauffin & Sundberg (1977) innehåller tonlösa ljud till skillnad från tonande ljud mycket energi på höga frekvenser, dvs. frekvenser över 4 kHz. Löfqvist & Mandersson (1987) analyserade normala röster med hjälp av LTAS och jämförde LTAS med och utan tonlösa ljud. De fann att tonlösa ljud påverkar främst energinivåer över frekvensen 5 kHz. Hos tonande ljud däremot minskar energin vid dessa högre frekvensvärden. En följd av detta blir att energin från tonlösa ljud till viss del maskerar energiutslaget från tonande ljud, och därmed går viss information om röstkällan förlorad i detta frekvensområde. Genom att jämföra funktionen som visar hur den genomsnittliga energin fördelas över frekvensen, för tonande och tonlösa ljud med funktionen som representerar enbart tonande ljud, kunde effekten av tonlösa ljud undersökas. Kurvorna visade sig sammanfalla bortsett från frekvensområdet 4-6kHz där de skiljer sig åt några decibel. Denna skillnad har emellertid ingen betydelse för resultaten eftersom det är energifördelningen mellan 0-4 kHz som är av intresse i denna studie. Värt att nämna i detta sammanhang är att Fant (1973) visade att det är formanterna i tonande ljud och inte högfrekvent brus från tonlösa ljud, som dominerar den spektrala energin i frekvensområdet 3-5 kHz.

Ur varje LTAS analyserades röstintensiteten fram som den genomsnittliga ljudnivån, mätt i den så kallade ekvivalentnivån(Leq). Även den genomsnittliga ljudnivån för specifika frekvensband räknades ut. Bandbreddens storlek var 400Hz och frekvensomfånget 0-6kHz. För de två omnämnda analysstegen användes ljudanalysverktygen "Hist" resp. "Sect" i datorprogrammet Soundswell Core™ Analysis (AB Nyvalla, DSP, 1996). Samtliga uträknade värden överfördes till kalkylblad i programmet Excel, där resterande analyssteg utfördes. För varje frekvensband avsattes sedan den genomsnittliga ljudnivån för respektive lässekvens, som en funktion av motsvarande Leq-värde (se figur 4).



Figur 4. *Figuren visar genomsnittlig ljudnivå som funktion av Leq för försöksperson K5, vid frekvensbandet 1625 Hz.*

Som vi ser i figuren ovan kan förhållandet mellan Leq och genomsnittlig ljudnivå beskrivas av en rät linje. Lutningskonstanten kallas för nivåökningsfaktor (K_i), och har värdet 1,85 i grafen ovan (figur 4). Nivåökningsfaktorn är specifik för ett frekvensområde och anger graden av ökning i den genomsnittliga ljudnivån vid en ökning av röstens intensitet (Leq). Genom att använda ekvationen för varje frekvensband kan den genomsnittliga ljudnivån beräknas för alla frekvensband vid ett valt Leq–värde. Därmed kan man få fram den genomsnittliga ljudnivåns fördelning över frekvensen vid ett givet Leq–värde, dvs. LTAS. En sådan beräkning gjordes i den här studien för alla försökspersoner vid Leq–värdet 70 dB. Medelvärdes–LTAS vid detta Leq-värde beräknades likaså för gruppen män och kvinnor. Slutligen avsattes K_i som en funktion av frekvensen, och medelvärdet av denna variation räknades ut för män respektive kvinnor.



Figur 5. *Ovan ses två grafer som visar nivåökningsfaktorns, K_i, variation med frekvensen för en manlig respektive kvinnlig försöksperson.*

3.4.1 Behandling av data

Statistikmåtten max.värde och min.värde, medelvärde och standardavvikelse (SD) användes för att beskriva resultaten i denna studie.

För att undersöka eventuella samband mellan två variabler, exempelvis genomsnittlig ljudnivå och Leq, beräknades korrelationskoefficienten enligt följande:

$$\rho_{x,y} = \frac{Cov\,(X, Y)}{\sigma_x \cdot \sigma_y}$$

där:

$-1 \leq \rho_{x,y} \leq 1$

och:

$$Cov\,(X, Y) = \frac{1}{n} \sum_{i-1}^{n} (x_i - \mu_x)(y_i - \mu_y)$$

4. FRÅGESTÄLLNINGAR OCH HYPOTESER

4.1 Frågeställningar

1. Hur varierar nivåökningsfaktorn med frekvensen hos män respektive kvinnor?
2. Vilka skillnader finns mellan män och kvinnor vad gäller nivåökningsfaktorn variation?

4.2 Hypoteser

1. Att nivåökningsfaktorn är >1,0 vid frekvenser ≥ 0,7 kHz.
2. Att nivåökningsfaktorn är densamma hos män och kvinnor.

5. RESULTAT

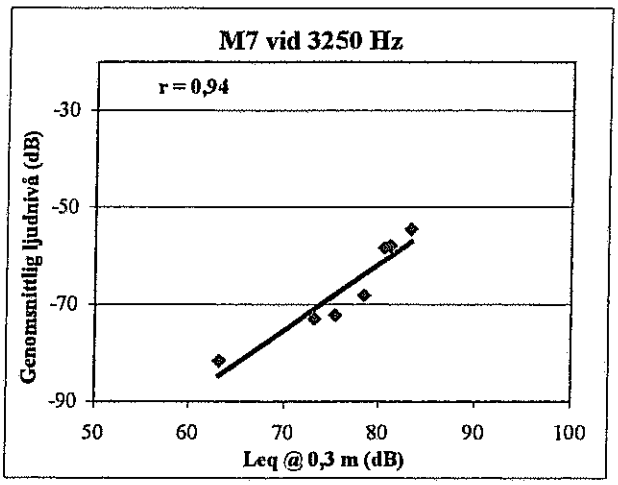
5.1 Extrema röststyrkor och korrelationen mellan dessa

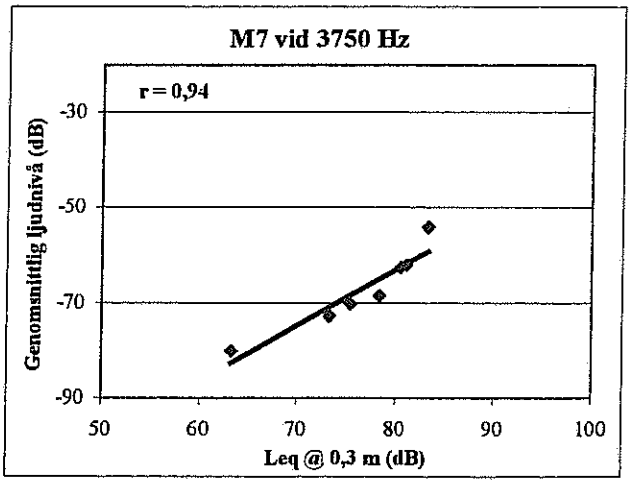
5.1.1 Medelvärden för extrema röststyrkor

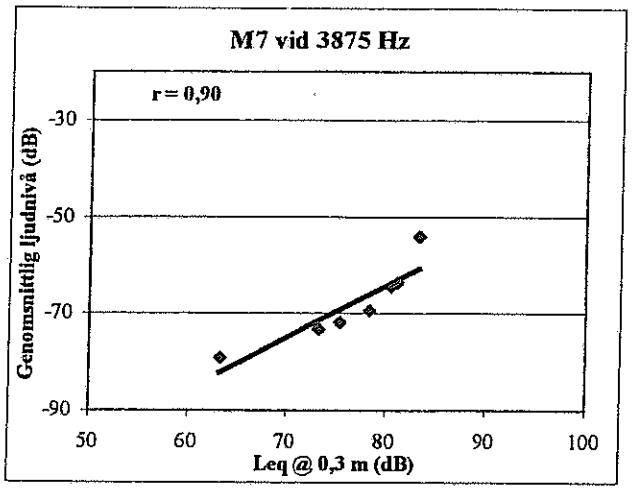
Medelvärdet för starkaste och svagaste röstintensitet (Leq), vid ett mikrofonavstånd på 30 cm, var för kvinnor 90 dB respektive 62 dB. Motsvarande medelvärden för män var 91 dB resp. 63 dB.

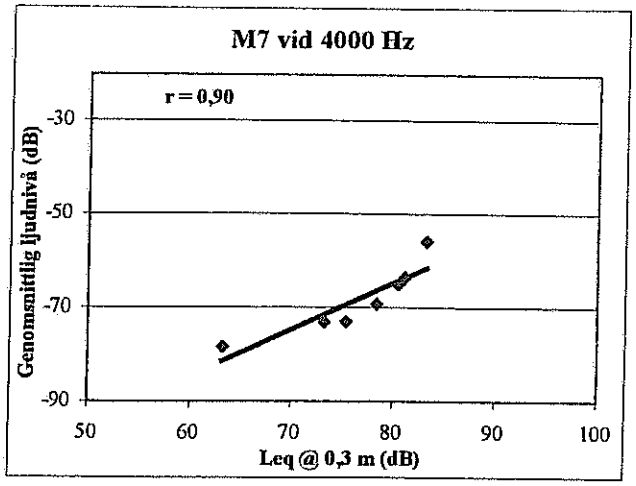
5.1.2 Korrelation vid extrema röststyrkor

Generellt var korrelationen, r, mycket hög för alla försökspersoner och aktuellt frekvensområde (0-4 kHz). Detta betyder att det rätlinjiga sambandet mellan Leq och genomsnittlig ljudnivå var mycket starkt. Trots en god korrelation avviker vanligen ett eller flera mätvärden något från den ekvation som beskriver sambandet. I denna studie visade det sig i de flesta fall vara så att punkterna, som representerar Leq och genomsnittlig ljudnivå, avvek mest vid extrema röststyrkor.



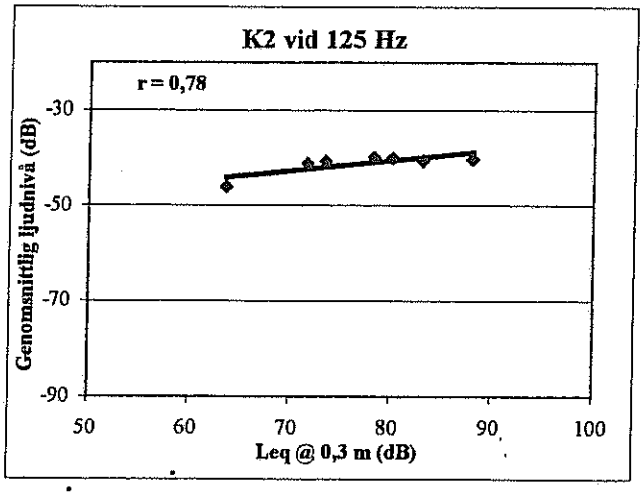


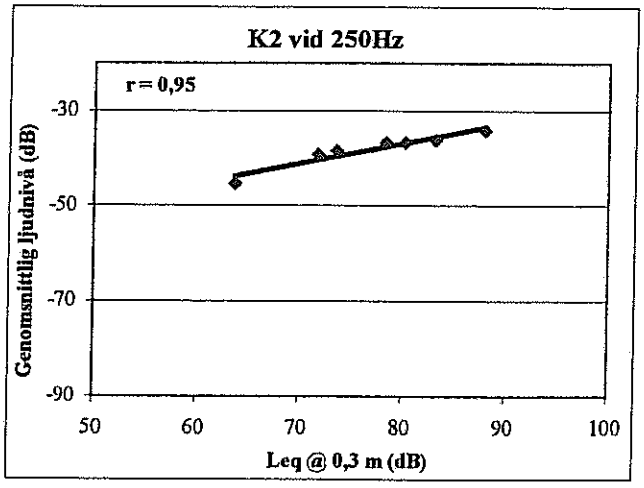


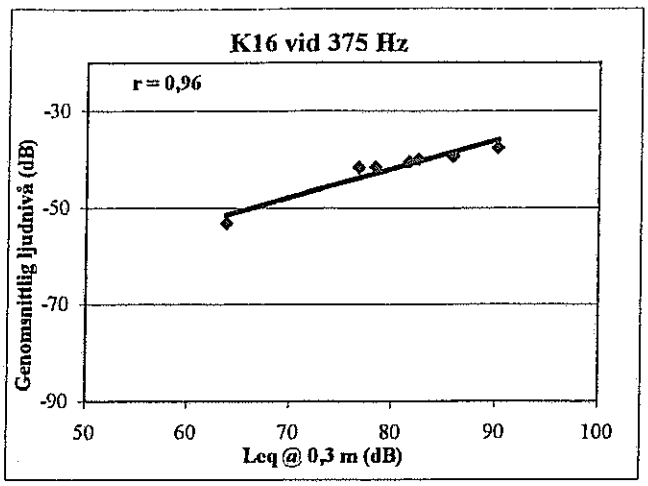


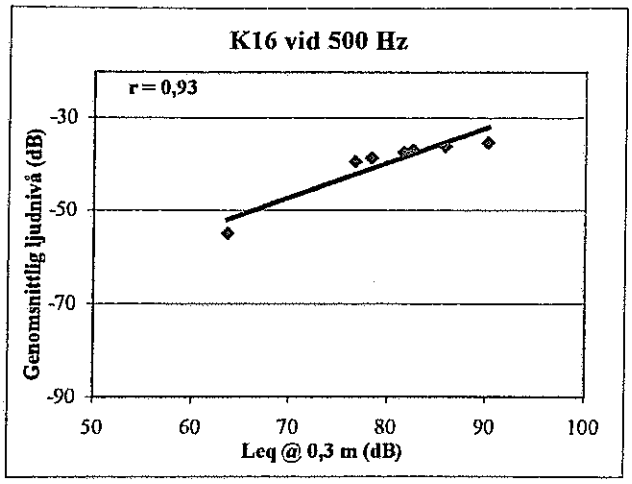
Figur 6. *Graferna ovan visar det rätlinjiga sambandet mellan Leq och den genomsnittliga ljudnivån, vid de fyra lägsta korrelationsvärdena i gruppen män.*

De fyra lägsta korrelationsvärdena tillhörde en och samma försöksperson (M7). Man kan tydligt se att högsta och lägsta röststyrka avviker från den räta linjen vid lägsta korrelationsvärdet för valt frekvensområde. Hos M7 tenderar också medelstarka röstintensiteter (Leq) att avvika relativt mycket. Det är särskilt tydligt vid 3250 Hz och 4000 Hz.









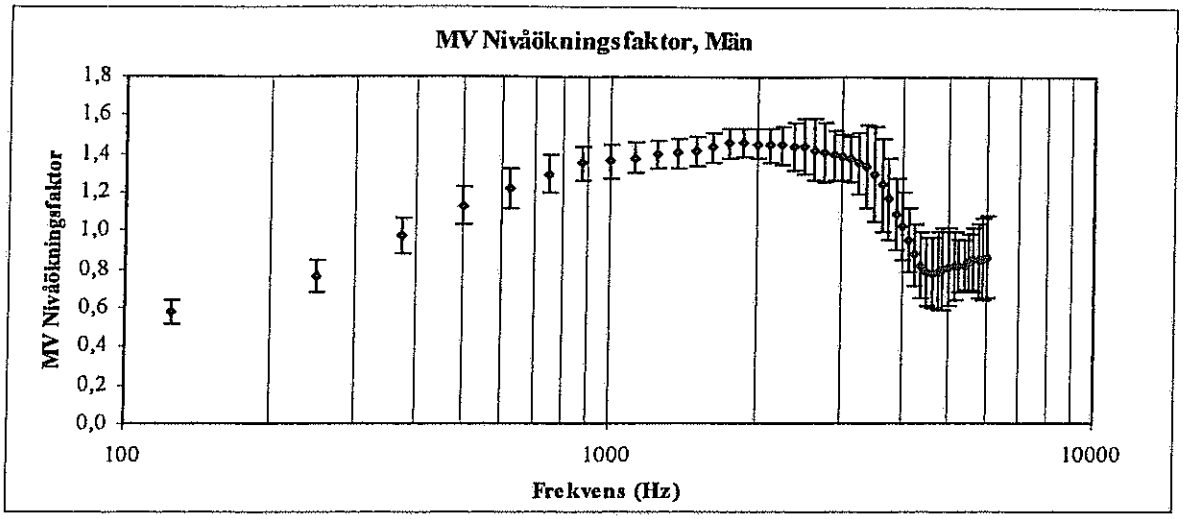
Figur 7. *Graferna visar det rätlinjiga sambandet mellan Leq och den genomsnittliga ljudnivån, vid de fyra lägsta korrelationsvärdena i gruppen kvinnor.*

Vid en jämförelse av de fyra graferna ovan ser vi att i det fall som ett värde avviker mest från den räta linjen, återigen utgörs av extremvärdena. Nämligen försöksperson K16 vid 500 Hz. De lägsta korrelationsvärdena för kvinnor befinner sig på låga frekvenser. I detta fall mellan 125-500 Hz. För männen däremot återfinns de lägsta r-värdena vid relativt höga frekvenser, nämligen mellan 3250-4000 Hz.

5.2 Nivåökningsfaktor, K_i

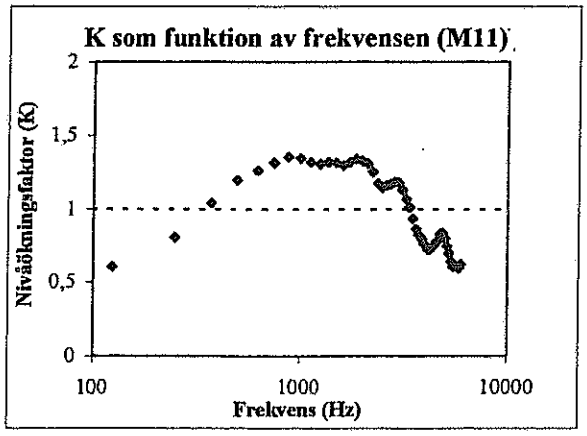
Det har tidigare nämnts under avsnittet "Metod" att förhållandet mellan den genomsnittliga röstintensiteten (Leq) och den genomsnittliga ljudnivån, i ett visst frekvensband, kan beskrivas av en rät linje. Ekvationen för en rät linje kan skrivas som "y = kx + m", där "y" i det här fallet representerar genomsnittlig ljudnivå i ett givet frekvensband. "k" är detsamma som nivåökningsfaktorn, K_i, och "m" är interceptet som anger funktionens skärningspunkt på y-axeln.

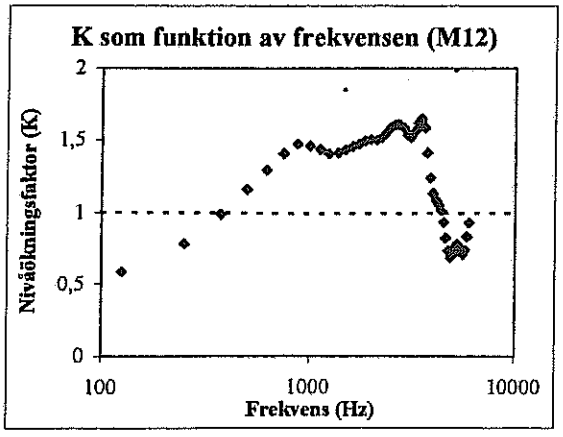
5.2.1 K_i hos män



Figur 8. *Figuren visar hur medelvärdet för K_i varierar med frekvensen, hos män. Vertikala linjer visar standardavvikelsen vid varje mätpunkt.*

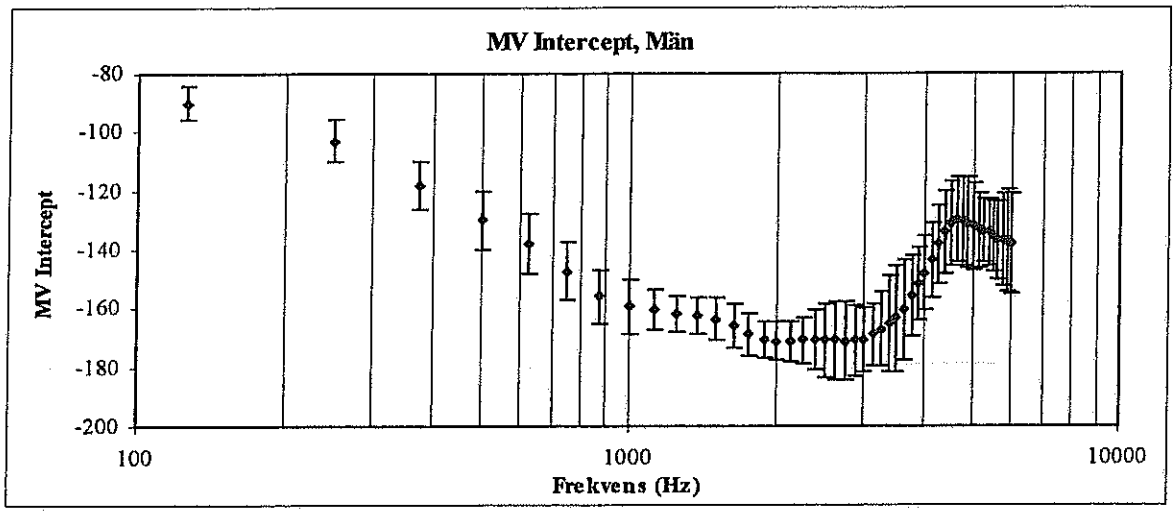
Frekvensen anges i logaritmisk skala vilket medför att små värden förstoras upp och stora värden förminskas. Grafen berättar hur mycket den genomsnittliga ljudnivån ökar vid specifika frekvensband, vid en ökning av den genomsnittliga röstintensiteten (Leq). K_i är >1,0 mellan 500 Hz och 3875 Hz. Dvs. i nämnt frekvensområde är ökningen av den genomsnittliga ljudnivån större än ökningen hos röstens genomsnittliga intensitetsnivå (Leq). Den mest omfattande ökningen av den genomsnittliga ljudnivån visar sig vara i frekvensområdet mellan 750 och 3500 Hz, då $K_i \geq 1,3$. Nivåökningsfaktorn ökar från 0,6 till 1,4 mellan frekvenserna 125–875 Hz, planar ut mellan 875 och 3500 Hz där värdet varierar mellan 1,3 och 1,5. Mellan 3500 Hz och 4625 Hz sjunker K_i's värde från 1,3 till 0,8, för att därefter åter stiga något. Vid 6000 Hz har K_i värdet 0,9. Nivåökningsfaktorns högsta och lägsta värde är 1,5 resp. 0,6, belägna vid frekvenserna 1750-2000 Hz respektive 125 Hz. Största standardavvikelsen har värdet 0,25 och återfinns vid 3500 Hz.





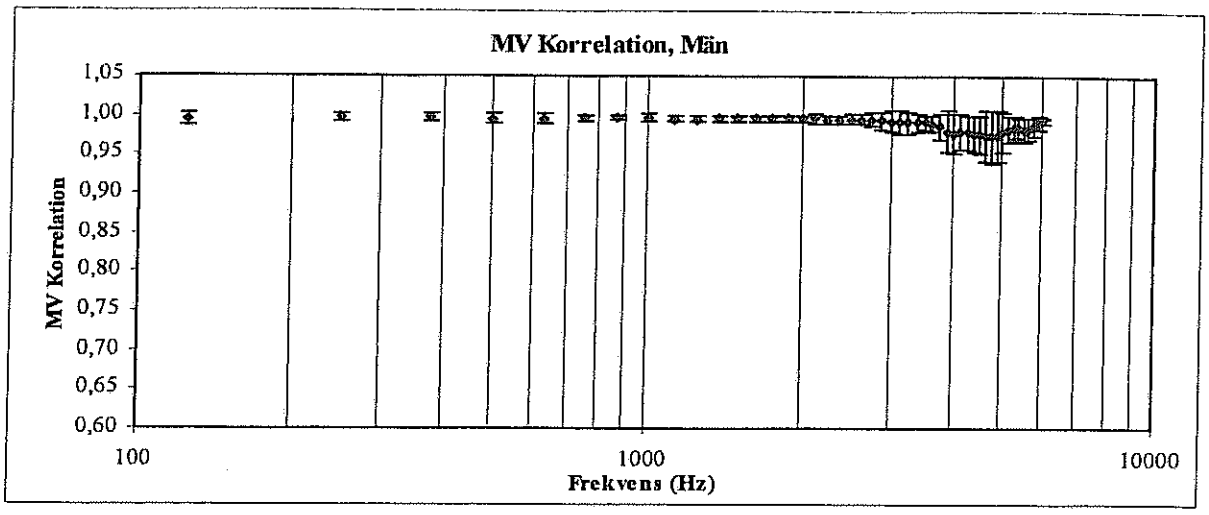
Figur 9. *Visar hur nivåökningsfaktorn varierar med frekvensen för två manliga försökspersoner.*

Graferna ovan visar hur olika K_i's variation med frekvensen kan se ut, vid en jämförelse mellan två manliga individer. Generellt sett är K_i större hos M12 med ett värde som är ≥ 1,4 mellan 750 Hz och 3750 Hz. Dessutom har M12 en distinkt topp vid 3500 Hz där K_i uppnår ett maxvärde på 1,6. M11 har värdet 1,3 vid 750 Hz, och uppnår maximum, 1,4, vid 1000 Hz. Vid 3750 Hz har K_i's värde sjunkit till 0,8. M11 har tre små toppar vid 1000,1875 samt 2875 Hz, med K_i–värdena 1,4; 1,3 respektive 1,2.



Figur 10. *Figuren visar interceptets variation över frekvensen. Varje punkt är ett medelvärde för alla män. Vertikala staplar visar standardavvikelsen vid varje mätpunkt.*

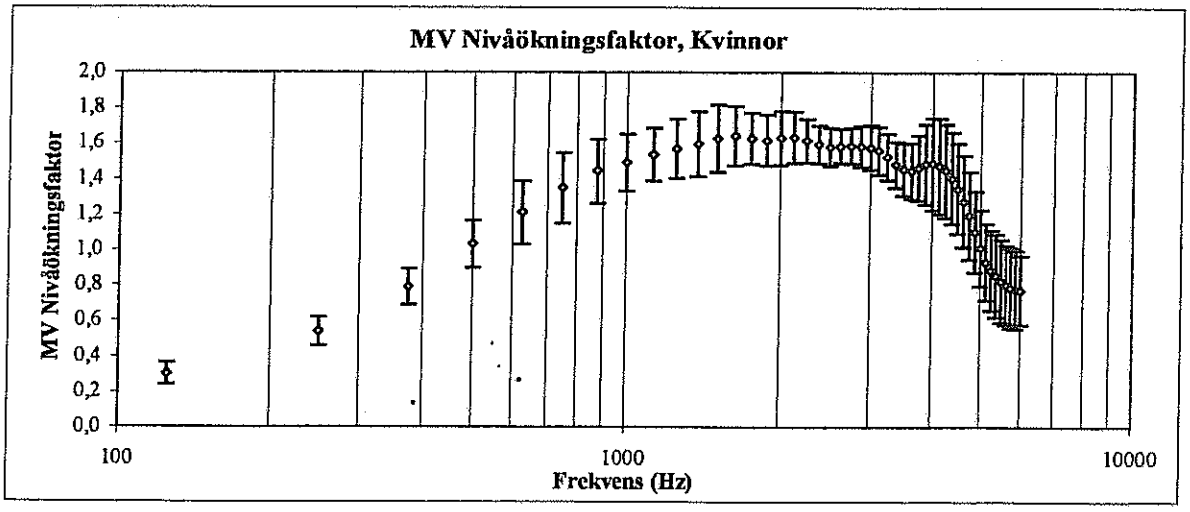
Grafen visar interceptets variation med frekvensen och förefaller vara en spegelbild av K_i's variation (se figur 8). Ett stort värde hos K_i, dvs. en brant lutning, motsvaras av ett lågt värde hos interceptet. Det senare innebär att räta linjen skär långt ner på y-axeln. Största värdet, -90, finns således vid frekvensen 125 Hz. Därefter minskar interceptets värde med ökande frekvens och mellan 1875 och 3000 Hz är interceptets värde i det närmaste konstant. Det minsta värdet, -171, återfinns mellan 1875 och 2875 Hz. Interceptet ökar sedan kraftigt med ett lokalt maximum vid 4625 Hz. Därefter minskar interceptets värde återigen och antar värdet –138 vid 6000 Hz. Största standardavvikelsen har värdet 18 och finns vid ca 3500 Hz.



Figur 11. Figuren visar medelvärdet för korrelationen. Vertikala staplar visar standardavvikelsen för varje mätvärde.

Grafen anger hur starkt sambandet är mellan genomsnittlig röstintensitet (Leq) och genomsnittlig ljudnivå vid en given frekvens. Korrelationen är mycket hög, vilket visar att nästan all variation i ett givet frekvensband kan förklaras med variationen av Leq. Lägsta korrelationen upp till och med 4000 Hz är 0,98 vid 4000 Hz. Största standardavvikelsen inom samma frekvensområde är 0,03, och återfinns vid 3875–4000 Hz.

5.2.2 K_i hos kvinnor

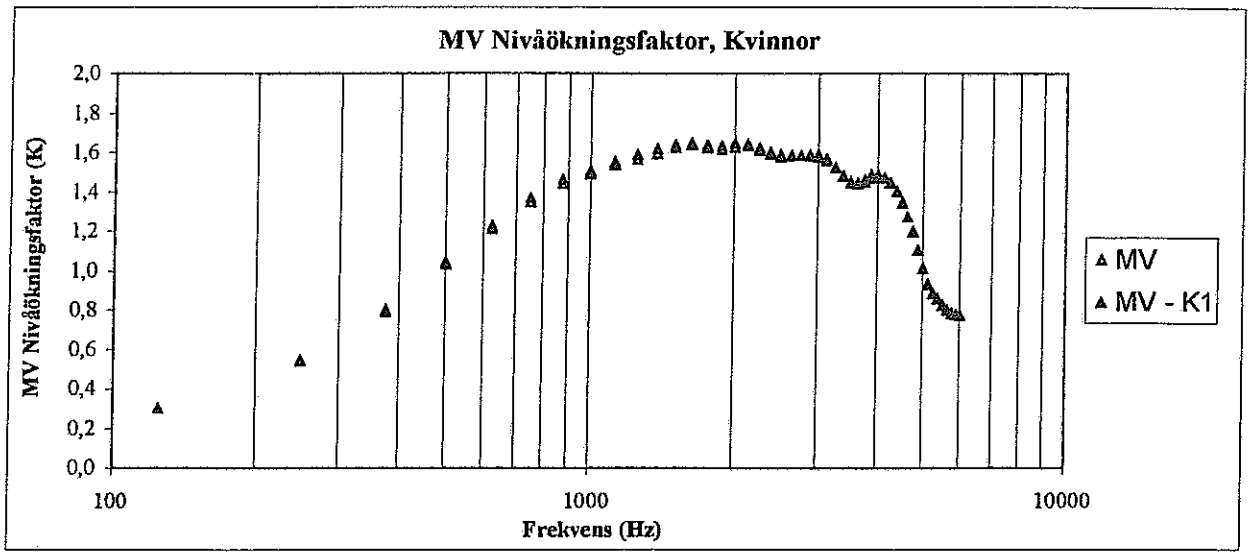


Figur 12. Figuren visar hur medelvärdet för K_i varierar med frekvensen hos kvinnor. Vertikala linjer representerar standardavvikelsen vid varje mätpunkt.

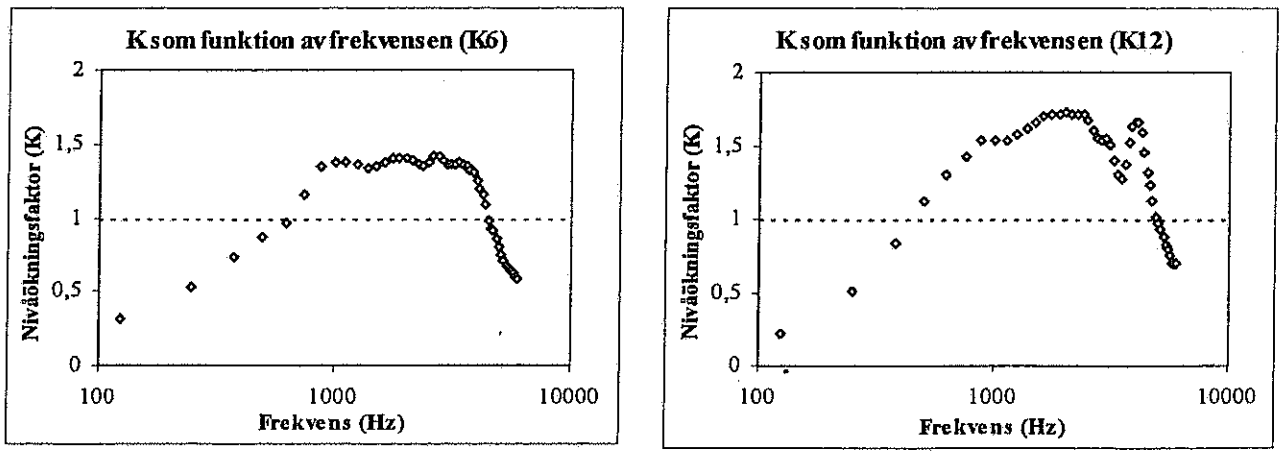
$K_i > 1{,}0$ mellan 625 Hz och 4875 Hz. Den största ökningen av den genomsnittliga ljudnivån ses i huvudsak mellan 750 Hz och 4625 Hz, då $K_i \geq 1{,}3$. Nivåökningsfaktorn ökar från 0,3 till 1,4 mellan frekvenserna 125–875 Hz. Därefter fortsätter K_i att öka till 1,6 vid 1375 Hz, och är i stort sett oförändrad till och med 3125 Hz. Från 3125 Hz minskar K_i's värde successivt, med undantag av ett lokalt maximum vid 4000 Hz. K_i's värde vid 6000 Hz är 0,8.

Nivåökningsfaktorns största värde är 1,6 vid 1375-3125 Hz. Minsta värde är 0,3 vid 125 Hz. Största standardavvikelsen är 0,27 vid 4000 Hz.

I avsnittet "Metod" nämndes tidigare att en kvinnlig försöksperson (K1) hade tränad talröst. För att undersöka K1's eventuella "inverkan" på medelvärdet jämfördes grafen som inkluderade K1, med grafen som exkluderade K1 (se figur 13 nedan).

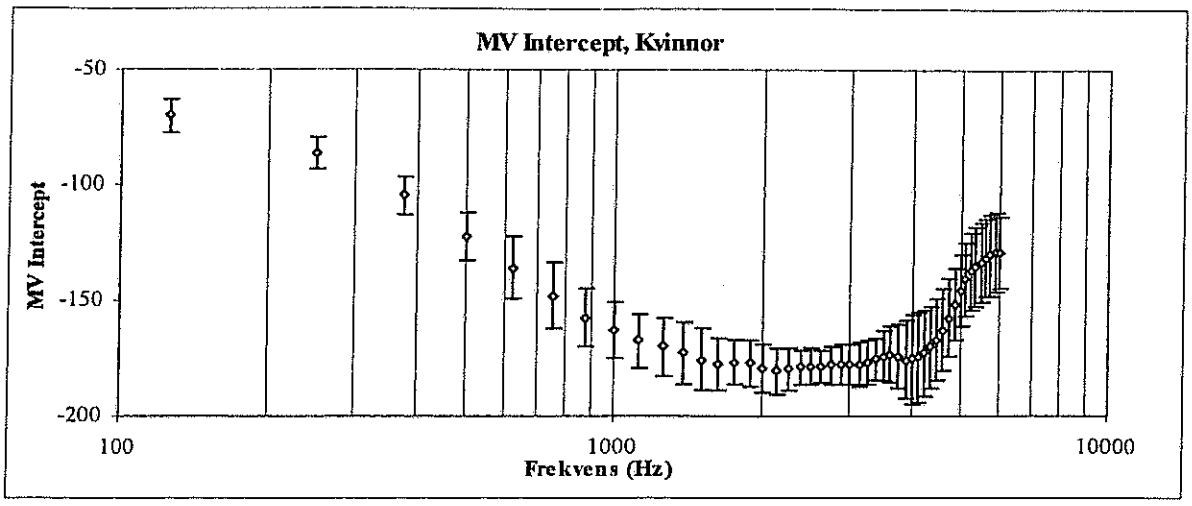


Figur 13. Grafen visar jämförelsen av K_i's medelvärde vid inklusion respektive exklusion av K1.



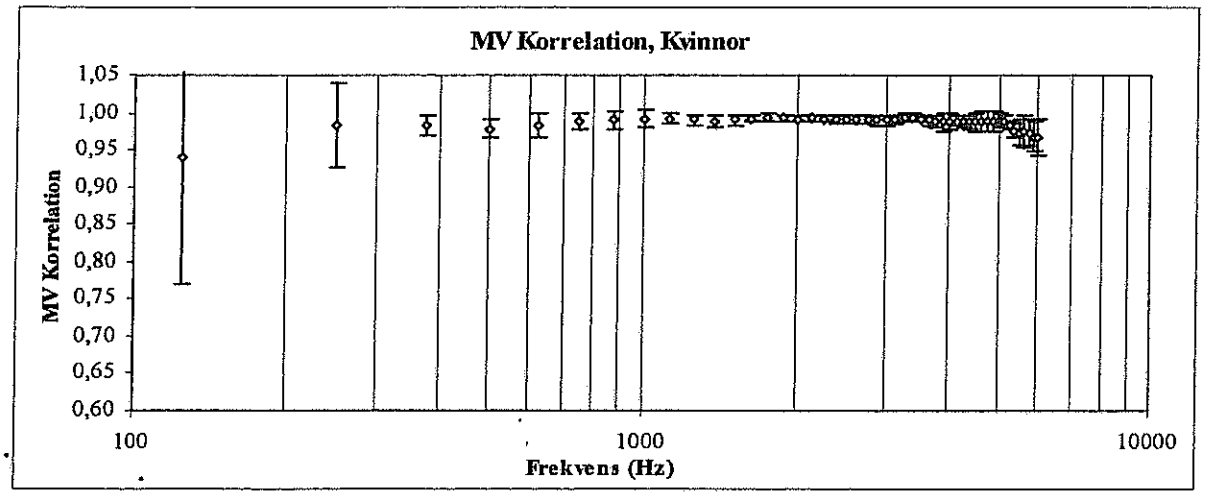
Figur 14. Visar hur nivåökningsfaktorn varierar med frekvensen för två kvinnliga försökspersoner.

Graferna ovan visar hur olika K_i's variation med frekvensen kan se ut vid en jämförelse mellan två kvinnliga försökspersoner. K_i ligger generellt på en högre nivå hos K12, där en "platå" på maxvärdet 1,7 fås mellan 1500-2500 Hz. En tydligt topp följer sen vid 4125 Hz. K6 däremot har tre mycket svaga toppar vid 1000 Hz, 1750 Hz samt 2625 Hz , där det sistnämnda frekvensbandet innehar maxvärdet på 1,4.



Figur 15. Figuren visar interceptets variation över frekvensen. Varje punkt är ett medelvärde för alla kvinnor. Vertikala staplar visar standardavvikelsen.

Även medelvärdet för kvinnornas intercept ser ut att vara en spegelbild av nivåökningsfaktorns variation för samma grupp. Interceptet minskar till en början vid en ökning av frekvensen och antar värdet –177 vid 1625 Hz. Interceptets värde är i stort sett oförändrat till och med 3250 Hz, men ökar därefter och uppnår värdet –129 vid 6000 Hz. Största och minsta intercept är –70 och –180 vid frekvenserna 125 Hz respektive 2125-2250 Hz. Maximala standardavvikelsen är 19 mellan 3875-4000 Hz.



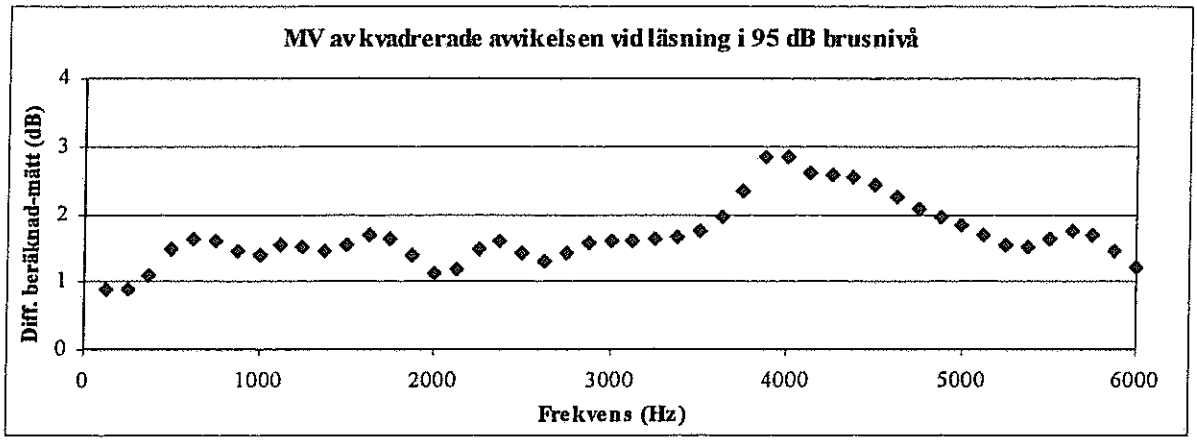
Figur 16. Figuren visar korrelationens medelvärde för kvinnor. Vertikala staplar visar standardavvikelsen för varje mätvärde.

Lägsta korrelationen har värdet 0,82 och återfinns vid 125 Hz. Korrelationen är mycket hög även för gruppen kvinnor och vi kan dra samma slutsats som vi gjorde för männen. Dvs. större delen av den variation vi får i ett givet frekvensband kan förklaras med en variation av Leq. Standardavvikelsens största värde är 0,17 vid 125 Hz. Även vid 250 Hz är värdet relativt högt, 0,06. Vid övriga frekvenser är standardavvikelsen mycket liten.

5.3 Differensen: beräknad–uppmätt

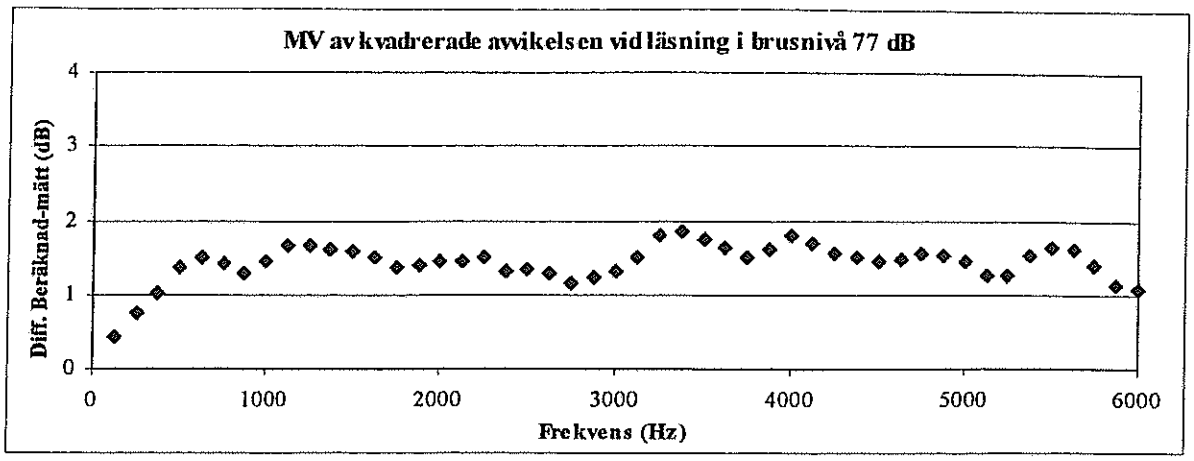
Det starka sambandet mellan genomsnittlig ljudnivå hos ett givet frekvensband och röstens intensitetsnivå (Leq) betyder att genomsnittlig ljudnivå hos varje frekvensband kan beräknas för varje värde på Leq, hos en röst. Hur väl beräknad och uppmätt ljudnivå stämmer överens kontrollerades genom att räkna ut differensen: beräknad-uppmätt värde. Beräknat värde syftar här på de värden som beräknats med hjälp av den ekvation som beskriver sambandet mellan röstintensiteten (Leq) och genomsnittlig ljudnivå vid en given frekvens. Dvs. den ekvation som nämndes tidigare. Differensen: beräknad – uppmätt värde räknades ut för läsningen i högsta och lägsta brusnivå samt för läsningen utan brus. Läsningarna valdes med tanke på att de största avvikelserna i flertalet fall uppstod vid extrema röststyrkor . Medelvärde för alla tre brusnivåer beräknades för gruppen män och kvinnor .

5.3.1 Differensvärde för män



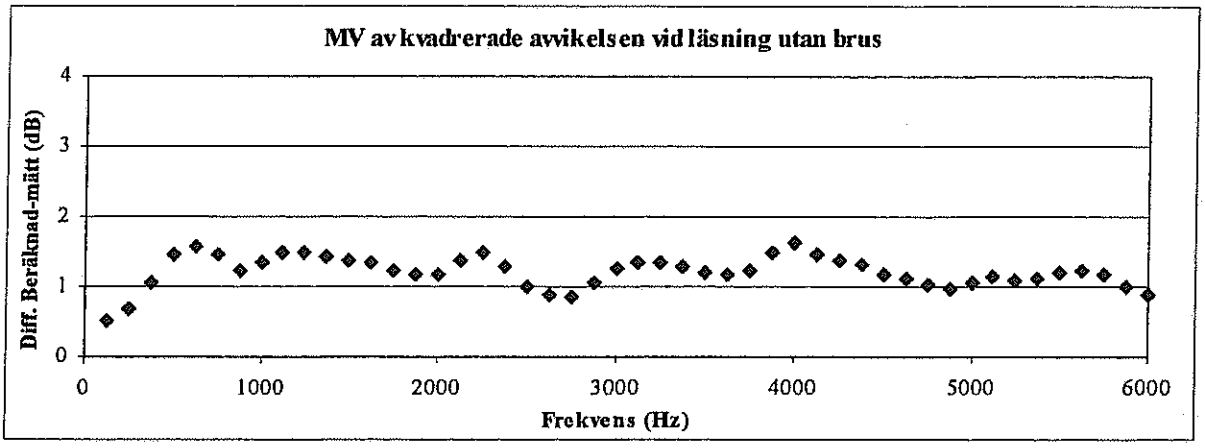
Figur 17. Grafen visar medelvärdet (MV) av den kvadrerade differensen för män vid läsning i brusnivå 95 dB.

Till och med 3625 Hz är den kvadrerade differensen mellan uppmätt och beräknat värde ≤ 2. Efter 3625 Hz stiger dock värdet och uppgår till 2,8 vid 3875-4000 Hz, som är största värdet.



Figur 18. *Grafen visar medelvärde av den kvadrerade differensen beräknad–mätt ljudnivå vid läsning i brusnivå 77 dB, för män.*

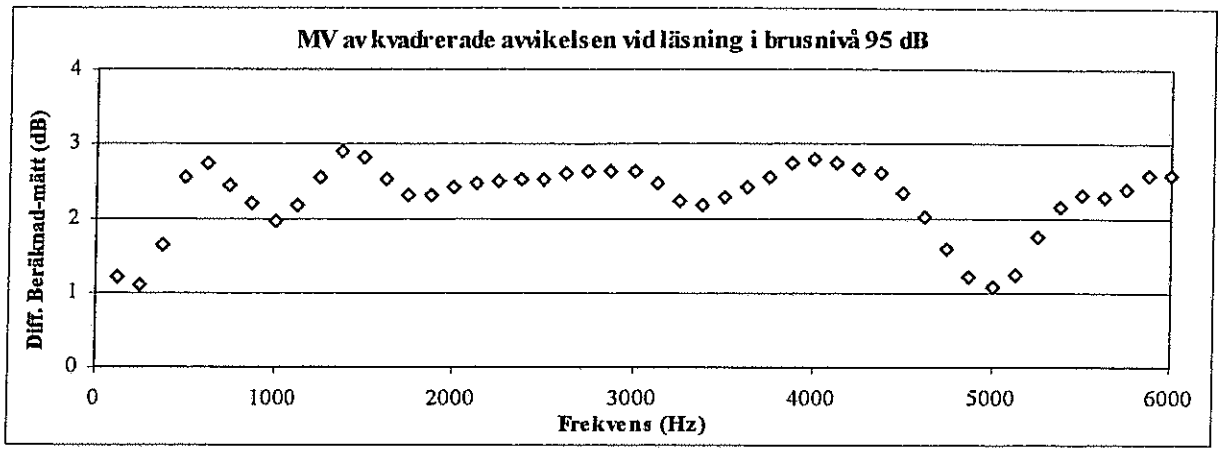
Den kvadrerade differensen överstiger aldrig värdet 2. Högsta värdet, 1,9, återfinns vid 3375 Hz.



Figur 19. *Grafen visar medelvärde av den kvadrerade differensen beräknad–mätt ljudnivå vid läsning utan brus, för gruppen män.*

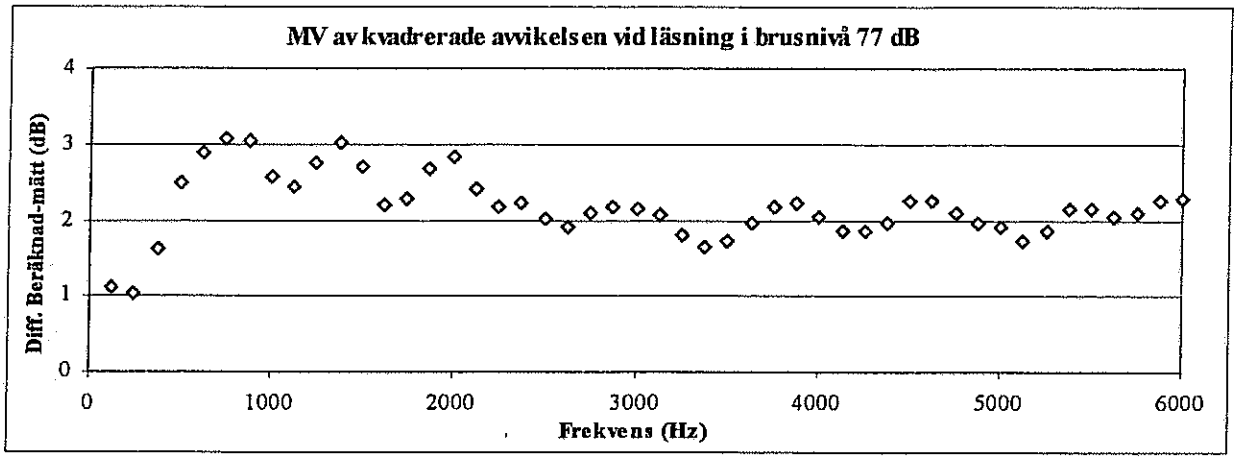
Högsta medelvärdet i grafen ovan återfinns vid 4000 Hz och har värdet 1,6. Vid en jämförelse av de tre graferna ovan kan man notera att differensen, generellt sett, är större för läsningarna i brus än vid läsningen utan brus. Dvs. vid högre röstintensiteter förefaller differensen: beräknad-uppmätt genomsnittlig ljudnivå vara större än vid låga röstintensiteter. Det kvadrerade differensvärdet är < 2 mellan 0–4 kHz, utom vid läsning i brusnivå 95 dB då värdet överstiger 2 efter 3625 Hz.

5.3.2 Differensvärde för kvinnor



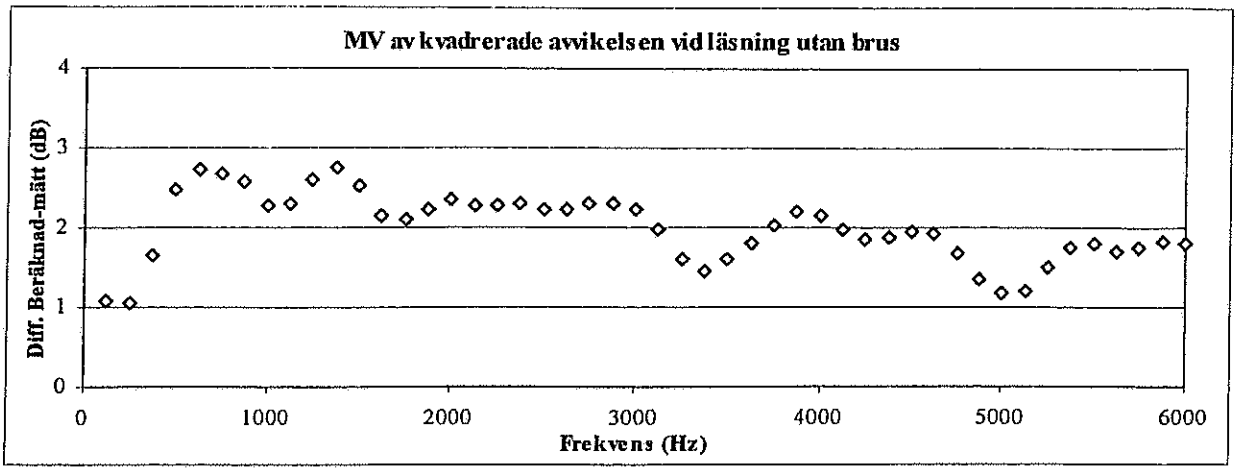
Figur 20. Grafen visar medelvärdet av differensen beräknad-mätt ljudnivå, för kvinnor, vid läsning i högsta brusnivå.

Vid 1375 Hz är den kvadrerade differensen störst, med ett värde på 2,9. Mellan 500 Hz och 4000 Hz varierar värdet mellan 2 och 3. Vid en jämförelse med motsvarande frekvensområde och läsning för gruppen män, blir det tydligt att differensen för kvinnor generellt sett är större. Mellan 500 Hz och 3625 Hz ligger differensvärdet mellan 1 och 2 hos gruppen män. Det ökar därefter upp till 3 vid 4000 Hz.



Figur 21. Grafen visar medelvärdet av den kvadrerade avvikelsen för kvinnor vid läsning i brusnivå 77 dB.

Största värdet, 3,1, finns vid 750 Hz. Mellan 500 Hz och 2375 Hz är alla värden > 2. Efter 2375 Hz följer en generell minskning. Vid en jämförelse med motsvarande läsning hos män finner man även här att differensvärdet generellt är högre hos kvinnor. Vid läsning i brusnivå 77 dB gäller det i synnerhet mellan frekvenserna 500 Hz och 2375 Hz, där differensvärdet aldrig överstiger 2 hos männen.



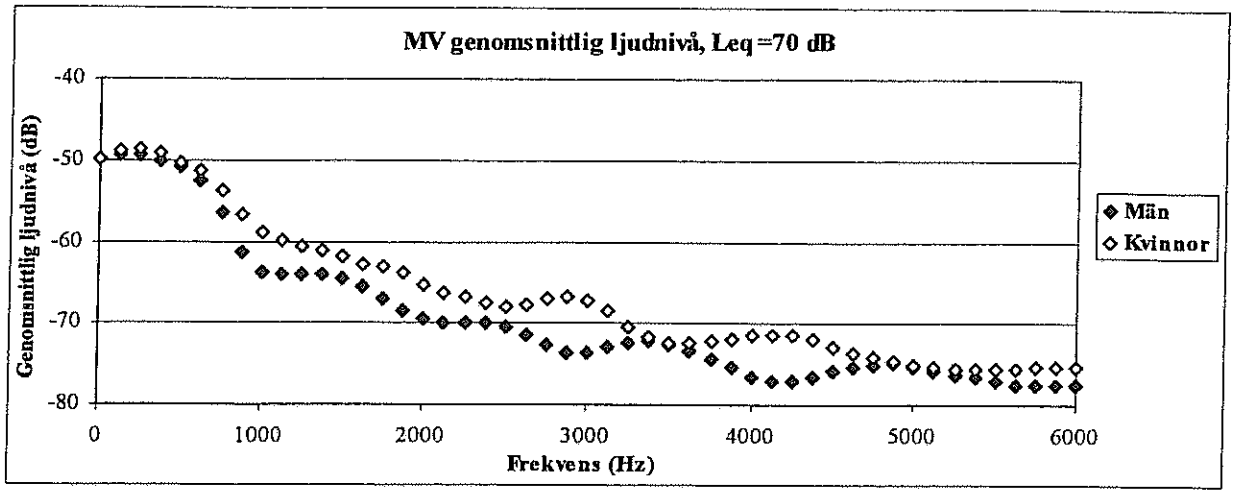
Figur 22. Grafen visar den kvadrerade avvikelsens variation med frekvensen vid lägsta röstintensitet, för gruppen kvinnor.

Vid 625 Hz, 750 Hz och 1375 Hz återfinns största värdet; 2,7. Ett liknande förlopp som det vid läsning i brusnivå 77 dB kan noteras. Dvs. vid de lägre frekvenserna, i detta fall mellan 500 Hz och 3000 Hz är differensvärdet högre än vid frekvenser över 3000 Hz. Återigen kan den tidigare nämnda skillnaden mellan kvinnor och män noteras. Nämligen att kvinnor generellt sett har en större differens än män vid motsvarande läsning. Det kvadrerade differensvärdet för kvinnor < 3 mellan 0–4 kHz utom vid läsning i brusnivå 77 dB då differensvärdet ≥ 3 vid tre tillfällen.

Vid en jämförelse mellan gruppen män och kvinnor visar det sig att största värdet återfinns på mycket högre frekvenser för män. Mellan 3375 Hz och 4000 Hz. Hos kvinnor däremot ligger alla maxvärden vid frekvenser ≤ 1375 Hz.

Högsta medelvärdet för den kvadrerade differensen beräknad–mätt är 3,1, vilket är mycket lite i förhållande till genomsnittliga ljudnivåns absoluta värde. Majoriteten värden är mindre än 3,1. Att avvikelserna är så små innebär att den linjära funktionen som tagits fram för att beskriva förhållandet mellan Leq och genomsnittlig ljudnivå stämmer mycket väl överens med verkligheten. Följaktligen kan funktionen användas för att räkna fram genomsnittlig ljudnivå vid ett godtyckligt valt Leq-värde. Det innebär att vi exempelvis kan välja Leq-värdet 70 dB, och med hjälp av den rätlinjiga ekvationen få fram genomsnittlig ljudnivå vid varje frekvensband. Dvs. vi får beräknat LTAS vid Leq = 70 dB.

5.4 Beräknad LTAS



Figur 23. *Medelvärde av den genomsnittliga ljudnivåns variation med frekvensen, för gruppen män respektive kvinnor.*

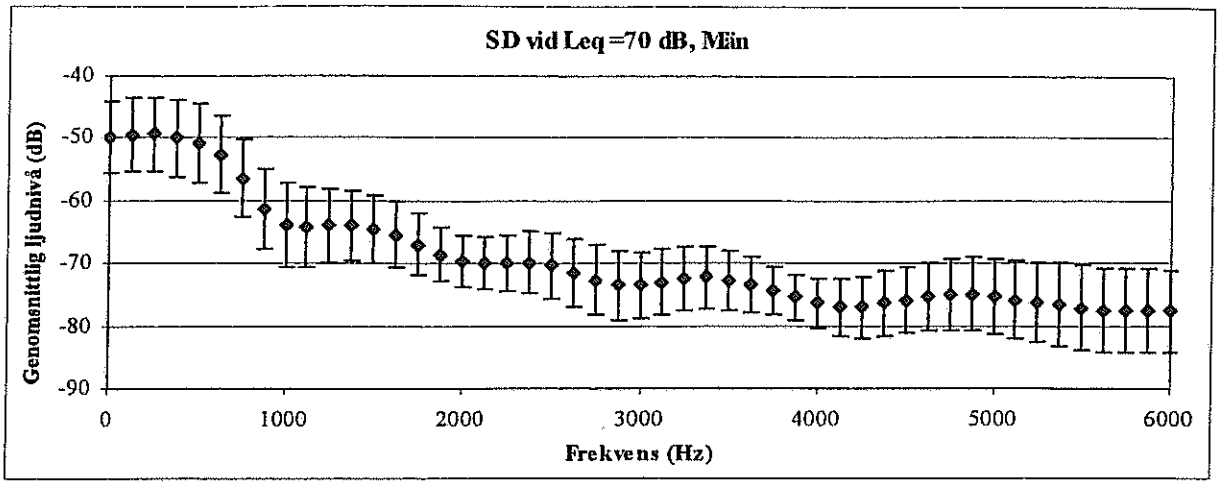
Med hjälp av erhållna data kunde ett "medelvärdes–LTAS" vid Leq–värdet 70 dB beräknas för män respektive kvinnor. Genomsnittlig ljudnivå har angetts som relativa värden till en referensnivå (0 dB). I de flesta fall har T1 en högre energinivå än F0. Det är också den spektrumtopp som dominerar den totala SPL-nivån. Under dessa förutsättningar kommer F0's ökade frekvensvärde inte att dominera eller påverka den första spektrumtoppen i samband med en ökning av röstintensiteten. Spektrumtopparna representerar medelvärden av formanterna och grafen visar hur kvinnornas fyra första spektrumtoppar (T1-T4) har ett högre frekvensvärde än motsvarande toppar hos männen, beroende på tidigare nämnda skillnader vad gäller formantfrekvenser. Spektrumtopparnas frekvensvärden och genomsnittliga ljudnivå anges i tabell 2 respektive 3 nedan. Värden i tabellerna har avrundats och därför har både män och kvinnor värdet –49 dB vid T1, trots att en skillnad kan noteras i diagrammet. Vidare kan noteras att kvinnornas T2 knappt är synbar och att kvinnorna har en högre genomsnittlig ljudnivå.

Tabell 2. Spektrumtoppar, kvinnor.

	Frekvens(Hz)	Ljudnivå(dB)
T1	250	-49
T2	1750	-63
T3	2875	-67
T4	4250	-71

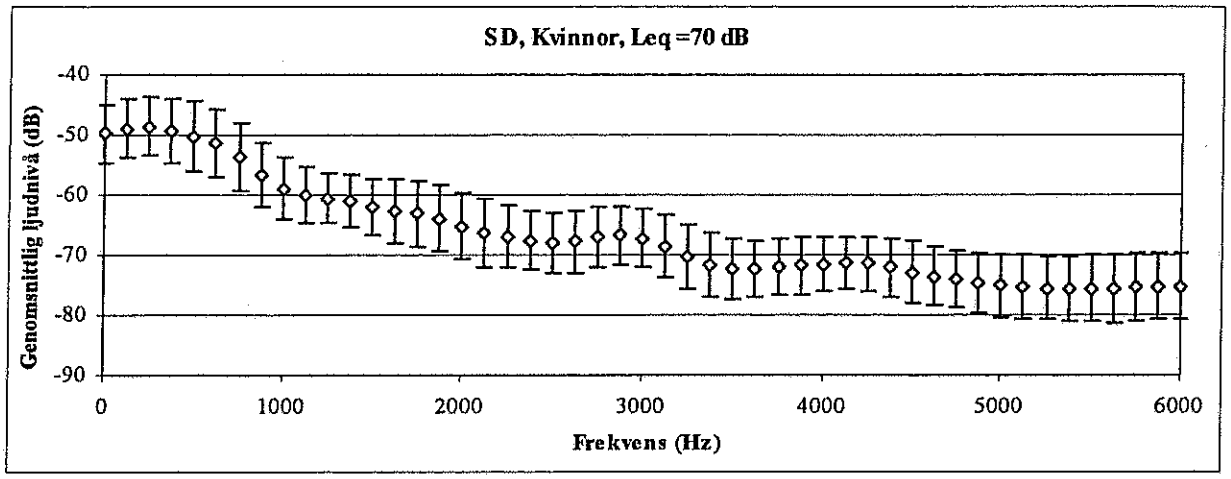
Tabell 3. Spektrumtoppar, män.

	Frekvens(Hz)	Ljudnivå(dB)
T1	250	-49
T2	1500	-65
T3	2375	-70
T4	3375	-72



Figur 24. *Figuren visar standardavvikelsen för män i form av en vertikal linje i varje mätpunkt.*

De högsta standardavvikelserna ses vid 375–1125 Hz och 5125-6000 Hz där värdet överstiger 6 och når maximum på 6,7 respektive 6,8.



Figur 25. *Figuren visar standardavvikelsen i varje mätpunkt för kvinnor.*

De största standardavvikelserna ses vid 500-750 Hz, 2125 Hz samt 625-5750 Hz där standardavvikelsen överstiger 5,5 och når ett maximum på 5,8; 5,6 samt 5,7.

6. DISKUSSION

6.1 K_i's variation

6.1.1 K_i's variation hos män

Hos männen är nivåökningsfaktorn > 1,0 mellan 500 - 3875 Hz. $K_i = 1$ innebär att den genomsnittliga ljudnivån för en specifik frekvens ökar lika mycket som Leq. Enligt uppställd hypotes skulle K_i överstiga 1,0 vid frekvensband ≥ 700 Hz. Vi ser utifrån studiens resultat att angiven frekvensgräns kan sänkas med 200 Hz. Det betyder att vid frekvenser < 500 Hz ökar genomsnittlig ljudnivå mindre än Leq, vid en ökad röstintensitet. Värdet 700 Hz bestämdes utifrån de data som framkom i Fants (1973) och Ternströms (1989) studier. Ternströms studie visade nivåökningsfaktorns variation vid körsång, där två av tre körer hade K_i–värdet 1 vid 700 Hz. Fant (1973) visade hur en ökning av röstintensiteten med 20 dB, hos uthållen fonation på vokalen [ɑ:], resulterade i en ökning med 23,5 och 26 dB hos F1 respektive F2. Det innebär att K_i's värde vid Fants studie var ca 1,2 vid 500 Hz respektive 1,3 vid 1000 Hz. Det bör emellertid framhållas att Ternström (1989) i sin studie undersökte körsång och Fant (1973) källspektrum vid fonation av en vokal. Följaktligen kan nämnda resultat ej helt likställas med LTAS för vanligt tal.

Resultat av den här studien säger att vid frekvenser ≥ 500 Hz ökar genomsnittlig ljudnivå mer än Leq vid ökad röstintensitet. Det sistnämnda samt det faktum att största ökningen av den genomsnittliga ljudnivån uppstod i frekvensområdet 750-3500 Hz, då $K_i \geq 1,3$, stämmer överens med det man funnit i tidigare studier, nämligen att energinivån vid höga frekvenser ökar i större utsträckning än energinivån vid låga frekvenser, vid en ökning av röstintensiteten (Fant, 1973; Gauffin & Sundberg, 1980; Sundberg m.fl., 1988; Ternström, 1989; Gramming, 1991).

Nawka m.fl. (1997) visade i sin studie att den spektrumtopp som kallats talarformanten ökar sin maximala energinivå med 17 dB vid en ökning av den totala röstintensiteten från normal konversationsnivå till starkt tal. I "Inledningen" (se sid. 4) nämndes tidigare att en förändring av talarformantens maximala energinivå uteblev vid övergången från starkt tal till skrik. Tidigare visades graferna för lägsta korrelationsvärdet mellan Leq och genomsnittlig ljudnivå. Punkterna vid extrema röstintensiteter avvek något mer än övriga från den räta linjen. Det skulle vara intressant att titta närmare på de manliga försökspersoner som har en talarformant i LTAS för att se huruvida punkten som representerar maximal Leq avviker nedåt från räta linjen vid talarformantfrekvensen. Om så vore fallet skulle Nawka m.fl. (1997) resultat bekräftas. Det sträcker sig dock utanför denna studies omfattning. Ytterligare forskning om hur talarformanten uppkommer och om den är mer framträdande vi låga röstintensiteter krävs.

Vid analys av materialet hos enskilda försökspersoner visade sig K_i ha en framträdande topp i frekvensområdet 3375-3750 Hz hos nio män. Ytterligare tre män har en synbar men mycket svag topp inom frekvensområdet 3000-3750 Hz. Hos resterande tre män finns ingen synlig topp i de nämnda frekvensområdena. Dessa toppar är alltså i frekvensområdet där man i LTAS normalt finner en talarformant. Att den genomsnittliga ljudnivån ökar särskilt mycket i detta område för ett flertal av männen är intressant. Man kan spekulera i om det är en starkt ökande talarformant som kan observeras.

Kitzing och Åkerlund (1993) omnämnde att medelvärdes-LTAS kan medföra att individuella data går förlorade. De menade vidare att den individuella variation som man ser i LTAS

förmodligen beror på röstintensiteten. Om man sen betänker att röstintensiteten bestäms av ett flertal komponenter, bl.a. slutningshastigheten hos stämläpparna, grundtonsfrekvensen och formanternas inbördes placering, förstår man att det lätt uppstår en variation mellan individer vad gäller LTAS' utseende. Två individer kan inte medvetet tala med exakt samma röstintensitet (Leq). Dessutom kan en och samma röstintensitet (Leq) påverkas av flera olika faktorer. Om man jämför K_i's variation med frekvensen hos två individer ser man också en variation som går förlorad när medelvärdet beräknas. I figur 9 åskådliggörs denna individuella variation genom en jämförelse av två manliga försökspersoner. M12 har en tydligt framträdande topp, medan M11 saknar en sådan.

6.1.2 K_i's variation hos kvinnor

För gruppen kvinnor är $K_i > 1$ mellan 625 Hz och 4875 Hz. Precis som för män var hypotesen att $K_i > 1{,}0$ vid frekvenser ≥ 700 Hz. Även för kvinnor överstiger K_i värdet 1,0 vid ett lägre frekvensvärde än vad som förutsågs. Medelvärdet för kvinnor stämmer också med de tidigare forskningsresultat som säger att energinivån ökar mer vid höga än låga frekvensområden, vid en ökning av Leq. Den största ökningen av genomsnittlig ljudnivå för kvinnor ses mellan 750 – 4625 Hz, då $K_i \geq 1{,}3$.

Ingen topp av nivåökningsfaktorn kan ses i det frekvensområdet runt 3500 Hz. Intressant att notera är emellertid att medelvärdet för kvinnornas K_i visar en lokal topp vid 4125 Hz. Vid en närmare analys av individuella grafer framkommer det att åtta kvinnor har en tydlig topp, eller ökning av K_i, i frekvensområdet 3875–4625 Hz. Ytterligare tre kvinnor har en synlig men svag topp i frekvensområdet 3875–4375 Hz. Förutom denna har även dessa tre kvinnor en synlig svag topp i frekvensområdet 3000–3125 Hz. Övriga fem kvinnor har ingen topp i de nämnda frekvensområden. Förmodligen är det så att toppen runt 4000 Hz, och strax över, hos vissa kvinnor motsvarar den topp vi tidigare noterade för vissa av de manliga försökspersonerna vid 3500 Hz. Är den starka tillväxt vi ser vid aktuella frekvensområden en stark tillväxt av F4?

I en studie av Löfqvist (1986) påvisades en intra- och interindividuell variation i spektral energifördelning hos kvinnliga, tränade talröster. Leino (1993) fann att spektrumlutningen, men framför allt nivån, hos talarformanten hade ett samband med uppfattad röstkvalitet hos 48 manliga professionella skådespelare. Försökspersonerna i Leinos studie hade alltså en tränad talröst, men trots detta uppvisade de skillnader i LTAS-konfiguration. I samma studie visade det sig emellertid också att en talarformant inte var ett nödvändigt villkor för att en röst skulle uppfattas som "bra". Det fanns alltså bra röster utan talarformant. Bele (2002) kunde i sin studie av rösten hos lärare och skådespelare bekräfta detta. Hon fann sålunda avsevärda skillnader mellan individer både i gruppen med tränade och i gruppen med otränade röster. Nämnda forskningsresultat visar på såväl intra- som interindividuell variation hos den normala, tränade talrösten. Likaså en variation mellan normala otränade och tränade talröster visar sig existera. Betydelsen av att inkludera den enda försökspersonen med tränad talröst (K1) i denna studie visade sig vidare vara försumbar (se "Resultat" avsnitt 5.2.2 "K_i hos kvinnor", figur 13). Med utgångspunkt från dessa fakta inkluderades kvinnan med tränad talröst i normalmaterialet.

6.1.3 Skillnader mellan män och kvinnor

En viss skillnad kan noteras mellan män och kvinnor vad gäller K_i's variation. Kvinnornas nivåökningsfaktor har ett något högre värde inom ett större frekvensområde. Kvoten: $\text{Frekvensvärde}_{\text{Män}}$ / $\text{Frekvensvärde}_{\text{Kvinnor}}$, då $K_i > 1$, visar sig vara 0,8 vid såväl den undre som övre frekvensgränsen. Enkelt uttryckt förefaller kvinnors K_i vara transponerad i förhållande till männens.

Tidigare nämndes att både kvinnor och män hade K_i-toppar i frekvensområdena runt ca 3500 Hz respektive 4000 Hz. Dessa värden är ungefärliga och inte exakta beräkningar. Författaren beräknade ändå kvoten av mäns frekvensvärde/kvinnors frekvensvärde (3500/4000), vid topparna, som visade sig vara ca 0,9.

6.2 Klinisk tillämpning

Utifrån den ekvation som beskriver det linjära förhållandet mellan genomsnittlig ljudnivå och Leq, vid en given frekvens, kan den genomsnittliga ljudnivån räknas ut för varje frekvensband vid valfri Leq. Resultatet av en sådan beräkning för män respektive kvinnor vid Leq = 70 dB presenterades tidigare i resultatsdelen (se sid. 27, figur 23).

För att visa på graden av överensstämmelse mellan beräknade och uppmätta värden hos LTAS beräknades medelvärdet för den kvadrerade differensen vid tre intensitetsnivåer (se sid. 23-26, figur. 17 –22). Medelvärdet överstiger aldrig 3 hos män, men befinner sig strax över 3 vid ett tillfälle hos gruppen kvinnor.

Möjligheten till att beräkna ett förväntat LTAS vid en vald röstintensitet (Leq) kan användas i det kliniska arbetet. Exempelvis skulle det kunna användas för att utvärdera effekten av röstbehandling. När en patient vid första besöket spelas in vid läsandet av en standardtext, görs det vid tre olika röststyrkenivåer. Mycket starkt, vanligt samtalston och mycket svagt. Utifrån inspelade lässekvenser skulle sedan LTAS och ekvationer med K_i som konstant fås fram på samma sätt som i denna studie. LTAS vid vald röstintensitet (Leq) skulle därefter kunna beräknas. Efter utförd behandling upprepas inspelningarna på samma sätt som före behandling. Ett LTAS som representerar rösten efter behandling kan beräknas för samma röstintensitet som valdes för beräknat LTAS före behandling. LTAS före och efter behandling kan sedan jämföras och därmed påvisa röstförändringar i form av akustiska korrelat. Man skulle till och med kunna jämföra patientens LTAS med de medelvärdes–LTAS som framkommit för män respektive kvinnor i denna studie. En sådan jämförelse är säkert värdefull för att beskriva stora röstavvikelser, men knappast användbar vid smärre avvikelser, eftersom en stor interindividuell variation förekommer inom respektive grupp i studien. Om en rösts LTAS emellertid avviker betydligt mer än en SD, några frekvensområden, från medelvärdes-LTAS kan detta troligen ses som en indikation på avvikande röst.

Exempel på en sådan stor röstavvikelse skulle kunna vara en kraftig topp i frekvensområdet 2-3 kHz, motsvarande en sångformant. Likaså en kraftig topp i frekvensområdet runt 3500 Hz för män och en förmodad sådan vid 4000 Hz för kvinnor, skulle kunna identifieras vid jämförelse mellan ett individuellt LTAS och medelvärdes-LTAS.

Den här studien genomfördes för att bidra med ökad kunskap om LTAS som röstanalysmetod. Strategin var att förklara ändringar i LTAS som uppstår till följd av en variation av

röstintensiteten. Denna kunskap liksom bästa sättet att extrahera information ur LTAS på, har efterlysts i tidigare studier (Kitzing & Åkerlund, 1993; Kitzing, 1986). Studien ger en uppfattning om hur variation av röstintensiteten tar sig i uttryck i LTAS, liksom förslag på hur framkomna data kan användas i den kliniska situationen.

Materialet från studien bidrar med etablerandet av normativa data för LTAS variation vid en förändrad röststyrka. Man kan vid användande av studiens data se om en röst är starkt avvikande eller inte, vad gäller tidigare diskuterade komponenter. Liknande undersökningar, på såväl normala som dysfunktionella röster, välkomnas eftersom de ytterligare skulle klargöra "röstbilden" vid en intensitetsvariation.

Nivåökningsfaktorn, K_i, som infördes av Ternström (1989), valdes för att beskriva konfigurationsförändringen hos LTAS i denna studie. Styrkan med K_i är den mycket höga korrelationen med Leq, vilket gör predicerande LTAS vid godtyckligt valda Leq relativt säkra. Om förändring i röstintensiteten är förknippad med förändringen hos ett eller flera avgränsade frekvensband kan det lätt dokumenteras. Det är dessutom möjligt att att frekvensvariationen hos K_i innefattar information av relevans för en beskrivning av röstfunktionen. Detta kan vara värt att undersöka i en ny studie.

Denna studie bidrar med fakta om hur vi kan utvinna information ur LTAS och vad denna information beskriver. Två olika LTAS kan nu jämföras utan att informationen om rösten till stor del maskeras av skilda intensitetsvärden (Leq).

7. SLUTSATSER

1. Vid frekvenser ≥ 0,5 kHz är $K_i > 1$, hos män.
   Vid frekvenser ≥ 0,625 kHz är $K_i > 1$, hos kvinnor.
2. Max $K_{i\ KVINNOR}$ > Max $K_{i\ MÄN}$.
   $K_{i\ KVINNOR} > 1$ vid 500–5000 Hz.
   $K_{i\ MÄN} > 1$ vid 500–4000 Hz.

SLUTORD

Ett varmt tack till min handledare Johan Sundberg, som med sin rika kunskap och erfarenhet väglett mig i mitt uppsatsarbete! Han har under arbetets gång uppmuntrat, stöttat och ödmjukt delat med sig av sin kompetens. Nu förstår jag vad forskning handlar om.
Tack för en lärorik och rolig tid!

Jag vill också tacka

Svante Granqvist på TMH (KTH) som utan förbehåll förklarade svårbegripligheter, hjälpte till med datorprogram och "experimentellt tänk".

Cathrin Dunger på TMH (KTH) för hjälp med knepiga koder och svårfunnen litteratur.

Alla på TMH (KTH) för ert varma omhändertagande och den generositet som ni visat mig.

Mina kursare och kollegor Anna Karlsson, Elisabet Sundewall, Britt Mattsson och Anna Henriksson för uppbackning med decibel, korrekturläsning och utskrift.

Alla försökspersoner utan vilkas medverkan studien inte existerat.

Min familj och mina vänner, som lyssnat och hjälpt mig under en tid som inneburit många förändringar.

REFERENSER

AB Nyvalla DSP (1996): Soundswell Signal Workstation, Users Guide.

Baken, R. J. (1987). *Clinical measurement of speech and voice.* Boston, Massachusetts, USA: College – Hill Press.

Bele, I. (2002). Professional speaking voice – a perceptual and acoustic study of male actors and teachers' voices, *Doctoral Thesis*, Department of Special Needs Education, Faculty of Education – University of Oslo, 2002.

Cleveland, T. F., Sundberg, J., & Stone, R. E. (2001). Long-term-average spectrum characteristics of country singers during speaking and singing. *TMH-QPSR*, TMH, KTH 2-3/2000: 89-94.

Coleman, R. F., Mabis, J. H., & Hinson, J. K. (1977). Fundamental frequency – sound pressure level profiles of adult male and female voices. *Journal of Speech and Hearing Research*, 20: 197-204.

Fahlstedt, E. & Morell, A. (1999). *Det subglottala tryckets inverkan på normalröstkällan vid konstant grundtonsfrekvens.* Examensarbete i logopedi. Institutionen för logopedi och foniatri, Lunds Universitet.

Fant, G. (1973). *Speech sounds and features.* Camebridge, MA: The MIT Press.

Fritzell, B., Hallén, O., & Sundberg, J. (1974). Evaluation of Teflon injection procedures for paralytic dysphonia. *Folia Phoniatrica*, 26: 414-421.

Gauffin, J., & Sundberg, J. (1977). Clinical application of acoustic voice analysis. Part II: Acoustic analysis, results and discussion. *STL-QPSR*, TMH, KTH 2-3/1977: 39-43.

Gauffin, J., & Sundberg, J. (1980). Data on the glottal voice source in vowel production. *STLP-QPSR*, TMH, KTH 2-3/1980: 61-70.

Gelfer, M. & Young, S. (1997). Comparison of intensity measures and their stability in male and female speakers. *Journal of Voice*, 11: 178-186.

Gramming, P. (1988). The phonetogram. An experimental and clinical study. *Doctoral thesis*, Dept. of Otolaryngology, University of Lund, Malmö General Hospital, Malmö, Sweden.

Gramming, P. (1991). Vocal loudness and frequency capabilities of the voice. *Journal of Voice*, 5: 144-157.

Hammarberg, B., Fritzell, B., Gauffin, J. & Sundberg, J. (1986). Acoustic and perceptual analysis of vocal dysfunction. *Journal of Phonetics*, 14: 533-547.

Hammarberg, B., Fritzell, J., Gauffin, J., Sundberg, J. & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica*, 90: 441-451.

Hammarberg, B., Fritzell, B. & Schiratzki, H. (1984). Teflon injection in 16 patients with paralytic dysphonia: Perceptual and acoustic evaluations. *Journal of Speech and Hearing Disorders,* 49: 72-82.

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *Journal of Acoustic Society of America,* 106: 1532-1541.

Kitzing, P. (1986). LTAS criteria pertinent to the measurement of voice quality. *Journal of phonetics,* 14: 477-482.

Kitzing, P. & Åkerlund, L. (1993). Long-time average spectrograms of dysphonic voices before and after therapy. *Folia Phoniatrica,* 45: 53-61.

Ladefoged, P. (1961). Subglottal activity during speech. *Proceedings of the Ivth International Congress of Phonetic Sciences, Helsingfors,* 73-91.

Leino, T. (1993). Long-term average spectrum study on speaking voice quality in male actors. In: Friberg A, Iwarsson J, Jansson E & Sundberg J, eds. *SMAC 93 (Proceedings of the Stockholm Music Acoustics Conference 1993),* Stockholm: Publication Nr 79 issued by the Royal Swedish Academy of Music: 206-210.

Leino, T., & Kärkkäinen, P. (1995). On the effects of vocal training on the speaking voice quality of male student actors. *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, 13-19 August 1995, K.Elenius and P.Branderud (eds.)* Vol. 3: 496-499.

Lindblad, P. (1992). *Rösten.* Lund: Studentlitteratur.

Löfqvist, A. (1986). The long-time-average spectrum as a tool in voice research. *Journal of Phonetics,* 14: 471-475.

Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice Analysis. *Folia phoniat.* 39: 221-229.

Mendoza, E., Valencia, N., Munoz, J. & Trujillo, H. (1996). Differences in voice quality between men and women: Use of long-term average spectrum (LTAS).

Nawka, T., Anders, L. C., Cebulla, M., Zurakowski, D. (1997). The speakers formant in male voices. *Journal of Voice,* 11: 422-428.

Sundberg, J. (1986). Session 5. Long-term average spectrum analysis. Chairman's Summary. *Journal of Phonetics. Theme Issue "Voice Acoustics and Dysphonia", Gotland, Sweden, August 1985* , 14: 493-494.

Sundberg, J. (2001). *Röstlära.* Stockholm: Proprius förlag.

Sundberg, J., & Gauffin, J. (1979). Waveform and spectrum of the glottal voice source. 301-320. *Frontiers of Speech Communication Research,* eds. Lindblom, B. & Öhman, S. Academic Press London.

Sundberg, J., Ternström, S., Perkins, W. H., & Gramming P. (1988). Long-term average spectrum analysis of phonatory effects of noise and filtered auditory feedback. *Journal of Phonetic,* 16: 203-219.

Ternström, S. (1989). Acoustical aspects of choir singing. *Doctoral dissertation*, Dept. of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Titze, I. R. (1988). Physiologic and acoustic differences between male and female voices. *Journal of Acoustic Society of America,* 85: 1699-1707.

Titze, I. R. (1994). *Principles of voice production.* New Jersey, U. S. A: Prentice Hall.

Titze, I. R., & Sundberg J. (1991). Vocal intensity in speakers and singers. *Journal of Acoustic Society of America,* 91: 2936-2946.

Wedin, S., Leanderson, R. & Wedin, L. (1978). Evaluation of voice training. Spectral analysis compared with listeners' judgements. *Folia phoniatrica,* 30:103-112.

White, P. (2001). Long-term average spectrum (LTAS) analysis of sex- and gender-related differences in children's voices. *Logopedics Phonetics Vocology,* 26: 97-101.

Bilaga 1

Frågor inför röstinspelning

Försöksperson nr:

Dagens datum:

Enkätsvaren kommer precis som inspelningarna att kodas och behandlas konfidentiellt.
Vid frågor där flera alternativ ges är det möjligt att välja flera alternativ.

1. Är du: ( ) Man ( ) Kvinna Födelseår_____

2. Har du nedsatt hörsel?
( ) Nej ( ) Ja

3. Röker du?
( ) Nej ( ) Ja, nämligen:
( ) "feströker" ( ) mindre än 5 cig/dag
( ) 5-10 cig/dag ( ) mer än 10 cig/dag

4. Har du någon lungsjukdom/astma?
( ) Nej ( ) Ja, nämligen_____

Om ja, medicinerar du?
( ) Nej ( ) Ja, nämligen_____

5. Upplever du någon form av problem med din röst?
( ) Nej (Hoppa till fråga 9) ( ) Ja

6. Sätt ett lodrätt sträck på skalan nedan så att den beskriver dina röstbesvär:

Inga röstbesvär ---------- Mycket stora röstbesvär

7. Hur vill du beskriva dina röstbesvär?
( ) heshet ( ) torrhet ( ) tryck över struphuvudet
( ) trötthet i rösten ( ) kliar ( ) klumpkänsla
( ) smärta/sveda ( ) harklingsbehov ( ) svårt att göra mig hörd
( ) "tappar rösten" ( ) torrhosta ( ) minskat i omfång (i sång)

( ) övrigt _____

8. När uppstår dessa besvär?
( ) Nästan alltid
( ) Efter röstansträngning, t ex _____
( ) I särskilda situationer, t ex _____

9. Håller du på med någon sång- eller talaktivitet på fritiden?
( ) Nej ( ) Ja, nämligen:
( ) körsång ( ) sjunger med barn ( ) sånglektioner
( ) solosång ( ) spelar teater ( ) tallektioner
( ) sjunger i band

( ) Annan tal-/sång aktivitet_____

Hur många timmar i veckan?
( ) 1 ( ) 2-3 ( ) 4-6 ( ) mer

Eventuell kommentar:_____

10. Har du tidigare ägnat dig åt någon av följande aktiviteter på arbetet eller fritiden?
( ) Nej ( ) Ja, nämligen
( ) körsång ( ) sjunger med barn ( ) sånglektioner
( ) solosång ( ) spelar teater ( ) tallektioner
( ) sjunger i band

( ) Annan tal-/sång aktivitet_____

Under hur lång period och hur ofta?_____

_____

11. Har du fått träning i, eller information om rösten?
( ) Nej ( ) Ja, nämligen_____

_____

Tack för din medverkan!

Standardtext

Ljuset kom från eldstaden. Gulvitt skarpt.

- Backa mamma! ropade jag.

Mamma tog ett kort steg bakåt med eldgaffeln. Men ljuset följde efter. Ut ur brasan svävade ett lysande klot.

Det gnistrade som vitt, smälthett järn och guppade svagt, som om det flöt i luften. Klotet stannade med en darrning, tätt intill eldgaffeln. Jag såg hur mamma började lysa. Ett blått sken kring huden. Hennes hår reste sig och började spreta åt alla håll.

- Släpp den! Släpp eldgaffeln!

Men hon tycktes som hypnotiserad. Backade bara steg för steg medan hon svängde gaffeln i sidled. Klotet följde efter. Hon svängde snabbare, men klotet drogs som till en magnet.

- Släpp den morsan!

I stället började hon snurra med gaffeln. Roterade som en släggkastare för att försöka skaka av sig främlingen.

Men den följde med. Hon ökade farten. Tog i så det frasade i luften. Små blixtar knastrade kring skaftet. Men klotet satt kvar vid spetsen. Hon flåsade, gapade och snurrade allt vildare. Snart omgavs hon av en ring, en gloria, en cirkel av elektricitet. Hon kunde inte sluta. Snurrade och snurrade tills det susade, sjöng, tills hela rummet sprakade av blåa lågor. Hon ökade ännu mer. Och ännu mer. Tills det borde varit omöjligt.

Då tappade hon greppet. Eldgaffeln och klotet for rätt ut. Smällde i väggen, kabraaaaakkk! Ett dån så öronen korvade sig, träflisor smattrade över oss. Och sedan tystnad.

Ur "Populärmusik från Vittula" av Mikael Niemi

MV nivåökningsfaktor (K_i), män

Frekvens	K_i	SD
125	0,6	0,06
250	0,8	0,08
375	1,0	0,09
500	1,1	0,10
625	1,2	0,10
750	1,3	0,10
875	1,4	0,09
1000	1,4	0,09
1125	1,4	0,08
1250	1,4	0,08
1375	1,4	0,08
1500	1,4	0,08
1625	1,4	0,08
1750	1,5	0,08
1875	1,5	0,07
2000	1,5	0,08
2125	1,4	0,09
2250	1,4	0,10
2375	1,4	0,12
2500	1,4	0,15
2625	1,4	0,16
2750	1,4	0,16
2875	1,4	0,13
3000	1,4	0,12
3125	1,4	0,12
3250	1,4	0,16
3375	1,3	0,22
3500	1,3	0,25
3625	1,2	0,24
3750	1,2	0,21
3875	1,1	0,19
4000	1,0	0,17
4125	1,0	0,16
4250	0,9	0,16
4375	0,8	0,17
4500	0,8	0,18
4625	0,8	0,18
4750	0,8	0,20
4875	0,8	0,21
5000	0,8	0,20
5125	0,8	0,17
5250	0,8	0,14
5375	0,8	0,13
5500	0,8	0,15
5625	0,8	0,16
5750	0,8	0,19
5875	0,9	0,21
6000	0,9	0,21

MV intercept, män

Frekvens	Intercept	SD
125	-90	6
250	-103	7
375	-118	8
500	-130	10
625	-138	10
750	-147	10
875	-156	9
1000	-159	9
1125	-161	7
1250	-162	6
1375	-162	6
1500	-164	7
1625	-166	8
1750	-169	7
1875	-171	6
2000	-171	7
2125	-171	7
2250	-171	8
2375	-171	10
2500	-171	12
2625	-171	13
2750	-171	13
2875	-171	12
3000	-170	11
3125	-169	11
3250	-167	13
3375	-165	16
3500	-163	18
3625	-160	17
3750	-156	14
3875	-151	12
4000	-148	13
4125	-144	13
4250	-138	13
4375	-134	14
4500	-131	14
4625	-130	14
4750	-130	15
4875	-131	15
5000	-132	15
5125	-133	13
5250	-134	11
5375	-134	11
5500	-136	12
5625	-137	13
5750	-137	16
5875	-137	17
6000	-138	17

MV korrelation (r), män

Frekvens	r	SD
125	0,99	0,01
250	1,00	0,00
375	1,00	0,00
500	1,00	0,01
625	0,99	0,01
750	1,00	0,00
875	1,00	0,00
1000	1,00	0,00
1125	1,00	0,00
1250	0,99	0,00
1375	1,00	0,00
1500	1,00	0,00
1625	1,00	0,00
1750	1,00	0,00
1875	1,00	0,00
2000	1,00	0,00
2125	1,00	0,01
2250	0,99	0,01
2375	0,99	0,01
2500	0,99	0,01
2625	0,99	0,01
2750	0,99	0,01
2875	0,99	0,01
3000	0,99	0,01
3125	0,99	0,01
3250	0,99	0,01
3375	0,99	0,01
3500	0,99	0,01
3625	0,99	0,01
3750	0,99	0,02
3875	0,98	0,03
4000	0,98	0,03
4125	0,98	0,02
4250	0,98	0,02
4375	0,98	0,02
4500	0,98	0,02
4625	0,97	0,03
4750	0,97	0,04
4875	0,97	0,03
5000	0,98	0,03
5125	0,98	0,02
5250	0,98	0,01
5375	0,98	0,01
5500	0,98	0,01
5625	0,98	0,01
5750	0,99	0,01
5875	0,99	0,01
6000	0,99	0,00

MV nivåökningsfaktor (K_i), kvinnor

Frekvens	K_i	SD
125	0,3	0,08
250	0,5	0,10
375	0,8	0,13
500	1,0	0,18
625	1,2	0,20
750	1,3	0,18
875	1,4	0,16
1000	1,5	0,15
1125	1,5	0,16
1250	1,6	0,18
1375	1,6	0,19
1500	1,6	0,17
1625	1,6	0,15
1750	1,6	0,14
1875	1,6	0,15
2000	1,6	0,15
2125	1,6	0,12
2250	1,6	0,11
2375	1,6	0,11
2500	1,6	0,10
2625	1,6	0,10
2750	1,6	0,12
2875	1,6	0,13
3000	1,6	0,13
3125	1,6	0,13
3250	1,5	0,13
3375	1,5	0,15
3500	1,5	0,16
3625	1,4	0,18
3750	1,5	0,22
3875	1,5	0,26
4000	1,5	0,27
4125	1,5	0,26
4250	1,4	0,26
4375	1,4	0,25
4500	1,3	0,26
4625	1,3	0,25
4750	1,2	0,23
4875	1,1	0,22
5000	1,0	0,22
5125	0,9	0,23
5250	0,9	0,24
5375	0,9	0,23
5500	0,8	0,22
5625	0,8	0,23
5750	0,8	0,22
5875	0,8	0,19
6000	0,8	0,17

MV intercept, kvinnor

Frekvens	Intercept	SD
125	-70	7
250	-86	8
375	-104	10
500	-122	13
625	-136	14
750	-148	13
875	-157	12
1000	-163	12
1125	-167	13
1250	-170	13
1375	-173	13
1500	-176	11
1625	-177	10
1750	-177	10
1875	-177	10
2000	-179	10
2125	-180	9
2250	-180	7
2375	-179	7
2500	-179	7
2625	-178	8
2750	-178	9
2875	-177	9
3000	-177	9
3125	-178	9
3250	-177	9
3375	-175	11
3500	-174	12
3625	-173	14
3750	-174	17
3875	-176	19
4000	-175	19
4125	-174	18
4250	-173	18
4375	-170	18
4500	-167	18
4625	-163	17
4750	-158	16
4875	-152	15
5000	-146	16
5125	-141	17
5250	-137	17
5375	-135	17
5500	-133	17
5625	-132	18
5750	-130	17
5875	-129	16
6000	-129	14

MV korrelation (r), kvinnor

Frekvens	r	SD
125	0,94	0,06
250	0,98	0,01
375	0,98	0,01
500	0,98	0,02
625	0,98	0,01
750	0,99	0,01
875	0,99	0,01
1000	0,99	0,01
1125	0,99	0,01
1250	0,99	0,01
1375	0,99	0,01
1500	0,99	0,00
1625	0,99	0,00
1750	0,99	0,00
1875	0,99	0,00
2000	0,99	0,00
2125	0,99	0,00
2250	0,99	0,01
2375	0,99	0,01
2500	0,99	0,01
2625	0,99	0,01
2750	0,99	0,01
2875	0,99	0,01
3000	0,99	0,01
3125	0,99	0,00
3250	0,99	0,00
3375	0,99	0,00
3500	0,99	0,01
3625	0,99	0,01
3750	0,99	0,01
3875	0,99	0,01
4000	0,99	0,01
4125	0,99	0,01
4250	0,99	0,01
4375	0,99	0,01
4500	0,99	0,01
4625	0,99	0,01
4750	0,99	0,01
4875	0,99	0,01
5000	0,99	0,01
5125	0,99	0,01
5250	0,98	0,01
5375	0,98	0,02
5500	0,97	0,02
5625	0,97	0,02
5750	0,97	0,02
5875	0,97	0,02
6000	0,97	0,03

MV av kvadrerad avvikelse hos män, vid läsning i brusnivå 95 dB, 77 dB samt utan brus.

Frekvens	95 dB	77 dB	Utan brus
125	0,9	0,4	0,5
250	0,9	0,8	0,7
375	1,1	1,0	1,1
500	1,5	1,4	1,5
625	1,6	1,5	1,6
750	1,6	1,4	1,5
875	1,5	1,3	1,2
1000	1,4	1,4	1,3
1125	1,5	1,7	1,5
1250	1,5	1,7	1,5
1375	1,4	1,6	1,4
1500	1,5	1,6	1,4
1625	1,7	1,5	1,3
1750	1,6	1,4	1,2
1875	1,4	1,4	1,2
2000	1,1	1,4	1,2
2125	1,2	1,4	1,4
2250	1,5	1,5	1,5
2375	1,6	1,3	1,3
2500	1,4	1,4	1,0
2625	1,3	1,3	0,9
2750	1,4	1,2	0,9
2875	1,6	1,2	1,0
3000	1,6	1,3	1,3
3125	1,6	1,5	1,3
3250	1,6	1,8	1,3
3375	1,7	1,9	1,3
3500	1,7	1,7	1,2
3625	2,0	1,6	1,2
3750	2,3	1,5	1,2
3875	2,8	1,6	1,5
4000	2,8	1,8	1,6
4125	2,6	1,7	1,5
4250	2,6	1,6	1,4
4375	2,6	1,5	1,3
4500	2,4	1,5	1,2
4625	2,2	1,5	1,1
4750	2,1	1,6	1,0
4875	1,9	1,5	1,0
5000	1,8	1,5	1,1
5125	1,7	1,3	1,2
5250	1,5	1,3	1,1
5375	1,5	1,5	1,1
5500	1,6	1,6	1,2
5625	1,7	1,6	1,2
5750	1,7	1,4	1,2
5875	1,5	1,1	1,0
6000	1,2	1,1	0,9

MV av kvadrerad avvikelse hos kvinnor, vid läsning i brusnivå 95 dB, 77 dB samt utan brus.

Frekvens	95 dB	77 dB	Utan brus
125	1,2	1,1	1,1
250	1,1	1,0	1,0
375	1,6	1,6	1,6
500	2,5	2,5	2,5
625	2,7	2,9	2,7
750	2,4	3,1	2,7
875	2,2	3,0	2,6
1000	2,0	2,6	2,3
1125	2,2	2,4	2,3
1250	2,6	2,8	2,6
1375	2,9	3,0	2,7
1500	2,8	2,7	2,5
1625	2,5	2,2	2,2
1750	2,3	2,3	2,1
1875	2,3	2,7	2,2
2000	2,4	2,8	2,4
2125	2,5	2,4	2,3
2250	2,5	2,2	2,3
2375	2,5	2,2	2,3
2500	2,5	2,0	2,2
2625	2,6	1,9	2,2
2750	2,6	2,1	2,3
2875	2,6	2,2	2,3
3000	2,6	2,2	2,2
3125	2,5	2,1	2,0
3250	2,2	1,8	1,6
3375	2,2	1,7	1,5
3500	2,3	1,7	1,6
3625	2,4	2,0	1,8
3750	2,6	2,2	2,0
3875	2,7	2,2	2,2
4000	2,8	2,0	2,2
4125	2,7	1,9	2,0
4250	2,7	1,8	1,8
4375	2,6	2,0	1,9
4500	2,3	2,2	1,9
4625	2,0	2,2	1,9
4750	1,6	2,1	1,7
4875	1,2	2,0	1,3
5000	1,1	1,9	1,2
5125	1,2	1,7	1,2
5250	1,7	1,9	1,5
5375	2,1	2,1	1,7
5500	2,3	2,1	1,8
5625	2,3	2,0	1,7
5750	2,4	2,1	1,7
5875	2,6	2,2	1,8
6000	2,6	2,3	1,8

MV av genomsnittlig ljudnivå hos män och kvinnor för beräknat LTAS vid Leq = 70 dB.

Frekvens	Genomsnittlig ljudnivå, män	SD	Genomsnittlig ljudnivå, kvinnor	SD
125	-49	5,8	-49	4,8
250	-49	6,0	-49	4,9
375	-50	6,2	-49	5,3
500	-51	6,4	-50	5,8
625	-53	6,3	-51	5,7
750	-57	6,2	-54	5,6
875	-61	6,3	-57	5,5
1000	-64	6,7	-59	5,1
1125	-64	6,4	-60	4,6
1250	-64	5,9	-61	4,3
1375	-64	5,5	-61	4,4
1500	-64	5,4	-62	4,7
1625	-65	5,2	-63	5,3
1750	-67	4,9	-63	5,5
1875	-69	4,4	-64	5,5
2000	-69	4,1	-65	5,5
2125	-70	4,1	-66	5,6
2250	-70	4,5	-67	5,3
2375	-70	5,0	-68	4,9
2500	-70	5,2	-68	4,9
2625	-71	5,4	-68	5,1
2750	-73	5,6	-67	5,1
2875	-74	5,5	-67	4,9
3000	-74	5,3	-67	4,9
3125	-73	5,2	-69	5,1
3250	-72	5,1	-70	5,4
3375	-72	4,9	-72	5,4
3500	-73	4,8	-72	5,0
3625	-73	4,4	-72	4,7
3750	-74	3,8	-72	4,7
3875	-75	3,7	-72	4,7
4000	-76	4,1	-72	4,5
4125	-77	4,6	-71	4,3
4250	-77	5,0	-71	4,4
4375	-76	5,1	-72	4,8
4500	-76	5,3	-73	5,0
4625	-75	5,3	-74	4,9
4750	-75	5,6	-74	4,7
4875	-75	5,9	-75	4,9
5000	-75	6,0	-75	5,2
5125	-76	6,2	-75	5,3
5250	-76	6,5	-76	5,3
5375	-77	6,8	-76	5,3
5500	-77	6,8	-76	5,5
5625	-78	6,7	-76	5,7
5750	-78	6,7	-75	5,7
5875	-77	6,7	-75	5,5
6000	-78	6,5	-75	5,4