

STM 1979:2

Om kongruens mellan text och musik i sång

Av Claes Witting

© Denna text får ej mångfaldigas eller ytterligare publiceras utan tillstånd från författaren.

Upphovsrätten till de enskilda artiklarna ägs av resp. författare och Svenska samfundet för musikforskning. Enligt svensk lagstiftning är alla slags citat tillåtna inom ramen för en vetenskaplig eller kritisk framställning utan att upphovsrättsinnehavaren behöver tillfrågas. Det är också tillåtet att göra en kopia av enskilda artiklar för personligt bruk. Däremot är det inte tillåtet att kopiera hela databasen.

Om kongruens mellan text och musik i sång

Av Claes Witting

I STM för år 1970 (s. 71 ff.) står att läsa en intressant artikel, *Towards a generative theory of melody*, författad av Björn Lindblom och Johan Sundberg (L & S). Som titeln antyder rör det sig om ett försök att till musikvetenskapen överföra en del begrepp och procedurer från generativ transformationsgrammatik (TG), däribland en viss regelmekanism, som utförligt beskrivits i en av TG:s mest kända urkunder, *The sound pattern of English* (New York, 1968) av Noam Chomsky och Morris Halle (Ch & H).

TG:s teoribyggnad, väsentligen ett verk av Chomsky, bjuder på många nya fruktbara tankar om människans språk. Man kan exempelvis peka på sådana teoretiskt centrala begrepp i densamma som djup- och ytstruktur, transformation och frasstruktur.¹ Annat i TG är förvisso lika nytt men inte fullt så användbart. Till denna senare kategori måste man tyvärr räkna nyssnämnda regelmekanism, vilken antagits vara tillämplig på fördelningen av "stress", dvs. ord- och satsbetoning (närmast i engelskan).

Några av bristerna hos ifrågavarande stressregler kommer att påpekas i denna artikels tidigare del. Genom att acceptera Ch & H:s idéer har L & S enligt mitt förmenande kommit att ge en felaktig bild av betoningens natur, och jag presenterar därför vad jag anser vara en mera verklighetstrogen motbild. Slutligen lämnas ett preliminärt förslag till metod för beräkning av samgången (kongruensen) mellan text och musik i sång.

Men först till en kortfattad beskrivning av vad frasstruktur och stressfördelning enligt TG innebär. Som demonstrationsexempel tjänar den barnvisa av Alice Tegnér som L & S utnyttjat men vars text av dem ej getts den behövliga satslösningen och därmed avhängiga betoningsstrukturen enligt TG. Frasstrukturen i meningen *Världen är så stor, så stor, Lasse, Lasse liten, större än du nånsin tror, Lasse, Lasse liten* återges sålunda i fig. 1. Förkortningarna är: S = mening, sats (eng. sentence), NP = nominalfras, VP = verbalfras, Cop = kopula, AdjP = adjektivfras, Adv = adverb, Subj = subjunktion (underordnande konjunktion), SubjP = subjunktionsfras (här = komparativ bisats), Vok = vokativ ("tillropsform", här betraktad som syntaktiskt jämställd med och relaterad till *du*), N = nomen, Attr = attribut. Den syntaktiska strukturen framträder med all önskvärd tydlighet i ett trädidiagram som detta. Är det dessutom så noggrant ritat att det syntaktiska konstruktionsdjupet bekvämt kan

¹ För viktigare transformationsgrammatiska termer och deras innebörd kan hänvisas till Peter A. Sjögren, *Termer i allmän språkvetenskap. Ett systematiskt lexikon*. Stockholm, 1978.

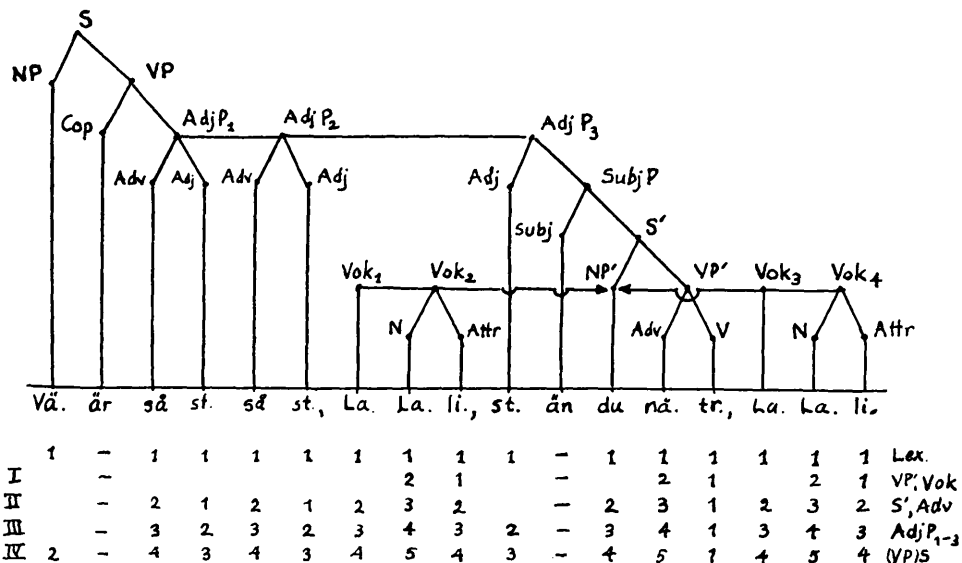


Fig. 1.

avläsas, måste trädigrammet anses vara ett mera praktiskt instrument för stressangivelse (enligt Ch & H:s modell) än etiketterade parenteser ("labeled bracketing"). Beträffande dessa senare se L & S.

En ortodox transformationsgrammatisk härledning av stressmönstret för texten "Lasse liten" lämnas i fig. 1, nedtill. Derivationen börjar med att de enskilda orden i meningen (med undantag för vanligen obetonade grammatiska funktionsord som *är* och *än*) alla tilldelas sin lexikaliska ordbetoning.² Denna utgörs av s.k. primary stress (primärbetoning) som betecknas med 1. Först därefter börjar tillämpninge av den märkliga satsbetoningsregel som Ch & H infört och kallat *Nuclear Stress Rule*. Regeln fungerar så att det "högraste" (Ch & H: rightmost) ledet inom ifrågavarande syntaktiska konstituent förlänas primary stress (dvs. behåller sin lexikaliska ursprungsbetoning) medan alla övriga led får sin betoning försvagad med en grad, så att primary stress blir secondary stress (markerat med 2), secondary stress blir tertiary (3) osv. Regeln tillämpas vid behov (som här) *cykliskt*, vilket innebär att den eller de djupast liggande konstituenterna i frasstrukturen bearbetas först, de övriga därefter i tur och ordning. I vårt aktuella fall kommer regeln alltså i första cykeln (I t.v. i fig.) att drabba andra och fjärde vokativen samt komparativsatsens nominalfras (vilket särskilt noterats längst t.h. i fig.). Så rullar regelmaskineriet vidare tills efter den fjärde transformationscykeln (IV) nominalfrasen i "matris-satsen" (S) vederbörligen åsatts sin secondary stress.

Sedan det språkliga stressmönstret för "Lasse liten" i överensstämmelse med TG:s principer blivit klarlagt, kan en mera ingående jämförelse göras mellan

² Enl. Ch & H skall kopula inte tilldelas någon lexikalisk betoning. Som ett vanligtvis obetonat funktionsord betraktas här även det satsinledande *än*. Vid emfas kan dock vilket ord som helst i en sats få betoning, varom mera nedan.

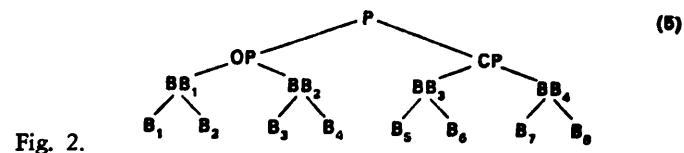


Fig. 2.



Fig. 3.

detta och det musikaliska stressmönster som L & S (s. 77 och 79) redovisar i två figurer, här återgivna som fig. 2 och 3.

Vissa likheter kan konstateras. Tydligt låter sig en sång av typ "Lasse liten" analyseras på det sätt L & S föreslår, dvs. med en första uppdelning av "perioden" (motsvarande S) i en "opening phrase" (motsv. NP) och en "closing phrase" (motsv. VP). Med likartade benämningar (labels) på övriga underordnade avsnitt fås en "frasstruktur", för vilken "frasstrukturregler" efter TG:s modell kan formuleras. Så långt är nog allt gott och väl.

Vid närmare betraktande framträder emellertid diskrepanser, som uppenbarligen har att skaffa med skillnaden mellan lingvistiskt och musikaliskt. Medan den språkliga frasstrukturanalysen naturligt går ned till de enskilda orden, stannar den musikaliska, enligt L & S, vid *takterna*. En fortsatt analys ner till de enskilda noterna och pausmomenten kunde ju annars te sig mera följdriktig. Vid en analys av sistnämnda slag skulle t.ex. den nod som L & S (se fig. 2) gett beteckningen B₁ bli försedd med fyra nedåtgående grenar, den därpå följande med tre osv. Varför har inte L & S löpt linan ut? Vilket deras motiv än må ha varit, *en* sak står fullt klar: en fortsatt tillämpning av Nuclear Stress Rule ner till de enskilda noterna skulle, som framgår av fig. 3, medföra att periodens slutgiltiga primärtryck kunde komma att falla på ett pausmoment (= fjärde slaget i takten) eller på den föregående fjärdedelsnoten. Denna fatalitet har L & S skickligt undgått genom att inom takten tillämpa en annan av Ch & H:s betoningsregler, den s.k. *Compound Stress Rule*. Enligt denna regel, vilken som namnet anger avser sammansatta ords accentuering, skall primary stress falla på första (med Ch & H:s uttrycksätt: "leftmost") komponenten i den ifrågavarande enheten. Härledningen av stressmönstret för de båda första takterna i "Lasse liten" sker sålunda enligt L & S (s. 82) genom att först lägga primärbetoning på samtliga taktslag, vilket ger åtta ettor i följd med den använda notationen (jfr den "lexikaliska" betoningen i fig. 1), därpå tillämpa Compound Stress Rule (vilket

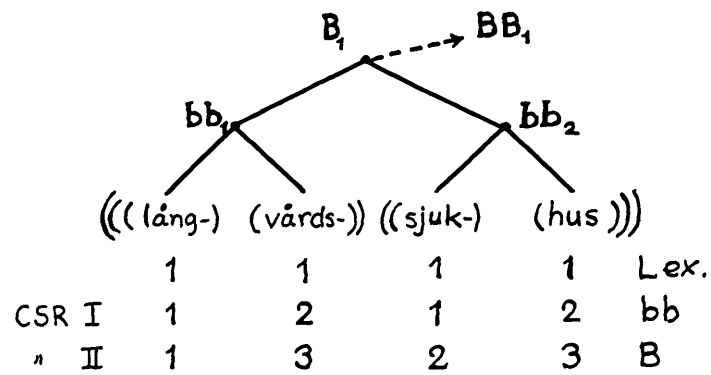


Fig. 4.

ger notationen 1 2 fyra gånger i följd), och till sist låta Nuclear Stress Rule sälla ut slutvärdena, som blir 23 23 23 13. Med hjälp av trädigrammet i fig. 2 kan vi alltså konstatera att Nuclear Stress Rule, dvs. satsbetoningsregeln, enl. L & S:s förslag, börjar arbeta på nivån med noderna BB_1 t.o.m. BB_4 . Musikens takter skulle därmed i princip svara mot språkets ord, vilket innebär att taktens underställda enheter, taktslagen, åtminstone grovt skulle komma att motsvara ordens *stavelser*. I och för sig skulle det synsättet kunna te sig fullt rimligt.

Emellertid, ser man något närmare på L & S:s förslag till stressgradering av takterna (i det aktuella fallet), finner man att varje taktslag motsvaras av ett ord. Detta ord måste antas ingå som led i ett fyrstavigt sammansatt ord av typen *långvårdssjukhus*, eftersom varje taktslag i utgångsläget tilldelats primärtryck. Den korrekta härledningen av stresskonturen för varje takt skall därför i enlighet med Ch & H:s principer utföras så att Compound Stress Rule i föreliggande exempelfall tillämpas två gånger, innan takten kan anses regelrätt genomgången. Proceduren finns åskådliggjord i fig. 4. För de båda närmaste konstituenterna under noden B_1 har jag infört beteckningen *bb* med indexsiffra. CSR = Compound Stress Rule. Regels tillämpning första gången (första cykeln, I) ger en stressgradfördelning som överensstämmer med den hos L & S. Sedan är blott att konstatera att L & S ej ansett sig böra applicera samma regel en andra gång. Som skäl härför anger de (s. 82) att slutresultatet av deras stressangivelse (för de båda första takterna) är förenligt med en av dem strax innan presenterad profil gällande fördelningen av noter med och utan ackord (i ifrågavarande takter) och att samma slutresultat "presupposes that there are no brackets for bars". Författarna menar vidare (ibid.) att deras analys står i överensstämmelse med iakttagelsen att en melodilinje normalt inte segmenteras i takter utan ofta rör sig kontinuerligt inom längre tidsintervall, t.ex. taktpar. Resonemanget verkar föga övertygande. Är inte taktmarkering och melodi klart skilda ting? Skulle det i sig ganska "mjuka" faktum att melodilinjerna ofta (men alltså inte alltid) sträcker sig över flera takter kunna motivera en så drastisk åtgärd som att definitivt utrangera takten som beståndsdel i den musikaliska frasstrukturen och därmed slopa dess egen parentes i "labeled bracketing"? Svaren på dessa frågor borde väl rimligen bli Jo, resp. Nej. För övrigt har detta frågekomplex ringa betydelse

för diskussionen i stort, då stressfördelningsmönstren enligt Ch & H:s teori saknar nödig förankring i fonetiska grundfakta — vilket jag närmast skall försöka bevisa.

Vår framställning har sålunda lett fram till spørsmålet: Vilken är den fonetiska verkligheten bakom en transformationsgrammatisk stresskontur som den som återges i fig. 1?

För att kunna besvara denna fråga måste man veta vad "stress" i fonetiskt sammanhang innebär. Tyvärr råder ingen allmän enighet om detta begrepp, och detta sakernas tillstånd beror inte enbart på att olika språk, ofta även olika dialekter, har olika system för sin stress, eller betoning som den bekväma svenska termen lyder. En fullständig redogörelse för divergerande åsikter skulle här föra för långt.³ Med en formulering, som torde ha stöd hos flertalet auktoriteter, kan dock sägas att betoning (stress) är den relativa framhävning (prominens) som kommer vissa stavelser till del i varje normalt stycke tal och att denna framhävning åstadkoms genom ett samspel mellan främst två fysiska variabler, grundtonsvariation (intonation) och ljudlängd (duration). En tredje variabel, intensiteten (ljudtrycket), är starkt korrelerad till intonationen och är också numera, liksom de båda förstnämnda, lätt tillgänglig för instrumentell akustisk analys. Emellertid är det mänskliga örat, som psykofysiska experiment tidigt ådagalagt, förvånansvärt okänsligt för intensitetsväxlingar.⁴ Den fysiska intensiteten kan därför lämnas ur räkningen vid behandlingen av den normala betoningens akustiska manifestation. Och detta trots att man i fonetik ofta talar om "tryckaccent" med därtill anslutna termer som "starktryck", "bitryck" osv. Jfr eng. "stress". Ursprungligen syftar dessa termer på de växlingar som trycket i utandningsluften, speciellt under glottis (det subglottala trycket), uppvisar vid växlande betoning. Eftersom betoningen från artikulatorisk synpunkt också innefattar variationer i stämbandets spänning och denna spänning dessutom kan ändras oberoende av det subglottala trycket, är det lätt att inse att en betoningsdefinition med referens till skeendet hos den talande måste bli ytterst komplicerad. Lyckligtvis kan en sådan definition här undvaras. Den omständigheten att Ch & H valt en artikulatorisk utgångspunkt för beskrivningen av sina särdrag ("features") skulle visserligen ha kunnat försvåra en bedömning av deras betoningsregler. Emellertid saknas i The sound pattern of English artikulatorisk definition av prosodiska drag, varför inget hinder föreligger att här tolka Ch & H:s stressbegrepp akustiskt-auditivt.

Betrakta nu stressgraderna för de enskilda orden i fig. 1 (nedersta raden). Det är tydligt att den syntaktiska strukturen inte behöver vara särskilt invecklad för att Nuclear Stress Rule skall generera ett stressmönster med upp till fem grader, utöver den obetonade stavelsens betoningsgrad. Tydligt är också att det betoningsbara ordet (här *nånsin*) närmast framför det ord som enligt regeln skall behålla

³ En översiktlig framställning av betoningsförhållandena, spec. ordaccenterna, i Skandinavien lämnas i Eva Gårdings *The Scandinavian word accents*. Lund, 1977. (Travaux de l'Institut de linguistique de Lund; 9).

⁴ Här avses i första rummet de övertygande experiment med "peak amplitude clipping" som redovisas i *Handbook of experimental psychology*, ed. S. S. Stevens, New York-London, 1951, s. 1040—1074.

primärtryck genom hela härledningsproceduren kommer att tilldelas en extremt svag betoning (dvs. hög gradsiffra). Denna tryckskillnad växer alltså i direkt proportion till konstituentens grad av s.k. inbäddning. Är nu detta fenomenologiskt rimligt? Svaret måste bli ett definitivt nej. Visserligen har det tidigare i fonetik lanserats betoningssystem med upp till tre och fyra grader (noll-tryck oräknat). Så har t.ex. det system som kommit till användning i Svenska akademiens ordbok en fyrgradig skala, och det amerikanska system som Ch & H närmast bygger på laborerar med tre grader, kallade just primary, secondary och tertiary stress.⁵ Det kan emellertid visas att det i båda dessa system rör sig om förvuxna teoretiska konstruktioner och att båda — utan förlust av någon relevant distinktion i den lingvistiska beskrivningen — kan reduceras till ett system med enbart två grader: betonad och obetonad, förutom en extra grad, starkt betonad, för emfas. Detta reducerade betoningssystem har stöd i utförda auditiva blindtest, som klarlagt att inte ens skolade lingvister kan säkert urskilja mer än två grader av betoning, när möjlighet till teoretiserande eftertanke uteslutits.⁶ Därmed torde under alla förhållanden stå klart att den *mångfald av stressgrader* som tillämpningen av Nuclear Stress Rule enligt Chomsky och Halle ger upphov till *saknar motsvarighet i den fonetiska empirin*.

Bristfälligheten hos Nuclear Stress Rule inskränker sig inte enbart till antalet stressgrader. I själva verket vilar regeln på ett oriktigt teoretiskt antagande, nämligen att betoningen skulle vara *syntaktiskt betingad* — en uppfattning som väl harmonierar med TG:s starka framhävnings av språklevets syntaktiska sida — ehuru det snarare är semantiken som behärskar betoningsförhållandena.

Att betoningen är *semantiskt* betingad innebär att det för stressplaceringen primärt avgörande momentet är vad den talande för tillfället uppfattar som viktigt ur betydelsynpunkt. Syntaxen spelar en helt underordnad roll. Vilket som helst ord i en sats kan tilldelas betoning. Om man t.ex. i satsen *Världen är så stor* lägger betoning på *är* så innebär det förslagsvis att talaren anser att världen faktiskt är av den storlek som han själv nyss hävdade men som ifrågasatts av hans samtalspartner. Traditionellt går en framhävnings av denna typ under namn av (kontrastiv) emfas. Förekomsten av emfas utgör måhända det starkaste argumentet mot påståendet att "syntactic structure determines stress" (L & S, s. 78).

Sakläget kompliceras emellertid av det förhållandet att satsen ofta är så uppbyggda att det innehållsligt viktigaste, det som även kallas *rema* (eng. comment), kommer sist i satsen, medan det varom något utsägs, det s.k. *temat* (eng. topic), uppträder i satsens tidigare del.⁷ Då temat ofta är något för de samtalande redan känt och sålunda mindre viktigt i det aktuella satssammanhanget, får dess uttryck ofta ingen betoning. Under normalt rema—tema-förhållande bör i satsen

⁵ Ifrågavarande amerikanska system beskrivs i G. L. Trager & H. L. Smith, *An outline of English structure*. Norman, Okl., 1951. (Studies in linguistics. Occasional papers; 3). Både detta system och det svenska i SAOB har ingående granskats i mitt arbete *Studies in Swedish generative phonology*. Uppsala, 1977. (Acta Universitatis Upsaliensis. Studia philologiae Scandinavicae Upsaliensis; 11).

⁶ Se P. Lieberman, *On the acoustic basis of the perception of intonation by linguists*. (Word 21, 1965.)

⁷ Här om närmare t.ex. i John Lyons, *Introduction to theoretical linguistics*. Cambridge, 1968, kap. 8. Mera elementärt hos Sjögren, op. cit.

Världen är så stor det sista ordet få betoning, övriga ord (stavelser) bli obetonade. Med det nyssnämnda reducerade betoningssystemet kan stressens fördelning på de enskilda stavelserna noteras så: 0 0 0 0 1, där 1=(normalt) betonad och 0=obetonad. Rent lingvistiskt är denna notation tillräcklig. Alternativt kan *Världen* också få betoning, nämligen — som fallet kan antas vara vid en uppläsning av dikten "Lasse liten" — ordet icke tidigare figurerat i "talakten".

Det ligger nära till hands att anta att det är det normala tema—rema-förhållandets betoningsstruktur som lagts till grund för Nuclear Stress Rule, och att syntaxen därmed förlänats en rangställning gentemot semantiken som den egentligen inte bör inneha inom prosodin.

En slutsats av vad som ovan anförts skulle kunna formuleras så: den regelmekanism som Ch & H antagit gälla för stressfördelning (spec. i engelska) är i väsentliga stycken felaktig. Då L & S direkt bygger på denna konstruktion, har deras framställning av stressen i talet blivit i motsvarande stycken lidande. Som en följd härav har deras musikaliska omtolkning av stresskonturen för "Lasse liten" i termer av *notlängd* (s. 78 f.) fått något förskruvat över sig. Till det intrycket bidrar deras påstående (ibid.) att stress i talat språk, spec. i språk som engelska, tyska och svenska, realiserar i artikulatorisk "timing", dvs. i längdrelationer. Som redan framhållits är stressens fysiska uttryck (åtminstone vad svenskan angår) tvåfaldigt: längd *plus* intonation (artikulatoriskt: variation i stämbandets svängningsfrekvens).

Om sålunda viss kritik kan riktas mot L & S:s försök att applicera regler som antagits gälla talat språk på musikaliska fenomen, så innebär detta förvisso *icke* att — som L & S tycks mena (s. 79) — det skulle saknas aprioriska skäl för en meningsfull sammanställning av språkligt och musikaliskt. Tvärtom förhåller det sig ju så att både talet och musiken är typiska mänskliga meddelelseformer, att båda kan analyseras med hänsyn till uttryck och innehåll (semantik), att båda främst appellerar till hörselsinnet, att båda kan ge uttryck för emotioner — för att nu bara nämna några av de viktigaste likheterna.

Men visst förefinns också stora olikheter. Språkets symboler (uttrycksenheter) har i stor utsträckning kognitiv (intellektuellt begreppslig) funktion, medan musikens symboler företrädesvis hänför sig till den emotionella och estetiska begreppssfären. I den mån språket antar klart estetiserande former närmar det sig, karakteristiskt nog, musiken. Så utnyttjas ju inom verskonsten musikaliska verkningssmedel som rytm och upprepning (*så stor, så stor, Lasse Lasse liten*), företeelser som man nogsam undviker i vanligt tal för att inte riskera att göra sig löjlig, verka tjatig osv.

Dessa och andra olikheter till trots är dock, som sagt, språk och musik inte mera artsilda än att fortsatta försök till parallell beskrivning kan förväntas ge god utdelning. Här skall endast en del av allt det som kunde förtjäna att tas upp till komparativ analys beröras, nämligen betoningen. Utgångspunkten är här, liksom hos L & S, lingvistisk; vi rör oss från språkligt till musikaliskt.⁸ Det bör

⁸ En omvänd analysgång skulle sannolikt visa på flera andra språk—musik-samband än dem som här kan behandlas. Betr. betoning i musikalisk mening se artikeln *Betoning* i Sohlmans musiklexikon (2. uppl.).

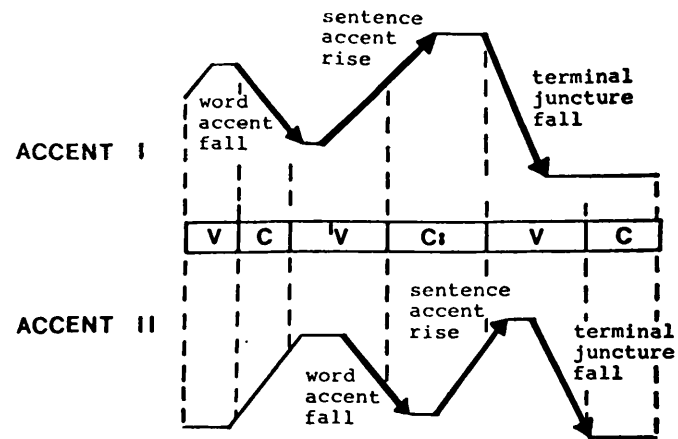


Fig. 5.

vidare framhållas att med "betoningen" då avses det internt lingvistiska fenomenet, ej de externt lingvistiska, "paralingvistiska", företeelser av prosodisk art som utgör icke-verbala, oftast väl rent reflexmässiga uttryck för sinnesstämningar och attityder.⁹

Först något om betoningsvariabeln *intonation*. Utmärkande för intonationen eller grundtonsrörelsen i talet är inte bara karakteristiska (och relativt konstanta) stigningar och fall som åtföljer de betonade orden i satsen (meningen) utan också en meningsavslutande tongång, s.k. *terminaljunktur*, vilken om meningen innehåller en fråga, består i en stigning (tonen blir s.a.s. hängande i luften) eller, om meningen är ett påstående, utgörs av ett fall i tongången ner till viloläge. Terminaljunkturen, som antagligen är djupt fylogenetiskt rotad i människans språk, har sin klara motsvarighet i musiken. Så är ju likheten mellan talets slutfall och musikens slutkadens (åtminstone i många musiktyper) särskilt påtaglig. Vad frågetonen beträffar finner man i sånger ofta goda exempel på musikens anslutning till talet. Lyssna — för att nu ta ett enda exempel ur den västerländska konstmusikrepertoaren — till melodistämmans stigningar i tenorarien nr 19 i Johannespassionen ("Ach, mein Sinn, wo willst du endlich *hin?*" osv.). I detta sammanhang bör det kanske påpekas att det är *melodistämmans* förlopp (snarare än t.ex. harmoniskt avstånd från tonikan, jfr L & S) som i första hand erbjuder det naturliga jämförelsematerialet, och detta gäller jämförelse inte bara med slutintonationen utan även ordens intonation i satssammanhanget.

För svenskans vidkommande har ganska nyligen mycket intressanta teorier och data rörande ord- och satsintonation lämnats i flera arbeten utgångna från Lunds universitet. Bland dessa är särskilt Gösta Bruce's doktorsavhandling, *Swedish word accents in sentence perspective* (Lund, 1977), värd uppmärksamhet, då därin ges en elegant analys av de båda ordaccenterna i svenskan — de som av gammalt kallas akut och grav accent, eller bättre: accent I och accent II. Dessa

Ord	Intonation (I)	Melodistämman (M)	Kongruens (I:M)
Världen	x → x	x → x	+
är	x	x	+
så	x	x	-
stor	→ x	→ x	+
så	x	x	+
stor	x	x	+
Lasse	x (2ggr)	x → x	--
Lasse	x (2ggr)	x → x	--
liten	x ← x	x (2ggr)	--
större	x → x	x ← x	--
än	x	x	+
du	x	x	+
nånsin	x → x	x ← x	--
tror	→ x	x ←	-
Lasse	x (2ggr)	x → x	--
Lasse	x (2ggr)	x → x	--
liten	x ← x	x (2ggr)	+ -
Tot.			8/25

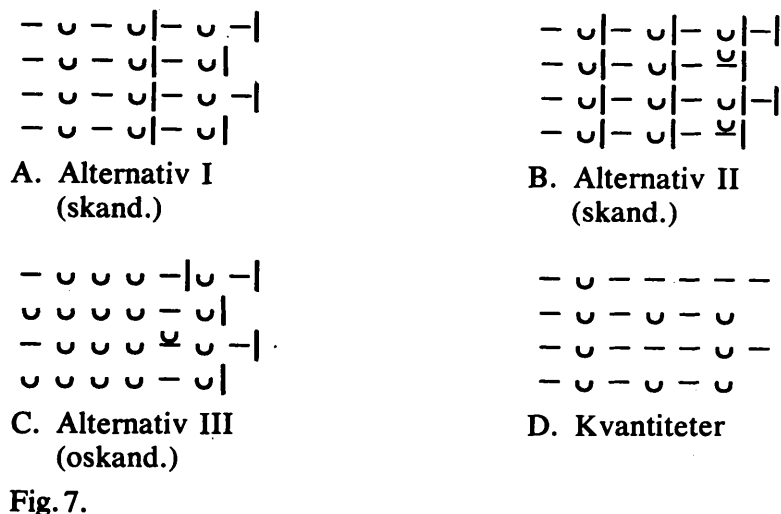
Fig. 6. Tonal kongruens.

accenter har undersökts med och utan satsaccent. Bruce skiljer i princip på tre lingvistiskt relevanta tonala förlopp: ordaccentfallet (som inträffar tidigare för accent I än för accent II), satsaccentstigningen samt terminalfallet. (Se fig. 5.) Diagrammet i figuren kan givetvis kompletteras med terminalstigning för frågesats.¹⁰

Redan med detta schema för ord- och satsintonation i svenska kan grundtonsförloppet exempelvis i "Lasse liten" i huvudsak prediceras. Fig. 6 visar ett efter Bruce's anvisningar uppgjort schematiskt intonationsmönster för texten i fråga. Tonrörelse uppåt anges med pil åt höger och vice versa. I kolumnen intill återfinns melodistämmans förlopp återgivet i en likartad, för ändamålet lämpad grov notation. Tilldelas accent I-ordet *Världen* normal betoning, skall grundtonen börja relativt lågt och stiga mot ordets slut. Detta uttal stämmer händelsevis med Alice Tegnér's tonsättning (fig. 3). Här föreligger sålunda en tonal samgång mellan tal och musik, vad man kan kalla en *tonal kongruens*, i figuren markerad med plus. Naturligtvis behöver inte denna samgång sträcka sig så långt som till full överensstämmelse i absoluta tonlägen. Brist på samgång, inkongruens, kan vi, utifrån den givna notationen, konstatera för ordet *större*, som skall ha accent I men som i tonsättningen erhållit en nedåtgående rörelse. Analysen fortskrider genom parvis jämförelse av intillstående stavelser, vilket med-

⁹ Bland arbeten rörande paralingvistik kan nämnas David Crystal & Randolph Quirk, *Systems of prosodic and paralinguistic features in English*. London—Haag—Paris, 1964. (Janua linguarum, series minor; 39). Författarna ger bl. a. en nyttig kritisk översikt över tidigare forskning på detta svårtillgängliga område.

¹⁰ Ett utförligare regelschema lämnas i Bruce, op. cit., s. 142. För de obetonade ordens intonation gäller som huvudregel att denna på enklaste sätt förbinder de kringstående betonade ordens maxima och minima i grundtonsfrekvens. Jfr "join rules" i kap. 8 hos Bruce.



för att textens allra första ord, som här råkar ha två stavelser, endast ger *en* markering i kongruenskolumnen. På detta sätt kan vilket sångstycke som helst undersökas med hänsyn till tonal kongruens. I vårt enkla exempelfall erhålls som synes en tonal kongruens på 8/25 eller 32 %.

Går vi så över till betoningsvariabeln *längd*, bör vi från början ha klart för oss inte bara att en sångtext kan vara mer eller mindre metriskt bunden utan också att analys av ett poems meter kan ge rätt skiftande resultat. I L & S:s rytmiska version av texten till "Lasse liten" — återgiven som Alternativ I i fig. 7 — är taktindelningen närmare ansluten till den musikaliska motsvarigheten än vad fallet är med den version som här presenteras som Alternativ II i samma figur. I Alternativ I laboreras med s.k. dipodier;¹¹ huruvida dipodier överhuvudtaget bör ingå i korrekt versanalys är en gammal tvistefråga. Hur som helst torde det vara riktigtast att vid en jämförelse mellan språkligt och musikaliskt ställa textens talspråkliga form i förgrunden.

Som bekant kan vers i muntligt framförande te sig rätt olika. Om versschemat strikt observeras, föreligger s.k. *skandering*. Vid uppläsning med anspråk på artisteri brukar emellertid taktmässigheten få ge vika för ett friare, dvs. mera tal-språksanpassat, föredrag. Ett rytmiskt schema som därvid kan framkomma återges som Alternativ III i fig. 7. Självfallet bör man räkna med ett visst variationsutrymme inom ramen för oskanderat alternativ. ("Taktstrecken" i Alternativ III står för de pauser som kan motiveras av textens innehåll och syntax.)

De i talet uppträdande (vid talregistrering direkt mätbara) ljudlängderna, durationerna, utgör realisationer av *kvantiteter*, dvs. lingvistiskt relevanta längdenheter, som varje ord tänks äga i sin ideala isolerade form, i språkets "lexikon".

¹¹ Termen dipodi ("tvåfot") avser verstant med två slag i takten. Betr. dipodibegreppets tillämplighet på modern vers se Bernhard Risberg, Den svenska versens teori I—II. Stockholm, 1932—1936, spec. II § 277.

Ord	Skand. (S)	Oskand. (O)	Notvärde (N)	Kongruens	
				S:N	O:N
Världen	— 2	— 2	♩. ♩	+	+
är	— 1	2 2 2	♩	+	—
så	2 2	2 2 2	♩	—	—
stor	— 1	2 2 2	♩	+	+
så	2 2	2 2 2	♩	—	—
stor	— 1	2 2 2	♩	+	+
Lasse	— 2	2 2 2	♩. ♩	++	—+
Lasse	— 2	2 2 2	♩. ♩	+-	—
liten	— 2	2 2 2	♩. ♩ }	++	++
större	— 2	2 2 2	♩. ♩ }	+-	+-
än	— 1	2 2 2	♩.	+	—
du	2 2	2 2 2	♩	+	+
nånsin	— 2	2 2 2	♩	+-	+-
tror	— 1	2 2 2	♩	+	+
Lasse	— 2	2 2 2	♩. ♩	++	—+
Lasse	— 2	2 2 2	♩. ♩	+-	—
liten	— 2	2 2 2	♩. ♩ }	++	++
Tot.				19/25	13/25

Fig. 8. Rytmisk kongruens.

Fig. 7 återger den aktuella textens kvantiteter (D). Påpekas bör att de kvantiteter som de enskilda orden tilldelats förutsätter att *varje* ord har fått sin "lexikaliska" betoning. En huvudregel för svensk kvantitetsspecifikation lyder: Varje betonad stavelse är lång, dvs. den innehåller (minst) en lång vokal eller en kort vokal följt av antingen en lång konsonant eller en konsonantförbindelse (ex. *så, vass, kvast*). Som en jämförelse mellan de olika alternativen i fig. 7 visar, blir flera stavelser som är lika långa i den lexikaliska djupstrukturen förkortade i den prosodiska ytstrukturen (ex. *är, så*). Detta är fysiskt möjligt tack vare talljudens, i synnerhet vokalernas, temporala elasticitet.

Förekomsten av två huvudtyper av textrytmisering, skanderad och oskanderad (se fig. 7), möjliggör iakttagelser av varierande *rytmisk kongruens* mellan dessa inbördes. Men en jämförelse mellan texten i olika rytmversioner och den musikaliska versionen är också möjlig, då kvantitet (eller stavelselängd) naturligt svarar mot musikens notvärden.

Beräkning av rytmisk kongruens mellan språk och musik, alltså med "Lasse liten" som exempel, åskådliggörs i fig. 8. Utgående från det kongruenta förhållandet i *Världen* med fjärdedelsnot för lång stavelse och åttondelsnot för kort, blir avvikelser från detta mönster givetvis inkongruens, såvida inte notvärdet ytterligare ökas för lång stavelse och vice versa. Orden *Lasse Lasse* uppvisar utifrån denna (obs. för just detta fall gällande) regel inkongruens, medan t.ex. *liten*

återigen blir så att säga inom sig kongruent. (Det sagda gäller jämförelsen mellan den oskanderade versionen och musiken.) Men analysen kan innefatta tillämpning av flera regler för kongruens än den nyss nämnda. Räknar man värdena inom frasen *Lasse Lasse liten*, kan de båda första orden i relation till det tredje betraktas som kongruenta, detta som följd av att den betonade stavelsen i *liten* fått halvnots värde. Möjligheten av att införa en specialregel som denna antyder vilka problem en mera nyanserad kongruensberäkning kan medföra. Den i fig. 8 framställda beräkningen är av det enklaste slaget, dvs. med endast en kriterieregel. Den rytmiska överensstämmelsen mellan text och musik i sången "Lasse liten" visar sig under dessa villkor uppgå till 52 % (13/25) för oskanderad och 76 % (19/25) för skanderad textversion. I jämförelse härmed är den tonala kongruensen påfallande låg. Möjligen ligger det någon allmängiltig sanning dold bakom dessa sifferförhållanden, men undersökningsmaterialet är givetvis alldeles för litet för vidare slutsatser.

Avslutningsvis kan dock sägas att det synes ligga inom det möjligas gränser att erhålla en någorlunda adekvat och t.o.m. kvantifierbar bild av den prosodiska samgången mellan tal och musik i sång genom att undersöka tonal och rytmisk kongruens såsom skisserats ovan. Att mycket arbete återstår innan metoden för kongruensberäkning är tekniskt och sakligt helt tillfredsställande säger sig självt. Så kan t.ex. nya variabler komma att visa sig behövliga för precisering av relationen språk—musik. Uppslag i den vägen står inte minst att hämta i den intresseväckande och idérika artikel av Lindblom och Sundberg som här ofta citerats.¹² Fortsatt forskning på området bör naturligen bedrivas under fortsatt interdisciplinärt samarbete.¹³

¹² Nyligen har Lindblom tagit upp ett hithörande problem till behandling i *Final lengthening in speech and music* (Nordic prosody. Papers from a symposium. Lund, 1978, s. 85 ff. (Travaux de l'Institut de linguistique de Lund; 13).)

¹³ Dyligt samarbete förekommer på flera håll numera, bl.a. vid institutionen för musikvetenskap i Uppsala, se t.ex. Ingmar Bengtsson, Alf Gabrielsson & Stig-Magnus Thorsén, *Empirisk rytmforskning*. (STM, 1969, s. 49—118.)

Summary

On congruence between text and music in songs

The present article consists of two major parts: (1) a critical discussion of a previous proposal for applying certain linguistic concepts and procedures to the analysis of musical compositions and (2) a counter-proposal, including the presentation of a new method of calculating so-called tonal and rhythmical congruence between text and music in songs.

The earlier attempt in question was made by Björn Lindblom and Johan Sundberg (L & S); their report, entitled *Towards a generative theory of melody*, appeared in the 1970 volume of this periodical. L & S have adopted the ideas concerning stress and stress assignment expressed by Noam Chomsky and Morris Halle — the leading exponents of the theory of generative transformational grammar (TG) — in their famous book *The sound pattern of English*. In particular, L & S advocate an application of the two rules of TG called the Nuclear Stress Rule (NSR) and the Compound Stress Rule. The present author tries to demonstrate, by means of a concrete example (see Fig. 1), a few deficiencies of these rules, particularly the unrealistic number of stress degrees resulting from the cyclical application of the NSR. The general validity of the TG dogma that "syntactic structure determines stress" is denied.

Instead, stress is here presented as basically determined by semantic conditions. Its essential physical correlates (in Swedish) are assumed to be intonation (fundamental-frequency variation) and syllable length. The occurrence of these two variables within the (spoken) text of a particular sample song (also used by L & S) is determined and typified according to the principles of recent phonetic theory and traditional metrics, respectively. The individual tonal and rhythmical occurrences are compared, syllable by syllable, with the corresponding items of the musical notation, i.e. with the relative pitch and time values of the separate notes of the melody line (Figs. 6 and 8). Finally, the number and percentages of positively correlated textual and musical values are computed. Naturally, this method of calculating tonal and rhythmical congruence is still at a rather primitive stage of development and will certainly need some revision.