

Att beskriva det som syns men inte hörs

Om syntolkning

Jana Holsanova

Mycket information i vardagen tar vi in visuellt, med hjälp av synen. Möjligheten att uppfatta visuella skeenden när man t.ex. tittar på en film är något som de flesta av oss tar för givet. Men för personer med synnedsättning och blindhet är inte detta möjligt. Man kan dock öka tillgängligheten av visuella skeenden genom syntolkning. Syntolkens främsta uppgift är att erbjuda personer med synnedsättningar och blindhet en rikare och mer detaljerad förståelse och upplevelse av film, teater eller ett annat evenemang genom att beskriva det som syns men inte hörs. Syntolkning handlar både om att förmedla *innehållet* så att personer med synnedsättning inte missar viktig information som bara är synlig och att förmedla *upplevelsen* genom att framkalla inre bilder och känslor hos målgruppen.

För att åstadkomma en effektiv syntolkning måste syntolken välja ut relevant information från den visuella scenen (miljön, händelser, personer, deras utseende, ålder, ansiktsuttryck, gester och kroppsrörelser) och uttrycka denna information språkligt med hjälp av målande beskrivningar på ett sätt som tillgodoser behoven hos den blinda och synskadade publiken (Holsanova 2016a,b). Syntolkens beskrivningar aktiverar inre bilder och föreställningar hos användare och gör det lättare att förstå och leva sig in i vad som händer (Johansson 2016). Syntolken måste även tidsanpassa informationen om det visuella skeendet, så att den på ett tillfredställande sätt synkroniseras med ljud- och talinformation från filmen. En mycket viktig princip är att syntolken inte pratar sönder dialogen utan beskriver helst i pauserna. På grund av tidspressen kan syntolken inte beskriva allt som syns utan måste fatta beslut om vad som behöver beskrivas och vad som är relevant för den fortsatta handlingen. Syntolken brottas hela tiden med frågan VAD som ska syntolkas, HUR det ska syntolkas, och NÄR det ska syntolkas (Holsanova 2016a).

Forskningsinitiativen kring syntolkning

Syntolkning är ett ungt forskningsfält. Internationellt finns så vitt vi vet inte mer än fyra doktorsavhandlingar, fem samlingsvolymmer och ett åttioatal publikationer som handlar om olika aspekter av syntolkning (Pedersen 2016). För bara några år sedan fanns ingen forskning alls om syntolkning i Sverige eller Skandinavien. De första svenska forskningsinitiativen startade för tio år sedan i form av workshops i syntolkning som anordnades av Jana Holsanova, Lunds universitet, och Cecilia Wadensjö, Stockholms

universitet, åren 2010–2014. Workshopen ledde fram till rapporten *Syntolkning: forskning och praktik* (2016) som publicerades som tryckt bok, talbok och i punktskrift och är den första boken av sitt slag i Skandinavien.

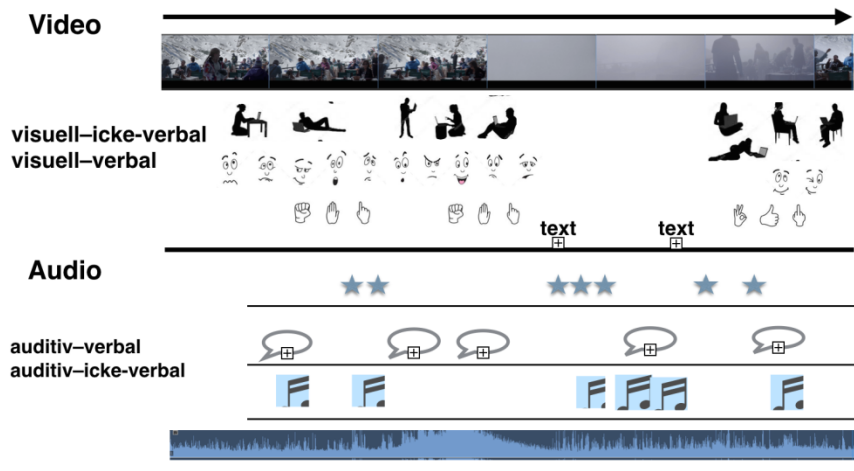
Bidraget fokuserar på syntolkning av film ur ett produktionsperspektiv (hur syntolkare gör) och ett receptions perspektiv (hur syntolkningen tas emot av användare). Meningsskapande i kontext illustreras med en syntolkad narrativ filmscen och kommenteras utifrån språkvetenskaplig och kognitiv forskning, i synnerhet teorier om relevans (Wilson & Sperber 2004; Forceville 2014), mentala bilder (Holsanova, Hedberg & Nilsson 1999; Holsanova 2008; Johansson et al. 2006, 2011; Johansson 2016) och multimodalitet (Kress & van Leeuwen 2006[1996]; Holsanova 2014a). Analysen fokuserar på vad, hur och när det syntolkas under tre faser i berättelsen: orientation, complication och resolution (Labov & Waletzky 1967).

Produktion och reception av komplexa multimodala texter

Inom multimodalitetsteori (Kress & van Leeuwen 2006[1996]) utgår forskare från ett vidgat textbegrepp (Gunnarsson & Karlsson 2007; Björkvall 2009; Holsanova 2010). Texter är inte enbart rent språkliga fenomen utan innehåller visuella element, t.ex. bilder, grafer, kartor, diagram och grafiska arrangemang, som skapar mening tillsammans med skriften. När man tillämpar multimodalitetsteori på film blir komplexiteten ännu större. Filmen består av olika uttrycksmodaliteter och innehållsliga spår som vävs samman för att berätta en historia. Som en komplex multimodal text består filmen av olika konfigurationer av fyra grundläggande uttrycksmodaliteter (semiotic modes) (Zabalbescoa 2008): (1) visuellt – icke-verbalt (t ex bilder, gester och kropps rörelser), (2) visuellt – verbalt (t ex text på skärmen), (3) auditivt – verbalt (t ex dialog) och (4) auditivt – icke-verbalt (t ex musik och ljudeffekter).

För att kunna kartlägga vad syntolken gör framställer jag filmen schematiskt som ett partitur med flera spår där olika uttrycksmodaliteter är tidskoordinerade och samspelar innehållsligt (Fig.1). För analys syften har jag delat in partituret i ett videospår och ett audiospår för att visa vad syntolken måste tillgängliggöra via syntolkning och vad den icke-seende publiken uppfattar från originalfilmen. Tiden går horisontellt, från vänster till höger. Videospåret i filmen har en stor betydelsepotential. Filmen innehåller visuella scener som visar miljöer, föremål och karaktärer. Karaktärerna har ett visst utseende, är involverade i olika aktiviteter och kommunicerar med varandra icke-verbalt via gester, blickar, ansiktsuttryck och kropps rörelser. Ibland visas även stillbilder och text i bild. Det första spåret representerar alltså det som syns i filmen (visuellt – icke-verbalt och visuellt – verbalt). Det auditiva spåret representerar det som hörs i filmen: dels bakgrundsljud, ljudeffekter (markerade med stjärnor), musik som stödjer handlingen (auditivt – icke-verbalt), dels dialoger där karaktärerna

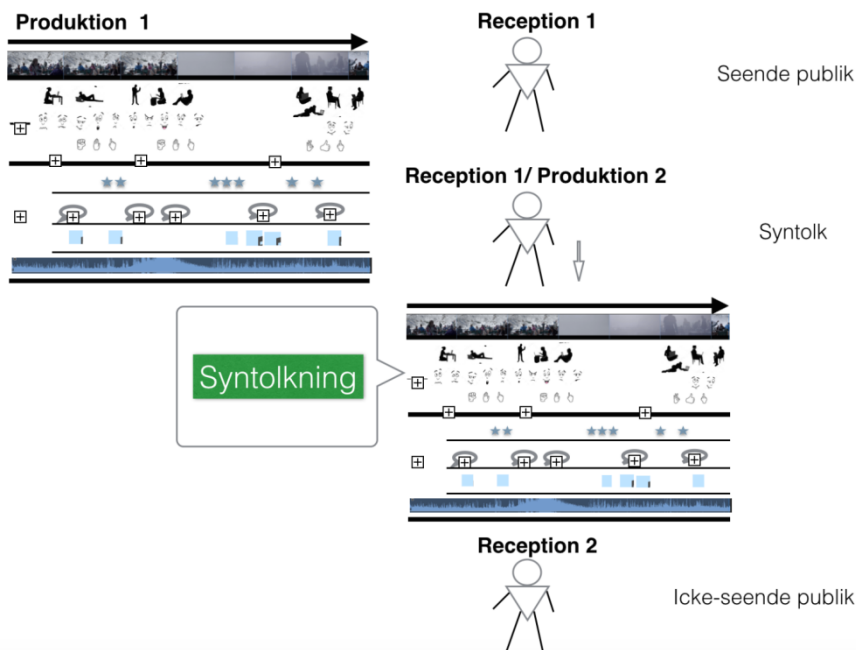
kommunicerar med varandra språkligt samt sång med text (auditivt – verbalt).



Figur 1. Schematisk bild över film som multimodal partitur.

I analysen av komplexa multimodala texter kan man utgå från produktions- eller receptionsperspektivet (Bucher 2010; Holsanova & Nord 2010; Holsanova 2014a). Produktionsperspektivet fokuserar på hur den komplexa multimodala texten presenteras med hjälp av olika uttrycksmodaliteter, såsom skrivet och talat språk, bilder och rörliga bilder, grafik, gester, ljud, musik, och hur dessa resurser samspelar för att bidra till en viss kommunikativ effekt (O'Halloran et al., 2012; Kress & Van Leeuwen 2006[1996]). Receptionsperspektivet fokuserar på hur den multimodala texten tas emot av användarna. Mottagandet är nära kopplat till användarnas förmåga att välja och sortera information, koppla samman information från olika källor och att fylla i information utifrån deras förkunskaper, erfarenheter och associationer (Holsanova 2014a,b).

Syntolkningen sker antingen live, direkt på plats i kontakt med användargruppen, eller är manusbaserad och förinspelad på en DVD-skiva eller i en app. I nästa steg kommer jag att visa hur produktions- och receptionsprocesser samverkar under live syntolkning och hur innehållet förmedlas till den seende och icke-seende publiken. Jag kommer att fokusera på syntolkning av film som sker offentligt på bio, där syntolken, via mikrofon och hörlurar, tolkar för en grupp av brukare som sitter i salongen tillsammans med den seende publiken.



Figur 2. Schematisk bild över produktions- och receptionsprocesser under syntolkning av film.

Resultatet av filmproduktionen (produktion 1, P1) presenteras för publiken i biosalongen. I biosalongen sitter både seende publik (inklusive den seende syntolken) som tar emot och tolkar budskapet (reception 1, R1), samt icke-seende publik som får syntolkning i hörlurar. Syntolken har i regel sett filmen en eller två gånger innan, antecknat ett antal nyckelord och producerar sedan en talad live syntolkning sittandes i biosalongen. Syntolkningen (produktion 2, P2) skräddarsys och anpassas till målgruppens behov och preferenser. Detta sker samtidigt med reception av originalfilmen (R1). Det innebär att syntolken utifrån det multimodala samspillet måste granska vilken information som förmedlas av ljuden, musiken och dialogerna (och kan uppfattas av den icke-seende publiken) och vilken information som enbart uttrycks visuellt (och således inte kan uppfattas av den icke-seende publiken). På basis av den bedömningen väljer syntolken vad som behöver syntolkas, hur och när. Det vore dock en förenkling att säga att syntolken enbart ersätter den visuella delen av partituret med den verbala. Man skulle snarare kunna påstå att syntolken kompletterar det som saknas i det multimodala samspillet för att uppnå en jämförbar förståelse och upplevelse hos den icke-seende publiken. Ibland behöver nämligen målgruppen även förklaring till vissa ljud som är svåra att tolka utifrån kontexten (t.ex. vad har hänt?, vems steg hörs?, vem öppnar dörren?, vem har skjutit vem?). Syntolkningen måste alltså tidsmässigt och innehållsligt

synkroniseras med de andra auditiva spåren. Syntolken skapar därmed en tillgänglig variant av produktionen (P2) som når den icke-seende publiken. Syntolkningen tas slutligen emot av den icke-seende publiken samtidigt med musiken, ljuden och dialogerna från originalfilmen (reception 2, R2).






I nästa steg kommer jag att analysera en narrativ filmscen i detalj med fokus på vad, hur och när det syntolkas. Jag kommer att synliggöra hur syntolken väljer ut relevant information, beskriver miljön, aktörer och händelser, synkroniserar beskrivningen med innehållet i film ljudet och skapar dramatik med hjälp av rösten.

Analys av en syntolkad filmscen: Vad gör syntolken?

Filmscenen som vi kommer att titta närmare på är ca 4 minuter lång och kommer från filmen *Turist* (2014) av Ruben Östlund¹. Vad gäller scenens uppbyggnad handlar det om en typisk berättelse med olika faser (Labov & Waletzky 1967). Den inleds med *orientation* där miljön, aktörerna och aktiviteterna introduceras (00:00–00:44), fortsätter med *complication* där en dramatisk händelse skildras som aktörerna reagerar på (00:44–01:51) och avslutas med *resolution* som visar hur problemet har lösts och hur den dramatiska händelsen påverkat aktörernas beteende (01:51–03:29). En explicit *coda*, dvs. generalisering eller moral, som ibland finns i slutet av berättelser formuleras inte utan moralen byggs upp stegvis under hela filmens lopp.

Vi ska nu närmare titta på hur syntolken gör för att skapa förståelse och inlevelse hos personer med synnedsättning och blindhet. För att kunna följa samspelet mellan alla innehållsspåren och syntolkningen har jag utformat en multimodal transkription (Figur 3). Tiden går vertikalt i transkriptionen. Horisontellt i transkriptionen visas de olika visuella och auditiva innehållsspåren. Längst till vänster finns ett videospår som representeras av en stillbild från filmen. Stillbilden är utvald så att den på bästa sätt fångar upp händelsen som beskrivs i syntolkningen. Sedan följer audiospåret med beskrivningen av de bakgrundsljud som hörs under tiden samt tidsangivelsen. Längre till höger syns dialogen som jag har transkriberat direkt från filmen. Repliker är försedda med namnförkortningar (T= pappa Tomas, E= mamma Ebba, H= son Harry, V= dotter Vera). Längst till höger följer transkriptionen av syntolkningen. Från relationen mellan syntolkningen i förhållande till dialogen (de sista två spalterna sedda uppifrån ner) kan det avläsas hur syntolken utnyttjat pauserna för att beskriva det visuella skeendet.

1 Syntolkningen som analyseras vann ett förstapris i tävlingen Guldtolken där elva syntolkningsföretag tävlade om att på bästa sätt tolka ett avsnitt ur Ruben Östlunds film *Turist*. Tävligen organiserades av Tillgänglig bio på Svenska filminstitutet 2015. Stort tack till vinnaren Per Lagergren från Ljudlager som gav mig tillstånd att använda hans syntolkning för vetenskapliga syften.

	Videospår	Ljudspår	Dialog	Syntolkning
Orientation		restaurangljud buller från en helikopter 00:11		En fullsatt uteservering med panoramat över berget. Det är soligt och klart, gäster i skidkläder. Tomas, Vera, Ebba och Harry sitter vid ett bord.
		restaurangljud buller från en helikopter steg tallrikar 00:16	T to H: Ge mig mobilen	En servitris kommer med tre tallrikar.
		restaurangljud bestick tallrikar 00:25	E: Merci beaucoup T: Hm. Det ser bra ut. E: Det ser väldigt bra ut.	Hon serverar Harry, Vera and Ebba och går iväg.
		restaurangljud bestick tallrikar steg 00:36	T: Merci	De börjar äta. Servitrisen kommer tillbaka med Tomas tallrik.
		restaurangljud bestick tallrikar röster 00:55	E: Detta var jättegott. Smaka. T: Hmm. Jättegott. V: Finns det ingen parmesanost?	Ebba sträcker fram gaffeln. Föräldrarna söker med blicken på bordet.

Figur 3. Multimodal transkription (Orientation).

Som vi har konstaterat tidigare måste syntolken hela tiden fatta beslut om VAD som ska syntolkas, HUR det ska syntolkas och NÄR det ska syntolkas. Jag ska kommentera några av dessa beslut i de tre faser av berättelse, orientation, complication och resolution. Vi ska även titta på hur syntolken arbetar med rösten.

Orientation. Första fasen av miniberättelsen (00:00–00:58) inleds med en introduktion som skapar sammanhang och ramar in berättelsen. Introduktionen ska vara kort och kärnfull och snabbt fånga in miljö, tid, karaktärer och deras aktiviteter eftersom tiden är begränsad. Filmscenen inleds med följande innehållstät syntolkning som presenteras med en lugn röst:

En fullsatt uteservering med panoramat över berget.
Det är soligt och klart, gäster i skidkläder.
Tomas, Vera, Ebba och Harry sitter vid ett bord.

I originalfilmen samspelar flera innehållsspår som hjälper publiken att tolka händelser. För att kunna välja ut relevant information från scenen måste syntolken analysera det multimodala samspelet och besvara en rad frågor: Vad saknas det för information som uttrycks enbart visuellt men inte artikuleras i språk eller ljud? Vad förmedlar ljuden? Vilka ljud kan inte

förstås från sammanhanget och måste förklaras? Vilken visuell sceninformation är relevant att förmedla? Vad är viktigt för fortsatt utveckling av handlingen? Vilken information är oviktig i sammanhanget och kan sorteras bort? Vad gäller valet av relevant information i den valda filmscenen står familjen i fokus av beskrivningen medan de övriga personerna på serveringen utelämnas. Familjemedlemmarna beskrivs dock inte i detalj eftersom de är kända sedan tidigare. Syntolkningen samspelar i många fall med film ljuden. Exempelvis stärks syntolkens formulering ”fullsatt uteservering” av restaurangljuden som hörs från filmen och utomhusmiljön bekräftas av en surrande helikopter. Även matserveringen och ätandet stöds av bakgrundsljudet med bestick och tallrikar som skramlar.

Förutom valet av relevant information (VAD som ska syntolkas) måste syntolken även bedöma NÄR den språkliga beskrivningen av det visuella innehållet ska ges. Redan i första fasen av berättelsen hittar vi exempel på hur syntolkningen tidsanpassas så att den samspelar med vad som sägs i replikerna och vad som hörs i ljuden. Synkroniseringen sker vid syntolkning av icke-verbala handlingar som visas i filmen och beskrivningen formuleras i pauserna mellan dialogerna. Exempelvis efterföljs beskrivningen ”Hon serverar Harry, Vera och Ebba ... och går iväg” av ljudet av servitrisens fotsteg och beskrivningen ”Servitrisen kommer tillbaka med Tomas tallrik” synkroniseras med Tomas svar: ”Merci”. Mammans kommenterar ”Detta var jättegott. Smaka.” följs av beskrivning av hennes icke-verbala handling ”Ebba sträcker fram gaffeln” och ryms exakt i pausen före Tomas svarsreplik: ”Hmmm. Jättegott.” Slutligen beskrivs föräldrarnas icke-verbala reaktion på Veras fråga ”Finns det ingen parmesanost?” direkt efter Veras replik: ”Föräldrarna söker med blicken på bordet.”

Complication. I andra fasen av miniberättelsen (00:58–01:44) skildras en dramatisk händelse som aktörerna reagerar på. Här skapas dramatik och intensitet med både visuella och akustiska uttrycksmedel. Publiken får en känsla av en kommande fara, det dramatiska händelseförloppet trappas upp och spänningen växer. Vad gäller syntolkning ska vi titta närmare på HUR sceninnehållet beskrivs och hur syntolken arbetar med rösten. Följande frågor blir intressanta: Hur skapar syntolken levande beskrivningar som stimulerar föreställningsförmåga och inlevelse hos mottagaren? Hur fångas stämningar och känslor? Hur åstadkommer syntolken en kärnfull beskrivning när det är ont om tid och mycket ljud och skrik? Hur detaljerad är beskrivningen? Hur väljer syntolken begrepp med tanke på mottagarens tolkning? Först tittar vi på den multimodala transkriptionen (Fig 4).







Complication		Smäll! 01:03	E: Oj! T: Kolla här nu. Kolla här.	Alla vänder sig ut mot berget där ljudet kom ifrån.
		Röster i bakgrunden Wow! 01:09	E: Är det en lavin eller? T: Ja men den är kontrollerad.	Tomas tar upp mobilen och filmar.
		01:15	T: Det är såna jäkla krafter! E: Är det säkert? T: De vet vad de gör.	Ett snömoln kommer farande ner för berget.
		Högt ljud av växande intensitet och volym 01:18	H: Pappa!	Det kommer närmare ...
		Högt ljud av växande intensitet och volym 01:20	H: Pappa!! T: Harry, Harry, det är lugnt. Det är lugnt.	... och närmare... ..
Complication		Högt ljud av växande intensitet och volym Människor skriker 01:25	H: Pappa! T: Harry!	Folk reser sig. Och flyr. I panik!
		Högt ljud av växande intensitet och volym Människor skriker Stolar faller omkull 01:33	H: Pappa!	Tomas knuffar skrikandes en bordsgranne framför sig och försvinner.
		Högt ljud av växande intensitet och volym Människor skriker! 01:44		Ebba och barnen blir kvar. Allt sveps in i den grå massiva snöröken.





Figur 4. Multimodal transkription (Complication).

Genom beskrivningen ”Alla vänder sig mot berget där ljudet kommer ifrån” uppmärksammar syntolken dels ljudet och dess ursprung, dels specificerar han restauranggästernas icke-verbala beteende. I dialogen nämns ”lavin” men syntolken använder begreppet ”snömoln” i den kärnfulla beskrivningen ”Ett snömoln kommer farande ner för berget.” för att fånga den annalkande faran. Tempot och volymen av hans röst ökar då lavinen kommer. Syntolken kan tygla sin röst men ändå behålla spänning. När volymen och intensiteten av film ljuden och människoskriken kulminerar håller syntolken tillbaka och låter ljuden och skriken bära dramatiken. Syntolken bidrar till stämningen och stegringen av spänningen genom att foga in mycket korta beskrivningar i de mikropauserna mellan människoskriken: ”Det kommer närmare ... och närmare. ... Folk reser sig. ... Och flyr. ... I panik!” Syntolken använder några sekunder av den tystare stunden som följer efter skriken till att i

efterhand beskriva den tumult som syntes i filmen men som inte kunde beskrivas under den mest högljudda, hysteriska delen av scenen: ”Tomas knuffar skrikandes en bordsgranne framför sig och försvinner. Ebba och barnen blir kvar. Allt sveps in i den grå massiva snöröken.” Valet av begreppen ”snömoln” och ”snörök” är intressant och viktigt. Dessa begrepp skapar andra associationer än *lavin* när personer med synned sättning och blindhet ska föreställa sig vad som försiggår. Ordvalet får konsekvenser för det scenario eller den inre bild som målgruppen bygger upp. Ordvalet påverkar också målgruppens förväntningar på vad som kommer att hända härnäst (t.ex. hur människor och miljön kommer att påverkas).

Resolution. Tredje fasen av miniberättelsen (01:44–03:29) skildrar hur den dramatiska händelsen har lösts och hur den påverkat miljön och karaktärernas agerande. Denna fas av berättelsen börjar med en lugnare passage då allting är vitt och inget syns i bilden under en lång tid (ca 20 sekunder). Man får höra några fotsteg och stolar som flyttas och spänningen kvarhålls. Vad gäller syntolkning ska vi titta närmare på HUR syntolken hanterar den långa tystnaden efter lavinen, hur han verbaliserar att man inte ser något på bild och hur han beskriver scenen när personer blir mer och mer synliga. Dessutom kommer vi in på hur det icke-verbala beteendet, karaktärernas blickar och känslor syntolkas. Vi utgår återigen från den multimodala transkriptionen (Fig 5).

		Tystnad Svagt tutande ljud 01:51		
		stolar flyttas steg 01:57		I diset. En gestalt. Något som ser ut som änglarvingar dyker upp.
Resolution		stolar flyttas 02:13	F: Sava? Vous bien? Are you ok? E: Är det bra? F: Safe to come up now.	Gestalten försvinner bort. Snöröken lättar.
		stolar flyttas steg 02:22	E: Är det ok? H: Hm.	Harry får hjälp att ställa sig upp.
		stolar flyttas steg bestick 02:36	E: Är det ok? H: Hm. E: Ja? ja? Det gick bra? H: Var är pappa?	Ebba och Vera reser sig från golvet.
		steg 02:46		Folk börjar komma tillbaka. Snöröken löser upp sig, fler och fler personer blir synliga.

Resolution		stolar flyttas steg bestick röster	T: Det var ...	De sätter sig vid sina bord. Tomas kommer tillbaka. Han skrattar generat till bordsgrannen ... och sätter sig.
		restaurangjud bestick röster stolar 03:01	T: Jäklar! Är det bra?!	Ebba borstar bort snön och sätter sig hon också. Ebba kastar en snabb blick på Tomas.
		restaurangjud röster tallrikar 03:27	T: ... Va fan! Det där var ... T: Där kommer solen.	Barnen tittar ner i bordet. Gestalt med änglarvingarna. En man med utlöst lavinrycksäck på ryggen.
Resolution		stolar flyttas steg bestick röster	T: Det var ...	De sätter sig vid sina bord. Tomas kommer tillbaka. Han skrattar generat till bordsgrannen ... och sätter sig.
		restaurangjud bestick röster stolar 03:01	T: Jäklar! Är det bra?!	Ebba borstar bort snön och sätter sig hon också. Ebba kastar en snabb blick på Tomas.
		restaurangjud röster tallrikar 03:27	T: ... Va fan! Det där var ... T: Där kommer solen.	Barnen tittar ner i bordet. Gestalt med änglarvingarna. En man med utlöst lavinrycksäck på ryggen.

Figur 5. Multimodal transkription (Resolution).

Skriken följs av tystnad och att inget syns i bilden. Hur beskriver man en sådan sekvens? Syntolken har redan i slutet av förra avsnittet sagt att ”*Allt sveps in i den grå massiva snöröken*” och markerat att inget syns i bilden. Därefter är han tyst och låter de låga bakgrundsljuden (tutande, fotstegen och stolar) spela sin roll och hålla kvar spänningen hos publiken som funderar på hur den dramatiska händelsen har slutat. Den seende publiken får en kontinuerlig visuell återkoppling om att inget syns medan den icke-seende publiken behöver påminnas om det då och då verbalt. Lite senare i pausen går därför syntolken in och uppdaterar den icke-seende publiken om att synligheten fortfarande inte är så bra: ”*I diset.*” I nästa steg beskriver han något som dyker upp: ”*En gestalt. Något som ser ut som änglarvingar.*” Lite senare återkommer han till gestalten igen och markerar på olika sätt synligheten blivit bättre: ”*Gestalten försvinner bort. Snöröken lättar. Folk börjar komma tillbaka. Snöröken löser upp sig, fler och fler personer blir synliga.*” Syntolken tar upp tråden igen i slutet av scenen med en förklaring av änglarvingarna: ”*Gestalt med änglarvingarna. En man med utlöst lavinrycksäck på ryggen.*” Syntolken har därmed inte bara beskrivit hur synligheten har förändras utan även indirekt förmedlat hur dynamiken i en

seendes perception fungerar: att man inte alltid känner igen alla objekt på en gång och att igenkänningen och kategoriseringen sker stegvis. Man börjar med en gissning eller hypotes, letar sedan efter kännetecknen som bekräftar den tolkningen och kan slutligen i bästa fall identifiera objekten. För personer som aldrig har sett kan detta vara intressant att veta.

Vad gäller hantering av den långa pausen kan vi konstatera att pausen blivit lite kortare för den icke-seende publiken. Början av pausen har fyllts, i efterhand, med en beskrivning av den tumult som syntes i bild men inte kunde beskrivas under den högljudda scenen och slutet av pausen har fyllts, på förhand, med beskrivningen av gestalten med änglarvingarna. Ibland behövs lite extra tid för att nämna viktiga händelser som man inte hunnit beskriva mellan dialogerna. Detta är dock en balansgång för syntolken. Pausen får å ena sidan inte vara för lång så att målgruppen tror att det är ett tekniskt fel när de under en lång tid inte hör någon syntolkning. Å andra sidan bör helst pausen inte fyllas helt då det hade kunnat förstöra den stämning som ljuden och musiken uppfyller.

Beskrivningen av scenen i denna tredje fas av berättelsen (resolution) står dels i relation till scenbeskrivningen i första fasen (orientation) och dels till skildringen av den dramatiska händelsen i andra fasen (complication). Medan stämningen i första fasen är lugn och trevlig – vilket också stöds av ljuden, dialogerna och syntolkningen – hörs i den andra, dramatiska fasen människoskrik och högt ljud av växande intensitet och volym som vanligtvis signalerar en stor fara. Det kan framkalla en (inre) bild full av rädsla, kaos och panik hos den icke-seende publiken. Utifrån syntolkningen i tredje avsnittet kan dock den icke-seende publiken dra slutsatsen att följderna efter den dramatiska händelsen inte blivit så allvarliga som befarat. Det var bara snörök och snödis och ingen stor katastrof har skett. Alla aktörer kunde återvända till sina platser: *"Folk börjar komma tillbaka. Snöröken löser upp sig, fler och fler personer blir synliga. De sätter sig vid sina bord."*

Under sista minuten sägs det mycket lite i dialogerna. I synnerhet Tomas repliker är korta och avbrutna: *"Det var ... Jäklar! Det där var ..."* Det mesta uttrycks implicit, genom skildring av karaktärernas agerande och icke-verbala beteende. Publiken har tidigare fått veta att Tomas har knuffat sig förbi bordsgrannen och försvunnit från serveringen medan Ebba och barnen stannat kvar. När allt lugnat ner sig frågar Harry: *"Var är pappa?"* Syntolken uppmärksammar barnens och Ebbas icke-verbala reaktion på Tomas återkomst: *"Tomas kommer tillbaka. Han skrattar generat till bordsgrannen...och sätter sig.... Ebba kastar en snabb blick på Tomas. Barnen tittar ner i bordet."* Det är en viktig markering för att beskriva hur spänd stämningen blivit mellan dem. Det som antyds genom blickar och med icke-verbalt beteende säger mycket om karaktärernas känslor och relationer och ska aktivt tolkas av publiken. I samband med formuleringen *"Han skrattar generat"* uppstår frågan: Ska syntolken uttrycka och värdera känslorna explicit? Ska syntolken berätta det som ska läsas mellan raderna

eller hörs känslor och sinnesstämningar alltid på karaktärernas röster? Det finns olika åsikter om det bland syntolkare. Det är ett område där syntolken måste använda sin fingertoppkänsla.

En annan klassisk frågeställning inom syntolkning är hur *objektiv* en bildbeskrivning ska vara respektive hur *subjektiv* den får vara. Idealet som formuleras i en del riktlinjer är ”beskriv bara det du ser”. Detta är dock inte realistiskt eftersom vi delvis ’ser’ olika saker. Vår perception färgas nämligen av den erfarenhet, kunskap och expertis som vi har förvärvat och samlat på oss (Holsanova 2014b). Det leder till att vi ’fyller i’ information på basis av våra förkunskaper, förväntningar och associationer. Det är alltså svårt att tala om en objektiv bildbeskrivning. Samtidigt får beskrivningen inte vara värderande eller dra över åt det spekulativa hållet. Det viktiga är att personer med synnedsättning och blindhet själva kan skapa mening i kontext, utifrån syntolkningen och andra uttrycksmodaliteter från originalfilmen.

Hur tas syntolkning emot av målgruppen?

Reception av filmberättelser är en mycket komplex och multimodal process där publiken måste hålla reda på vem som gjorde vad mot vem, var, när och varför för varje händelse i historien och ständigt uppdatera denna information i enlighet med hur berättelsen utvecklas (Vercauteren & Remael 2014; Zwaan & Radvansky 1998). Jag har inledningsvis nämnt att syntolkningen aktiverar inre bilder och föreställningar hos den blinda publiken vilket gör det lättare att leva sig in i handlingen. Frågan är hur personer med synnedsättning och blindhet förstår språkliga beskrivningar? Vilka beskrivningar föredrar de? Hur föreställer de sig miljöer, karaktärer, händelser? Vilka inre bilder skapar de?

Det finns i dag en hel del forskning som visar att även personer som är blinda från födseln skapar inre föreställningar och upplever en sorts mentala bilder. Forskningen visar dock att blindas mentala föreställningar har en annorlunda karaktär än de som seende har (Cattaneo & Vecchi 2011, Johansson 2016). De blindas inre föreställningar är ofta mer rumsligt orienterade och bygger på känsel. Vad gäller informationsbearbetning har seende möjlighet att bearbeta stora bitar av informationen samtidigt för att förstå större helheter, globala strukturer i en visuo-spatial miljö. Blinda kompenserar för den här typen av begränsningar genom att kombinera sina mentala bilder med abstrakt semantisk kunskap. De föredrar i stället en sekventiell metod för informationsbehandling och tar ett första-persons-perspektiv när de förflyttar sig från en plats till en annan. Denna navigationsstrategi bygger på att identifiera hur viktiga landmärken längs vägen är placerade i förhållande till den egna kroppen och syftar till att knyta ihop hållpunkter till en meningsfull enhet. Alla dessa ovannämnda faktorer har implikationer för syntolkning. Kunskap om hur seende och blinda personer bearbetar visuell och rumslig information är avgörande för

kommunikation i allmänhet samt för produktion och mottagning av syntolkning i synnerhet (Johansson 2016; Holsanova 2016a,b).

Även rösten och bakgrundsljuden är viktiga i syntolkning och dess mottagande (Fryer 2010). Forskning visar att rösten och andra akustiska egenskaper av talet spelar en viktig roll för hur syntolkning tas emot. Talhastighet och pauser har exempelvis stor inverkan på kognitiv bearbetning och förståelse av talad information (Eklund, 2004, Haake et al. 2014, Lyberg Åhlander et al. 2015). Intonation, tempo och pauser hjälper mottagaren att känna igen när en viss informationsbit börjar och slutar och vilken information som är viktig. Rösten kan antingen underlätta kognitiv bearbetning av det talade budskapet eller försvåra den och göra budskapet mer ansträngande att lyssna på. Med röstens hjälp kan man förtydliga innehållet så att lyssnarens uppmärksamhet fångas upp. Rösten är betydelsebärande vilket får konsekvenser för mottagning av syntolkning. I den dramatiska scen som analyserats ovan får förmodligen inte rösten låta för neutralt. Det måste finnas spänning i rösten men det får inte bli skådespeleri. Dessa aspekter behöver testas och utvärderas av målgruppen.

I våra användarundersökningar om icke-verbalt beteende ansåg personer med synnedsättning och blindhet att ansiktsuttryck, blickar och känslor ska syntolkas (Strukelj 2016). I en studie där olika grupper av användare fick bedöma olika syntolkningsversioner ansåg såväl de som varit blinda från födseln som de som haft synrester eller förvärvat synskada att de bättre hängde med i handlingen om ansiktsuttryck, blickar, kroppsrörelser och känslor beskrivits. De får bättre grepp om sammanhang, bättre möjlighet till inlevelse och förstår relationer bättre (Holsanova et al. 2015). Det behövs dock en systematisk undersökning av hur gester och kroppsrörelser syntolkas (Andrén 2016). Informanter i våra studier har betonat att syntolken måste använda sin fingertoppkänsla för att förmedla tillräckligt mycket men inte för mycket information. Dessa frågor går tillbaka till konversationsmaximerna och relevansteorin (Grice 1975; Wilson & Sperber 2004; Forceville 2014).

Diskussion

Braun (2007) definierar syntolkning som en komplex kognitiv-språklig och intermodal medierande aktivitet där kreativa meningsskapande processer under produktion och reception sammanfaller. Vad menas med det? Vad kännetecknar syntolkning och vilka utmaningar står syntolkare inför? Vilka teoretiska frågeställningar väcks i samband med syntolkning?

Syntolkning kallas för *intermodal översättning* eftersom syntolken överför innehåll från bild till ord (Jakobson 1959; Reivers 2017). En viktig teoretisk fråga i det sammanhanget är hur det visuella från en filmscen ska överföras till det språkliga i syntolkning. Svårigheten beror delvis på att språk och bild är två betydelskapande system som fungerar olika (Holsanova 1996, 1999). Bilder är analoga, har en rumslig dimension och

innehåller mycket information samtidigt. En bild säger mer än tusen ord, enligt det gamla ordspråket. Språket är linjärt och innehållet formuleras sekventiellt, i en viss temporär följd. Det finns inga självklara vägar genom bilden som kan jämföras med raderna i den skrivna texten avsedda för linjär läsning eller med strömmen av språkliga formuleringar som produceras och tas emot sekventiellt. Språk och bild har olika inneboende styrkor och svagheter och är därmed lämpliga att uttrycka olika typer av innehåll. Språket lämpar sig att uttrycka abstrakta tillstånd, relationer och begrepp (*sanning, rättvisa*), villkor, hypotetiska tillstånd (*om...så*), temporala relationer (*samtidigt, alltid, tre dagar senare*) och kvantifiering (*de flesta, flera, några*) som är svåra att uttrycka visuellt (Bernsen 1994; Lu & Bernsen 1995; Maybury 1995; Mullet & Sano 1995, Telemann 1989, 1994). Medan det generiska begreppet *träd* med fördel kan uttryckas språkligt visar bilden/filmscenen alltid ett specifikt träd (*tall, ek, björk*). Bilder är bra på att uttrycka spatiala relationer, att ge överblick, och att åskådliggöra ett föremåls storlek, utseende och läge (Holsanova 2010). Därför är det mycket svårt att språkligt beskriva innehållet av ett spridningsdiagram eller en karta (Holsanova 2019). Bilder påkallar uppmärksamhet, är lätta att känna igen och komma ihåg och används därför med fördel för omedelbar perceptuell verkan och identifikation. Det är bl.a. därför polisen föredrar ett foto på en rånare framför en verbal beskrivning.

En annan utmaning med intermodal översättning består i att bilder tillåter tolkning på olika nivåer (Holsanova 2011; Remael et al. 2016). Beskrivningen kan bli mer eller mindre abstrakt eller mer eller mindre detaljerad. Detta blir tydligt vid jämförelse av den vinnande syntolkningen av Lavinscenen med de andra tävlingsbidragen. Det faktum att åskådare vid tolkning av bilder påverkas av bl.a. förkunskaper, tidigare erfarenheter, intressen, associationer, förväntningar och kontexten har konsekvenser för syntolkning (Holsanova 2014a,b, 2016). En och samma scen kan uppfattas, tolkas och beskrivas olika av olika åskådare. Om vi zoomar in på Lavinscenen i ögonblicket då Tomas kommer tillbaka (resolution, ca 2:50 in i scenen) ser vi att olika aspekter fångas i de olika syntolkningarna. Beskrivningen varierar från en kort neutral beskrivning av Tomas ankomst (*"Pappan dyker upp", "Och där kommer Tomas"*) till en mer utvecklad beskrivning av hans sätt att gå, hans ansiktsuttryck, känslor och skor (*"Med flackande blick släntrar Tomas tillbaka till sin familj vid bordet", "Tomas kommer stolpande tillbaka i sina slalompjäxor", "Tomas kommer tillbaka. Han skrattar generat till bordsgrannen och sätter sig"*). En del beskrivningar omfattar redan vid denna tidpunkt familjens reaktioner på hans ankomst (*"Tomas närmar sig. Ebba vänder genast bort blicken", "Pappan letar sig tillbaks ut igen. Ebba vänder honom ryggen"*). Även abstraktionsnivån varierar när vi tolkar bilder. Medan en syntolk exempelvis beskriver en scen som *storstadsmiljö*, kan en annan specificera den som *Stockholm* och en tredje känna igen miljön som *Östermalm*. Eller ta

syntolkning av blickar, ansiktsuttryck och kroppsrörelser som är lätta att känna igen för seende personer men svåra att beskriva. Hur beskriver man exempelvis en karaktärs ansiktsuttryck då han precis fått höra något som han inte tror på? *Han tittar misstroget? Han sneglar på henne och ler skeptiskt? Han kisar misstänksamt?* Eftersom en rad individuella faktorer påverkar tolkningen är det oundvikligt att beskrivningen är färgad av den aktuella syntolken och kan variera. Frågan kring en objektiv respektive subjektiv syntolkning har vi tagit upp tidigare. Det är viktigt att beskrivningen inte är värderande eller drar över åt det spekulativa hållet samt att det finns utrymme för användares egen tolkning.

Den tredje utmaningen hänger ihop med *anpassningen* till användare. Syntolkar har en *medierande roll* mellan producenten och användaren. De måste respektera det kommunikativa syftet med filmen som producenten avsett och översätta budskapet på ett adekvat sätt. Samtidigt måste de tillgodose de kommunikativa behoven hos mottagarna och anpassa syntolkning till det (Reviere 2017: 34). Det kan tyckas vara självklart att vi anpassar formen och innehållet av språkliga yttranden till mottagaren och kontexten. Det sker i många situationer: under talproduktion, under interaktion i samtal, under språktolkning. Det specifika med syntolkningen är att syntolken måste leva sig in i blindas situation och beskriva bild-/sceninnehållet med tanke på målgruppens behov och preferenser, så att de förstår och kan leva sig in i handlingen. Att förmedla bilder till någon som inte ser är ingen enkel uppgift. Vissa begrepp och sakförhållanden måste förklaras med hänsyn till blindas referensramar och konceptualiseringar. Metaforer och liknelser tycks stärka specificiteten hos språket och väcka inre föreställningar hos mottagaren. Beskrivningen *hon som ser ut som en pudel i håret* gör det exempelvis lätt att identifiera en viss person i en större grupp. Frågan är dock om denna bildliga beskrivning väcker samma associationer och inre bilder hos personer som aldrig har sett.

Nästa utmaning vid överföringen mellan ord och bild hänger ihop med *tidsbegränsning*. Bilden/scenen innehåller mycket information men syntolken hinner inte beskriva allt man kan se. Idealet är att beskrivningen ska infogas i naturliga pauser mellan dialogen så att den inte stör viktiga repliker, ljud- och musikeffekter (Benecke 2004; Remael et al. 2014). Denna begränsning påverkar syntolkars beslut om vad, när och hur ska syntolkas. Syntolken måste välja visuell information som är viktig för att förstå handlingen. Sker filmberättelsen i hög takt, med snabba scenbyten och mycket dialog, blir det extra svårt. Syntolkar måste då prioritera information och överväga noga vad de ska inkludera för att uppnå en effektiv syntolkning som samtidigt inte belastar användare kognitivt (Vercauteren 2016).

Ytterligare en utmaning är kopplad till det *multimodala samspelet* (cf Figur 1). De olika uttrycksmödeliteterna är tidskoordinerade och samspekar innehållsligt. De är orkestrerade på ett bestämt sätt för att åstadkomma en

viss kommunikativ effekt. Varje uttrycksmodalitet har en viss betydelse och en 'relativ vikt' i textens övergripande betydelse. Betydelsen skapas inte på basis av de individuella modaliteterna, utan genom deras samspel. Syntolken måste därför ställa frågan: Vad saknas för information som uttrycks enbart visuellt men inte artikuleras i språk eller ljud? Vad förmedlar ljuden? Vilka ljud kan inte förstås från sammanhanget och måste förklaras?

Syntolkning kallas ibland för *partiell översättning* eftersom den endast översätter delar av filmen, d.v.s. bilderna, medan delar av originalet, d.v.s. ljud, musik och dialoger, förblir tillgängliga för publiken i sin ursprungliga form (Reviere 2017). Jämfört med språktolkning är alltså resultatet av syntolkningssprocessen inte en fristående produkt utan syntolkningen tas emot och bearbetas av användare tillsammans med de verbala yttrandena och ljudeffekterna från originalet. En av konsekvenserna är att syntolkare måste vara medvetna om samspelet mellan de olika uttrycksmodaliteterna i originaltexten och måste ta hänsyn till innebörden i ljuden, musiken och dialogen när de bestämmer vad som behöver beskrivas och hur det behöver beskrivas. Syntolken engageras därmed i en *multimodal medierande aktivitet*.

En intressant fråga är om det sker meningsförluster och meningsförändringar i syntolkning. En möjlig förlust gäller *redundans*. Filmmediet använder ofta flera, överlappande uttrycksmodaliteter. Det bidrar till att publiken lättare kan ta till sig betydelsen av en scen, eftersom flera semiotiska resurser ger ledtrådar till dess tolkning. I den syntolkade, tillgängliga, versionen av filmen går dock en stor del av denna redundans förlorad. Å andra sidan måste syntolkare i vissa situationer komplettera med ytterligare information för att förhindra informationsförlust. Det gör de genom att skapa *relevanta innehållsliga länkar* mellan syntolkningen, ljudet och dialogen så att publiken ska kunna rekonstruera berättelsen korrekt och effektivt (Reviere 2017; Remael & Reviere 2015). Till exempel, efter repliken "Vad är det?", bör syntolkningen innehålla en beskrivning av objektet som karaktären hänvisar till och som syns i bild. Likaså bör repliken "Gör så här" följas av beskrivning av handlingar och rörelser som visas. Om repliken "Vilket härligt väder" avser en ironisk tolkning måste den följas av en beskrivning av karaktärens blick mot fönstret där man kan se ihållande regn utomhus. Ironi och överdrift kan visserligen ibland höras på rösten men beskrivningen ger den icke-seende publiken tydliga ledtrådar till kontexten som är självklara för de seende. Tack vare denna komplettering kan samma kommunikativa effekt uppnås. Innehållsliga länkar omfattar även förklaring av ljud som inte är uppenbara och som måste kompletteras med informationen om vem gör vad eller vad som förorsakat ljuden: "Hanna ger Olga en örfil". "Marcus har vält stolen".

Slutsatser

Syntolkning har definierats som en språklig-kognitiv intermodal medierande aktivitet där kreativa processer i produktion och reception sammanfaller (Braun 2007). Syntolkning har också – enligt Roman Jakobsons taxonomi (Jakobson 1959) – karakteriserats som en intersemiotisk översättning mellan två teckensystem (bilder och språk). En del forskare karakteriserar syntolkning som partiell översättning eftersom den är tänkt att bearbetas av mottagare tillsammans med de andra uttrycksmodaliteterna från originalfilmen (Reviere 2017). Med analysen av en syntolkad filmscen ville jag illustrera att syntolkning handlar om en komplex meningsskapande process som inte låter sig begränsas till att syntolken enbart ersätter den visuella delen med en verbal del. Syntolken kompletterar snarare det som saknas i det multimodala samspelet för att uppnå en jämförbar förståelse och upplevelse hos den icke-seende publiken. För att kunna göra det måste syntolken göra en rad bedömningar och fatta en rad beslut, bl.a. om informationens relevans, val av språkliga uttryck och begrepp, tidsanpassning av informationen samt mottagarens behov och preferenser.

Forskningen är både teoretiskt intressant och samhällsrelevant². Forskningen om syntolkning ger oss nya kunskaper om kopplingen mellan språk, tänkande och meningsskapande och kan berika modeller av kognition och kommunikation. Den berör frågan om hur vi tänker, hur vi tar in information via olika sinnen, hur vi översätter mellan olika sinnen, hur vi formulerar visuellt innehåll språkligt samt hur vi föreställer oss saker och händelser och skapar inre bilder. Forskningsresultaten kan dessutom direkt tillämpas i syntolkningspraktiken till att utbilda fler syntolkare och vidareutbilda erfarna syntolkare. Sist, och viktigast av allt, har forskningen om syntolkning direkta implikationer för användare. Den bidrar till att öka tillgänglighet för personer med synnedsättning och blindhet och möjliggör för dem att bättre förstå och uppleva omvärlden.

Det behövs dock mycket mer forskning på området syntolkning, ur både produktions- och receptionsperspektiv. Ett område som vi fortfarande vet för lite om är syntolkars förståelse och bearbetning av det multimodala originalet. Framtida forskning behöver därför kartlägga syntolkars meningsskapande och medierande aktiviteter i detalj. Hur förstår syntolkare meningspotentialen hos verbala, visuella och auditiva uttrycksmodaliteter samt deras samspel? Hur väljer syntolkare relevant information som inte hörs men som behövs för förståelse av helheten? Vilka bedömningar gör de i frågan när och hur denna information ska syntolkas, i synnerhet när det är ont om tid? Ett annat område som vi fortfarande vet för lite om är hur

2 Under de senaste tio åren har författaren byggt upp ett stort nätverk bestående av forskare, syntolkare, syntolksutbildare, personer med synnedsättning och blindhet och företrädare för statliga myndigheter och intresseorganisationer (MTM, SPSM, SRF, US) som engagerar sig i syntolkning. Författaren är också koordinator för samverkansinitiativet "Syntolkning och bildbeskrivning för tillgänglig kommunikation och social inkludering" som nyligen har fått medel från Lunds universitet.

syntolkningen tas emot av målgruppen. Framtida forskning behöver kartlägga användares behov, preferenser, associationer och föreställningar och systematiskt undersöka hur personer med synnedsättning och blindhet förstår och upplever syntolkning. Vilka beskrivningar föredrar de? Hur föreställer de sig miljöer, karaktärer och händelser? Vilka inre bilder skapar de? Dessa undersökningar behövs för att kunna formulera språkliga och kommunikativa strategier för en effektiv syntolkning. Ett viktigt steg i denna riktning görs inom ett pågående forskningsprojekt vid Lunds universitet³.

Referenser

- Andrén, Mats (2016), "Att beskriva gester och handlingar", i Jana Holsanova, Cecilia Wadensjö & Mats Andrén (2016) (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier, rapport nr. 4, s. 39 – 45.
- Benecke, Bernd (2004), "Audio-Description", *Meta: Translators' Journal*, 49(1): 78–80.
- Bernsen, N.O. (1994), "Foundations of multimodal representations. A taxonomy of representational modalities", *Interacting with Computers* Vol. 6 No. 4/1994: 347–71.
- Björkvall, Anders (2009), *Den visuella texten - Multimodal analys i praktiken*. Lund: Studentlitteratur.
- Braun, Sabine (2007), "Audio description from a discourse perspective: a socially relevant framework for research and training", *Linguistica Antverpiensia* NS 6: 357–369.
- Bucher, Hans-Juergen (2010), "Multimodalität – eine Universalie des Medienwandels: Problemstellungen und Theorien der Multimodalitätsforschung", i Hans-Jürgen Bücher, Thomas Gloning, Kathrin Lehnen (red), *Neue Medien - Neue Formate. Ausdifferenzierung und Konvergenz in der Medienkommunikation*. Frankfurt am Main, New York: Campus Verlag, s. 41–79.
- Cattaneo, Zaira & Vecchi, Tomaso (2011), *Blind vision. The Neuroscience of Visual Impairment*. Cambridge: MIT Press.
- Eklund, Robert (2004), *Disfluency in Swedish human-human and human-machine travel booking dialogues*. Doctoral dissertation. Linköping University Electronic Press.
- Forceville, Charles (2014), "Relevance theory as a model for analysing visual and multimodal communication", I David Machin (red), *Visual communication*, Mouton - De Gruyter, s. 51–70.
- Fryer, Louise (2010), "Audio Description as Audio Drama — a practitioner's point of view", *Perspectives: Studies in Translatology*, 18(3): 147–172.

³ Författaren har tillsammans med två forskarkollegor från Lunds universitet tilldelats medel från Forskningsrådet för hälsa, arbetsliv och välfärd (FORTE) för ett treårigt tvärvetenskapligt projekt "Hur den blinda publiken förstår och upplever syntolkning av visuella händelser" (2019-2021).

- Grice, Paul H. (1975), "Logic and Conversation", i Peter Cole & Jerry L.Morgan (red), *Syntax and Semantics* Vol. 3. Speech Acts. New York: Academic Press 1975, s. 41–58.
- Gunnarsson, Britt-Louise & Karlsson, Anna-Malin (red) (2007), *Ett vidgat textbegrepp*. TeFa rapport 46. Uppsala universitet: Uppsala.
- Haake, Magnus, Hansson Kerstin, Gulz Agneta, Schotz Susanne & Sahlen Birgitta (2014), "The slower the better? Does the speaker's speech rate influence children's performance on a language comprehension test?", *Int J Speech Lang Pathol*, 16(2): 181–90.
- Holsanova, Jana (2019), *Bildbeskrivning för tillgänglighet - riktlinjer, forskning och praktik*. Myndigheten för tillgängliga medier, rapport nr. 6.
- Holsanova, Jana (2016b), "Cognitive approach to audio description", i A. Matamala & P. Orero (red), *Researching audio description: New approaches*. London: Palgrave Macmillan, s. 49–73. ISBN 978-1-137-56916-5.
- Holsanova, Jana (2016a), "Kognitiva och kommunikativa aspekter av syntolkning", i Jana Holsanova, Cecilia Wadensjö & Mats André (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier, rapport nr. 4, s. 17–27. ISBN:978-91-981060-9-1.
- Holsanova, Jana (2014a), "Reception of multimodality: Applying eye tracking methodology in multimodal research", i Jewitt Carey (red), *Routledge Handbook of Multimodal Analysis*. Second edition, s. 285–296.
- Holsanova, Jana (2014b), "In the mind of the beholder: Visual communication from a recipient perspective", i David Machin (red), *Visual communication*, Mouton - De Gruyter, s. 331–354.
- Holsanova, Jana (red) (2012), *Methodologies for multimodal research*, Special issue of Visual communication Vol. 11(3). Sage.
- Holsanova, Jana (2011), "How we focus attention in picture viewing, picture description, and during mental imagery", i Sachs-Hombach, K. & Totzke, R. (red), *Bilder, Sehen, Denken*. Herbert von Halem Verlag: Köln, s. 291–313.
- Holsanova, Jana (2010), *Myter och sanningar om läsning. Om samspelet mellan språk och bild i olika medier*. Norstedts/Språkrådet. ISBN 978-91-1-302841-5.
- Holsanova, Jana (2008), *Discourse, vision, and cognition*. John Benjamins Publishing Company: Amsterdam/Philadelphia.
- Holsanova, Jana (2001), *Picture Viewing and Picture Description: Two Windows on the Mind*. Doctoral dissertation. Lund University Cognitive Studies 83.
- Holsanova, Jana (1999), Olika perspektiv på språk, bild och deras samspel. Metodologiska reflektioner. i Inger Haskå & Carin Sandqvist (red), *Alla tiders språk*. Lundastudier i nordisk språkvetenskap A 55. Lund University Press, Lund, s. 117–126.
- Holsanova, Jana (1996), "När säger en bild mer än tusen ord? Om samspelet mellan text och bild i gränssnitt", i Löwgren, J. (red.), *Teman i*

- människa-datorinteraktion*. Institution för Informatik, Lunds universitet, s. 65–76.
- Holsanova, Jana, Hedberg, Björn & Nilsson, Nils (1999), "Visual and Verbal Focus Patterns when Describing Pictures", i W. Becker, H. Deubel & T. Mergner (red), *Current Oculomotor Research: Physiological and Psychological Aspects*. Kluwer Academic & Plenum Publishers: New York.
- Holsanova, Jana, Hildén, Anita, Salmson, Maria, & Kesen Tundell, Veronika (2015), *Syntolkning och uppläst text*. En studie om hur användarna vill ha det med riktlinjer för audiovisuella medier. Stockholm: Tundell & Salmson.
- Holsanova, Jana & Nord, Andreas (2010), "Textens fragmentering och läsares meningsskapande", i Gunilla Byrman, Anna Gustafsson & Henrik Rahm (red), *Svensson och svenskan: med sinnen känsliga för språk*. Lund: Författarna, s. 110–123.
- Holsanova, Jana, Wadensjö, Cecilia & Andrén, Mats (2016) (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier, rapport nr. 4.
- Jakobson, Roman (1959), "On linguistic aspects of translation". i R. Brower (red), *On translation*. Cambridge, MA: Harvard University Press, s. 232–239.
- Johansson, Roger (2016), "Mentala bilder hos seende och blinda", i Jana Holsanova, Cecilia Wadensjö & Mats Andrén (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier. Rapport nr. 4, s. 29–38.
- Johansson, Roger, Holsanova, Jana & Holmqvist, Kenneth (2006), "Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness", *Cognitive Science* 30(6): 1053–1079.
- Johansson, Roger, Holsanova, Jana, Dewhurst, Richard & Holmqvist, Kenneth (2011), "Eye Movements During Scene Recollection Have a Functional Role, but They Are Not Reinstatements of Those Produced During Encoding", *Journal of Experimental Psychology: Human Perception and Performance* 2012, Vol. 38, No. 5: 1289–1314.
- Johansson, Roger, Holsanova, Jana, & Holmqvist, Kenneth (2013), "Using Eye Movements and Spoken Discourse as Windows to Inner Space", i Carita Paradis, Jean Hudson, & Ulf Magnusson. (red), *The Construal of Spatial Meaning: Windows into Conceptual Space*, Oxford: Oxford University Press, s. 9–28.
- Kress, Gunther & van Leeuwen, Theo (2006[1996]). *Reading images: The grammar of visual design*. London: Routledge.
- Kruger, Jan-Louise & Pilar Orero (2010), "Audio description, audio narration – a new era in AVT", *Perspectives: Studies in Translatology* 18(1): 141–142.
- Labov, William & Joshua Waletzky (1967), "Narrative analysis", i J. Helm (red), *Essays on the Verbal and Visual Arts*, Seattle: U. of Washington Press, p. 12–44.

- Lu, Shijian & Bernsen, Niels Ole (1995), "The taxonomy workbench. A multimedia database system for analysing representational modalities", *Proceedings of the 4th International Conference on Interface to Real and Virtual Worlds*, Montpellier, June 1995.
- Lyberg-Åhländer V., Brännström K. J. & Sahlén B. S. (2015), "On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition", *Frontiers in Psychology* 6: 1–5.
- Maybury, M.T. (1995), "Research in Multimedia and Multimodal Parsing and Generation", i P. Mc Kevitt (red), *Integration of Natural Language and Vision Processing*. Springer: Dordrecht, s. 31–55.
- Mullet, Kevin & Sano, Darrell (1995), *Designing Visual Interfaces. Communication Oriented Techniques*. Prentice-Hall: Upper Saddle River, NJ, USA.
- O'Halloran, K. L., A. Podlasov, A. Chua & Marissa K. L. E (2012), "Interactive Software for Multimodal Analysis", i Jana Holsanova (red), *Methodologies for multimodal research*. Special issue of Visual communication 11 (3). Sage, s. 363–381.
- Pedersen, Jan (2016), "Översättningsvetenskaplig forskning om syntolkning", i Holsanova, Jana Andrén, Mats & Wadensjö, Cecilia (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier, rapport nr. 4, s. 53–59.
- Remael, Aline, Reviere, Nina & Vercauteren, Gert (2014), "Pictures painted in words: ADLAB audio description guidelines". Trieste: Edizioni Università di Trieste.
- Remael, Aline, Reviere, Nina & Vandekerckhove, R. (2016), "Audiovisual Translation. Theoretical and methodological challenges: some research trends", i Y. Gambier & S. Ramos Pinto (red), *Audiovisual Translation. Theoretical and methodological challenges* [Special Issue], *Target* 28(2): 248–260.
- Remael, Aline & Reviere, Nina (2015), "Recreating multimodal cohesion in audio-description: The case of audio-subtitling in Dutch multilingual films". *New Voices in Translation Studies* 13: 50–78.
- Reviere, Nina (2017), *Audio Description in Dutch. A corpus-based study into the linguistic features of a new, multimodal text type*. [PhD thesis]. University of Antwerp.
- Strukelj, Alexander (2016), "Praktiska erfarenheter av syntolkning - en intervjustudie", i: Jana Holsanova, Mats Andrén & Cecilia Wadensjö (red), *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/Myndigheten för tillgängliga medier, rapport nr. 4, s. 47–51.
- Teleman, Ulf (1989), "The World of Words - and Pictures". i Sture Allén (red), *Possible Worlds in Humanities, Arts and Sciences. Proceedings of Nobel symposium 65*. Walter de Gruyter: Berlin, New York, s. 199–208.
- Teleman, Ulf (1994), "Ordens och bildernas språk", *Svenska i skolan* 1994:1, s. 6-15.
- Vercauteren, Gert (2016), *A Narratological Approach to Content Selection in Audio Description* [PhD thesis]. University of Antwerp.

- Vercauteren, Gert, & Remael, Aline (2014), "Spatio-temporal Settings", i Anna Maszerowska, Anna Matamala, Pilar Orero (red), *Audio Description: New Perspectives Illustrated*. Amsterdam/Philadelphia: John Benjamins, s. 61–80.
- Wilson, Deidre & Sperber, Dan D. (2004), "Relevance Theory", i L.R. Horn & G. Ward (red), *The Handbook of Pragmatics*. Oxford: Blackwell, s. 607–632.
- Zabalbeascoa, Patrick (2008), "The nature of the audiovisual text and its parameters", i J. Díaz Cintas (red), *The didactics of audiovisual translation*. Amsterdam: John Benjamins Publishing Company, s. 21–37.
- Zwaan, Rolf A., & Radvansky, Gabriel A. (1998), "Situation models in language comprehension and memory", *Psychological bulletin*, 123(2): 162–185.