

In: Inger Haska & Carin Sandqvist (eds.): *Alla tiders språk*. Lundastudier i nordisk språkvetenskap A 55. (pp. 117–126). Lund University Press, Lund.

Olika perspektiv på språk, bild och deras samspel

Metodologiska reflektioner

Jana Holánová

Vi möter dem dagligen och tar det för självklart att de består av en kombination av text och bild: tidningar, tidskrifter, reklambroschyrer, affischer, sagoböcker, läroböcker, encyklopedier och instruktionsböcker. Alla dessa tryckta medier är alltså multimodala - om man vill använda en term som har fått en stor spridning under det senaste decenniet. Multimodalitet betyder att informationen uttrycks på olika sätt, med olika modaliteter (skrift, tal, bild och andra visuella medel) som sedan kombineras till en "sammansatt grafisk produkt" (Melin 1995). Därför känns det också naturligt att språkvetare under den senaste tiden försökt utöka analysen från texten till bilden och börjat studera sambandet mellan språklig och grafisk form.

Multimodalitet har studerats inom olika discipliner. Inom semiotiken har det länge funnits en tradition att analysera språk, bilder och andra teckensystem. Man har bl.a. studerat relationen mellan bilden och den verbala komponenten i reklamtexter (Barthes 1980, Eco 1971) eller i utformningen av tidningssidor (Kress & van Leeuwen 1990, 1996). Det finns arbeten som jämför verbalspråk och bildspråk och pekar på deras likheter (Kjørup 1980), men också andra som polemiserar mot denna uppfattning (Ulbæk 1997). När det gäller uppslagsböcker, har man i Sverige i slutet av 50-talet skapat en s.k. lexivisuell berättarteknik som sedan använts i uppslagsverket Focus (Pettersson, under utg.). I lexivisionen betonades samverkan av text och bild och helhetssynen på den visuella formen och det tematiska innehållet i uppslagsboken. Lexivisionen har senare blivit en inspiration för visuella tekniker för till exempel hypermedia (Kindborg 1991). Intresset för multimodalitet har dessutom varit stort inom gränssnittsdesign (Bernsen 1994b, Bernsen & Verjans 1995, Mullet & Sano 1995) och människa-dator-interaktion (Maybury 1996, Preece 1994).

Ändå tror jag att steget utanför textens ramar inte varit enkelt för språkvetarna. Hur skulle man gå tillväga? Trots att vissa semiotiska arbeten har inspirerats av den allmänna lingvistik, kan det ha varit svårt att hitta en enhetlig modell som är teoretiskt förankrad, som är lätt att använda empiriskt, som omfattar både text och bild och som passar för olika genrer i olika tider. Men samtidigt är det svårt att bortse från multimodaliteten. I det följande kommer jag att ge en överblick över några lösningar i arbeten som på olika sätt utforskat multimodali-

teten. Målet är alltså inte att noga redovisa resultaten utan snarare att granska de olika sätten hur författarna gått till väga för att få grepp om text och bild. Jag kommer i första hand att peka på olika möjliga perspektiv som man kan utgå ifrån när man analyserar text och bild. Jag kommer att exemplifiera vart och ett av perspektiven och föra en metodisk diskussion kring analysen av text och bild i olika genrer.

När det gäller de olika tillvägagångsätt som utkristalliserar sig i de enskilda arbetena om multimodalitet har jag lagt märke till åtminstone sex möjliga sådana: Perspektiv som fokuserar modaliteternas skillnader, modaliteternas likheter, samspillet mellan modaliteterna, layoutens formella aspekter, multimodalitetens perception samt multimodalitetens reproduktion i samtal. (Perspektiven ska uppfattas som tendenser och som ett första försök att systematisera möjliga tillvägagångsätt.)

(1) Ett av perspektiven är att analysera text och bild för sig som två betydelseskapande system och att titta på det som är utmärkande för det ena och det andra. Vilka funktioner fyller text och bild? Vilka semantiska relationer uttrycks med fördel textuellt och vilka piktorialt (med hjälp av bild)? Med denna ansats fokuserar man alltså modaliteternas **skillnader** (för en mer utförlig diskussion kring detta perspektiv se Holmström & Nová 1996). Grundtesen är att språk, bild och andra modaliteter har olika inneboende styrkor och svagheter och därmed är lämpliga att uttrycka olika typer av information. Vad gäller människa-dator-interaktion tänker Maybury (1996) i följande banor: När man har bestämt vilket budskap som ska förmedlas till adressaten, måste man som nästa steg bestämma i vilket medium det ska ske. Valet är inte enkelt. Några medier och modaliteter är mer effektiva än andra för vissa uppgifter, användare och kontexter. Kartor lämpar sig för att förmedla spatial information, medan talat språk är bäst för att kontrollera apparater i kontexter där händer och ögon är fullt upptagna. I andra fall stödjer en kombination av olika medier en mer naturlig och effektiv interaktion (t. ex. *lägg det där* som blandar talat språk och gester). Vad gäller gränssnitts-design ger Mullet & Sano (1995) praktiska råd om hur man ska använda visuella och språkliga representationer. Bernsen (1994a, b) utgår ifrån att ju bättre informationsmappning det finns mellan det som representeras och dess representation, desto lättare är det för användaren att kommunicera med systemet.

Företrädare för detta perspektiv bedömer det naturliga språkets versus bilders och analoga grafiska representationers egenskaper. De ställer ofta upp en taxonomi över modaliteterna, både inom gränssnittdesign (Bernsen 1994b, Lu & Bernsen 1995) och inom språkvetenskap (Teleman 1989, 1994).

Det råder enighet om att semantiska relationer som orsak/följd, händelse/resultat, problem/lösning, villkor samt kvantifiering (*de flesta, flera, några*) kan uttryckas bättre språkligt (Maybury 1996, Mullet & Sano 1995, Teleman 1989, 1994). Man anser att språket kan fokusera på en aspekt av det som representeras och därför är lämpligt för att uttrycka abstrakta tillstånd, relationer och

begrepp (t.ex. *sanning, rättvisa*). Enligt föreläsarna är språk lämpligt för att uttrycka temporal överlappning och temporal kvantitet (*alltid*), temporala förändringar (*tre dagar senare*) och temporal sekventialitet. Språket innehåller dessutom viktiga logiska och epistemiska operatörer (*icke, eller, om/så*). Med hjälp av språket kan man hänvisa till *jag, du, här*.

Grafiska representationer karakteriseras som specifika. De erbjuder en ungefärlig föreställning om egenskaper och kvaliteter hos det som representeras, och kan därför användas som ersättning för direkt perceptuell erfarenhet och för omedelbar identifikation. (Det är bl.a. därför polisen föredrar ett foto på rånaren framför en verbal beskrivning). Bilder är dessutom dominanta, lätta att känna igen och komma ihåg. Om de är relativt konstanta, kan man kommunicera med dem tvärs över sociala och språkliga gränser (Mullet & Sano 1995).

Enligt Bernsen (1994b) och Teleman (1999) har både språket och analoga representationer utvecklat en rad mekanismer för att komma över sina egna nackdelar. Genom att använda korta, klara och innehållsrika instruktioner, sammanfattningar eller upprepningar av viktiga nyckelord kan man öka språkets specificitet. Språket blir också mer specifikt om vi använder oss av indexikal referens, dvs. kopplar språket till den icke-språkliga erfarenheten, pekar och använder väl definierade beskrivningar. Att lägga till exakta metrisk egenskaper hjälper till att reducera tolkningsutrymmet. Och slutligen tycks metaforer och analogier stärka specificiteten hos språket (beskrivningen *hon som ser ut som en pudel i håret* gör det lätt att identifiera en viss person i en större grupp).

Grafiska representationer har i sin tur utvecklat möjligheter till fokusering. Ju mer abstrakt bilden är desto mer minskar dess analogkaraktär och samtidigt höjs dess fokuseringsförmåga. I abstraktionsprocessen bevaras bara objektets mest karakteristiska delar. Det gör att bilden börjar närma sig konventionella representationer. Skillnaden mellan foton och schematiska representationer är alltså att schematiska representationer tack vare en selektiv utelämnning av detaljer blir mer kanoniska och generiska och lämpar sig som representationer av bredare klasser av objekt.

Man skulle kunna sammanfatta detta perspektiv med slagorden *informationsmappning och multimodalitet* eftersom det i dessa sammanhang ofta handlar om att hitta rätt *förpackningsformat* för informationen. Modalitetsinneboende styrkor och svagheter samt frågan vilken modalitet som är överlägsen har diskuterats rätt extensivt. Jag tror dock att lösningen inte är antingen textuell eller grafisk. Ska taxonomin användas för en genrespecifik beskrivning måste också hänsyn tas till kombinationen av dessa modaliteter, till deras samspel. Målet skulle då snarare vara att rekonstruera vissa för genren relativt stabila konstellationer av text och bild.

(2) Ett annat perspektiv är att beskriva bildens och textens karaktär, stil och uppbyggnad på ett relativt högt plan och fokusera på deras **likheter**. Likheterna kan bestå i att båda modaliteterna har generisk karaktär, uttrycker

tillstånd, har samma specificitetsgrad och kan framhäva vissa aspekter. Denna syn på språk och bild företräds av Kress & van Leeuwens (1990, 1996) sociosemiotiska analysapparat som antar att bilder och visuella element är regelbaserade tecken och formar en visuell grammatik. Författarna utgår från Hallidays funktionella grammatik och utvidgar hans multifunktionella synsätt till bilder. De gör fortlöpande jämförelser mellan hur en grammatisk regel uttrycks visuellt och språkligt. (Frågan är dock i vilken relation denna grammatik står till språkets grammatik). Bilder behandlas som ett slags språk och båda modaliteterna bedöms med samma mått. Inom detta perspektiv medges bildens *töjbarhet*: bilder uttrycker inte bara tillstånd utan kan uttrycka även dynamiska processer, rörelse, riktning och förflyttning. Man kan *läsa* bilder och det finns en predefinierad *väg* att gå när man läser bilder. Med bilder kan man t.o.m. uttrycka modalitet (även om det inte är självklart vilken den språkliga motsvarigheten till en viss bildmodalitet är). Visserligen är en av Kress & van Leeuwens poänger att bilder är bättre på att uttrycka konceptuella och analytiska processer men i analysen handlar det mycket om likheter. Författarna använder alltså samma analytiska kategorier när de analyserar språk och bild och på så sätt likställer språk och bild som uttrycksmodaliteter.

Denna analysapparat har introducerats och tillämpats av Ledin (1997) i en spännande analys av veckopressens föregångare och har sedan delvis använts för att beskriva språk och bild i andra genrer (Pettersson 1997, under utg.). Kress & van Leeuwens analysapparat verkar ha flera fördelar: dels ansluter den sig till Hallidays funktionella grammatik som är välkänd för språkvetare, dels verkar den vara lätt att tillämpa empiriskt. Den erbjuder dessutom en enhetlig modell som omfattar både text och bild och som är tillräckligt allmän för att passa för olika genrer i olika tider. Nackdelen kan vara att just likheter i betydelseskapandet lyfts fram och att Kress & van Leeuwens analys ofta arbetar på en generell nivå. Här kan en allmän kritik riktas mot den semiotiska ansatsen som vanligtvis börjar och slutar med att likheter i betydelseskapandet mellan olika uttrycksmodaliteter fokuseras. Det som är unikt för en genre riskerar då att gå förlorat. Vill man komma ner till den genrespecifika nivån kan en komplettering med modaliteternas skillnader och deras samverkan behövas. På så sätt kan man kanske komma åt konventionaliserade multimodala kombinationer som kännetecknar en viss genre. (3) Ett tredje perspektiv som är nära kopplat till de två föregående fokuserar **samspelet** mellan text och bild. Närmare bestämt undersöker man vilken funktion bilden har i förhållandet till texten. Här befinner vi oss på ny mark. Någon enhetlig teori eller terminologi för samverkan verkar inte finnas till hands. Det finns dock förslag till en funktionell beskrivning av samspelet mellan text och bild som baseras på empiriska studier från olika genrer. Vi kan börja med Barthes (1980) som urskiljer två relationer: *ancrage* och *relais*. *Ancrage* innebär att bilden är beroende av texten och texten styr läsaren i tolkningen av bilden (t.ex. reklamens undertext). *Relais* innebär att bild och text

är komplementära och meddelandet realiseras genom deras koppling på en högre nivå (i t.ex. bildserier). Dane%o(1995) undersöker relationen mellan text och illustration i skönlitterära böcker. Han frågar sig: Dominerar den ena modaliteten eller kompletteras modaliteterna? Upprepar eller styrker bilden informationen som ges i texten, reducerar den textens innehåll eller lägger till ytterligare detaljinformation? Liknande frågor - dock med hänsyn till bildtexter - ställer Melin (1990) i en analys av samhällskunskapsböcker, Sandqvist (1995) i en analys av historieböcker och Pettersson (under utg.) i en analys av uppslagsböcker. Följande funktioner av bildtext har fastställts av forskarna (även blandformer kan förekomma):

- Bildtexten anger bildens innehåll och har en etiketterande funktion.
- Bildtexten upprepar något som omtalas i texten.
- Bildtexten upprepar information som ges i bilden.
- Bildtexten tillför ny information som inte ges i texten.
- Bildtexten innehåller yttre upplysningar om bilden eller om konstnären.
- Bildtexten har en instruerande och stämninggivande funktion. (Denna funktion är specifik för samhällskunskapsboken).

Teleman (1999) nämner också samspelet mellan bild och text och nyanserar därmed sin taxonomi. Han anser att ordtexten ger konkretion åt bilden, kontextualiserar bilden genom deixis, typologiserar verkligheten, karakteriserar referenter och platsen. Enligt Teleman kan bilden, tack vare kompletteringen med texten, på så sätt överskrida sina egna styrkor och svagheter.

(4) Analys av **formella aspekter på layouten** är kännetecknande för det fjärde perspektivet. Här fokuseras textens och bildens kontakt, yta, frekvens och position på sidan. Dels används kvantitativa metoder för att beräkna bildfrekvensen, bildtexternas del i textmassan, bildernas yta och mängd, dels undersöks bildernas placering, textens kontakt med bilden och hänvisningar mellan text och bild (Sandqvist 1995, Pettersson, 1997, under utg.). Att lägga märke till återkommande mönster i layouten, såsom till textens och bildens närhet, proportion eller deras relativa självständighet, kan vara viktigt för en genrespecifik beskrivning.

Skillnader som ligger i layouten kan också vara informationsbärande och kan studeras kvalitativt. Kress & van Leeuwen (1990, 1996) hävdar att bild och text står i relationen generell-specifik om de placeras i relationen överst-underst, i relationen viktig-oviktig om de placeras i relationen centrum-marginal och i relationen känd-ny om de placeras i relationen vänster-höger. Det framgår inte vilken grund denna regel vilar på. Man kan gissa att den härstammar från perceptionspsykologin, framförallt från Arnheim (1974) och Gombrich (1977), som undersöker hur balans, form, ljus, färg och spatiala förhållanden i bilden förmedlar vissa effekter och påverkar vår perception av bilden. Frågan är sedan om redaktörerna medvetet utnyttjar dessa innehållskriterier i layouten och om läsare uppfattar dem på samma sätt.

I de ovannämnda tillvägagångssätten 1-4 har text och bild behandlats som en artefakt, som en färdig grafisk produkt, och analyserats ur forskarens synvinkel. De återstående två perspektiven 5-6 tar däremot upp aspekter knutna till användningen av dessa grafiska produkter.

(5) Ur ett femte perspektiv kan man studera multimodalitetens **perception**. Följande frågor är centrala: Hur inhämtas informationen från den multimodala produkten? Vilken modalitet fokuseras först, vilken sedan? Hur uppfattar läsare sambandet mellan språklig och grafisk form? Vilken roll spelar samverkan mellan språkliga och grafiska medel för textens läsbarhet och begriplighet? **Danes** (1995) tar upp receptionssidan i analysen av illustrerad skönlitteratur och frågar sig: Vad händer när läsarens tolkning av text och bild möter illustratörens tolkning av texten? En annan aspekt han nämner är illustrationers olika reception över tiden: Den lokala miljön i Ladas illustrationer i boken om soldaten *• vejk* var redundanta strax efter 1:a världskriget men är berikande för dagens läsare.

En möjlighet att ta reda på hur läsare uppfattar multimodala sidor är att intervjua dem som **Melin & Pettersson** (1991) har gjort. Enligt vad eleverna själva uppger att de har uppmärksammat i läroböckerna följer läsningen *normalbanan* rubrik - brödtext - faktaruta - bild - bildtext - uppgift. I samband med en undersökning av hur användare beter sig när de läser hemsidor och letar på nätet har de **Léon & Holánová** (1997) analyserat användarnas retrospektiva kommentarer medan de tittade på inspelningen av den egna aktiviteten. En ytterligare metod är att ta fasta på ögonrörelser. Blicken kan avslöja vägar genom texten, vägar genom bilden samt de kopplingar som skapas mellan dem. Man skulle också kunna kombinera ögonrörelsemätningen med läsarens simultana kommentarer (**Holánová** 1998, **Holánová**, Hedberg & Nilsson 1998). Slutligen skulle man kunna kombinera ögonrörelsemätningen med retrospektiva kommentarer: Spela in läsarnas ögonrörelser över text och bild och låta läsarna titta på inspelningen och kommentera vad de har uppmärksammat och varför. En sådan metod skulle kunna användas för att studera mönster i informationsinhämtning från de olika modaliteterna, för att evaluera layoutens formella egenskaper eller för att ta reda på genreigenkänningen.

Ett problem som nämns i samband med analys av text och bild är att texten anses vara ett sekventiellt, lineärt fenomen medan bilder är tvådimensionella och inte självklart erkänner en enda privilegierad läsriktning (**Teleman** 1989, **Ledin** 1999). Detta kommer att kvarstå om vi undersöker text och bild ur analytikerns perspektiv. Om vi dock utgår från läsarens perspektiv och undersöker perception av bilden finns det ögonrörelsemätning att tillgå. Genom att blicken fixerar vissa bildportioner i taget kommer en väg genom bilden skapas och bildinnehållet lineariseras.

(6) Slutligen skulle jag vilja lägga till ett sjätte perspektiv som fokuserar multimodalitetens **reproduktion** i samtal. De dynamiska momenten i användningen av en multimodal grafisk produkt slutar inte vid dess läsning. De sträcker

sig även till dess kreativa rekonstruktion för samtalspartner i en annan kommunikativ situation. Det kan vara spännande att se hur överföringen sker: Bibehålls den ursprungliga "informationsförpackningen" eller förskjuts informationen mellan modaliteterna? Försvinner modaliteterna? Till förmån av vad?

Som ett exempel på detta kan nämnas en undersökning (Holmström & Nekvapil 1995) där talaren berättar för sina vänner en roande historia om en folkgrupp, men denna historia närmar sig en skröna och utlöser en häftig reaktion hos samtalsdeltagarna som frågar efter informationskällan. Talaren hävdar att historien är sann och legitimerar den genom att hänvisa till en tidningsartikel och ett tidningsfoto försett med bildtext. Intertextualitet skapas genom att informationen från dessa källor rekonstrueras och används i en annan situation. Den ursprungliga texten kan här uppfattas som "mediational means - the tools by which people undertake mediated action. The purpose is not the production of the text but the production of the action which the text makes possible". (Scollon 1998:14). Nästa spännande moment är att informationen från en massmedial diskurs överförs till en privat diskurs. Sist men inte minst är den dynamiska kombinationen av olika modaliteter relevant här. Nekvapil (in preparation) har granskat tidningsartiklar från den perioden när samtalet skedde och upptäckt att talaren kombinerat information från flera verbala och icke-verbala källor: Talarens kreativa reproduktion av historien har egentligen baserats på innehållet av två tidningsartiklar som fanns i två olika tidningar och kombinerats med tidningsfotot och bildtexten. Ett annat exempel är situationer där samtalspartner spontant ritar bilder medan de berättar och beskriver saker (Holmström 1999:63f.). Med detta perspektiv har vi förflyttat oss från tryckta media till talat språk och undersökt återspeglingsen av multimodaliteten i samtal.

I flera undersökningar om text och bild i olika genrer har man utgått från mer än ett av de ovanstående perspektiven. Ekvall (1997) diskuterar relationen mellan bild, bildtext och text samt bildtyper i naturkunskapsböcker. Wendt (1998) analyserar bildförekomst, bildtyper, typografisk utformning av texten, samverkan mellan text och bild och bildernas funktion i soldathandbäckernas utveckling. Pettersson (under utg.) använder sig av både kvantitativa och kvalitativa metoder. Hon modifierar delar av Kress & van Leeuwens analysapparat, kompletterar den med en analys av samspelet mellan artikeltext, bildtext och bild samt studerar formella aspekter av layouten. Byrman (under utg.) studerar språk och bild i preventivmedelsbroschyrer. Deras gemensamma fråga är vilken information som förs fram vid olika tidpunkter och med vilka språkliga och grafiska medel det sker. Jag anser att det kan vara en fördel att kombinera flera ansatser för att komma ner till den genrespecifika nivån och för att hitta återkommande multimodala konstellationer som kännetecknar en viss genre.

Vad gäller perception av multimodala produkter bör det tilläggas att det vore mycket intressant att testa resultaten från de ovannämnda studierna med hjälp av en ögonrörelsemätare. Exempelvis skulle man få veta hur läsaren

uppfattar skillnaden mellan den pedagogiskt upplagda presentationen av TV-artikeln i Focus och uppläggningsordet i NE. Skulle blickbeteendet bekräfta slutsatsen att det i artikeln i NE för den tekniskt okunnige läsaren är ”mycket svårt att riktigt förstå kopplingen mellan artikeltext och bild/bildtext” (Pettersson, under utg.)?

Litteratur

- Arnheim, R. (1974): *Art and Visual Perception. A Psychology of the Creative Eye*. University of California Press: Berkley, Los Angeles, London.
- Barthes, Roland (1980): Billedets retorik, In: *Visuel kommunikation I*. Fausing, B. og B. Larsen (red.), København: Medusa, s. 42-57. Översättning efter *Rhétorique de l'image*, in: *Communication*, no. 4, (1964).
- Bernsen, N.O. (1994a): Why are Analogue Graphics and Natural Language both Needed in HCI? *Working Papers in Cognitive Science and HCI*. Centre for Cognitive Informatics. Risø National Laboratory, Roskilde University, 20s.
- Bernsen, N.O. (1994b): Foundations of multimodal representations. A taxonomy of representational modalities. *Interacting with Computers* Vol. 6 No. 4/1994, 347-71.
- Bernsen, N.O. and Verjans, S. (1995): Designing interfaces by information mapping. A case study of the CERD design. *CCI Working Papers in Cognitive Science and HCI*, WPCS-95-1. Centre for Cognitive Science, Roskilde University 1995.
- Buswell, G. T. (1935): *How people look at pictures*. A study of the psychology in perception of art. The university of Chicago Press: Chicago.
- Byrman, Gunilla (under utg.): Försiktighetsmått - text, bild, genus i preventivmedelsbroschyren. Utkommer i: *Svensk sakprosa*. Institutionen för nordiska språk. Lund.
- Daneš, František (1995): Text a jeho ilustrace. (Text och dess illustration). I: *Slovo a slovesnost*, 56, 1995. Praha.
- de Léon, David & Holmström, Jana (1997): Revealing User Behavior on the World-Wide Web. *LUCS 60*. Lunds universitet.
- Ekvall, Ulla (1997): Formativt, figurativt, operativt i läroböcker för barn. Del 1. Utvecklingen under 1800-talet. *Svensk Sakprosa* 12. Lund.
- Gombrich, E. H. (1977): *Art and Illusion: A Study of the Psychology of Pictorial Representation*. Fifth edition. Phaidon: London.
- Holmström, Jana (1996): När säger en bild mer än tusen ord? Om samspelet mellan text och bild i gränssnitt. In: Löwgren, J. (red.) *Teman i människa-datorinteraktion*. Institution för Informatik, Lunds universitet, s. 65 - 76.
- Holmström, Jana (1998): *Picture descriptions: Comparing verbal and visual "areas of interest"*. Presentation at the Second Swedish Conference in Multimodal Communication. Lund, Oktober 1998.
- Holmström, Jana (1999): På tal om bilder. Om fokusering av uppmärksamhet i och strukturering av talad beskrivande diskurs. *LUCS 78*. Lunds universitet.
- Holmström, Jana, Hedberg, Björn & Nilsson, Nils (1998): Visual and Verbal Focus Patterns when Describing Pictures. I: Becker, W., Deubel, H. & Mergner, T. (eds.): *Current Oculomotor Research: Physiological and Psychological Aspects*. Plenum: New York, London, Moscow.
- Holmström, Jana & Nekvapil, Jiří (1995): Menschen "fünfter Klasse": Reden über Abwesende in der Alltagskommunikation am Beispiel tschechischsprachiger Daten. I:

- Czyzewski/Gülich/Hausendorf/Kastner (Hrsg.): *Nationale Selbst- und Fremdbilder im Gespräch. Kommunikative Prozesse nach der Wiedervereinigung Deutschlands und dem Systemwandel in Ostmitteleuropa*. Westdeutscher Verlag: Opladen 1995, S. 145-179.
- Kindborg, M. (1991): *Visual techniques for orientation in hypermedia structures*. Department of Computer & System Science, The Royal Institute of Technology, Stockholm university, 67s.
- Kjørup, Søren (1980): Billedkommunikation, I: *Visuel kommunikation 1*. Fausing, B. og B. Larsen (red.), København: Medusa, s. 58-82.
- Kress, G. & van Leeuwen, T. (1990): *Reading images*. Deakin University: Deakin University Press.
- Kress, G. & van Leeuwen, T. (1996): *Reading images. The Grammar of Visual Design*. Routledge: London.
- Ledin, Per (1997): "Med det nyttiga skola vi söka att förena det angenäma...". Text, bild och språklig stil i veckopressens föregångare. *Svensk sakprosa* 14. Lund.
- Ledin, Per (1999): Texter och textslag - en teoretisk diskussion. *Svensk sakprosa* 27. Lund.
- Lu, S. and Bernsen, N.O. (1995): The taxonomy workbench. A multimedia database system for analysing representational modalities. I: *Proceedings of the 4th International Conference on Interface to Real and Virtual Worlds*, Montpellier, June 1995.
- Maybury, M.T. (1996): Research in Multimedia and Multimodal Parsing nad Generation. To appear in *Journal of Artificial Intelligence Review: Special Issue on the Integration of Natural language and Vision Processing*.
- Melin, Lars (1990): *Text, bild, lexivision*. En studie i text-bildsamverkan i en lärobok. Nordiska språk. Stockholms universitet.
- Melin, Lars (1995): Grafisk pyttipanna. Om text och grafisk form i läroböcker. I: *Läroboksspråk. Ord och Stil. Om språk och layout i svenska läroböcker*. Hallgren & Fallgren. Uppsala, s. 77-123.
- Melin, Lars & Pettersson, Rune (1991): *Bild och bildtextexter i läroböcker*. Stencil. Nordiska språk. Stockholm.
- Mullet, K, and Sano, D. (1995): *Designing Visual Interfaces*. Communication Oriented Techniques. SunSoft Press, kap. 6: Image and Representation.
- Nekvapil, Jifi (in prep.): On the emergence of dialogical networks in Czech mass media at the beginning of 1993: Dispute over migration law.
- Pettersson, Gertrud (1997): Uppslagsböckerna och världen. Rapport om en pilotundersökning. *Svensk sakprosa* 11. Lund.
- Pettersson, Gertrud (under utg.): Håller bilden på att ta över? En studie av samspelet mellan text och bild i uppslagsböcker under 1900-talet. Utkommer i: *Svensk sakprosa*. Institutionen för nordiska språk. Lund.
- Pettersson, Rune (1992): Bilden och bildens roll i läroböcker. I: Melin, Lars, & Ekvall, Ulla (red.) *Språk och bild i läroböcker*. The state of art. Dokumentation från en konferens. s. 47-53.
- Preece, J. m fl (1994): *Human-computer interaction*, Addison-Wesley, Wokingham.
- Sandqvist, Carin (1995): Från 50-tal till 80-tal. Om form och innehåll i historieläroböcker. I: *Läroboksspråk. Om språk och layout i svenska läroböcker. Ord och stil* 26. Uppsala. s. 124-188.
- Scollon, Ron (1998): *Mediated Discourse as Social Interaction*. A Study of News Discourse.
- Teleman, Ulf (1989): The World of Words - and Pictures. I: Possible Worlds in Humanities, Arts and Sciences. *Proceedings of Nobel symposium 65*. Edited by Sture Allén. Walter de Gruyter: Berlin, New York, s. 199-208.

- Teleman, Ulf (1994): Ordens och bildernas språk. I: *Svenska i skolan* 1994:1, 6-15.
- Teleman, Ulf (1999): Om olika samband mellan ord och bild i text. Otryckt uppsats.
- Ulbæk, Ib (1997): Er billeder sprog? *ROLIG papir* 59/97. RoskildeUniversitetsCenter Lingvistgruppen.
- Wendt, Bo-A. (1998): Krigsmans läsning. Textmönster och språkbruk i svenska soldathandböcker 1769-1994. *Svensk sakprosa* 23. Lund.