

AUTOMATISK TEXTUPPLÄSNING PÅ WEBBEN – TILL NYTTA FÖR VEM?

Av Rickard Domeij och Ola Karlsson, Språkrådet, april 2015

INLEDNING

SYFTE OCH MÅL

Syftet med arbetet bakom den här rapporten är att inventera utbudet av automatiska uppläsningssfunktioner för webben och gå igenom forskningen om deras användbarhet med fokus på personer med andra modersmål än svenska. Här presenteras resultatet av det arbetet. Målet är att öka vår egen kunskap om användningen och användbarheten av hjälpmedel för textuppläsning, bl.a. för att kunna följa upp och ge vägledning till myndigheter och andra organisationer i deras tillgänglighetsarbete. Det kan ses som en del i vårt arbete med uppföljningen av språklagen.

PROBLEMFÖRMULERING

Det praktiska ansvaret för uppföljningen av språklagen ligger på Språkrådet. Som en del i det arbetet har vi tidigare undersökt språksituationen på internet i Sverige (Domeij 2010) och utarbetat en vägledning med praktiska råd till myndigheter och andra organisationer om hur de kan arbeta med flerspråkig webbinformation (Språkrådet 2012). Vi håller nu på att ta fram metoder för att undersöka den flerspråkiga tillgängligheten på myndigheters webbplatser (Domeij & Spetz 2012, 2014) med utgångspunkt i språklagen och vägledningen.

I tidigare projekt har det framkommit att många myndigheter har en funktion som gör att besökarna automatiskt kan läsa upp texter på webbplatsen. Vilka krav på sådana uppläsningssfunktioner som ställs och vilka de tänkta målgrupperna är förefaller i de flesta fall vara oklart. Det finns en utbredd förhoppning om att inte bara dyslektiker och personer med synnedsättningar har nytta av textuppläsning, utan också personer med andra modersmål än svenska. Användningen av mobiltelefoner och små skärmar har samtidigt gett ett ökat behov av att lyssna på snarare än att läsa texter. Men det saknas kunskap om hur användbara sådana hjälpmedel är för webbesökarna. Därför är det högst angeläget att undersöka användbarheten av uppläsningssfunktioner på myndigheternas webbplatser som en del i Språkrådets arbete med att följa upp och ge vägledning till myndigheter och andra organisationer i deras tillgänglighetsarbete. (Domeij och Karlsson, 2013; Domeij och Spetz, i tryck). Som ett första steg vill vi titta på vad som tidigare gjorts för att undersöka saken.

RAPPORTENS INNEHÅLL

Rapporten är uppdelad i fyra delar. I den första delen redogör vi för vad talsyntes är och hur det används som hjälpmedel vid läsning på nätet. I del två går vi igenom vilka hjälpmedel som finns för uppläsning av text på webben och vilken funktionalitet de har. Vi ser också på vilka råd det finns för användningen av dem. Vi berör även kort de skärmläsare som finns på marknaden, dvs. uppläsningshjälpmedel som användaren själv installerat på sin dator. I del tre redogör vi för den forskning som finns om användbarheten av uppläsningshjälpmedel för olika grupper av användare med särskilt fokus på personer med andra modersmål än svenska. I del fyra drar vi slutsatser och blickar framåt.

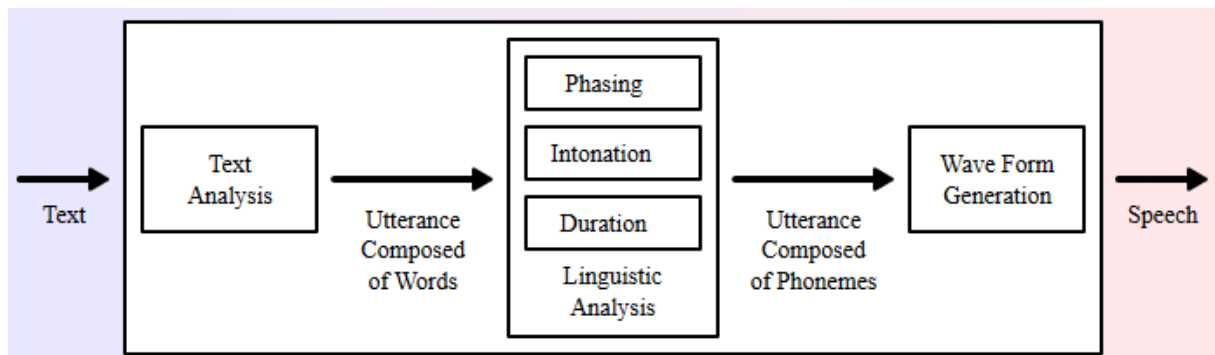
DEL 1: TALSYNTESSOM HJÄLPMEDEL VID LÄSNING PÅ NÄTET

UPPLÄSNING AV TEXT MED TALSYNTESS

För att läsa upp text i datorn behövs ett program som omvandlar text till tal. För det behövs en talsyntes som framställer tal på elektronisk väg. Ordet talsyntes kan i allmänspråket användas om allt tal från maskiner som inte är ren ljudåtergivning. Vid egentlig talsyntes utnyttjas en modell för talproduktion som gör att systemet kan uttala i princip vilket yttrande som helst, till skillnad från många telefontjänster med förinspelade repliker. Många av dagens kommersiella talsyntesprogram arbetar med att sätta ihop korta ljudavsnitt till önskade

yttranden med s.k. ljudvalssyntes (se Olika slags talsynteser). Under senare år har intresset för multimodal talsyntes ökat. Ljudsyntesen kompletterats då med en animerad ansiktsbild. (Granström u.å.).

Talsyntes avser ofta hela system för omvandling av text till tal, se figur 1. Men förutom att skapa tal gör ett text-till-talsystem (TTS) först en analys av texten som identifierar ord och meningar. Sedan görs en lingvistisk analys som resulterar i en fonetisk och prosodisk representation av orden och meningarna som slutligen omvandlas till ljudvågor i själva talgenereringen (talsyntes i snäv mening). I avancerade sådana system ingår språklig kunskap av många slag, t.ex. grammatisk analys, uttalslexikon och uttalsregler. Text-till-talsystemet används ofta i kombination med en skärmläsare som styr vad som ska bli uppläst. Med skärmläsaren kan användaren få datorn att läsa upp text både i programmenyer och dokument. Stora operativsystem som Windows och Mac OS har både text-till-talsyntes och skärmläsare inbyggda.



Figur 1. Text-till-talsystem. (Speech synthesis u.å.).

En av de tidigaste tillämpningarna av talsyntes i Sverige var för omvandling av text till tal som läshjälpmiddel och konstgjord röst (talprotes) för personer med syn- och talnedsättningar. På senare tid har det också kommit att användas som kompensatoriskt läshjälpmiddel personer med dyslexi och andra läs- och skrivsvårigheter. I och med att kvaliteten utvecklats har talsyntesen fått en mer allmän användning i t.ex. persondatorer och automatiska informationssystem. TTS används till exempel i telefontjänster för uppläsning av information från databaser, e-brev, sms och liknande. Många taltjänster är kopplade till dialogsystem som man både kan tala till och få talade svar ifrån, så kallade talsvarstjänster. I många situationer där ögon och händer är upptagna av annat är talkommunikation effektivare och säkrare än text, som vid bilkörning. Dessutom används TTS i många multimedieprodukter, till exempel i talande böcker och leksaker, interaktiva spel och språkinlärningsprogram. I framtiden kommer det också bli möjligt att översätta mellan språk från tal till tal.

OLIKA SLAGS TALSYNTESE

Det finns olika slags talsynteser. Här beskrivs kort parametrisk talsyntes, konkatentativ talsyntes och ljudvalssyntes (Engvall 2008). Historiskt har de vuxit fram i nämnd ordning.

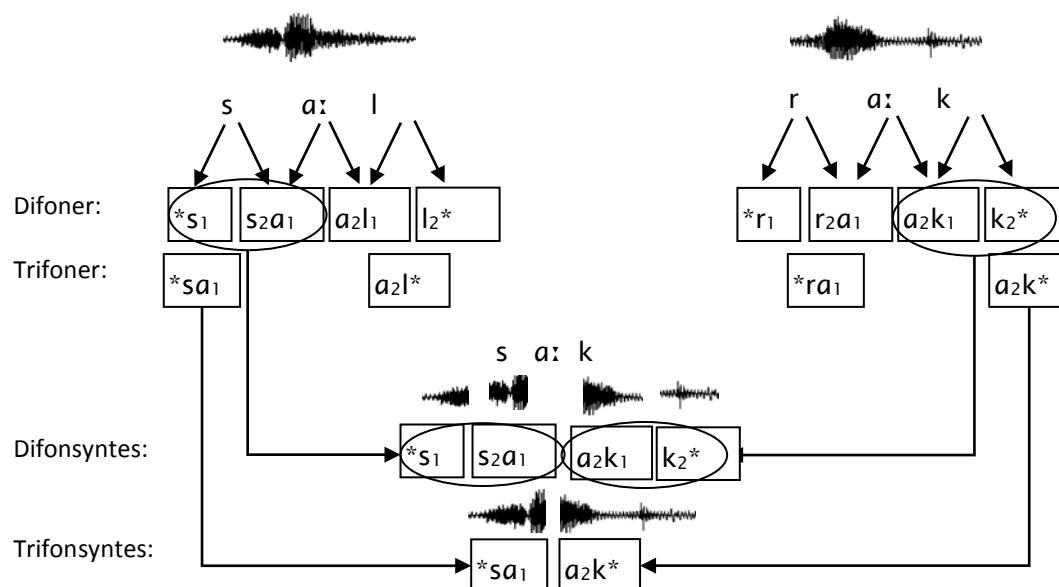
Parametrisk talsyntes kallas den klassiska formen av talsyntes. I den skapas ljud helt på elektronisk väg, och regler används för att beskriva hur de påverkas av koartikulation och meningsbetoning. Eftersom samma regler används på samma sätt varje gång låter den parametriska talsyntesen enformig i jämförelse med mänskligt tal. Bristen på variation gör att lyssnaren lätt tröttnar. Därför används parametrisk talsyntes sällan längre i kommersiella produkter. En fördel med parametrisk talsyntes är att den är begriplig vid höga talhastigheter vilket blinda och personer med synnedsättning har nytta av vid snabbläsning av text.

I *konkatentativ talsyntes* används förinspelade ljudenheter som sätts ihop till ord och hela yttranden. Ofta används difoner (två halva fonem) eller trifoner (en följd av ett halvt, ett helt och ett halvt fonem), se figur 2. Det gör att skarvarna mellan enheterna hamnar mitt i fonemen och passar bättre ihop, vilket får resultatet att låta naturligare. Man utgår från inspelat tal som innehåller alla ljudkombinationer som kan behövas.

Inspelningen klipps sedan upp i enheter som används om och om igen vilket gör att rösten blir onaturligt variationslös.

En modernare form av konkatentativ talsyntes kallas *ljudvals syntes* (eng. unit selection synthesis) och låter naturligare, både för att det krävs mindre signalbehandling och för att variationen i rösten blir större. Den använder betydligt större databaser med ljudenheter som gör att det går att ha flera varianter av varje difon eller trifon och vid syntesen välja dem som passar bäst ihop. Man lägger också till vanliga ord och fraser så att syntesen kan använda längre fraser från originalinspelningen. Det gör att en ljudvals syntes i kortare avsnitt kan låta helt naturlig medan det förekommer skarvar eller feluttalade ord på andra ställen.

Förutom ovan nämnda typer av akustisk talsyntes som modellerar talsignalen finns också artikulatorisk talsyntes som använder en modell av talorganen för att skapa tal. Artikulatorisk talsyntes används främst för forsknings- och utbildningsändamål.



Figur 2. Uppdelning av orden sal och rak i difoner och trifoner. Tecknet * anger här inledande eller avslutande tystnad och index 1 och 2 första respektive andra delen av fonemet. Ordet sak kan syntetiseras av difonerna eller trifonerna i sal och rak. (Figuren är hämtad ur Engvall 2008)

DEL 2: INVENTERING AV HJÄLPMEDEL OCH RÅD FÖR ANVÄNDNING

Här följer en genomgång av de uppläsningshjälpmedel som finns på den svenska marknaden i dag. Vi har i första hand letat efter talsyntesprogram, i andra hand har vi letat fram några av de skärmläsare användaren själv kan installera. Genomgången gör inga anspråk på att vara fullständig; vi har försökt ge en ungefärlig bild av produktutbudet och av de tongivande tjänsterna. Vi har utgått från företagen bakom tjänsterna i listningen.

INLEDNING TALSYNTER

Talsyntesprogrammen har utvecklats en hel del senaste åren, och många program har i dag rätt stora möjligheter till individuella inställningar av sådant som läshastighet, läsmarkör/linjal, röst, tonhöjd, volym, markering av text som ska läsas upp osv. De flesta talsynteser kan i dag även spela in den upplästa texten och spara ner den som ljudfil, så att användaren kan lyssna på den senare i t.ex. en daisyspelare eller mobiltelefon. Att inställningsmöjligheterna är flexibla är bra, eftersom det i undersökningar av särskilda behovsgrupper som använder hjälpmedel har visat sig att behov och önskemål kring funktioner till stor del är individuella snarare

än generella för olika användargrupper (läs mer nedan och under "Om brukarsamverkan" hos Fungerande medier, <http://fm.darub.org/?q=om-brukarmedverkan>).

Dessvärre kan flexibiliteten vara sämre när det gäller programmens interoperabilitet. Bristen på standardisering mellan program och skrivmiljöer kan t.ex. innebära att skolor som byter ut program möter praktiska hinder i användningen av talsyntes.

Vi har delat in programmen och tjänsterna efter huruvida de är kommersiella eller fria samt efter huruvida de i huvudsak vänder sig till informationsutgivare eller slutanvändare. Grupperingen är överlappande och kan förstås diskuteras, syftet är mer att skapa överskådlighet i presentationen.

Vi talar också om talsyntestjänster och skärmläsare för sig, även om gränsen kan vara något oklar ibland. Talsyntestjänster innehåller en skärmläsare, och vice versa, lite förenklat.

KOMMERSIELLA TALSYNTEPROGRAM

Utgivarbaserade program

Denna grupp av tjänster är normalt tänkta att användas av webbplatsutgivare på deras webbplatser, som ett led i att göra webbplatserna mer tillgängliga för användare: personer med nedsatt syn, dyslektiker, personer med annat modersmål, personer med kognitiva funktionshinder, användare med små skärmar osv. Många företag har produkter även för andra tillämpningar, som uppläsning av dokument och e-publikationer, och med en rad olika röster i utbudet.

Acapela: Internationellt företag som även har svenska produkter, bl.a. för uppläsning av publikationer, presentationer och ljudfiler: <http://www.acapela-group.com/?lang=sv>. Exempel: demo på ingångssidan.

Insipio: Read IT och Speak IT är avsedda för talsyntes på utgivares webbplatser. Med plusversionen får man en verktygslist med separat uppläsning av länkar, ordförklaringar m.m. Med Speak IT Online kan uppläsningen breddas till e-böcker, dokument osv.: <http://www.insipio.se/>. Exempel: Sunne kommuns webbplats, <http://sunne.se/> (välj Lyssna). Uppläsning med egna tjänster på Insipios ingångssida.

Läs Web 2.0: Stor leverantör av text-till-tal lösningar i Norge och Danmark men finns också på den svenska marknaden: <http://www.lasweb.se/>. Exempel: uppläsning integrerat på företagets webb.

Nuance: Internationellt företag som även har talsyntesprodukter med svenska röster, som Vocaliser: <http://www.nuance.com/for-business/text-to-speech/vocalizer/index.htm>. Exempel: demo på ingångssidan.

Pratsam: Talsynteslösning för "läshandikappade" med utgångspunkt hos synskadade finlandssvenskar (systemet har en finlandssvensk röst) och med tillämpning på t.ex. taltidningar och talböcker.: <http://www.pratsam.org/>. Exempel: lyssna på "talsyntesen Samuel" hos Pratsam, http://www.pratsam.org/samuel_sv.php.

Readspeaker: Brett produktutbud för webb, talböcker, upplästa dokument etc. Stor leverantör för talsyntes på webbplatser: <http://www.readspeaker.com/sv/>. Exempel: regeringens webbplats, <http://www.regeringen.se/> (välj Lyssna). Demo på Readspeakers ingångssida.

Veridict: Svenskt företag som i huvudsak erbjuder taligenkänning, men även talsynteslösningar för olika tillämpningar, t.ex. inom sjukvården, industrin och transportväsendet: <http://www.veridict.com/>. Exempel: Vomote, <http://www.veridict.com/solutions/assistive-technology>.

Användarbaserade program

Andra program är mer tänkta för användning av enskilda, för olika slags dokument och i olika miljöer (dator, mobil, datorplatta osv.). Vanliga målgrupper är dyslektiker och barn och unga som behöver extra stöd i skrivandet och läsandet. Observera att det finns en överlappning med program och leverantörer ovan.

Insipio: Med prenumerationstjänsten Text IT kan man få alla slags texter på sin dator, surfplatta eller mobil upplästa: webbtexter, dokument osv. Man kan även få uppläsning när man skriver, kompletterat med stavningskontroll, ordprediktion och ordförklaringar: <http://www.insipio.se/>.

MV-Nordic:

- CD-Ord är pc-baserad och riktar sig till skolan och dyslektiker. Uppläsning av inmatade texter liksom av bildinskannade texter via OCR, även för engelska, tyska, franska och spanska. Ordprediktion och ordförklaringar som stöd vid skrivande och läsning: <https://www.mv-nordic.com/se/produkter/cd-ord>. Exempel: Demofilm på CD-Ords webbsida.
- Into Words för Apple-enheter och webben (men via avtal även för pc och Android) läser upp texter på svenska eller engelska och ger skrivstöd: ordprediktion, stavningshjälp och ordförklaringar: <https://www.mv-nordic.com/se/produkter/intowords/>. Exempel: Demofilmer på programmets webbsidor för olika enheter.

Oribi: Svenskt företag som utvecklar olika typer av språkstöd för olika plattformar och enheter, inklusive rättstavningsstöd och talsyntes. För talsyntes finns Speak for Chrome som i Chrome kan läsa upp text i t.ex. Docs, pdf-dokument och epub-filer (e-böcker): <http://www.oribi.se/>. Test och nedladdning: <https://chrome.google.com/webstore/detail/oribi-speak-for-chrome/limnpfdlpifbklhnhjmapnhlpgkhhbakf?hl=sv>.

Readspeaker: Readspeaker Textaid liknar Insipios Text IT ovan men har fler möjligheter för språkval, textformatering, textsparande, översättning etc.: <http://www.readspeaker.com/sv/readspeaker-textaid/>.

Scandis: Dansk producent som framför allt ger ut programmet Vital och underprogram som Visaord och Vitex. Fungerar i pc-miljö för Microsoft Office, Open Office och Star Office. Läser upp dokument, webbsidor, mejl osv. En mängd kompletterande tjänster som att spara till mp3-fil, "Talande tangentbord" med uppläsning av bokstavs- och siffertangenter, inskanning med läspenna osv.: <http://www.scandis.dk/da>, <http://www.frolundadata.se/butik/vital-talsyntes/>.

Svensk Talteknologi: Program både för skolan och enskilda.

- Claro Read Plus är ett talsyntesprogram som läser upp text när orden skrivs, bokstav för bokstav, ord eller mening. Ljudning, stavningskontroll, homofonkontroll, ordprediktion, ordförklaring. Spara text till ljudfil eller till video, skanning från papper, pdf eller andra bildformat. Fungerar i flera format, som MS Word och i webbläsare: <http://www.svensktalteknologi.se/?artnr=ClaroRead+Plus+v5>. Exempel: Demo kan laddas ner från produktens webbsida. Programmet finns för både pc och Mac och med tillägget Word Read och usb-version.
- Claro Cloud är ett liknande webbaserat programpaket i två delar: Claro Webb är en ordbehandlare med stöd för uppläsning och skrivande, Claro Chrome ett tillägg till Chrome med lite större möjligheter att lägga till röster och läsa upp texter i Google docs: <http://svensktalteknologi.temp.box.se/download/produktblad/ClaroCloud.pdf>.
- På appfronten finns Clarospeak svenska för Iphone och Ipad som även läser dokument (den dyrare plusversionen har även OCR-funktion): <http://claro-apps.com/>. Captura Talk är ett liknande program för Android (men ska även finnas för nyare IOS-versioner): <http://www.svensktalteknologi.se/capturatalk>. Demo och demofilmer finns på produktens webbsida: <http://www.capturatalk.com/121/resources>.

Svox: Schweizisk leverantör med appen Svox för Android, inklusive svensk röst:

<https://svoxmobilevoices.wordpress.com/>.

Talande webb (Browse aloud): Ett mellanting mellan de utgivar- och användarbaserade grupperna ovan.

För att tjänsten ska fungera måste webbplatsutgivaren ha anslutit sig till systemet och betalat en avgift, men utgivaren behöver inte installera tjänsten. Omvandling från text till tal sker i stället lokalt på besökarens enhet. Besökaren installerar bara en programfil som placeras på t.ex. skrivbordet eller på ett usb-minne, och inställningar för läshastighet, typsnitt, färger, läslinjal osv. kan sedan göras för alla anslutna webbplatser man besöker. Det går också att använda en enklare version, Talande webb plus, som inte kräver nedladdning utan bara en aktivering av tjänsten via en länk. Fördelen med Talande webb är enkelheten och flexibiliteten liksom inställningsmöjligheterna för användaren. Nackdelen är att det bara fungerar på webbplatser som anslutit sig: <http://www.funkaportalen.se/Tema/Ljud/Sa-funkar-Talande-Webb/>. Exempel: SCB:s webb, http://www.scb.se/Pages/List_276697.aspx.

Tor Talk: Avskalat och användarvänligt talsyntesprogram för Windows som läser upp all text på datorn. Man markerar bara text och klickar på play. Har ett fåtal knappar i en verktygslist med de nödvändigaste funktionerna. Genom att placera en enkel OCR-ram kring t.ex. e-böcker läses även sådana låsta format enkelt upp: <http://www.tortalk.se/>. Demofilmer på <http://www.tortalk.se/produkten/>.

Voxit: Speak App är huvudsakligen avsett för användning med ordbehandlingsprogram som MS Word. Det omvandlar text till tal i realtid när man skriver och innehåller svensk rättstavning och ordprediktion samt pdf- och OCR-hantering. Programmet finns för ett antal olika språk: <http://voxit.se/>. Exempel: demo kan laddas ner från företagets webb, <http://www.voxit.se/budgiepro/default.asp?s=voicex>.

FRIA TALSYNTEPROGRAM

Det finns framför allt vissa fritt tillgängliga insticksprogram som vanliga användare kan ladda ner och använda för att lyssna på webbsidor som inte har inbyggd talsyntes.

Chrome: Ett exempel för Chrome är Select and Speak, som finns för ett stort antal språk inklusive svenska: <http://www.ispeech.org/>. Ett annat är Speak it: <https://chrome.google.com/webstore/detail/speakit/pgeolalilifpodheeocmbhehgnkbbak?hl=sv>. Ytterligare ett är Read & Write for Google: <https://chrome.google.com/webstore/detail/readwrite-for-google/inoeonmfapjbbkmdafoankkfajkcphgd>.

Firefox: Read 2 Me för engelska: <https://addons.mozilla.org/en-US/firefox/tag/speech%20synthesis>.

IOS: Voice Over är både inbyggd skärmläsare och talsyntes för Mac OS X och IOS under Systeminställningar > Hjälpmedel. Man kan även få text i t.ex. anteckningar och mejl upplästa: <https://www.apple.com/se/accessibility/ios/voiceover/>

Appar: Gratisappar är t.ex. Spesynto speech synthesis för IOS: <http://spesynto.com/>. För Android finns bl.a. Acapela: <http://www.acapela-for-android.com/>, liksom Red Ninja Studio: <http://www.redninjastudio.com>.

INLEDNING SKÄRMLÄSARE

En skärmläsare är ett program som användaren själv installerar på sin dator eller enhet. Den läser av informationen på bildskärmen och presenterar den som syntetiskt tal, i förstorat format eller i punktskrift. Det kan handla om programmenyer, rubriker, länklister och andra navigeringsobjekt, men också om

ordbehandlingsdokument, mejl etc. Skärmläsare används i första hand av blinda och synskadade för att kunna navigera utan mus; skärmläsaren läser i stället upp information om navigeringen, som då normalt styrs via tangentbordet. En skärmläsare strukturerar informationen, t.ex. så att alla länkar på sidan presenteras i en samlad lista.

Några program nedan är framför allt avsedda för webbplatsutgivare som vill testa hur deras webbplatser fungerar för personer som navigerar utan tangentbord. Andra kan ha både vanliga slutanvändare och testare som målgrupp. Det är vanligt att testa med Jaws på Windows och Voice Over på IOS.

KOMMERSIELLA SKÄRMLÄSARE

Det finns ett antal skärmläsare på marknaden. En välkänd sådan är Jaws, som har många inställningsmöjligheter. Man kan komplettera programmet med en svensk talsyntes:
<http://www.freedomscientific.com/Products/Blindness/JAWS>. Den kan även testas gratis:
<http://www.freedomscientific.com/Downloads/ProductDemos/#JAWS>.

Supernova är en annan spridd produkt. Den inkluderar styrning av punktskriftsskärm:
<http://www.yourdolphin.com/productdetail.asp?id=5&lang2=sv>. Demoversion:
<http://www.yourdolphin.com/tryitnew.asp?id=5>.

FRIA SKÄRMLÄSARE

Firefox: För Firefox finns en fri skärmläsare, Firevox: <http://firevox.clcworld.net/>.

I Firefox kan man även testa det inbyggda tillägget Firefox Screen Reader Simulator, som ger en textpanel där man kan se vad en skärmläsare skulle läsa upp på en webbsida: <https://addons.mozilla.org/en-US/firefox/addon/screen-reader-simulator/>.

Fangs är en annan sådan emulator, som visar det programmet Jaws (se ovan) skulle läsa upp:
<https://addons.mozilla.org/en-US/firefox/addon/fangs-screen-reader-emulator/>.

Chrome: Chromevox är en skärmläsare specifikt för Chrome: <http://www.chromevox.com/>.

Windows: I Windows finns en inbyggd funktion som helt enkelt kallas Skärmläsaren. Den läser texten på skärmen högt och beskriver t.ex. navigeringsval och felmeddelanden som visas. Man aktiverar skärmläsaren via startmenyn: <http://windows.microsoft.com/sv-se/windows/hear-text-read-aloud-narrator#1TC=windows-7>.

NVDA är en gratis skärmläsare för Windows med inbyggt stöd för ett flertal språk, bl.a. svenska:
<http://www.nvaccess.org/download/>.

Om man betalat för MS Office går även Window Eyes att ladda ner gratis: <http://window-eyes.se/>.

IOS: Voice Over är en inbyggd skärmläsare för både Mac OS X och IOS under Systeminställningar > Hjälpmedel. Navigeringen av enheten kan då styras med t.ex. fingerrörelser på pekskärmen, så att man får olika objekt på skärmen upplästa. Under inställningar kan man ladda ner svenska röster. Med den s.k. rotern kan man få t.ex. rubriker, länkar eller bilder beskrivna, och välja hur man vill navigera på sidan:
<https://www.apple.com/se/accessibility/ios/voiceover/>.

LÄS MER

Allmänt och teknik

- Begripsam, nätverk för personer med behov av läs- och skrivstöd: <http://www.fungerandemedier.se/?q=begripsam>
- Fungerande medier, "Teknik – Ljud": <http://www.fungerandemedier.se/?q=!%C3%A4sa-och-lyssna>
- Funka, "Om datorhjälpmedel": <http://www.funkanu.com/sv/Design-for-alla/Information-webb-och-IT/Datorhjalpmedel/>
- Inuse, "Kom igång med att testa i skärmläsare": <http://www.inuse.se/blogg/kom-igang-med-att-testa-i-skarmlasare-inuseful/>
- PTS, "Projektet Läsaren": <http://www.pts.se/sv/Bransch/Internet/Anvandbarhet-och-tillganglighet-/Projektet-Lasaren-/>
- Ung kompensation, "Talsyntes": <http://www.ungkompensation.se/lasmeromtalsyntes.aspx>, <http://www.ungkompensation.se/talsyntes.aspx>
- Wikipedia, "Talsyntes": <http://sv.wikipedia.org/wiki/Talsyntes>

Råd och instruktioner

- Tommy Sundström (2005), Användbarhetsboken. Studentlitteratur.
- Synkodat, "Vad en skärmläsare ska klara": <http://www.synkodat.se/viewpage.php?page=4>
- WCAG 2.0 (2009), "Providing a spoken version of the text": <http://www.w3.org/TR/2014/NOTE-WCAG20-TECHS-20140916/G79>
- Webbriktlinjer.se: "Ge all information på begriplig svenska": <http://webbriktlinjer.se/r/10-ge-all-information-pa-begriplig-svenska/>, "Kombinera skrift med ljud, bild och film": <http://webbriktlinjer.se/r/11-kombinera-skrift-med-ljud-bild-och-film/>, "Gör det möjligt att hoppa förbi delar på sidorna": <http://webbriktlinjer.se/r/75-gruppera-och-skapa-mojlighet-att-hoppa-forbi-delar-pa-sidorna/>

DEL 3: STUDIER OM ANVÄNDBARHETEN AV TEXTUPPLÄSNINGSHJÄLPMEDEL FÖR PERSONER MED LÄSSVÅRIGHETER OCH ANDRASPRÅKSINLÄRARE

Vi har redan konstaterat att text-till-talsyntes har haft en revolutionerande betydelse som läshjälpmedel och konstgjord röst (talprotes) för personer med syn- och talnedsättningar. På senare tid har text-till-talsyntes också kommit att användas som kompensatoriskt läshjälpmedel för personer med dyslexi och andra läs- och skrivsvårigheter. Ulla Föhrer og Eva Magnusson drar slutsatsen att talsyntes som lässtöd kan vara ett effektivt kompensatoriskt hjälpmedel för många elever med ordavkodningsproblem med hänvisning till Elkind, Cohen & Murray 1993, Higgins & Zvi, 1995 och Montali & Lewandowski, 1996.

Här sammanfattas de viktigaste forskningsresultaten om användningen av text-till-talsystem som hjälpmedel för personer med lässvårigheter med utgångspunkt i en forskningsöversikt av Elkind och Elkind (2007):

- **Långsamma läsare läser snabbare med talsyntes.** Detta är ett genomgående resultat i flera av deras studier där nästan alla långsamma läsare förbättrade sin läshastighet när de läste med talsyntes. (Elkind, Black & Murray 1996; Elkind 1998; Heckner, Elkind, Elkind & Katz 2002.)
- **Man kan läsa under längre tid.** De allra flesta uppger att de blir mindre stressade, mindre trötta och kan läsa under längre tid med TTS. Över 90 % av studenterna med inlärningssvårigheter uppgav det. I genomsnitt uppgav studenterna med koncentrationssvårigheter en ökad läsuthållighet med 60 %. (Elkind, Cohen & Murray 1993; Elkind et al 1996; Elkind 1998).

- **De som har koncentrationssvårigheter kan koncentrera sig under längre tid.** Hecker et al (2002) fann att studenter med koncentrationssvårigheter gör färre avbrott i läsningen (> 50%) och kan koncentrera sig under längre tid med hjälp av TTS, vilket gör att de i genomsnitt kan läsa 50 % mer text på samma tid i längre läspass.
- **Många men inte alla svaga läsare förbättrar sin läsförståelse.** Flera studier har visat att många svaga läsare förstår mer av det de läser när de läser med talsyntes men att det inte gäller alla (Montali & Lewandowski 1996; Elkind, Cohen & Murray 1993; Disseldorp & Chambers 2002; Higgins & Raskind 1997; Elkind, Black & Murray 1996; Elkind 1998).
- **Den individuella variationen är stor.** Personer som har svag läsförståelse utan hjälpmedel förbättrar sitt resultat mest, medan många av de som annars har god läsförståelse till och med kan försämra sitt resultat. Forskarna har identifierat två bidragande faktorer bakom resultaten: auditiv förmåga och förmåga att samtidigt ta till sig visuell och auditiv information (Elkind, Black & Murray 1996; Elkind 1998). Se även Balajthy 2005.
- **Läsförståelsen förbättras vid bimodal läsning, inte vid monomodal läsning.** Läsförståelsen ökar inte om deltagarna enbart får texten uppläst med talsyntes utan att se den, däremot gör den det när visuell läsning kombineras med auditiv uppläsning (Montali & Lewandowski 1996). Detta förklaras med att bimodal läsning stärker associationerna mellan ljud och skrift. Det antas också frigöra kognitiva resurser till högre förståelseprocesser.
- **Läsförståelsen kan försämrars när annat visuellt material presenteras samtidigt.** Läsförståelsen förbättras vid bimodal läsning förutsatt att inget annat visuellt material presenteras samtidigt (Moreno & Meyer 1998; Mayer & Johnson 2008¹). Enligt tvåprocessmodellen (dual processing model of working memory; Mayer & Moreno 1998) behandlar arbetsminnet visuellt och auditivt material separat vilket gör att information som presenteras både auditivt och visuellt kan bearbetas samtidigt i olika kanaler utan att arbetsminnet överbelastas. Däremot kan flera visuella representationer i samma kanal överbelasta arbetsminnet. Detta förklaras av den kognitiva teorin om multimodalt lärande (cognitive theory of multimedia learning) som får konsekvenser för multimedial design (t.ex. Mayer & Moreno 2002, 2003; Sorden 2012). Här är en schematisk beskrivning av teorin och dess tre grundantaganden hämtad från Mayer & Moreno 2003:

¹ Mayer & Johnson reviderar den så kallade redundansprincipen ("People learn more deeply from graphics and narration than from graphics, narration and on-line text") med följande tillägg: "except when the on-screen text is short, highlights the key action described in the narration, and is placed next to the portion of the graphic that it describes".

TABLE 1
Three Assumptions About How the Mind Works in Multimedia Learning

<i>Assumption</i>	<i>Definition</i>
Dual channel	Humans possess separate information processing channels for verbal and visual material.
Limited capacity	There is only a limited amount of processing capacity available in the verbal and visual channels.
Active processing	Learning requires substantial cognitive processing in the verbal and visual channels.

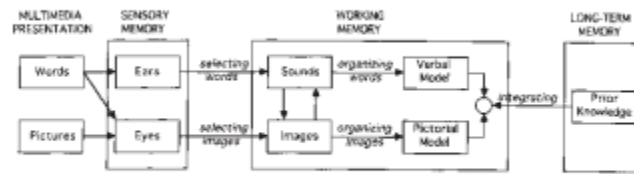


FIGURE 1 Cognitive theory of multimedia learning.

De flesta av ovanstående studier är utförda på studenter i laboratorieliknande sammanhang vilket gör att man kan ifrågasätta hur relevanta resultaten är i praktiken. Projekt PC-läsning är namnet på en stor dansk undersökning om lässvårigheter och it-hjälpmedel som gjordes ganska nyligen. Projektet har undersökt hur, och i vilken omfattning, IT-stöd i form av särskilda uppläsningsprogram kan vara till nytta för personer med dyslexi. Delundersökningarna omfattar mellan 60 och 90 vuxna med dyslexi i utbildning, arbetsliv och privatliv. Dessa intervjuades och testades vid olika tillfällen med ca ett års mellanrum. Syftet var att undersöka om, och i så fall hur, lässtödjande teknik kan kompensera för otillräcklig läsförmåga. Beteckningen pc-läsning används om läsning vid skärm där läsaren både ser texten och samtidigt lyssnar till en uppläsning av texten med syntetiskt tal.

Lästesten i undersökningen visade att pc-läsning har en ganska betydande kompensatorisk effekt, i synnerhet för de svaga läsarna. Ju svagare läsförmåga, desto större omedelbart kompensatoriskt värde har uppläsningen. Andra deltagare förbättrar i stort sett inte sitt resultat i lästestet och några gör sämre ifrån sig vid pc-läsning än vid läsning utan talsyntes. I intervjuerna framgick det dock att deltagare med relativt god läsförmåga tyckte sig ha stor hjälp av uppläsningen vid läsning av längre och mer komplexa texter. Som pc-läsare läser ett stort flertal av deltagarna mer, och de läser svårare texter. De förstår också texterna bättre. De behöver andras hjälp i mindre utsträckning och de har större lust att läsa.

Projektets resultat sägs kunna vara relevanta för en mycket större grupp av människor. Det gäller främst personer med mindre grad av läs- och skrivsvårigheter, men också människor helt utan dessa svårigheter som rent erfarenhetsmässigt kommer ha nytta av den nya tekniken. Det gäller till exempel läsning och skrivning på främmande språk där talsyntes kan vara ett komplement till andra program, till exempel stavnings- och grammatikkontroll. (Hjälpepedagogiska institutet 2010; Specialpedagogiska skolmyndigheten 2011).

Vi har hittat få studier med andraspråksperspektiv på läsning med talsyntes vilket gör det svårt att dra några generella slutsatser om det. Dela Rosa, Parent och Eskenazi (2010) redogör för två relaterade studier som pekar på att andraspråksinlärare av engelska får signifikativt bättre resultat i ett vokabulärtest om orden läses upp med talsyntes än utan. Det pekar på att bimodal presentation också är bra vid läsning och inläring av andraspråk. I en studie av Neuman och Koskinen (1992) lärde sig andraspråkstalare oavsiktligt fler nya ord och minde mer av faktainnehållet när de fick information i flera former genom olika modaliteter (rörlig bild, ljud

och text) än när de fick information i färre former (rörlig bild med ljud, tal och skrift, enbart skrift) utan att deras kognitiva kapacitet överbelastades (i strid med den kognitiva teorin om multimodalt lärande).

Det har funnits en hel del frågetecken kring om befintliga talsyntesprogram har tillräckligt god kvalitet för att användas av inlärare. En omfattande utvärdering av Handley (2009) pekar på att kvaliteten på de bästa nuvarande TTS-systemen är tillräcklig för att användas för textuppläsning inom datorstödd språkinläring. Ovan nämnda studie av Dela Rosa m.fl. bekräftar den slutsatsen.

DEL 4: SLUTSATS OCH FRAMÅTBlick

Som framgått i många studier är de individuella skillnaderna stora när det gäller effekten av läsning med talsyntes. Också Mayer (2009) menar att vissa designprinciper för multimodalt lärande har visat sig vara mer effektiva för inlärare med mer kunskap än inlärare med mindre kunskap (low-knowledge and high-knowledge learners). I jämförelse med personer med kognitiva funktionsnedsättningar kan skillnaderna vara ännu större, vilket också framgår i en nyligen gjord kunskapsöversikt av vetenskapligt utvärderade och rapporterade empiriska studier av tillgänglighet till elektronisk kommunikation för personer med kognitiva funktionsnedsättningar (Borg, Lanz & Gulliksen 2014). Där hänvisas till två studier där personer med dyslexi eller autism kunde lära sig mer eller förstå bättre när de använde enbart text. Text tillsammans med bilder eller talad text gav inte bättre resultat. I en annan redovisad studie kunde personer med inlärningssvårigheter lära sig mer och förstå bättre när de använde text, talad text och symboler samtidigt. De slutsatser man drar av denna forskningsgenomgång är att användargränssnitt behöver vara anpassningsbara till enskilda individers behov och att det finns behov av mer forskning som kan ligga till grund för standarder och riktlinjer (se även Anderson-Inman & Horney).

Mayer betonar att forskningen om multimodalt lärande och design måste vara teorigrundad och evidensbaserad. Ett exempel på evidensbaserade riktlinjer är Meyers 12 multimodala designprinciper (se fig. 3) som har utvecklats på basis av nästan 100 studier (Meyer 2009):

1. Coherence Principle – People learn better when extraneous words, pictures and sounds are excluded rather than included.
2. Signaling Principle – People learn better when cues that highlight the organization of the essential material are added.
3. Redundancy Principle – People learn better from graphics and narration than from graphics, narration and on-screen text.
4. Spatial Contiguity Principle – People learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.
5. Temporal Contiguity Principle – People learn better when corresponding words and pictures are presented simultaneously rather than successively.
6. Segmenting Principle – People learn better from a multimedia lesson is presented in user-paced segments rather than as a continuous unit.
7. Pre-training Principle – People learn better from a multimedia lesson when they know the names and characteristics of the main concepts.
8. Modality Principle – People learn better from graphics and narrations than from animation and on-screen text.
9. Multimedia Principle – People learn better from words and pictures than from words alone.

10. Personalization Principle – People learn better from multimedia lessons when words are in conversational style rather than formal style.
11. Voice Principle – People learn better when the narration in multimedia lessons is spoken in a friendly human voice rather than a machine voice.
12. Image Principle – People do not necessarily learn better from a multimedia lesson when the speaker's image is added to the screen.

Några av principerna har direkt relevans för läsning med talsyntes, som den tidigare nämnda redundansprincipen. Mayers principer är framför allt inriktade på lärande men kan utgöra en utgångspunkt för utformning av praktiska råd och fortsatt forskning om elektronisk läsning i arbetslivet och på fritiden också med hänsyn till personer med lässvårigheter. I synnerhet behövs mer forskning om effekten av läsning med talsyntes för andraspråksanvändare och hur läsning med talsyntes kan användas i kombination med andra stöd vid läsning på dator, som textförenkling, översättning och ordförklaringar (se t.ex. Eskenazi m.fl. 2013 och Anderson-Inman & Horney 2005).

REFERENSER

Anderson-Inman, Lynne and Horney, Mark A. 2007. Supported eText: Assistive technology through text transformations. *Reading Research Quarterly* 42.1: 153–160.

Balajthy, Ernester. 2005. Text-to-speech software for helping struggling readers. *Reading Online* 8.4 (2005): 1–9.

Rosa, Kevin Dela, Gabriel Parent, and Maxine Eskenazi. 2010. Multimodal learning of words: A study on the use of speech synthesis to reinforce written text in L2 language learning. *Proceedings of the 2010 ISCA Workshop on Speech and Language Technology in Education*.

Domeij, R. 2010. *En språkpolitik för internet*. Rapporter från Språkrådet 2. Institutet för språk och folkminnen, Språkrådet.

Domeij, Rickard och Karlsson, Ola. 2013. Delaktig i det digitala samhället – ett flerspråkigt perspektiv. I: Hyltenstam, K. och Lindberg, I. (red). *Svenska som andraspråk – i forskning, undervisning och samhälle*. Studentlitteratur.

Domeij, Rickard. och Spetz, Jennie. 2012. *Flerspråkig tillgänglighet på webben? Metoder för att mäta webbtillgänglighet i flerspråkigt perspektiv*. Slutrapport till .SE för projektet Metoder för mätning av webbtillgänglighet i språkpolitiskt perspektiv.

Domeij, Rickard och Spetz, Jennie. 2014. *Begripliga myndigheter på nätet – hur skapar man tillgänglighet för alla i ett flerspråkigt samhälle?* Konferensvolymen för Svenskans beskrivning 33.

Disseldorp, Bronwyn, and Dianne Chambers. 2002. Independent access: Which students might benefit from a talking computer?. *Untangling the Web: Establishing Learning Links*. Proceedings ASET Conference.

Elkind, Jerome. 1998. Computer reading machines for poor readers. *Perspectives* 24.2: 9–13.

- Elkind, Kenneth, and Jerome Elkind. 2007. Text-to-speech software for reading. *Perspectives* 33.3 (2007): 11–16.
- Elkind, Jerome, Karen Cohen, and Carol Murray. 1993. Using computer-based readers to improve reading comprehension of students with dyslexia. *Annals of Dyslexia* 43.1: 238–259.
- Elkind, Jerome, Molly Sandperl Black, and Carol Murray. 1996. Computer-based compensation of adult reading disabilities." *Annals of Dyslexia* 46.1: 159–186.
- Engvall, O. 2008. *Bättre tala än texta – talteknologi nu och i framtiden*. I: Domeij (red) *Tekniken bakom språket*. Norstedts förlag.
- Eskenazi, Maxine, Yibin Lin, and Oscar Saz. 2013. Tools for non-native readers: the case for translation and simplification. *NLP4ITA 2013*: 20.
- Granström, Björn. [u.å.] Talsyntes. *Nationalencyklopedin*.
<http://www.ne.se/upplagsverk/encyklopedi/lang/talsyntes> (hämtad 2014-10-24).
- Handley, Zöe. 2009. Is text-to-speech synthesis ready for use in computer-assisted language learning? *Speech Communication* 51.10 (2009): 906–919.
- Higgins, Eleanor L., and Jennifer C. Zvi. 1995. Assistive technology for postsecondary students with learning disabilities: From research to practice. *Annals of Dyslexia* 45.1: 123–142.
- Higgins, Eleanor L., and Marshall H. Raskind. 1997. The Compensatory Effectiveness of Optical Character Recognition/Speech Synthesis on Reading Comprehension of Postsecondary Students with Learning Disabilities. *Learning Disabilities: A Multidisciplinary Journal* 8.2: 75–87.
- Hjælpemiddelsinstituttet 2010. *Pc-læsning. Ordblindhed og it-hjælpemidler*. Projektrapport utarbetad av Erik Arendal, Bent Sabye Jensen og Åse Brandt.
- MacArthur, Charles A., et al. 2001. Technology applications for students with literacy problems: A critical review. *The Elementary School Journal*: 273–301.
- Mayer, Richard E. 2009. *Multimedia learning* (2nd edition). New York: Cambridge University Press.
- Mayer, Richard E., and Cheryl I. Johnson. 2008. Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology* 100.2: 380.
- Mayer, Richard E., and Roxana Moreno. 1998. A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of educational psychology* 90.2: 312.
- Mayer, Richard E., and Roxana Moreno. 2002. Aids to computer-based multimedia learning. *Learning and instruction* 12.1: 107–119.
- Mayer, Richard E., and Roxana Moreno. 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist* 38.1: 43–52.
- Montali, Julie, and Lawrence Lewandowski. 1996. Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities* 29.3: 271–279.
- Sorden, S. 2012. The cognitive theory of multimedia learning. *Handbook of educational theories*. Charlotte, NC: Information Age Publishing. (Nedladdningsbart utkast:
http://sorden.com/portfolio/sorden_draft_multimedia2012.pdf)

Specialpedagogiska skolmyndigheten 2011. *Projekt PC-läsning. En sammanfattning av ett danskt projekt om effekter av IT-stöd för personer med lässvårigheter.*

Språkrådet 2012. *Vägledningen för flerspråkig information – praktiska riktlinjer för flerspråkiga webbplatser.* Rapporter från Språkrådet 5.

Speech synthesis [u.å.]. Wikipedia. http://en.wikipedia.org/wiki/Speech_synthesis (hämtad 2001-10-21).