

Lunds universitet
SOL Lingvistik och fonetik
Handledare Gösta Bruce
2006-09-20

Malin Svensson
kandidatuppsats i fonetik

En fonetisk analys av uppläst fri vers

de tonala utspelen hos Kristina Lugn

Abstract

Uppläst fri vers upptar mycket liten plats inom den fonetiska och talteknologiska forskningen. Syntetiska modeller har ännu inte utarbetats för uppläst fri vers, troligen på grund av att det är svårt att fastställa poetiska effekter. Låg talhastighet och särskilda accent- och betoningsmönster är en del av en litterär konvention för den upplästa fria versen. Dessa parametrar är troligen knutna till den specifika talsituationen. Även poetiska effekter i form av unika prosodiska mönster för varje dikt tillhör konventionen för uppläst fri vers.

Auditiv, fonologisk och akustisk analys görs av en svensk poets upplästa fria vers; Kristina Lugns *Om ni hör ett skott...* Tidigare undersökning fastställer denna talares låga talhastighet, vilken är enligt konventionen för uppläst fri vers. I denna undersökning upplevs intonation vara den mest prominenta prosodiska parametern i form av tonala utspel. Analysen visar att det globala tonomfånget varierar i storlek i samspel med emfas hos den fokala accenten i frasen. En poetisk effekt hos det undersökta materialet kan antas vara stor variation i emfas mellan fraser. Detta realiserar i närvaro och frånvaro av tonala utspel. Talaren visar sig använda korta fraser. I materialet påträffas inte sällan mer än en fokal accent inom vissa fraser, samt frånvaro av fokala accenter i andra fraser, vilket går emot en allmän förväntan om EN fokal accent per fras. Att medvetet undvika samt framhäva fokala accenter kan även det ses som en poetisk effekt.

Innehåll

Abstract.....	2
Innehåll.....	3
Introduktion.....	4
Syfte och tidigare forskning.....	4
Teoretisk bakgrund.....	4
Prosodin i talstilar.....	4
Talsyntes: poesiuppläsning.....	5
Poetiska effekter i tal.....	6
Frågeställning.....	7
Material.....	8
Metod.....	8
Initial analys.....	8
Fonologisk transkription.....	8
Akustisk analys.....	9
Resultat.....	10
Initial analys.....	10
Fonologisk transkription.....	10
Akustisk analys.....	10
Diskussion.....	14
Källförteckning.....	17
Bilaga	

Introduktion

All uppläst poesi försöker fånga sin publiks uppmärksamhet. Poeten/artisten som framför verket trollbinder sin publik med hjälp av bl.a. rytmiska och systematiska grepp i talet. Olika versformers uppläsning har genom olika tider följt vissa konventioner. Accentversen och metrisk vers hade t ex under sin storhetstid skaller som sjöng och läste högt. De använde versrytm och förutbestämda prominenser som rytmiska grepp. Ett sådant grepp innebär regelbundna starka och svaga enheter, där de starka enheterna uppfattas som mer prominenta och i kontrast till de svagare (Bruce, 1998, Lilja, 2003). Rytmiken kan innebära att ord förlorar sin betoning eller att betoningsmönster förändras (Lilja, 2003). Detta för att undvika betoningskrock när versrytmen ska bibehållas.

Eva Lilja, professor i litteraturvetenskap, menar att liksom det finns en konvention för rytmiska versformer finns det också en konvention i uppläsandet av obunden fri vers. Uppläsning av fri vers ser inte likadan ut som den av rytmisk vers. Till skillnad från rytmisk vers läses fri vers efter prosaaccenter. Uppläsning av den fria versen liknar därför på många sätt uppläsning av prosa. Men fri vers tar längre tid att läsa upp än prosa vilket kan betyda att talaren av uppläst fri vers har lägre talhastighet, kanske i form av längre pauslängd eller fler pauser. Dessutom styrs uppläsandet av den individuella versens unika tematik, rytmik, repetitioner, fraslängd o s v, vilket är knutet till versförlagan (Lilja, 2003). Det står inte helt klart hur konventionen för uppläst fri vers ser ut. Emellertid tycks uppläst fri vers kännetecknas av låg talhastighet, fler/längre pauslängder samt betoningsmönster som följer prosaaccenter. Unika prosodiska mönster tycks dessutom vara en del av varje dikt.

Syfte och tidigare forskning

Det finns mycket lite fonetisk forskning om poesiuppläsningar. Viss forskning om talstilar berör uppläsning av text, men inte specifikt för fri vers. Inom TTS, text-till-tal-syntes, används prosodiska modeller för att skapa tal av inmatade texter. Vid KTH har man arbetat på prosodiska modeller för läst tal av skönlitterär text (prosa). Modeller för talhastighet, frasering, intonation och timing har utarbetats för läst prosa och för rytmisk vers (Fant et al, 1989, 2002). Eva Lilja har hänvisat till resultat från Fant et al i sina teorier om olika konventioner för uppläsning (Lilja, 2003). Det finns dock ännu inga prosodiska modeller utarbetade specifikt för uppläst fri vers.

Syftet med denna uppsats är att se hur man fonetiskt kan beskriva vad som utmärker en poesiuppläsning och om det finns prosodiska mönster i talarens sätt att läsa den fria versen. Syftet är också att se på vilket sätt en sådan beskrivning kan bidra till fonetisk forskning. Denna undersökning är alltså deskriptiv och ett första försök till en fonetisk analys av uppläst fri vers.

Teoretisk bakgrund

Prosodin i talstilar

Olika talstilar förmodas verkställas prosodiskt på olika vis. Traditionellt sett har talstilar delats in i spontan talstil respektive läst talstil. Medan spontant tal kännetecknas av bl.a. stor fonetisk variation, flexibilitet, oväntade pauseringar, repetitioner och reparationer, brukar läst tal beskrivas som dess motsats. En läst talstil är istället mer formell och mindre variabel. Medan den spontana talaren kanske tappar några språkljud här och där, är en uppläsande talare ofta välartikulerad. En läst talstil följer också gärna interpunktion hos textförlagan och har helst ingen större variation i tontoppar eller tondalar (Bruce, 1998).

Distinktioner mellan spontan, planerad, konversationell och formell talstil har också gjorts. Forskning om vilka prosodiska drag som kännetecknar vilka talstilar har lett till diskussioner om huruvida dessa kategoriseringar är lämpliga. Prosodin varierar inom olika talstilar bl.a. beroende på relationen mellan talare och lyssnare, deras sociala status, talarens intentioner och talsituationen (Hirschberg, 2000). Julia Hirschberg påpekar att just distinktionen mellan läst och spontant tal hellre skall ses som ett kontinuum mellan dessa båda poler, mellan vilka det förekommer ytterligare talstilar. Genom en undersökning baserad på korpus av spontant tal och läst tal visar hon att prosodiska parametrar kan variera inom de båda talstilarna. Parametrarna disfluens, intonation, talhastighet och diskurs skiljer sig något från tidigare uppfattningar om spontant och läst talstil. Disfluens visade sig till exempel inte vara utbredd i lika hög grad i spontan talstil som förut antagits. Talhastigheten var dessutom oväntat högre för läst tal än för spontan talstil (Hirschberg, 2000).

Även andra undersökningar visar på svårigheter att knyta ett prosodiskt drag till en viss talstil. Lilja påpekade just skillnaden i talhastighet mellan uppläst prosa och uppläst fri vers, vilken förmodades vara högre för uppläst prosa (Lilja, 2003). Både uppläst prosa och uppläst fri vers ses allmänt som lästa talstilar. Att talhastigheten skiljer de två lästa talstilarna åt kan tyda på en skillnad mellan till exempel två talsituationer. En social gemenskap knuten till en talsituation bestämmer vilka prosodiska variationer och även vilka paralingvistiska drag som den talsituationen belönar (Gussenhoven, 2004). I det här fallet kan det alltså även finnas en betydelsefull skillnad mellan två litterära traditioner. För den konventionella upplästa fria versen innebär detta låg talhastighet hos talaren. Att en specifik parameter ingår i konventionen för uppläst fri vers, tyder på att konventionen även troligen inbegriper en viss talsituation.

Hirschberg (2000) tar också upp förmodandet att diskursen är relaterad till en global prosodisk variation. Denna tycks ha viss betydelse för perceptionen av en talstil. Hon menar också att medan globala prosodiska parametrar i talet varierar mellan talsituationer hos samma talare, är det troligt att lyssnare går efter lokala parametrar när de analyserar diskursen semantiskt och pragmatiskt. Man kan alltså säga att för en specifik diskurs och talsituation ser de globala prosodiska dragen ut på ett visst sätt. I det här fallet består diskursen av fri vers. Talsituationen består av uppläsning inför publik. Genom att analysera korpus av uppläst fri vers skulle man kunna hitta globala prosodiska parametrar för den specifika diskursen och talsituationen. Dessa parametrar kan sedan antas vara representativa för uppläst fri vers.

Talsyntes: poesiuppläsning

Om man identifierar viktiga prosodiska drag, globala som lokala, kan man skapa en prosodisk modell av ett yttrande. Detta kan sedan få ett praktiskt användningsområde inom den talsyntetiska röstforskningen. Det finns olika modeller för olika prosodiska parametrar. Intonationsmodeller inom TTS utvecklades först från enkla akustiska modeller, t e x Fujisakimodellen, samt perceptoriska modeller med deklinationskurva hos intonationsmodellen (Dutoit, 1997). Numer inbegriper intonationsmodeller även den lingvistiska informationen hos intonationskurvan. Till exempel signaleras frågande eller osäker lingvistisk information med hög final gränston, medan låg final gränston förmedlar säker eller påstådd information (Dutoit, 1997, Bruce, 1998, Gussenhoven, 2004). Inom TTS har det också utarbetats prosodiska modeller för frasering och talhastighet, främst för den formella lästa talstilen.

Att talteknologisk forskning ännu inte tagit fram prosodiska modeller för uppläst fri vers kan bero på att den syntetiska rösten allmänt sett inte anses vara lämplig för poesiuppläsningar. Blankinship och Beckwith har emellertid gjort ett försök att utarbeta en användarvänlig mjukvara, *Poet shop*, för en befintlig syntetisk röst. Man ville förbättra förutsättningarna för en rullstolsbunden amerikansk poet med talsvårigheter. Poeten i fråga

använde redan en syntetisk röst vid poesiuppläsningar. Utvalda dikter kunde där matas in och verkställas i tal. Syftet med *Poet shop* var att poeten manuellt på scen vid en uppläsning skulle kunna tillföra poetiska effekter till den syntetiska rösten under talets gång. Verktøygen i *Poet shop* bestod av höjning eller sänkning av talhastigheten, höjning eller sänkning av F_0 , samt höjning eller sänkning av ljudnivån. Eftersom poeten hade full rörelsefunktion enbart i höger fot, tillverkades en joystick lämpligt placerat på rullstolen. Med den kunde han manövrera parametrarna och uppnå poetiska effekter i talet (Blankenship och Beckwith). *Poet shop* lyckades dessvärre inte tillföra talet naturlig personlig kvalitet, vilket av forskarna ansågs ha stor betydelse för det konstnärliga framförandet. Vad naturligt personligt utförande innebär står dock inte helt klart. Avsaknad av paralingvistiska och biologiska särdrag hos en syntetiskt simulerad röst får oftast den välkända ”robotrösten” som följd, förutom vare sig ålder, kön, känsloutspel eller attityder (Dutoit, 1997, Schötz, 2006, Gussenhoven, 2004). Personliga särdrag och poetiska effekter tycks vara betydelsefulla för uppläst poesi. De poetiska effekter som eftersträvades i *Poet shop* bestod till stor del av paralingvistisk information, d v s information som ligger utanpå den lingvistiska. Den paralingvistiska informationen i just *Poet shop* hade emellertid inte fonologisk anknytning via en prosodisk modell utan manövrerades på plats av användaren, vilket innebar att accentueringen kunde bli felaktig. Men paralingvistisk information bör samspela med fonologiska funktioner för att låta naturlig, vilket inte kunde säkerställas med verktøygen i *Poet shop*. För att få ett ”naturligt personligt utförande av poetiska effekter” hos en syntetisk röst är det därför önskvärt att poetiska effekter istället ingår i prosodiska modeller.

Poetiska effekter i tal

Vad är då poetiska effekter? Lilja säger att uppläst fri vers innebär unika prosodiska mönster (2003). De effekter som *Poet shop* eftersträvade bestod av paralingvistisk information. Poetiska effekter kan alltså antas vara mönster av paralingvistisk information. De är medvetna effekter som vi tillför den lingvistiska informationen i talet. Det finns paralingvistisk information i talet som vi med svårighet eller inte alls kan kontrollera (Gussenhoven, 2004). Att förstå en rösts ålder tillhör nog en av de svårare utmaningarna. Men det finns också paralingvistisk information som vi kan kontrollera och medvetet producera. Carlos Gussenhoven har sammanställt forskning om en prosodisk parameter, intonationskurvan, i sin *Tone and Phonology of Intonation* (2004). I ett avsnitt behandlas forskning om paralingvistisk information som är möjlig för oss att kontrollera i talet. Gussenhoven har delat upp den tonala forskningen i denna del efter tre biologiska koder: *produktionskod*, *frekvenskod* och *ansträngningskod*. De biologiska koderna förknippas med olika aspekter av vår talproduktion och påverkar såväl emfas hos fokala accenter som information om attityd bl.a. hos finala gränstoner.

Produktionskoden beror på det subglottala trycket. En effekt av produktionskoden kan man till exempel finna hos frasinitiala gränstoner. Hög ton där signalerar ett nytt ämne och låg ton en fortsättning på ämnet i föregående fras. Det omvända förhållandet gäller för frasfinal gränston. *Frekvenskoden* kan beskrivas vara förknippad med fysisk storlek. Till exempel beror lågt F_0 -register på stor larynx, vilket ger en viss affektiv information till lyssnaren. Troligen tolkas lågt F_0 -register av lyssnaren som information om maskulinitet, men det finns även andra möjliga tolkningar beroende på bl.a. andra parametrar i talet, talsituationen och sociala kontexten. Högt F_0 -register kan å andra sidan leda till andra associationer för lyssnaren. *Ansträngningskoden* beror på talarens högre ansats. Detta realiseras bland annat i fokus av vad talarens anser som viktiga delar i talet. En högre ansträngning kan också ge mer emfas. De biologiska koderna samspelar också på så vis att prominens perceptoriskt relateras till förväntat register, varför liknande tonkurva hos olika talare kan uppfattas som olika prominenta (Gussenhoven, 2004). Vi använder oss av våra biologiska koder i talet för att

påverka lyssnaren på olika vis Vad vi upplever som poetiska effekter är alltså beroende av paralingvistisk information som produceras med dessa koder.

Emotioner är även de paralingvistisk information som vi delvis kan kontrollera. Försök har gjorts att fastställa vilka prosodiska parametrar som förmedlar talarens olika känslor till lyssnaren, såsom ilska, glädje, sorg, rädsla och avsmak. Forskningen har gått isär om det verkligen finns globala tendenser i talet för en viss känsla (Bätzing och Scherer, 2005). Det tycks inte troligt att ett specifikt prosodiskt drag bär på ett visst emotionellt uttryck. Affektiv information såsom feminin och maskulin kan man hitta i intonationskurvan. Hög ton står för vänlighet men också för försvar. En final hög ton ger emellertid intryck av undergivenhet. Låg ton står för förtroendeskapande attityd, men också aggression. (Gussenhoven, 2004) Samma parameter kan alltså bära på olika information. Hur en parameter uppfattas beror bl.a. på talsituation, talstil och andra parametrar i talet.

I en förundersökning till denna uppsats (B-poster i prosodi) använde jag mig av ett sätt att fastställa emotioner i tal. Tre binära semantiska positionspar hämtade från psykologin samspelar för att bygga upp en känsla i talet. De olika paren är *aktivitet* (activation), *valens* (valence) och *dominans* (dominance) (Kehrein, 2002). Varje par verkställs i talet via en eller flera prosodiska parametrar, vilka fungerar utmed en skala mellan binära semantiska positioner. Aktivitet signaleras till exempel via talhastighet och utmed skalan *upprymd – lugn*, där upprymdhet verkställs med hjälp av högre talhastighet och lugn med lägre talhastighet. Dimensionen valens innebär en skala mellan valörerna *positiv – negativ*. En positiv attityd signaleras av samtidig förlängd stavelse, hög F0 och lägre intensitet, vilka tillsammans utgör en lokal enhet. Det finns ännu så länge ingen motsvarig enhet för negativ attityd. För dominans är de binära positionerna *dominant – undergiven*. Dessa urskiljs genom prominens. Det tycks finnas ett samband mellan global prominens, tonomfång och intensitet. En undergiven attityd har lägre tonomfång och lägre intensitet än en dominant attityd (Kehrein, 2002). Liknande modell av semantiska positionspar har bl.a. använts för att utveckla emotioner i talsyntesen MARY, *Emospeak*. I deras mjukvara kan man genom två skalor enkelt välja en kombination av de binära semantiska positionsparen, som sedan verkställs i syntetiskt tal av en inmatad text.

I förundersökningen till denna uppsats testade jag aktivitet hos en talare av uppläst fri vers. Poeten Kristina Lugns medeltalhastighet hos ett prosodiskt yttrande fastställdes till 3 stavelser/sekund. Vid ett perceptionstest i samma undersökning med 6 lyssnare visade sig den låga talhastigheten bära på den binära positionen *lugn* inom den emotionella dimensionen aktivitet (Svensson, 2005). En global prosodisk parameter i just Kristina Lugns tal kan sägas vara den låga talhastigheten, vilken enligt Lilja ingår i konventionen för uppläst fri vers (2003). Enligt resultatet i postern bör alltså konventionen för uppläsning av fri vers bära på den binära positionen *lugn*.

Frågeställning

Följande frågor preciserar syftets olika delar. Frågorna besvaras i undersökningens efterföljande diskussion.

- * Hur kan en fonetisk analys användas för att beskriva uppläst fri vers hos en talare?
- * Kan man hitta globala och lokala prosodiska parametrar som kan antas vara poetiska effekter?
- * Kan uppläst fri vers utgöra en egen talstil?
- * På vilket sätt kan en fonetisk analys bidra till att utveckla syntetisk simulering av uppläsning av fri vers?

För den fortsatta analysen i uppsatsen är utgångspunkten att det finns en variationsrik men konventionell långsam uppläsning av fri vers och att den verkställs fonetiskt unikt för varje

dikt. I nästa avsnitt görs en fonologisk och akustisk analys i ett försök att hitta poetiska effekter hos en talare av uppläst fri vers.

Material

Materialet som undersöks består av en uppläsning av poeten Kristina Lugn. Denna utgavs 1988 av Bokbandet, och är en uppläsning av diktverket "Om ni hör ett skott..." (Lugn, 1979, 1980). Diktverket är författat av Kristina Lugn och utgavs i sin tur 1979 av förlag Alba, numera Bonnier. Materialet valdes utifrån inspelningskvalitet och tillgänglighet (att det inspelade verket är komplett och tillgängligt för allmänheten), samt utifrån kriterierna för fri vers. Poeten Kristina Lugn beskrivs skriva i en satirisk realistisk litterär tradition. Som föregångare i denna tradition nämns bl.a. Majken Johansson och Sonja Åkesson, och vid beskrivning av den senares poesi används ofta begreppet "diskbänksrealism". Kristina Lugns ironiska stil har också liknats vid Anna Maria Lenngrens (Beckman, 2002, Elleström, 1990). Diktverket är i sin helhet ca 45 minuter långt. En initial analys gjordes först av hela verket. De påföljande analyserna begränsades till 6 minuters tal, d v s ca 1/7 av den ursprungliga uppläsningen.

Metod

Initial analys

En initial auditiv och delvis akustisk analys av materialet utförs i Praat. Två fonetiskt tränade personer anser följande fonetiska och prosodiska kategorier vara prominenta i materialet: dialekt, röstgolv, röstregister, röstkvalitet, talhastighet, frasering, intensitet och intonation. En stor variation i det tonala mönstret påträffas; så kallade tonala utspel. Eftersom det tycks förekomma en prominent tonal variation, undersöks denna parameter vidare.

Fonologisk transkription

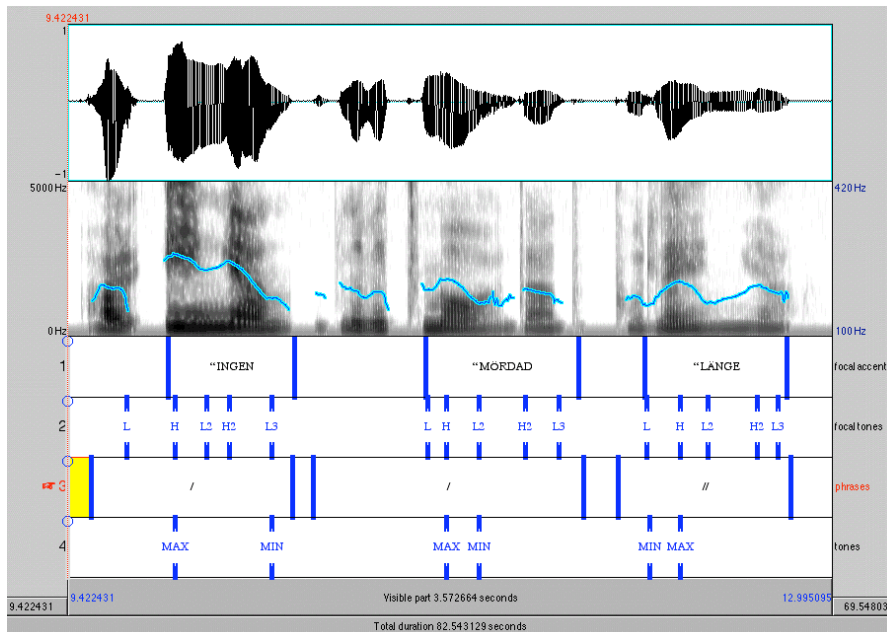
Som underlag för vidare tonala undersökningar transkriberas materialet. En variant av prosodisk bastranskription för svenska används (Bruce, 1998). Den ortografiska fonologiska transkriptionen betecknar diskursprosodi beträffande två prosodiska drag; gruppering och prominens. I transkriptionen har symboler för fyra grupperingsnivåer och fyra prominensnivåer införts. Grupperingarna är i nedstigande storleksordning; talparagraf [///], prosodiskt yttrande [//] och prosodisk fras [/]. Ingen gränssymbol betecknar en frasintern relation. Grupperingsgränserna har urskiljts efter perceptuell pausduration. Prominensnivåerna är i sin tur fokal accent [ˈ], accent [ˑ] och betoning [ˌ], där obetonat ord inte har någon beteckning. De fokala orden urskiljs genom tonala gester. Fokala accenter är de ord som, enligt talarens centralsvenska dialekttyp, har en extra tonal gest lokalt. Denna oberoende av accenttyp i form av en extra tontopp eller i form av höjd topp vid enstaviga ord. Distinktionen vid enstaviga ord mellan fokuserat eller ickefokuserat accentuerat ord har gjorts hos tonomfånget. Vid ensam tonal gest utan stort tonomfång ses orden som endast accentuerade. Övriga prominenta ord ses som betonade. Fokala ord skrivs med versaler i transkriptionen (se bilaga).

å ˈKURT / e ˈGLADARE / än ˈnänsin ˈnu för ˈtiden //

å sätter ˈin ˈFEMTI ˈtusen ˈKRONOR / på ett ˈhemligt ˈBANK,KONTO / ˈvarje ˈmånad ///

Akustisk analys

Det transkriberade materialet (sex talparagrafer, ca 1/7 av ursprungliga materialet) förbereds för mätningar i Praat. Modellen för datanalen är utförd med hjälp av Johan Frid (SOL, Lunds Universitet). Mätdata från olika parametrar kan insamlas samtidigt genom "rutsystem" (*textgrids*), t ex till en lista för vidare bearbetning i Excel. Rutsystemet införs genom annotering av utvalda parametrar. Annotering utförs manuellt för att hitta korrekt fonetisk parameter, bl. a. F₀ fri från mikroprosodisk påverkan. Rutsystemet utformas till att bestå av fyra skikt (*tiers*). Två av dessa är intervallskikt (*interval tiers*) som markerar grupperingar, en global och en lokal. (figur 1) Rutsystemets övriga två skikt är punkter (*point tiers*) som mäter toner vid exakt ögonblick, en för globala toner och en för lokala toner.



Figur 1. Praat-fönster som visar ett spektrogram med fyra *textgrids*. Två *intervall tiers* och två *point tiers* är annoterade. Talsignalen är från slutet av ett prosodiskt yttrande i materialet: <har ingen blivit mördad på länge>. Intervallskiktet *phrases* visar här tre intonationsfraser (/) inom ett yttrande (//). Initial låg ton har hamnat utanför det fokuserade ordet <ingen>.

Det ena intervallskiktet (skikt 1, *focal accents*) utgörs av ord som är extra framhävda genom fokal accent; transkriptionens högsta prominensnivå. Det andra intervallskiktet (skikt 3, *phrases*) utgörs av frasindelningar, d v s transkriptionens tre högsta grupperingsnivåer: talparagraf, yttrande och fras. De återstående två skikten i rutsystemet annoteras genom punkter. Det ena skiktet betecknar låga och höga toner hos ord med fokal accent; L, H, L2, H2, L3 (skikt 2, *focal tones*). H2 urskiljs här som fokal tontopp (hos enstaviga ord urskiljs emellertid H som fokal tontopp). Det andra skiktet för annoterade punkter (skikt 4, *tones*) visar en global tonal prominensnivå. Den markerar grundtonens max- och minvärde hos varje fras.

Ett *script* i Praat möjliggör automatiska mätningar, vilken samlar in värden i Hertz för vidare bearbetning i Excel. Värdet hos en valfri prominensnivå inom en valfri grupperingsnivå kan registreras. Den initiala analysen uppfattade prominent global tonal variation. Eftersom diskurs och talsituation är knutna till globala prosodiska parametrar, bestäms det att mätningarna ska koncentreras på globala tendenser. Mätningar görs därför på F₀-max och F₀-min i prosodiska fraser. Den tonala variationen realiseras hos lokala tonala utspel. Eftersom emfas realiseras hos den fokala accenten görs också mätningar på den fokala tontoppen H2 eller H, d v s den fokala tonala gesten, inom prosodiska fraser. Detta ger mätresultat för globalt tonomfång och globala fokala tonomfång.

Resultat

Mätningarna av materialet i Praat visar att talaren läser poesin med stor tonal variation. Både lokalt hos fokala accenter och globalt inom successiva fraser varierar tonomfånget stort. Emfas hos de fokuserade orden samspelar med det globala tonomfånget. Vid lite emfas hos den fokala accenten förekommer lågt tonomfång i resten av frasen, vid stor emfas hos fokal accent har resten av frasen även stort tonomfång. Det förekommer också fraser med stort globalt tonomfång, det vill säga emfas, även vid frånvaro av fokal accent. Transkriptionen av den fonologiska frasindelningen visar att uppläsningen består av korta fraser. Hos dessa fraser förekommer flera fokala accenter tätt ihop medan andra fraser helt saknar fokala accenter. Nedan följer ingående resultat av undersökningens olika delar.

Initial analys

Den initiala auditiva och delvis akustiska analysen gav åtta prominenta fonetiska kategorier. Dessa ansågs som framhävda på ett eller annat vis i materialet. Det sätt som poeten framhäver *dialekt, röstgolv, röstregister, röstkvalitet, talhastighet, frasering, intensitet* och *intonation* antas vara unikt för hennes uppläsning av dikten. Talarens *röstgolv* ligger vid 160 Hz och hennes *röstregister* sträcker sig normalt mellan 160 och 300 Hz (ca en oktav). Vid ett antal tillfällen når dock tontopparna över 400 Hz. *Röstkvaliteten* är ofta knarrig vid frasslut. Talaren tycks även ha en benägenhet för preaspiration och vokaler avslutas därför ibland tonlöst. *Talhastigheten* är låg. Talaren läser långsamt, ca tre stavelser/sekund. Intrycket av *intensiteten* är att Kristina Lugn talar med relativt jämn styrka. Initialt hos talparagrafen tycks ljudnivån dock vara något högre. *Fraseringen* är tydligt markerad med pauseringar. Talarens *dialekt* är av en centralsvensk dialekttyp med en tidig timing av tontopp i ordet och en tvåtoppig lokal intonationskurva hos fokuserade ord. Den extra tonala gesten i denna dialektvariant är tydligt framträdande i materialet. Den fonetiska kategori som upplevs som mest prominent i materialet är *intonationen*, närmare bestämt tonomfånget, då oväntat stort tonalt utspel varvas med partier med närmast minimalt tonalt utspel.

Fonologisk transkription

Accentuering och frasering undersöktes ytterligare genom den fonologiska och ortografiska transkriptionen. Transkriptionen visar att materialet består av korta fraser. Storleken på yttrande varierar från en fras till ca tio stycken. Talparagrafers storlek varierar mellan ett yttrande och ca tio stycken (se bilaga).

Frasgräns förekommer ibland direkt efter fokal accent och en fokal accent kan också ensam utgöra en fras.

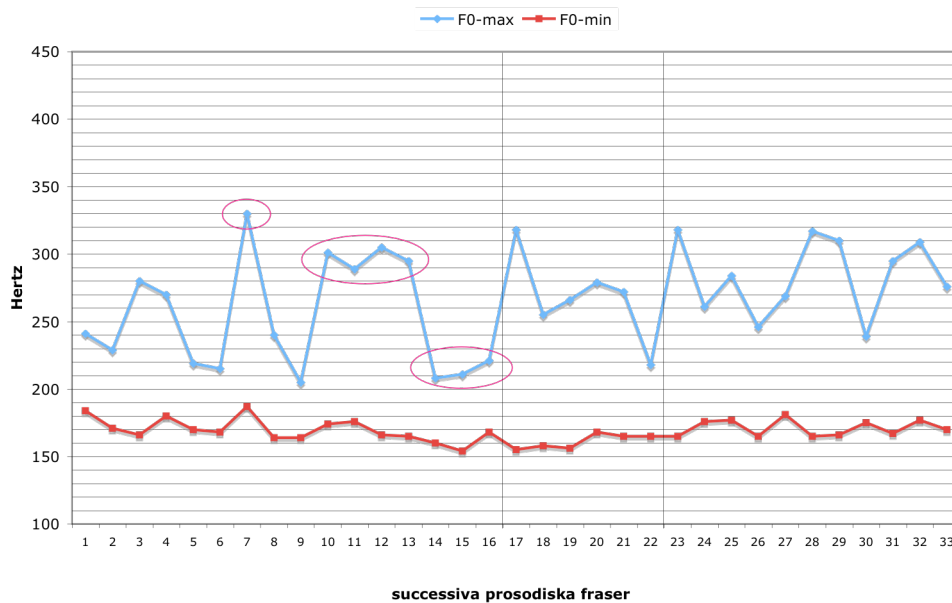
```
o ca'millas "PAPPA / "AVLED 'nyligen / till 'följd av ma'ligna TU"MÖRER / i "STRUPEN / o  
"TJOCK,TARMEN / o "KURT "ÖNSKAR / att CA"MILLA / ville 'tvätta 'fötterna / lite 'oftare //
```

Inbördes avstånd mellan fokala accenter varierar. De fokala accenterna kan stå flera i tät följd efter varandra, ibland flera inom en fras. Men de kan också förekomma långt ifrån varandra, till följd att flera fraser i följd är utan fokala accenter.

Akustisk analys

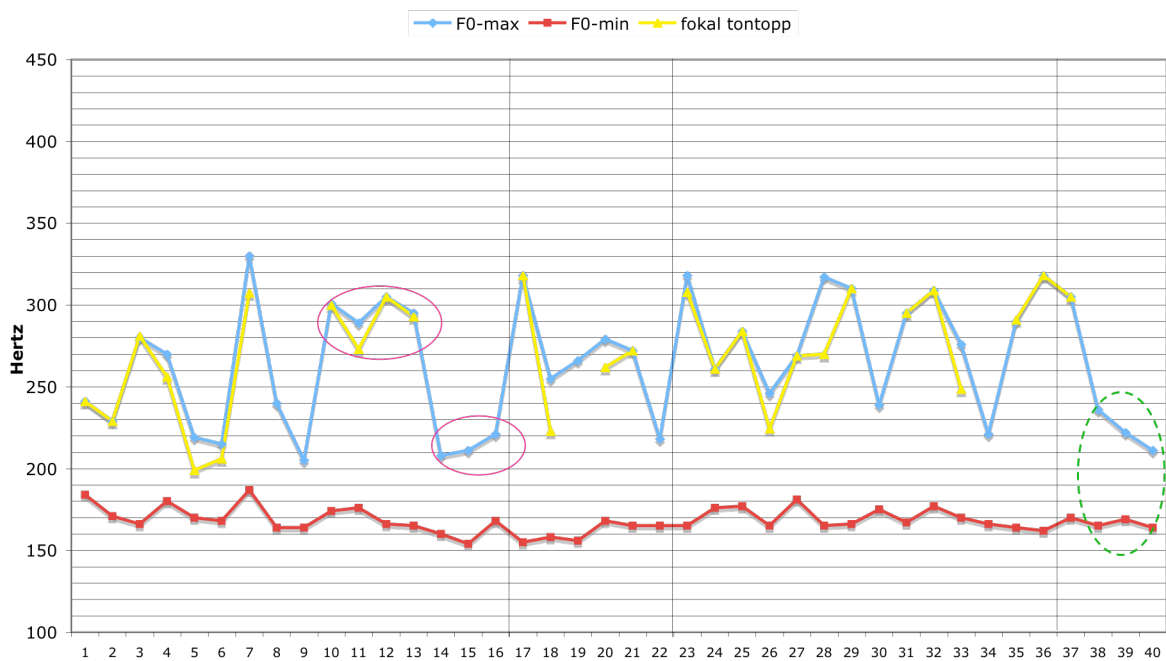
Den akustiska analysen i Praat ger globala mätvärden. De successiva frasernas globala F0-värde gestaltas i linjediagram. Detta diagrams x-axel betecknar successiva prosodiska fraser och dess y-axel värdet i Hertz hos F0-max (blå kurva) respektive F0-min (röd kurva). Figur 2 visar t ex högt F0-max hos sjunde frasen (rosa markering). Flera tonala utspel hos fraser i tät

följd kan man senare se hos höga F0-max i fraserna 10-13. Efterföljande fraser (14-16) är istället ett exempel på flera fraser i följd utan tonala utspel.

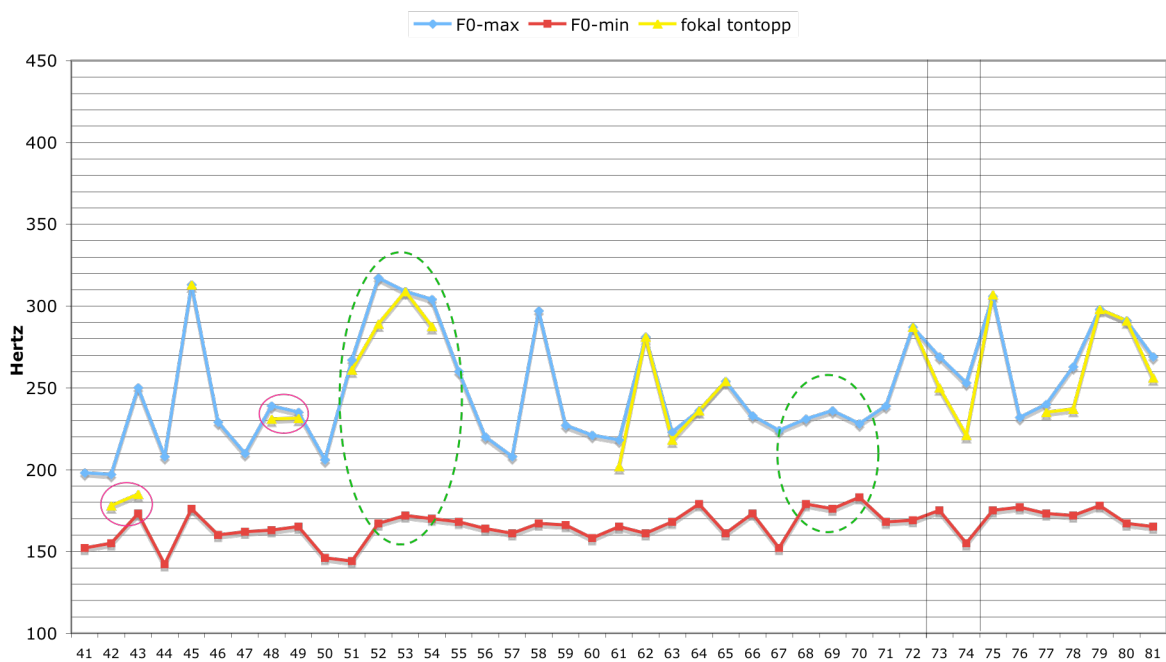


Figur 2. Grundtonens max- och minvärde i Hertz. Värde på x-axel visar successiva prosodiska fraser, lodräta streck visar talparagrafgränser. Rosa markering visar globala tonala utspel.

Akustiska mätningar på de tonala utspelen gjordes vid ytterligare successiva fraser vilket representeras av ett nytt linjediagram. (figur 3 och 4) Ytterligare en kurva har här lagts till, värde i Hertz hos fokala accenters tontoppar, d v s den extra tonala gesten i det fokuserade ordet. I figur 3 och 4 kan man se när värdena hos den fokala accenten och det tonala utspelet sammanfaller. Frånvaro av tonalt utspel i materialet sker oftast i fraser utan några fokala accent. Se i figur 3 och 4 åter på tonala utspel i fras 10-13 och jämför dem med frånvaro av fokala accent i fras 14-16 (rosa markering). Mätresultatet visar dock att även fraser utan tonalt utspel kan ha fokala accent, t e x fras 42-43 och 48-49, men dessa har då lågt tonomfång (rosa markering).



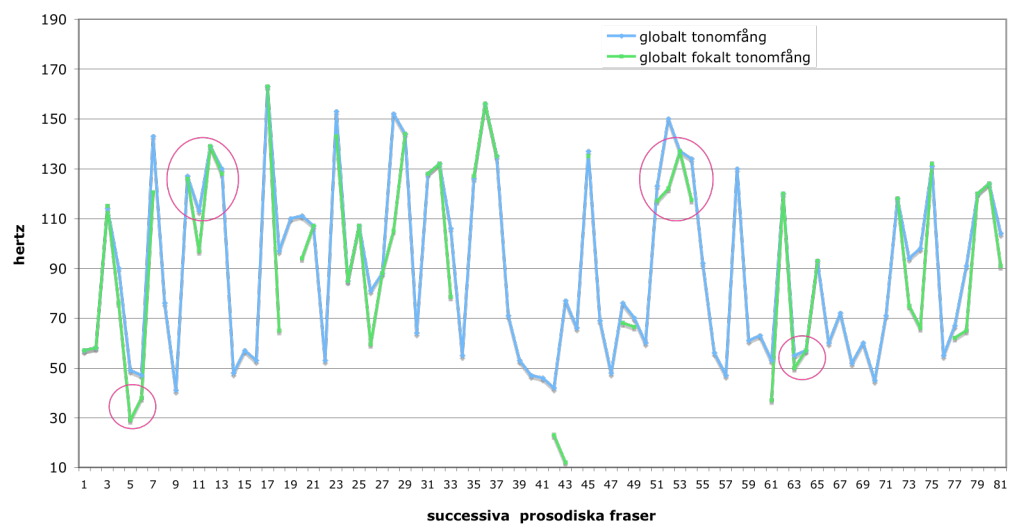
successiva prosodiska fraser



successiva prosodiska fraser

Figur 3 och 4. Successiva prosodiska fraser i hela det undersökta materialet. Grundtonens max- och minvärde i Hertz, samt grundtonens värde hos fokala accenters fokala tontopp. Värde på x-axel visar successiva prosodiska fraser, lodräta streck visar talparagrafgränser. Rosa markering visar globala tonala utspel, grön streckad markering visar globalt tonomfång.

Den röda F0-min-kurvan ger en bild av hur röstgolvet varierar successivt. Kombinerat med F0-max-värdet kan det globala tonomfånget inom en fras skönjas i linjediagrammet (grön prickad markering). Flera fraser i följd med lågt globalt tonomfång varvas med passager med stort globalt tonomfång. En bättre bild av variation hos det globala tonomfånget presenteras i det linjära linjediagrammet i figur 4.



Figur 5. Det globala tonomfångets storlek i Hertz hos prosodiska fraser (blå kurva). Förekomst av globalt fokalt tonomfång markeras med grön kurva. Rosa markering visar stora och små globala fokala omfång.

Det finns förutom en variation i storlek hos de globala tonomfången, också en variation i storlek hos de globala fokala tonomfången (rosa markering). Figur 5 tydliggör växlingen mellan stora fokala tonomfång och små fokala tonomfång i successiva fraser.

Diskussion

I introduktionen ställdes frågor som var vägledande för undersökningen:

- * Hur kan en fonetisk analys användas för att beskriva uppläst fri vers hos en talare?
- * Kan man hitta globala och lokala prosodiska parametrar som kan antas vara poetiska effekter?
- * Kan uppläst fri vers utgöra en egen talstil?
- * På vilket sätt kan en fonetisk analys bidra till att utveckla syntetisk simulering av uppläsning av fri vers?

Diskussioner förs under respektive fråga.

- * Hur kan en fonetisk analys användas för att beskriva uppläst fri vers hos en talare?

Resultaten av den initiala auditiva och delvis akustiska analysen visar att det är fullt möjligt att fonetiskt beskriva en poets individuella uppläsning. De mest prominenta dragen i Kristina Lugns tal gällde talhastighet och intonation, därmed bör de lyftas fram i en beskrivning av hennes poesiuppläsning. För den individuella diktens karaktär behövs dock en närmare beskrivning. En fonetisk analys i kombination med en textanalys, samt en fonologisk transkription i den form som här beskrivits, borde ge en god beskrivning av även den individuella diktens särdrag. Enbart en fonetisk analys är därför inte tillräcklig för att karaktärisera en uppläsning av poesi.

Kristina Lugn läser långsamt och det är enligt konventionen för fri vers (Lilja, 2003). Hon gör konsekventa pauseringar under uppläsningen, förmodligen styrda av interpunktionen i textförlagan. Eftersom uppläsningen sker med låg talhastighet förekommer det troligen fler pauser än vad det skulle göra hos en motsvarande uppläsning med hög talhastighet. Detta kan man se hos de korta fraserna som uppmärksammades i den fonologiska transkriptionen. När det gäller intonationen, så visar resultaten att talaren läser flera fraser utan att fokusera orden tonalt, medan fokuserade ord i andra fraser står tätt tillsammans. Vissa frasers fokala accenter uttrycks med oväntat stort emfas. Att poeten skulle välja att inte alls fokusera ord i vissa fraser med lågt tonomfång stämmer inte eftersom det finns fokala accenter med också litet emfas. Därför kan man anta att den fonologiska funktionen av emfas, dvs prominens, bibehålls i möjligaste mån av poeten, medan den fonetiska eller paralingvistiska funktionen av emfas systematiskt återhålls och framhävs av poeten. Man kan även tänka sig att en variation mellan stora tonala utspel och små tonala utspel - stor variation i storlek hos emfas mellan fraser - uppfattas som rytmisk av lyssnaren.

- * Kan man hitta globala och lokala prosodiska parametrar som kan antas vara poetiska effekter?

Ja och nej. Det förekommer globala tonala variationer, men man kan inte säkert säga om de är poetiska effekter. Den stora variationen i globalt tonomfång mellan fraser och de tonala utspelen funna i vissa fraser avviker emellertid från normen för formell läst talstil och liknar mer den spontana talarens fonetiska variation. Enligt Eva Lilja (2003) skulle konventionen för en uppläst fri vers, förutom den låga talhastigheten, också innebära poetiska effekter i form av stor prosodisk variation hos någon parameter. Troligen kan den globala tonala, och eventuellt fokala, variationen i denna undersöknings resultat tolkas som en poetisk effekt. Om dessa poetiska effekter beror på den publika talsituationen eller den specifika talarens och diktens karaktär står inte helt klart. De tonala utspel som hittades i materialet kan heller inte säkert säga avvika från uppläst fri vers hos andra talare.

En textanalys kan visa att vissa ord som fokuseras är oväntade och att repetition av fraser verkställs fonetiskt och prosodiskt varierande. En poetisk effekt kan även därmed vara att välja att inte fokusera när det är väntat, t ex vid innehållsord. En poetisk effekt kan också vara

att nästan alla ord i frasen får fokal accent, vilket transkriptionen visade att Kristina Lugn gör i sin uppläsning. Den fonologiska transkriptionen visade att vissa fraser i materialet saknade fokala accenter medan andra fraser hade fler än en fokal accent. En prosodisk bastranskription för svenska ger EN fokal accent per fras. På så vis avviker Kristina Lugns tal också här från normen. Frånvaro av fokala accenter i vissa fraser behöver dock inte betyda att orden inte fokuseras, bara att tonalt fokus inte är den mest prominenta parametern just där. Detsamma gäller när flera fokala accenter förekommer i en fras. Forskning visar att det är möjligt att ord kan vara ofokuserade men ändå bära på en tonal gest (Heldner, 2001). Närvaro och frånvaro av fokala accenter sker emellertid med en sådan systematik att man kan anta att talaren gör det medvetet. Även systematisk variation hos det globala tonomfånget tyder på medvetna effekter i talet.

* Kan uppläst fri vers utgöra en egen talstil?

Tidigare forskning har pekat på global prosodisk variation som en grundläggande faktor för perceptionen av en talstil. Den globala parametern är troligen knuten till talsituationen, vilket förmodas även av Hirschberg (2002). Det kan finnas en global prosodisk variation hos uppläst fri vers. Slutsatser kan dock endast dras om den specifika talaren, det går inte att säga om denna undersöknings resultat gäller uppläst fri vers hos andra talare. Det finns konventioner, men konventionens relation till poetiska effekter står inte helt klart. Stora tonala variationer är emellertid inte typiskt för läst tal (Bruce, 1998). Resultaten visar att tonala variationer även går att finna i läst tal, närmare bestämt uppläst fri vers.

* På vilket sätt kan en fonetisk analys bidra till att utveckla syntetisk simulering av uppläsning av fri vers?

I prosodisk bastranskription för svenska förekommer högre prominensnivå [ˈ] endast en gång inom en fras (Bruce, 1998). Frasindelningen i denna undersöknings transkription har skett efter upplevd prominent pauslängd, och inte efter gränstoner. Den fokala tonala gester har å andra sidan varit riktgivande vid beteckning av prominensnivåer. Detta har lett till att fokala tonala gester visat sig förekomma flera gånger inom en fras. I denna undersökning har därför flera fokala accenter påträffats inom en fras.

Efter en genomförd fonologisk transkription av ett material kan man anta att man har den som underlag för en syntetisk modell. Det tillvägagångssättet skulle i detta fall ge fel utslag hos de fokala accenterna. Förutom att transkriptionen skulle erkänna alltför många fokala accenter per fras, kan transkriptionen dessutom bara visa distinktionen mellan närvaro och frånvaro av fokala accenter. Den visar inte hur stort fokalt omfång och inte heller hur mycket det tonala utspelet varierar. Tidigare syntetiska modeller har antagit samma prominenta status, dvs samma tonomfång, för alla fokala accenter i ett yttrande (Dutoit, 1997, Bruce, 1998). Men här uppvisas alltså stor emfatisk variation hos fokala accenter i yttranden. Det tycks finnas en distinktion mellan stort och litet fokalt tonomfång i materialet. Dessutom skiljer sig fraser åt i globalt tonomfång även där fraserna är utan fokala accenter.

Undersökningen visar inte vilken fonologisk funktion de tonala utspelen har inom frasen, dvs om det gäller fokus eller gränston. Därför vet man inte heller om den fonologiska funktionen är oväntad. Däremot indikerade den initiala analysen att det tonala omfånget ibland upplevdes som oväntat stort, särskilt det fokala omfånget.

Både hos Kristina Lugn och den syntetiska rösten i *Poet shop* förmodades talhastighet och intonation vara viktiga för poesins effekter under talets gång. Resultaten i denna undersökning visar att det finns en global variation hos den prosodiska parametern intonation, till synes systematisk. Istället för att manuellt frambringas som i *Poet shop*, skulle denna variation därför kunna införlivas i en prosodisk modell av intonation, vilket används till syntetisk simulering av en röst.

Detta är en deskriptiv uppsats inom ett ämne som ännu inte är etablerat. En möjlig benämning skulle kunna vara *fonetisk poetik* när det gäller beskrivningen av de individuella poeterna och deras verk. Resultaten här är inte tillräckliga för att dra slutsatser om en existerande talstil specifikt för poesi, till exempel ”poetiskt tal” eller ”uppläst poesi”.

Det skulle förstås vara intressant att skapa en syntetisk röst av en uppläsare av poesi, förslagsvis Kristina Lugn. Emellertid behövs det ses närmare på lokala parametrar än vad denna uppsats haft möjlighet till. Det tycks förekomma globala emfatiska tendenser eftersom inte bara den fokala accenten bär på emfas. En vidare analys av lokala variationer hos fokala accenter i relation till globala tendenser, är önskvärt. Forskning om prominensmönster skulle kunna klargöra kring förekomst av globala tonala variationer. Kristina Lugns tonala utspel i vissa fraser är onekligen prominenta i materialet. Men vilken prosodisk parameter är prominent i de fraser där det tonala utspelet uteblir?

Man behöver också titta närmare på hur det kommer sig att vissa fokuseringar kan få stor emfas. Det måste ju möjliggöras genom andra fonetiska parametrar. En hög tonkurva behöver t e x beredas plats både före och efter tontoppen; en viss tidsintervall för stigning och sänkning av F0. Hur kan flera fokala accenter ”få plats” i en fras? I detta avseende är det därför önskvärt att se närmare på fraseringen i Kristina Lugns tal.

Vad som skulle kunna utgöra en poetisk talstil behöver tittas på mer, förslagsvis genom att jämföra flera talare i samma situation. Ett framtida projekt, *Public speech*, vid Umeå universitet kommer att undersöka hur olika talare gör för att fånga en publik (Strangert och Deschamps, 2006). Är samma parametrar prominenta hos alla talare inför publik, även för poesiuppläsaren? Och vad har ”poetiska effekter” för funktion i talarens relation till lyssnaren? Vad som upplevs som poetiska effekter av lyssnarna skulle kunna klargöras genom en perceptorisk undersökning.

Semantiska förhållanden mellan fraser är också intressant; hur kan Kristina Lugns fraseringar och fokuseringar uppfattas som oväntade av lyssnaren? Kan det kanske finnas någon semantisk och även syntaktisk avvikelse i Kristina Lugns poesi? Om så är fallet, finns det då anledning att anta även prosodisk avvikelse i hennes uppläsning? Kanske är det sådana avvikelser som ligger till grund för det vi lyssnare upplever som ”poetiska effekter”.

Källförteckning

- Beckman, Åsa (2002) *Jag själv ett hus av ljus: 10 kvinnliga poeter*. Stockholm: Bonnier
- Blankinship, E. & Beckwith, R. *Tools for expressive text-to-speech markup*.
<http://www.media.mit.edu/~erikb/papers/uist01-tools.pdf>.
- Bruce, Gösta (1998) *Allmän svensk prosodi*. Lund: Institutionen för Lingvistik
- Bänzinger, Tanja och Klaus R. Scherer (2005) "The role of intonation in emotional expressions" s. 252-267. I *Speech Communication* 46
- Dutoit, Thierry (1997) *An introduction to Text-to-Speech Syntheses*. Dordrecht: Kluwer Academic Publishers
- Elleström, Lars (1990) "Tryggare kan ingen vara. Om Kristina Lugn" s.82-106. I Elleström, Lars och Cecilia Hansson, *Samtida: essäer om svenska författarskap*. Stockholm: Alba
- Fant, Gunnar, Anita Kruckenberg, Kjell Gustafsson och Johan Liljencrants (2002) "A new approach to intonation analysis and synthesis of Swedish" s. 161-164. I *vol. 44 Fonetik 2002*
- Fant, Gunnar och Anita Kruckenberg (1989) "Preliminaries to the study of Swedish prose reading and reading style" s. 1-80. I *STL-QPSR* 30(2)
- Gussenhoven, Carlos (2004) *The Phonology of Tone and Intonation*. Cambridge: University Press
- Heldner, Mattias (2001) *Focal accent – f₀ movements and beyond*. Umeå: PHONUM 8
- Hirschberg, Julia (2000) "A corpus-based approach to the study of speaking style" s. 335-350. I Horne, Merle (ed.) *Prosody: theory and experiment. Studies presented to Gösta Bruce*. Dordrecht: Kluwer Academic Publishers
- Kehrein, Roland (2002) "The Prosody of Authentic Emotions" s. 423-426. I *Speech Prosody 2002, Proceedings of the 1st International Conference on Speech Prosody*.
- Lilja, Eva (2003) "Åtta punkter till en förnyelse av versläran. I dialog mellan historia och teori, muntligt och skriftligt, form och betydelse". I Bäckman, Sven, Mattias Hansson och Eva Lilja. *Rytm och dialog. Studier framlagda vid Åttonde nordiska metrikkonferensen Umeå 4-7 oktober 2001*. Göteborg: Centrum för Metriska Studier 14
- Lugn, Kristina (1988) *Om ni hör ett skott...* Stockholm: Bokbandet
- Lugn, Kristina (1979) *Om ni hör ett skott...* Stockholm: Förlag Alba
- Schötz, Susanne (2006) "Data-driven formant synthesis of speaker age" s. 105-108. I *Working papers* 52. 2006. Lund: Centre for Languages and Literature, Dept. Of Linguistics and Phonetics.
- Svensson, Malin (2005) *Kristina Lugn och den prosodiska emotionella kategorin sorg*. B-poster i Prosodi vid Lunds Universitet
- Strangert, Eva och Thierry Deschamps (2006) "The prosody of public speech – A description of a project" s. 121-124. I *Working papers* 52. 2006. Lund: Centre for Languages and Literature, Dept. Of Linguistics and Phonetics.
- Övriga källor: <http://mary.dfki.de/online-demos/emospeak>

Bilaga

Ortografisk Transkription Kristina Lugn Om ni hör ett skott... 1988

” fokal accent (för tydlighetens skull markeras fokuserade ordet dessutom med versaler)
, accent
, betonad

/// talparagraf
// yttrande
/ intonationsfras

”O / ”NEJ //

i ”DENNA ’borgerligt ’styrda ’kommun / har ”INGEN / blivit ”MÖRDAD / på ”LÄNGE //
”INTE ’ens CA”MILLA / har blivit ’knivskuren / på ’hemväg från ’centrum //
hennes ’yngste ”SON / har ”FORT,FARANDE ”INTE blivit ’kvävd / av en ”PLAST,PÅSE //
som en ”OFATT,BART ’elak ’människa / ’trätt över ’hans ’huvud / och ’knutit ’hårt / om
’hans ’spänsliga ’hals ///

å ”KURT / e ”GLADARE / än ’nånsin ’nu för ’tiden //

å sätter ’in ”FEMTI ’tusen ”KRONOR / på ett ’hemligt ”BANK,KONTO / ’varje ’månad ///

när ”KURT ”GIFTE ’sig / med CA”MILLA / för ’tio ”ÅR sedan //

var ca’milla ”UPP över ”ÖRONEN / FÖ”RÄLSKAD / i ”SVEN ’bertil ”TAUBE //

för ”TIO ’år sedan / stod ca’milla / ”JUST i be’grepp / att säga ”UPP sig / från sitt

”ANSVAR,S,FULLA ’arbete som ”HEMSAMA,RIT //

och skriva ’under / ett ’mycket ’fördel,aktigt ”KONTRAKT / med bibi ”ANDERS,SON ///

”EN ’ljummen ’försommar,eftermiddag / för ’tio ’år sedan / par’kerade ’kurt sin ’svarta

’lång,tradare / framför ca’millas fö’räldrars ’välpla,nerade / ’röda /

”SEXVÅNINGSTRÄ,VILLA //

’han par’kerade ’sin ’snygga ’svarta ”LÅNG,TRADARE / framför den ’handsmida

’järn,grinden //

’kurt ’visslade på en ”BARN,VISA / när ’han ’öppnade ’grinden / till ca’millas fö’räldrars
impo’nerande plan’tering //

han ”SPARKADE ”INTE / på en ”JÄTTE,GULLIG ”KATT,UNGE / som ’lekte i ’gräset //

CA”MILLA ’var ’nitton ’år //

hon ’kunde ’helt ’enkelt ’inte låta ”BLI / att ’stjåla dom gamla ’damernas

FOTOGRA”FI,ALBUM //

Bilaga

'hon 'kunde "HELT 'enkelt 'inte låta "BLI / att 'stjäla / dom gamla 'damernas /
för'lovnings,ringar //

hon 'kunde 'helt 'enkelt 'inte låta 'bli / att 'stjäla dom gamla 'damernas 'såp,bubblor / o
'polkagris,klubbor //

ca'milla var 'nästan 'alltid "HUNGRIG //

hon hade 'fått "TRE 'dikter / REFU"SERADE / av 'hemmets "JOURNAL //

o hon tyckte "OM / att 'vandra i 'skogen / i 'ensam,heten / i 'kvällningen / under 'himlen /
under 'månen / under 'kråkornas 'vingar / o 'tänka på 'sin "MAMMA ///

o "KURT har varit "GIFT med ca'milla / i "TIO 'år nu ///

"KURT 'känner sin ca'milla 'lilla //

'han 'känner 'henne 'utan o 'innan ///

'kurt är 'mycket "SKICKLIG //

han e 'mycket "FLITIG //

han e 'mycket "VÄLUT,BILDAD //

o "JAG vill 'lova / att "KURT e "HYGGLIG mot ca'milla //

o "JAG vill 'lova / att ca'millas "MAMMA / tycker att det är "TRÅKIGT / att ca'milla 'saknar
SEXAP"PEAL //

o ca'millas "PAPPA / "AVLED 'nyligen / till 'följd av ma'ligna TU"MÖRER / i "STRUPEN / o
"TJOCK,TARMEN / o "KURT "ÖNSKAR / att CA"MILLA / ville 'tvätta 'fötterna / lite 'oftare
//

"KURT o ca'milla / 'sitter 'mitt emot 'varandra / vid 'köksbordet / o 'äter "PRINS,KORV //

"DE e 'inte "FARLIGT //

det 'hänger en CARL"LARSSONREPRODUK,TION / bakom 'kurts 'huvud //

CA"MILLA 'har ett "SÅR /