



ANNOTERINGSKONVENTIONER

för den finlandssvenska teckenspråkskorpusen CFSTS

Version 1.0



Karin Hoyer, Juhana Salonen & Lena Wenman
Teckenspråkscentret
Institutionen för språk- och kommunikationsvetenskap
Jyväskylä universitet
December 2024

INNEHÅLL

1 INLEDNING	4
2 PRINCIPER FÖR ANNOTERINGEN	8
2.1 Annoteringscellens längd	10
2.2 Annotering med ID-glossor	11
3 KLASSIFICERING AV TECKEN PÅ RADEN FÖR ID-GLOSSAN	16
3.1 Lexikala tecken	16
3.1.1 Numeraler (_num)	18
3.2 Pekningar	19
3.3 Avbildande tecken (_at)	19
3.4 Gester (_gest)	20
3.5 Bokstavering (_bokst)	21
4 ANNOTERING AV GRAMMATISK OCH ANNAN SPRÅKLIG INFORMATION	23
4.1 Negation (@neg)	24
4.2 Upprepning (@upprep)	25
4.3 Sammansättning (@ssg)	26
4.4 Språkkontakt (@sk)	27
4.5 Tecken som innehåller bokstaverade element (@bt)	30
4.6 Listboj/Listkonstruktion (@bojl)	31
4.7 Huvudskakning (@hs) och huvudnickning (@hn)	31
4.8 Oklart fall (@oklar)	32
4.9 Söker tecken (@st)	33
4.10 Avbrott (@avbr)	33
4.11 Handen blir "kvar" (@k)	33
4.12 Raden <i>ID_observationer</i> och principer för att skriva in observationerna	34
5 ÖVERSÄTTNING	35
5.1 Allmänna översättningsprinciper	35
5.2 Parenteser och hakparenteser	36
5.3 Bindestreck	36
5.4 Val av tempus	38
5.5 Om citering och användning av citattecken	38
5.6 Översättning av ofullständigt tecknande	39
5.7 Bokstavering med handalfabet och användning av versaler i översättningen	39
5.8 Sifferuttryck i översättningen	40
5.9 Kommentrar på raden <i>Översättning_observationer</i>	40

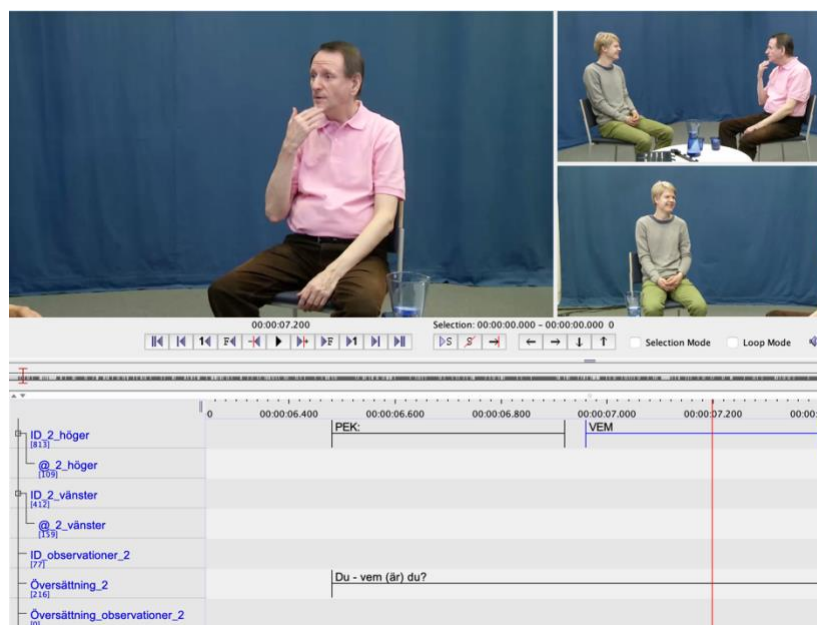
6 DEN LEXIKALA DATABASEN SIGNBANK	41
6.1 Vad är Signbank och hur fungerar den?	41
6.2 Vilken information finns för ID-glossorna i Signbank?.....	44
6.2.1 Videofilmer och namnet på videon	44
6.2.2 Översättningsmotsvarigheter (svenska)	45
6.2.3 Mer information.....	46
6.2.4 Relationer till andra glossor	47
6.2.5 Webbadresser	47
KÄLLOR	48

1 INLEDNING

Som en del av revitaliseringsåtgärderna för att stärka det hotade finlandssvenska teckenspråket (FSTS)¹ fick teckenspråkscentret vid Jyväskylä universitet 2021–2024 delat ansvar för forskning i språket. Undervisnings- och kulturministeriet beviljade både Jyväskylä universitet och Helsingfors universitet forskningsfinansiering, som vid teckenspråkscentret har använts för att skapa en (första del av en) finlandssvensk teckenspråkskorpus. Dessa annoteringskonventioner är en del av resultatet av korpusarbetet. I konventionerna beskrivs hur man på ett systematiskt sätt behandlat det teckenspråkiga materialet så att det resulterat i en korpus.

En korpus är en samling skrivet, talat eller tecknat språk i elektronisk form som man kan göra sökningar i med dator. En korpus behövs för att kunna forska i det finlandssvenska teckenspråket – hur ser teckenförrådet ut, hur är språket uppbyggt och hur varierar och används språket? Samtidigt utgör korpusen ett verktyg för undervisning i språket och är viktig med tanke på dokumentering och bevarande av detta hotade språk.

För teckenspråkens del startade traditionen med korpusarbete i Australien där korpusprojektet för Auslan påbörjades 2004. Tätt efter följde korpusar bland annat för holländskt, svenskt och brittiskt teckenspråk (Johnston & Schembri 2013). I Finland publicerade Finlands Dövas Förbund rf (FDF rf) år 2015 Kipo-korpusen för finskt teckenspråk (SVK, suomalainen viittomakieli). Korpusen är baserad på det språkpolitiska programmet för teckenspråken i Finland (Nordlund & Savolainen 2017). På teckenspråkscentret vid Jyväskylä universitet pågick 2014–2018 ett korpusprojekt för de finländska teckenspråken (CFINSL) då man gjorde teckenspråkiga videoinspelningar av våra båda inhemska teckenspråk (Salonen m.fl. 2019b). Korpusen för SVK byggdes upp först och en första del kunde publiceras redan 2019. Nu, 20 år efter att världens allra första korpusprojekt för ett tecknat språk startade, har den första finlandssvenska teckenspråkskorpusen (CFSTS) publicerats (Figur 1).



Figur 1. Skärmdump från den finlandssvenska teckenspråkskorpusen CFSTS.

¹ År 2013 konstaterades FSTS vara allvarligt hotat i en kartläggning utförd vid University of Central Lancashire i samarbete med UNESCO (Safar & Webster 2014).

Korpusmaterialet består av videoupptagningar med 12 finlandssvenska teckenspråkiga informanter. Ett viktigt kriterium då man sammanställer en korpus är korpusens representativitet (se till exempel Hanke & Fenlon 2022 och Wehrmeyer 2023). Det innebär att materialet skall bestå av språkbrukare med olika bakgrund för att kunna ge en allmängiltig bild av språket under en viss tidpunkt. I den finlandssvenska teckenspråskorpusen varierar bakgrundsfaktorer så som tecknarnas ålder, kön, boplat, skolgång och hörselstatus. Trots att ett antal på 12 informanter kan upplevas som litet utgör det ändå procentuellt ett skäligt sampel i förhållande till språkgemenskapens storlek. Det totala antalet språkbrukare i Finland uppskattas till ca 200 personer (Andersson-Koski, 2024). Videoinspelningarna gjordes under åren 2015 och 2017 som en del av ovannämnda CFINSL-projekt. Inspelningarna utfördes i Jyväskylä och i Helsingfors, och i samband med dem samlade man också in metadata, så som bakgrundsinformation om informanterna (Salonen m.fl. 2022). Denna information kompletterades med ytterligare intervjuer om informanternas språkliga bakgrund under våren 2024.

Den finlandssvenska teckenspråskorpusen CFSTS finns publicerad i Språkbanken². Språkbanken är FIN-CLARIN-konsortiets serviceenhet där det finns språkkorpusar i över 60 olika språk. Publiceringen av den finlandssvenska teckenspråskorpusen sker i enlighet med det samtycke som informanterna har givit för användning av videomaterialet.

Eftersom teckenspråk saknar etablerad skriven form måste de teckenspråkiga videoinspelningarna koda om till textsträngar på ett annat språk som kan bearbetas i dator. Att vara tvungen att använda ett annat språk – i vårt fall svenska – som metaspråk innebär en extra belastning. Det kan vara utmanande att minnas att det är det tecknade och inte det skrivna språket som står i fokus. Det här är bra att hålla i minnet också när man använder korpusen.

Att koda de videoinspelade sekvenserna med textsträngar kallas att *annotera* (också *märka upp* eller *tagga*). När annoteringen görs konsekvent och systematiskt möjliggör den således att materialet blir sökbart. Korpusmaterialet är annoterat med hjälp av annoteringsverktyget ELAN (EUDICO Linguistic Annotator)³. I ELAN är de skrivna annoteringarna tidsmässigt kombinerade med videofilmerna på skärmen. ELAN är via nätet kopplad till den lexikala databasen Signbank (se kapitel 6) där grundformer till korpusens annoterade tecken finns dokumenterade.

Korpusen har skapats i samarbete mellan många olika aktörer. Eftersom ingen av de tre anställda inom korpusbygget (författarna till dessa konventioner) har FSTS som modersmål har samarbetet med finlandssvenska teckenspråkiga språkbrukare varit oersättligt. Samarbetet har behövts för att säkerställa korpusannoteringar av hög kvalitet och garantera tillförlitligheten i videomaterialet i Signbank. För att kunna definiera de annoterade tecknens grundformer och grundbetydelser behövs språkbrukare med modersmålskunskaper och metalingvistisk kännedom om sitt språk. De existerande ordbokspublikationerna på nätet (Suvi⁴ och Signwiki⁵) har inte varit täckande nog för korpusbyggets behov. Dessutom finns det ur ett forskningsetiskt perspektiv inget alternativ till att bygga upp en korpus för ett teckenspråk annat än i samarbete med språkgemenskapen (se till exempel Hochgesang & Palfreyman 2022). I praktiken har samarbetet handlat om att på regelbunden basis konsultera språkrådgivarna vid FDF rf och sammankomma till

² <https://www.kielipankki.fi/aineistot/cfsts/>

³ <https://archive.mpi.nl/tla/elan>

⁴ <https://suvi.teckensprak.net/>

⁵ <https://finssl.signwiki.org>

korpusworkshoppar med språkbrukare. Dessa workshoppar har också arrangerats i samarbete, där både FDF rf och Finlandssvenska teckenspråkiga rf (FST rf) deltagit med personresurser och delfinansiering (Bild 1 och 2).



Bild 1. Korpusworkshop 20.5.2022 i Borgå. Deltagarna runt bordet (från vänster till höger): Karin Hoyer, Eva Mäkinen, Alf Ekström, Brita Peura, Birgitta Höysti, Maj-Len Karlsson-Pelkonen, Håkan Westerholm, Åke Uusimäki och Janne Kankkonen. (Fotograf: Juhana Salonen)



Bild 2. Korpusworkshop 06.06.2024 i Helsingfors. Stående i bakersta raden från vänster: Alf Ekström, Elin Westerlund, Karin Hoyer, Rolf Westerlund, Juhana Salonen. Stående i mittersta raden från vänster: Monica Fyrqvist, Brita Peura, Leila Sanssi, Birgitta Höysti. Längst framme från vänster: Lena Wenman och Emmy Nyström. (Fotograf: Magdalena Kintopf-Huuhka)

Ett tema där språkbrukarnas insats verkligen har varit oundgänglig är när det kommer till att identifiera språklig variation och språkkontakt. Finlandssvenska teckenspråkiga är flerspråkiga och har en bred språklig kompetens. Språklig variation och språkkontakt i finlandssvenskt teckenspråk har konstaterats och beskrivits redan tidigare (Hoyer 2000, 2004, 2005, 2012). Finlandssvenska teckenspråkiga har fortfarande ofta i sin vardag fem språk: kontakt med både SVK och svenskt

teckenspråk (STS) och med talad och skriven svenska och finska. Dessutom är engelska och International Sign/ASL också språk som finlandssvenska teckenspråkiga använder alltmer i och med personliga kontakter och sociala medier. Det här framkom tydligt ur den språkliga bakgrundsinformation, och språkkontaktfenomen syns också i korpusen (se annotering av språkkontakt i 4.4).

Korpusen kommer förhoppningsvis att stöda den pågående revitaliseringsprocessen av FSTS. Både i FST rf:s *Språkplan för finlandssvenskt teckenspråk* (2017) och i utkastet till det kommande revitaliseringsprogrammet för FSTS (Finlands Dövas Förbund rf & Finlandssvenska teckenspråkiga rf 2024) poängteras att forskning är viktigt för språkets överlevnad. Genom att språkbrukarna själva aktivt medverkar i språkliga forskningsaktiviteter ökar den språkliga medvetenheten och kunskapen om språket. Då man via korpusbygget och korpusforskning blir medveten om särdragen i FSTS ökar i bästa fall motivationen att värna om och använda FSTS. Korpusen är en resurs att använda även i språkundervisning, och också i lexikografiskt arbete drar man nytta av att använda en korpus⁶. Förhoppningsvis kommer korpusen och databasen Signbank senare att utökas med både nyinspelningar och språkligt material inspelat sedan förr som är relevant för revitaliseringssträvandena. Också ett framtida webbaserat korpusverktyg i stil med STS-korpus⁷ som fungerar utan att användaren separat behöver ladda ner program och videofiler skulle öka korpusens tillgänglighet.

⁶ Se Mesch m.fl. (2023) om samverkan mellan teckenspråkslexikon och teckenspråkskorpus.

⁷ <https://teckensprakskorpus.su.se>, se också Mesch (2023)

2 PRINCIPER FÖR ANNOTERINGEN

Den allra största utmaningen under annoteringsfasen i korpusarbetet har varit att det saknas grundforskning i hur FSTS är uppbyggt. Det få forskningsbaserade studier och språkliga resurser som finns från förr är lexikografiska studier (Hoyer & Kronlund-Saarikoski 2002), nätordboken Suvi och några pro gradu- och doktorsavhandlingar (eller i dem ingående artiklar). Avhandlingarna har haft fokus på (en del av) teckenförrådet (Hoyer 2000, 2004, 2005) och den icke-dominanta handens funktioner och listkonstruktioner (Siltaloppi 2016, 2023) i FSTS. Men då korpusbygget startade fanns inga studier i till exempel olika teckentyper i finlandssvenskt teckenspråk⁸ (lexikaliserade tecken vs. produktiva tecken så som avbildande tecken, eller pekningar), eller i grammatiska fenomen så som negation. Dessa är exempel på kategorier vi har varit tvungna att märka upp i korpusmaterialet baserat på kategorier använda i andra teckenspråk.

Vi har varit medvetna om dilemmat med att kategorisera enheter i språket utan forskning från förr. Den korpus vi skapat kommer att utgöra en förutsättning för vidare forskning i det finlandssvenska teckenspråkets uppbyggnad. Därför är det viktigt att vara medveten om att hur man väljer att kategorisera med högsta sannolikhet kommer att påverka framtida analyser av språket. En diskussion vi många gånger fört under vårt arbete är om det överhuvudtaget är möjligt att, före man känner till hur FSTS är strukturerat, bygga en korpus för FSTS och samtidigt undvika påverkan från kategorier hämtade utifrån. Det här gäller speciellt eftersom språkkontakt tydligt syns i korpusen och också framkom i bakgrundsintervjuerna: språkbrukarna har aktiv kontakt med andra teckenspråk så som SVK och STS. Det här väcker frågor så som: Vilka språkliga fenomen hör till vilket språk och varför är det relevant att försöka definiera det?

Korpusannotering kan utföras på många olika sätt och på olika språkliga nivåer. Den annotering som presenteras i dessa konventioner utgör ett första steg: det handlar om en första grundannotering. Det innebär att man identifierar och märker upp de tecken i teckenspråkstexten som artikuleras med händerna. Dessa tecken har annoterats med hjälp av *ID-glossor* som utgörs av svenska ord skrivna med versaler (se närmare 2.2) vilket är ett vanligt sätt att starta annotering av teckenspråkskorpusar (Johnston 2010, 2016; Hodge & Crasborn 2022). Beroende på vilket språkligt fenomen man skall studera kan det senare behövas ytterligare kompletterande annotering så som syntaktisk eller prosodisk annotering.

Annoteringen har skett manuellt med genomgång i många omgångar, vilket gör arbetet väldigt tidskrävande (jämfört med arbetsskeden vid automatisk annotering, till exempel den finlandssvenska talspråkskorpusen *Talko* som beskrivs av Södergård & Leinonen 2017). På grund av begränsade tids- och personresurser har vi vid annoteringen varit tvungna att bortse från språkets icke-manuella signaler och de är således inte medtagna i dessa konventioner. Icke-manuella signaler så som huvudets och kroppens rörelser, blickriktning, ansiktets mimik och munrörelser är viktiga betydelseförmedlande element i teckenspråkstexten. Trots att icke-manuella signaler inte annoteras i detta första skede⁹ beaktas de naturligtvis i korpusöversättningarna till svenska och vid indelning av tecken under olika ID-glossor i Signbank.

⁸ Oili Heimonens pro gradu-avhandling (2024) om teckentyper i FSTS baserar sig på den nästan färdigt annoterade korpusen.

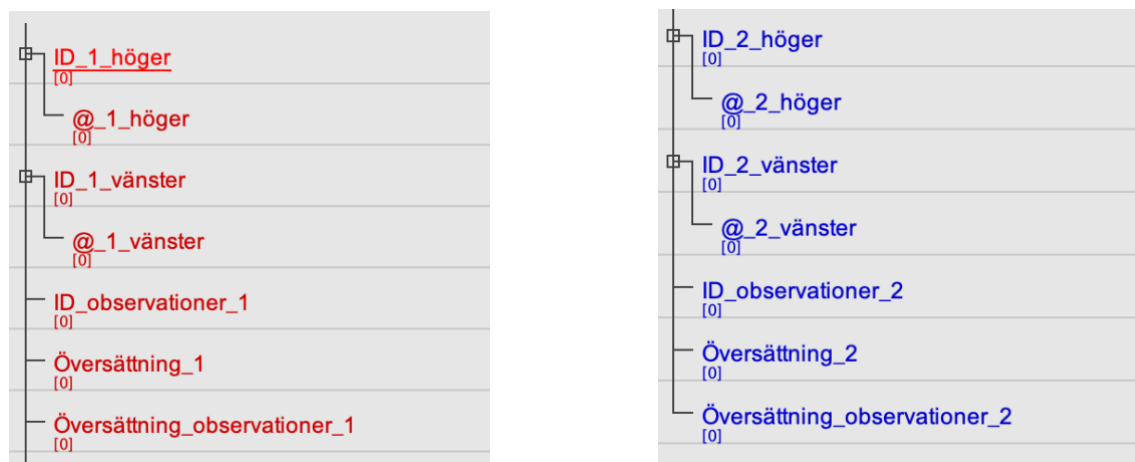
⁹ Undantag är de fall där det icke-manuella uttrycket utförs utan samtidigt manuellt tecken (och uttrycket inte utgör en uppbackning i dialogen). I dessa fall annoteras de med ZZZ (se 4.7 om huvudskakning och -nickning). Det andra undantaget gäller språkkontakt: finskinlånade munrörelser annoteras med @sk (se 4.4).

Som jämförelse kan nämnas att annoteringskonventionerna för STS har utvecklats under ett längre tidsspann och beaktar också icke-manuella komponenter (se Mesch & Wallin 2021). Vi hoppas att annoteringsarbetet för FSTS kommer att fortsätta så att konventionerna utvecklas och också icke-manuella signaler senare kan inkluderas i annoteringen.

Annoteringsanvisningarna i korpusarbetet för SVK vid JYU (Salonen m.fl. 2019a) har varit utgångspunkten för det finlandssvenska korpusarbetet (se Keränen m.fl. 2016 för en diskussion om annoteringskonventionernas användbarhet). Dessutom har konventionerna för STS (Mesch & Wallin 2021) fungerat som ett stöd i utarbetandet av FSTS-konventionerna. Genom att välja likartad grundläggande annotering underlättas också jämförande studier i språken. Som exempel kan nämnas att vi följt de allmänna klassificeringsriktlinjerna för tecken enligt modell i SVK (se kapitel 3). Indelningen torde vara tillräckligt grov för att inte kategorisera "för mycket". Också de fem kategorier av avbildande tecken som vi övertagit från konventionerna för SVK stöder forskningens behov. Detta trots att det är helt möjligt att framtida forskning i FSTS kommer att dela in den produktiva delen av lexikonet med hjälp av andra slags kategorier.

Johnston (2010, 2019) har definierat att en korpusannotering för tecknade språk bör bestå av ett minimum på tre annoteringsrader: en ID-rad för var hand och en rad för översättning. I vår FSTS-korpus har vi inalles 14 rader, vilket innebär 7 rader var för de två tecknarna i videon (Figur 2). Raderna med röd text syftar på tecknare 1, dvs. tecknaren till vänster i videorutan. Den blåa texten står för tecknare 2 som syns till höger i rutan.

På ID-glossans rader **ID_1_höger** och **ID_2_höger** annoteras höger hands aktivitet för respektive tecknare. Vänstra handens aktivitet annoteras på **ID_1_vänster** och **ID_2_vänster**. Dessa rader är överordnade rader som har underordnade rader (till exempel **@_1_höger**) anknutna till sig. På dessa underordnade rader annoteras grammatisk och annan språklig information (se kapitel 4) som berör tecknet som ID-glossan representerar. På **ID_observationer**-raden markeras eventuella observationer och kommentarer till tecknet som ID-glossan representerar. En översättning på satsnivå till svenska skrivs in på **Översättning**-raden och eventuella observationer eller kommentarer till översättningen skrivs in på raderna **Översättning_1_observationer** eller **Översättning_2_observationer**.



Figur 2. Annoteringsrader i ELAN.

Dessa annoteringskonventioner är sammanställda med två målsättningar. Dels är de ämnade att klargöra betydelsen av korpusannoteringarna och därmed stöda användningen av korpusen. Dels kan de vara till nytta vid annotering av ytterligare material vilket förhoppningsvis bidrar till att konventionerna utvecklas och uppdateras i framtiden.

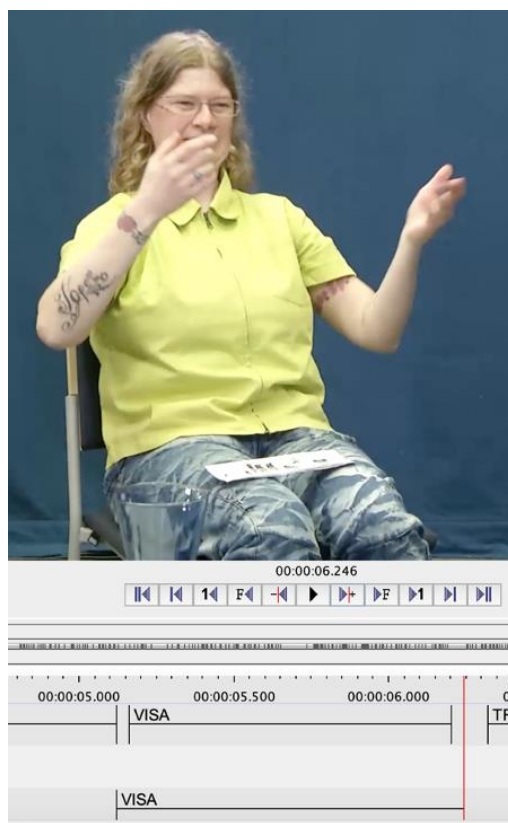
Konventionerna innehåller inte information om användningen av programmet ELAN. Sökfunktionen i ELAN är en av korpusens viktigaste funktioner och är en förutsättning för att hitta och studera språkliga fenomen. ELAN möjliggör sökning både i en enskild fil eller i flera filer samtidigt. Se FSTS videoinstruktion om hur korpusen och ELAN kan användas på webbsidan <https://vk-oppimateriaali.jyu.fi/fsts/>. En mer utförlig ELAN-handledning på svenska finns i [Bruksanvisning för annotering av teckenspråkstexter i ELAN](#) (Mesch & Cortes 2021) och på finska [CFINSL:n ELAN-ohjeet](#) (Keränen m.fl. 2017).

2.1 Annoteringscellens längd

Vid teckenspråkig artikulation består händernas rörelser dels av de rörelser som produceras då man utför tecknet, och dels av transportrörelser mellan två olika tecken. För att kunna annotera en teckenspråkig text måste man alltså avgöra var tecknet börjar och var tecknet slutar. Det här är inte alltid en entydig uppgift, eftersom tecknen i teckenspråkig artikulation inte utgör välavgränsade enheter utan gränserna suddas ut då tecknens artikulation går in i varandra.



Figur 3a. Annoteringscellens längd: exempel på början av annoteringscellen VISA för vänster hand.

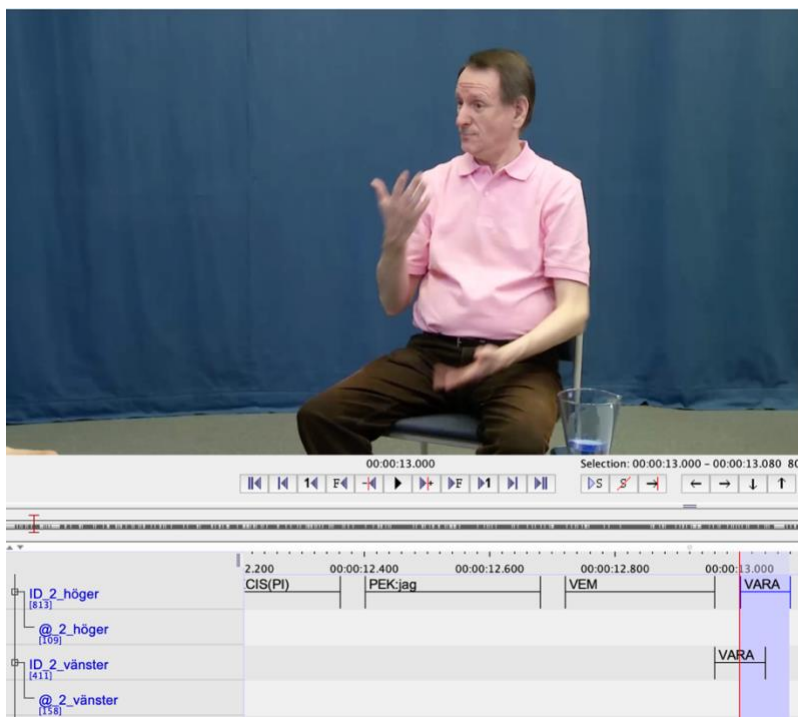


Figur 3b. Annoteringscellens längd: exempel på slutet av annoteringscellen VISA för vänster hand.

Vi följer kutymen i korpusarbetet för SVK (Salonen m.fl. 2019:6–7¹⁰) där man i stället för det enskilda tecknets längd vid annoteringen definierar **annoteringscellens längd** för ett tecken enligt följande (se exempel i figur 3a. och 3b.):

- Annoteringscellen **börjar** från den bildruta (videoruutu/frame) där man första gången kan identifiera en parameter av tecknet (i allmänhet handformen eller handflatans eller fingrarnas orientering).
- Annoteringscellen **slutar** vid den bildruta där något av tecknets parametrar (i allmänhet handformen eller handflatans eller fingrarnas orientering) för sista gången går att identifiera.
- För att sökningen i ELAN skall fungera skall man lämna minst en bildruta mellan två olika annoteringsceller.

Eftersom höger och vänster hand annoteras på var sin rad i ELAN kan annoteringscellerna för händerna sinsemellan vara olika långa. Det här gäller till exempel om händerna samtidigt producerar två olika tecken eller vid artikulation av ett tvåhandstecken (figur 4).



Figur 4. Exempel på skillnad i annoteringscellens längd för höger och vänster hand i tvåhandstecknet VARA.

2.2 Annotering med ID-glossor

När tecknet väl är identifierat och avgränsat ur teckenspråkstexten och segmenterat i en annoteringscell gäller det att kategorisera och namnge tecknet. Denna process att namnge de språkliga enheterna (tecknen) liknar processen med lemmatisering i lexikografiskt arbete då uppslagsformen ska definieras. De glossor¹¹ som används för att på ett systematiskt sätt

¹⁰ Diskussionen om tecknets längd för SVK baserar sig på Jantunen (2013, 2015).

¹¹ Vi har valt att hålla oss till ordformen "glossa", som användes i tidig teckenspråskorpuskontext på svenska (se Lyxell 2009). I STS korpusarbete i dag används "glosa".

identifiera och namnge varje tecken i korpusen kallas ID-glossor (Johnston 2010, 2019). ID-glossan utgörs av ett svenskt ord i grundform skrivet med versaler, och kan ha en efterföljande tilläggsdel. ID-glossan fungerar som en individuell namnskylt som identifierar vart och ett tecken som förekommer i korpusen.

Tecknet kan i korpusen förekomma i modifierad form, och också olika teckenvarianter kan stå under samma ID-glossa. Tecknets eller teckenvarianternas grundform(er) finns dokumenterade i den lexikala databasen Signbank (Signbank presenteras i kapitel 6). ID-glossor används av tradition för teckenspråkskorpusar eftersom de på ett effektivt sätt möjliggör den grundläggande idén med en korpus: sökfunktionen. När alla tecken är annoterade konsekvent och systematiskt hittar man deras förekomst i korpusmaterialet, och korpusen kan då utgöra en grund för vidare språklig analys och forskning i språket¹².

ID-glossan är tecknets synliga "etikett" och kunde i princip bestå av någon annan symbol, till exempel av en viss siffra. Men eftersom sifferkombinationer är svåra att minnas består varje ID-glossa av ett svenskt ord. Ordet som valts som ID-glossa för tecknet representerar (en av) tecknets mest grundläggande betydelser. Det är ändå skäl att minnas att tecknets betydelse kan avvika från det svenska ordet som syns som ID-glossans namn. Det är tecknets form som i första hand avgör hur tecknen indelas under olika ID-glossor (se 3.1 *Lexikala tecken* om principerna för hur tecknen grupperas under olika eller samma ID-glossa).

En tilläggsdel inom parentes kan ingå i ID-glossan. Detta tillägg skrivs ut för att skilja på tecken som har samma svenska ord som ID-glossa. Symbolerna inom parentes hänvisar till tecknets parametrar, och det är i första hand handformen som anges. Se tabell 1 för handformer och deras symboler.

Tabell 1. Symboler och bilder på handformer.¹³

A		Bc	
Al		Bq	
Ax		C	
B		D	

¹² Se till exempel Hodge & Crasborn (2022) för god annoteringspraxis.

¹³ Tabellen följer indelningen och namngivningen av handformer använd i korpusarbetet för SVK vid JYU (Salonen m.fl. 2019). Bilderna på handformerna ägs av Finlands Dövas Förbund rf och vi vill rikta ett tack till förbundet för tillståndet att använda bilderna här.

E	
F	
G	
Gi	
H	
I	
K	
L	
Lc	
Lo	
Lq	
M	
O	
7	

R	
S	
T	
V	
Vc	
VI	
W	
X	
Y	
2	
2o	
4	
5	
5c	

Om parenteserna innehåller två handformssymboler handlar det om ett tvåhandstecken. Exempel på ID-glossor i korpusen som innehåller en tilläggsdel inom parentes med information om handform:

BARA(AI)	BARA(BB)			
FARA-HEM(B)	FARA-HEM(5)			
FÖRSÖKA(5c)	FÖRSÖKA(VcVc)			
NU(H)	NU(BB)			
PASSA(BqBq)	PASSA(FF)			
TVINGA(AA)	TVINGA(GB)			
FLYTТА(55)	FLYTТА(BqBq)	FLYTТА(SS)		
ÅR(AI)	ÅR(Bq)	ÅR(GB)	ÅR(G)	
PRATA(5)	PRATA(55)	PRATA(B)	PRATA(BqBq)	PRATA(G)

Om handformen är likartad då skillnaden mellan två tecken skall markeras kan också andra parametrar användas. Dessa är handflatans (eller fingrarnas) orientering, tecknets rörelseart eller antal stavelser, och tecknets läge (artikulationsställe eller artikulationsplats). Dessutom kan tilläggsdelen bestå av en markering om enhands- eller tvåhandstecken. Se tabell 2 för exempel på dessa parametrar som används i ID-glossornas namn.

Tabell 2. Symboler för tecknets parametrar som används i namnet på ID-glossan.

Parameter som beskrivs i ID-glossan inom parentes	Exempel på markering	Exempel på ID-glossa
ORIENTERING (O_)		
	O_bakåt	SE(O_bakåt) FÖRSTA(O_bakåt) FRAMTID(O_bakåt)
	O_framåt	SE(O_framåt) FÖRSTA(O_framåt) FRAMTID(O_framåt) INTE(O_framåt)
	O_uppåt	INTE(O_uppåt)
	O_neråt	MÅSTE(O_neråt)
	O_mot-varandra	MÅSTE(O_mot-varandra)
RÖRELSE (R_)		
	R_framåt	ARG(R_framåt)
	R_mot-kroppen	ARG(R_mot-kroppen)
	R_förbi	BYTA(R_förbi)
	R_cirkel	BYTA(R_cirkel)
	R_turvis	JÄMFÖRA(R_turvis)
	R_närmande	JÄMFÖRA(R_närmande)
Antal stavelser i rörelsen	1S	BÖRJA(1S) KVAR(1S) SITTA(1S)
	2S	BÖRJA(2S) KVAR(2S) SITTA(2S)
LÄGET eller PLATSEN för artikulation (P_) ¹⁴		
	P_kind	OFTA(P_kind)
	P_hand	OFTA(P_hand)
	P_näsan	ONÖDIG(P_näsan)
	P_hakan	ONÖDIG(P_hakan)
	P_handflata	SKRIVA(P_handflata)
	P_neutral-utrymmet	SKRIVA(P_neutral-utrymmet)
Antal händer		
Enhandstecken	1-handstecken	SKRIKA(1-handstecken) HEM(1-handstecken)
Tvåhandstecken	2-handstecken	SKRIKA(2-handstecken) HEM(2-handstecken)

¹⁴ Vi har valt att använda symbolen (P_) för läget eller platsen där tecknet artikuleras i likhet med SVK-korpusen (Salonen m.fl. 2019a). Användningen av (L_) för läget undveks för att inte sammanblanda med symbolerna i SVK-korpusen där (L_) står för rörelsen (fi. liike).

3 KLASSIFICERING AV TECKEN PÅ RADEN FÖR ID-GLOSSAN

Vid annoteringen identifieras fem olika teckenkategorier: **lexikala tecken** (där **numeraler** utgör en undergrupp), **pekningar**, **avbildande tecken**, **gestlika tecken** och i korpusen förekommande **bokstaving**. Kategorierna markeras i korpusen på följande sätt:

- Lexikala tecken saknar i regel kategorimarkering (med undantag av _STS och _SVK)
 - Numeraler _num
- Pekningar PEK
- Avbildande tecken _at
- Gester _gest
- Bokstaving _bokst

En utmaning vid annoteringen är att det inte alltid är entydigt att identifiera ett tecken i en teckenspråkstext. Två tecken kan i teckenspråkstexten se likartade ut trots att de till sina grundformer är olika. Tecknets form kan modifieras – till exempel rörelsen eller platsen för artikulation ändras, upprepning utebli, tvåhandstecken artikuleras som enhandstecken – och (hand)formen kan påverkas av föregående eller efterkommande tecken.

För att avgöra vilket tecken det är fråga om kan man få hjälp av kontexten där tecknet förekommer, egenskaper hos tecknet så som munrörelsen och icke-manuella signaler eller tecknandets rytm och pausering. Exempel på tecken som i korpusen kan se likartade ut och kan vara svåra att åtskilja är: HANDFLATA-UPP_gest, VARA, PEK(B);, VARSÅGOD, ELLER(5), NU(BB) och GE(BB). I de fall det är osäkert vilket tecken det handlar om kan man också skriva in en kommentar på raden ID_observationer.

3.1 Lexikala tecken

Lexikala (eller fasta) tecken är sådana tecken som är lexikaliserade dvs. etablerade till både form och betydelse. De tilldelas alla en unik ID-glossa, så som:

BRO GE(Bq) LUGN ÖVER BORTSKÄMD TVINGA(AA) TILL

Ibland kan glossan bestå av flera ord som är kombinerade med bindestreck:

MISSTA-SIG MED-FLIT FINNAS-INTE FRAM-OCH-TILLBAKA

För ID-glossan väljs som princip det svenska ordets grundform. Ibland används i ID-glossan former som beaktar svenskans grammatik och läsbarheten (**EN-GÅNG_num** men **ETT-ÅR-SENARE_num**).

Ifall tecknet kan ha både en nominal och en verbal betydelse väljs ett svenskt verb i första hand som ID-glossa:

(FÖRSLAG, FÖRESLÅ) => FÖRESLÅ

Tecken med olika betydelser kan kombineras under samma ID-glossa i korpusen. Det här kan av korpusanvändaren kännas ovant till en början, då ett tecken som betyder 'glad' är annoterat med VILJA(R_mot kroppen) eller ett tecken med betydelsen 'sova' annoteras som HEM(1-handstecken). Det gäller alltså att minnas att ID-glossan inte alltid är den samma som tecknets betydelse. Tecknets olika betydelser syns i Signbank under *Översättningsmotsvarigheter* men lättast är det att kolla raden för översättning i korpusen i ELAN. Ur översättningen framkommer vad tecknet betyder i detta specifika sammanhang och man kan därmed få en helhetsbild av vad som tecknas och tolka betydelsen av enskilda tecken rätt. Principerna för hur ID-glossorna är indelade presenteras i figur 5.

- Tecken med **olika form** (skiljer sig åt med minst en parameter) och **olika betydelse** annoteras med olika ID-glossor:

ANVÄNDA	KÄNNA(HH)
FAVORIT	BE(BB)
VETA(R_mot-pannan)	HATT(1-handstecken)
- Tecken med **olika form** (minst 2-3 parametrar skiljer sig) och **samma eller likartad betydelse** annoteras med olika ID-glossor:

BETYDA(VV)	BETYDA(GB)
BOLL(BB)	BOLL(5c5c)
NÄR(P_kind)	NÄR(H)
- Tecken med **likartad form** (fonetisk variation i 1-2 parametrar) och **samma betydelse** kombineras under en ID-glossa:

HEMSK(5), HEMSK(B)	=> HEMSK(P_bröstet)
BARA(BB-1S), BARA(BB-2S)	=> BARA(BB)
KLÄDER(O_uppåt), KLÄDER(O_bakåt)	=> KLÄDER(2S)
- Tecken som har **samma form** (munrörelsen beaktas inte) men **olika betydelser** kombineras under en ID-glossa:
 - homonyma tecken

VÄNTA, NÖJD	=> VÄNTA(P_bröstet)
STÄMMA, BERÄTTA, VARJE	=> STÄMMA(BB)
 - polysema tecken med besläktad betydelse

VILJA, HOPPAS, TYCKA, GLAD	=> VILJA(R_mot-kroppen)
BEKVÄM, AUTOMATISK	=> BEKVÄM

Figur 5. Principer för indelning av tecken i ID-glossor.

Ibland är det andra än de manuella parametrarna som avgör att vi ändå skapat två olika ID-glossor för likartade tecken. I GÖRA(BB) och VARFÖR finns det skillnader i rytmen som gör att tecknen inte kombinerats. Också andra undantag finns.

Övriga annoteringskonventioner (för lexikala tecken)

- De tecken som språkbrukarna upplevt att inte hör hemma i FSTS har fått kategorimarkeringen **_STS** (tecken ur svenskt teckenspråk) eller **_SVK** (tecken ur finskt teckenspråk). Läs mera om dessa under *4.4 Språkkontakt*.
- Persontecken skrivs på raden för ID-glossan endast i formen *persontecken*.

3.1.1 Numeraler (_num)

Numeraler (siffertecken) kan antingen vara grundtal, det vill säga tecken för siffror i grundform, eller inkorporerade numeraler. Numeralerna markeras med kategoribeteckningen **_num**. Ordningsföljden för annotering av ID-glossans delar följer principen:

SIFFRA – STAM – TIDSMARKERING _num.

Bindestreck används i glossans namn i långa grundtal också för att förbättra läsbarheten:

- ELVA_num
- HUNDRA-FEMTIO_num
- TVÅTUSEN-TOLV_num
- TRE-VÅNING_num
- EN-GÅNG_num
- TVÅ-DAG-TIDIGARE_num
- FYRA-ÅR-SENARE_num

Grundtal

Grundtalen som förekommer i korpusen annoteras alla med kategorimarkeringen **_num**. Alla grundtal ur korpusen finns inte medtagna i Signbank. I korpusen finns ett begränsat antal förekomster av grundtal och vi har valt att överföra följande ID-glossor till Signbank:

- Jämna grundtal mellan NOLL_num och TJUGO_num, HUNDRA_num och TUSEN_num. Dessutom finns det i Signbank exempel på en del andra numeraler så som TVÅHUNDRA_num och TRETUSEN_num för att visa hur grundtal produceras i FSTS. Alla dessa markeras med **@numg** i Signbank.
- Eftersom FSTS avviker från SVK när det kommer till hur en del årtal uttrycks¹⁵ överförs hela årtal från korpusen till Signbank där de markeras med **@numår**. Till exempel:

NITTONHUNDRA-SJUTTIOSJU_num

Inkorporerade numeraler

Inkorporerade numeraler har en stam som består av ett visst tecken (till exempel VECKA, MÅNAD, VÅNING). Alla inkorporerade numeraler överförs till Signbank där de markeras med **@numi**. Till exempel:

FEM-TIMME_num
EN-MÅNAD_num
FYRA-VÅNING_num
TVÅ-VECKA-SENARE_num

¹⁵ Skillnaderna mellan hur SVK och FSTS uttrycker årtal hänger högst sannolikt ihop med kontakt med omgivningens talade språk, eftersom svenskan och finskan har olika sätt att uttrycka tid (sv: nittonhundra, fi: tuhat yhdeksänsataa).

ID-glossan för *ordningstal* skrivs ut i den form de har på svenska (inklusive eventuell tilläggsdel):

FÖRSTA(O_bakåt)_num
FÖRSTA(O_framåt)_num

FJÄRDE_num
FEMTE_num

ANDRA(R_neråt)_num

3.2 Pekningar

Pekningar kan ha olika funktioner i tecknade språk. Pekningar utgör en speciell grupp som inte uppfyller kriteriet för lexikaliserade tecken – att vara etablerade både till form och till betydelse. Se till exempel kap. 3.2 i Siltaloppi (2023) för en översikt över studier i pekningar i olika teckenspråk.

- Pekningar i korpusen annoteras med PEK:
- Pekning med tilläggsmarkering:
 - pekningar mot tecknarens kropp PEK:jag
 - pekningar med B-handform PEK(B):
 - pekning mot en listboj (se 4.6)
- Tecknet EGEN som uttrycker ägande annoteras EGEN:
 - EGEN mot tecknarens kropp markeras EGEN:min
- Tecknet SJÄLV annoteras SJÄLV:
 - SJÄLV mot tecknarens kropp markeras SJÄLV:jag

3.3 Avbildande tecken (_at)

Avbildande tecken annoteras med kategoribeteckningen **_at**. Indelningen av avbildande tecken i undergrupper följer Salonen m.fl. 2019a (kuvailevat viittomat **_kv**). Se tabell 3.

Tabell 3. Indelning av avbildande tecken.

Typ av avbildande tecken	Annotering i FSTS_JYU-korpusen (_at) ¹⁶	SVK_JYU-korpus (Salonen m.fl. 2019a)
Grepphantering Representerar alla slag av hantering – också de fall då handformen inte direkt "greppar" utan t.ex. formar en snögubbe, stryker med handen mot väggen eller knådar en deg	_at_gh	Käsittely (_kvkt)

¹⁶ Exempelsatser med avbildande tecken av olika typer finns presenterade i Signbank (<https://signbank.csc.fi/>) då man söker på **_at**.

Form och storlek	_at_fs	Muoto ja koko (_kvmk)
Hela objektet	_at_ho	Kokonainen kohde (_kvkk)
Tid och plats	_at_tp	Aika ja paikka (_kvap)
Abstrakt fixpunkt	_at_af	Abstraktinen kiintopiste (_kvak)
Oklart eller icke-kategoriserat fall	_at_x	Epäselvät tapaukset (_kvx)

Att vid annotering avgöra om tecknet är lexikalt eller avbildande har ibland sina utmaningar. Det är inte alltid helt klart när ett tecken är så pass etablerat till sin form och sin betydelse att det kan ses som ett lexikalt tecken, och när det borde annoteras som ett produktivt avbildande tecken. Speciellt krävande är det för FSTS då det inte finns någon forskning i avbildande tecken att tillgå. De lexikografiska studierna och publikationerna i FSTS har hittills fokuserat på att beskriva lexikala tecken. Många lexikala tecken kan ha sitt ursprung i avbildande tecken, och kan även i vissa kontexter "återta" sin avbildande funktion.

Till exempel gäller det för annoteraren att avväga ifall ett uttryck som beskriver en persons rörelse är en modifierad form av GÅ(R_fingrarna-vibrerar) eller om tecknet borde annoteras som avbildande _at_ho. Ibland kan en munnrörelse från ett talat språk utgöra en antydning om att det handlar om ett lexikalt tecken. Ibland måste avgörandet fattas endast baserat på annoterarens språkkänsla. Om tolkningen är oklar kan en kommentar skrivas in på raden för ID_observationer.

3.4 Gester (_gest)

Gester utgör en del av teckenspråkig diskurs, och de annoteras med kategoribeteckningen **_gest**. Det kan handla om allt från gestikulering till mer etablerade emblematiske gester. Att särskilja mellan tecken och gester i korpusen utan tidigare forskning i språket har varit utmanande. Som riktlinje har vi valt en bredare kulturell kontext: Ifall det handlar om en mer etablerad gest som används också av icke-teckenspråkiga ges gesten en egen ID-glossa. Exempel på sådana gestlika tecken (emblem) som har tillagts i Signbank är:

BRA(AI)_gest (figur 6)
SKRIKA_gest

HANDFLATA-UPP_gest
HEJ-HÖRDU_gest

HURRA_gest

Övriga gester och gestikulering markeras endast med beteckningen **_gest** och deras form finns inte närmare beskriven i Signbank. Också i en del oklara fall då ett tecken varken kan annoteras som lexikalt, avbildande eller bokstaverat med handalfabetet annoteras det med beteckningen **_gest**. Kategoribeteckningen **_gest** gör det möjligt att använda korpusen för att studera gesters funktion och användning närmare.



Figur 6. Skärmdump från Signbank BRA(AI)_gest

3.5 Bokstaving (_bokst)

Bokstaving med handalfabet annoteras med små bokstäver med ett bindestreck mellan varje bokstav. Bokstavingen markeras med _bokst i slutet av bokstavingen.

ID_1_höger: l-e-n-a_bokst

Annoteringen kan sökas endast i ELAN vilket betyder att markeringar med bokstaving inte överförs till Signbank. Här skiljer sig konventionerna för _bokst och kategoritagggen @bt (se 4.5). @bt markerar lexikala tecken och de införs därför i Signbank.

Bokstavingen annoteras i den form den förekommer i korpusmaterialet. Ifall bokstavingen avbryts markeras @avbr på ID-glossans underordnade rad. Hela ordet skrivs då ut på ID_observationer-raden. Om bokstavingen är oklar kan den markeras med zzz förutom de bokstäver som är urskiljbara, exempelvis:

zzz_bokst r-o-s-zzz_bokst

I fall där bokstavingen endast innehåller ordets första bokstav eller bokstäver, medan slutet av ordet artikuleras med munrörelser, skrivs en kommentar om munrörelsen på ID_observationer-raden:

ID_1_höger:	s_bokst
@_1_höger:	@avbr
ID_1_observationer:	[mun: snappertuna]
Översättning_1:	<i>Snappertuna</i>

I korpusen använder tecknarna tre olika handalfabet, både i renodlad och i blandad form. Vilket handalfabet som används markeras inte. Enda undantaget är de fall där informanten bokstaverar samma ord, namn eller term två eller flera gånger efter varandra med olika handalfabet. Handalfabeten namnges då enligt följande på ID_observationer-raden (taggen @sk används inte om bokstaving, se 4.4):

[gamla_bokst] = gamla handalfabetet som varit i bruk före 1965 i Finland

[nya_bokst] = det nya (internationella) handalfabetet som infördes 1965 i Finland

[STS_bokst] = handalfabetet som används i STS

Ifall valet av handalfabet har relevans för innehållet i det som tecknas och därmed påverkar översättningen, kan en kommentar om det här skrivas in på raden för översättningsobservationer.

4 ANNOTERING AV GRAMMATISK OCH ANNAN SPRÅKLIG INFORMATION

I korpusen annoteras förutom manuella tecken med ID-glossor också en del grammatisk information och vissa språkliga fenomen. Den här informationen markeras med hjälp av taggar som alla börjar med **snabel-a (@)** i likhet med de annoteringskonventioner som tillämpas i korpusarbetet för både STS och SVK. Se FSTS-korpusens alla kategoritaggarna i tabell 4.

Medan ID-glossorna skrivs in på överordnade rader i ELAN, markeras @-kategoritaggarna på ID-glossaradens underordnade rad. Det enda undantaget är taggen @ssg som annoteras på observationsraden för sammansättnings ID-glossor.

Ifall ID-glossan taggas för flera olika kategorier skrivs dessa in på den underordnade raden efter varandra utan mellanslag (till exempel @hs@neg@k) utan någon bestämd inbördes ordning.

Tabell 4. @-kategoritaggarna.

Grammatisk kategori eller språkligt fenomen	Kategoritagg i FSTS_JYU-korpusen	Motsvarande tagg i SVK_JYU-korpus
Negation	@neg	@neg
Upprepning	@upprep	@toisto
Sammansättning	@ssg	@y
Språkkontakt 1. Tecken ur andra teckenspråk 2. Gränsfall: tillfällig kodväxling eller redan etablerat lån? 3. FSTS-tecken i STS-betydelse, dvs. tecken som finns i båda språken men har olika betydelser 4. Finskinlånade munrörelser	@sk	saknas (i vissa fall @lv "lainaviittoma", dvs. 'lånetecken')
Tecken som innehåller bokstaverade element	@bt	@sv
Listboj/Listkonstruktion	@bojl	@poijul
Huvudskakning	@hs	@pp
Huvudnickning	@hn	@pn
Oklart fall (tecken eller betydelse)	@oklar	@epä
Söker tecken	@st	@hv
Avbrott	@avbr	@kes
Handen blir "kvar"	@k	@j

4.1 Negation (@neg)

I korpusen finns olika uttryck för negation som markeras med **@neg** på ID-glossans underordnade rad:

a) **Tecken som till sin betydelse är nekande** får en ID-glossa av typen TECKNETS GRUNDBETYDELSE-INTE. På det här sättet kan man som sökresultat få träff på både jakande och nekande form av tecknet då man söker på ID-glossan (till exempel TYCKA-OM). På den underordnade raden annoteras @neg.

ID_1_höger: TYCKA-OM-INTE(5)
@_1_höger: @neg
Översättning_1: *tycker-inte-om*

b) På motsvarande sätt markeras **sammansatta tecken som består av ett verbalt tecken som tillsammans med en negativ markör** (till exempel INTE) **skapar en sammansmält enhet**. På ID-glossans underordnade rad annoteras @neg.

ID_1_höger: VETA-INTE(BB)
@_1_höger: @neg
ID_1_vänster: VETA-INTE(BB)
@_1_vänster: @neg
Översättning_1: *vet-inte*

Andra exempel är:

HINNA-INTE	KUNNA-INTE(P_huvud)	BEHÖVA-INTE
------------	---------------------	-------------

c) **Negation som uttrycks endast med en huvudskakning** som utförs i samband med ett tecken annoteras både med @hs och @neg på ID-glossans underordnade rad.

ID_1_höger: FUNGERA
@_1_höger: @hs@neg
ID_1_vänster: FUNGERA
@_1_vänster: @hs@neg
Översättning_1: *fungerar-inte*

d) Lexikala tecken som till sin **grundbetydelse uttrycker negation men vars form inte explicit innehåller en negativ markör** (jämför med SVK "kieltopartikkelit", Salonen m.fl. 2019a) annoteras också med @neg på ID-glossans underordnade rad. Exempel på dessa är

NEJ	ALDRIG	INTE(O_uppåt)	INGENTING
KUNNA-INTE(P_näsan)	OMÖJLIG(X)	FATTAS(2S)	

4.2 Upprepning (@upprep)

Taggen **@upprep** används i de fall då tecknet utförs många gånger efter varandra eller rörelsen upprepas. Grundformer som i sig innehåller en upprepad rörelse (till exempel ALLTID) skall inte markeras med @upprep. Taggen används endast om tecknet har en tillagd rörelse. Korpusen visar att upprepning i FSTS bland annat kan användas för att uttrycka plurala former av nominala tecken (figur 7) eller uttrycka modifierade former av verbala tecken (figur 8).

ID_2_höger:	KOMPIS
@_2_höger:	@upprep
ID_2_vänster:	KOMPIS
@_2_vänster:	@upprep
Översättning_2:	<i>kompisar</i>

Figur 7. Exempel på annotering med taggen @upprep



ID_1_höger:	HOS
@_1_höger:	@upprep
ID_1_vänster:	HOS
@_1_vänster:	@upprep
Översättning_1:	<i>besöker-flera-olika-platser</i>

Figur 8. Exempel på annotering med taggen @upprep

Som det framgår ovan ändras inte ID-glossans form, utan den kvarstår i singular form. Undantag utgör sådana ID-glossor där ordet i svenskan används i pluralform. Dessa skrivs i sin etablerade pluralform trots att tecknet inte är i pluralform. Dessa ID-glossor skall inte markeras med taggen @upprep. Exempel är:

FÖRÄLDRAR	GLASÖGON	BYXOR	FINLANDSSVENSKA-TECKENSPRÅKIGA
-----------	----------	-------	--------------------------------

Observera att trots att det finns olika typer av upprepning annoteras de alla med samma kategoritagg @upprep. Det kan till exempel handla om PEK: som upprepas på samma plats i teckenrummet eller om en upprepning som förflyttas längs en tidslinje i teckenrummet.

4.3 Sammansättning (@ssg)

Avsaknaden av forskning i FSTS medför bland annat att sammansättningar utgör en stor utmaning att annotera. Det saknas konsensus också inom forskningen i tecknade språk om vad som utgör en sammansättning, definitionerna varierar¹⁷. Det gäller därför för annoteraren att ta ställning till när två på varandra följande tecken utgör en egen enhet. För att garantera att alla potentiella sammansättningar lätt kan hittas i korpusen och studeras närmare markeras dessa kandidater med @ssg på ID_observationer-raden. Denna rad väljs eftersom @ssg hänvisar till flera än en ID-glossa. Då man söker efter sammansättningar framkommer deras betydelse på översättningsraden ifall betydelsen inte är uppenbar via dess två delar.

Sammansättningar kan kännas igen både till betydelsen och/eller formen. Till sin betydelse står sammansättningen för ett eget nytt begrepp. Till sin form kan den utgöra en sammansmält enhet där pausen mellan de två tecknen har krympt eller fallit bort. Ofta har det skett fonologiska förändringar när de två tecknen bildat en enhet. Det kan till exempel handla om handformen där förleden kan ha assimilerats med efterledens handform som i en av varianterna i DÖVFÖRENING. Förledens rörelse kan vara komprimerad som i TECKENSPRÅK. Ibland kan en munrörelse som sträcker sig över två teckensekvenser visa på att det handlar om en sammansättning. Ett exempel är tecknet som betyder 'berg' och består av STEN +_at_fs som utförs simultant med munrörelsen [be].

En stor del av de sammansättningar som markerats i korpusen förs inte in som en egen ID-glossa i Signbank utan de annoteras med två skilda ID-glossor. Förutom att sammansättningar är svåra att definiera har detta varit ett sätt att spara på de knappa annoteringsresurserna i det första skedet av grundannoteringen. Trots att sammansättningarna annoteras med två separata ID-glossor kan de alltså ändå lätt hittas i korpusen genom en sökning med taggen @ssg. Många av sammansättningarna handlar om översättningslån (från svenskan) av typen:

HÄLSA + STATION	'hälsostation, hälsocentral'
ÅLDER(P_bröst) + GRÄNS	'åldersgräns'
KANIN + FÖRSÖKA(VcVc)	'försökskanin'
KAMERA + BUTIK	'Photoshop'

¹⁷ För jämförelsens skull kan det konstateras att i annoteringskonventionerna för STS skiljer man vid annotering av sammansatta tecken mellan *översättningslån* (^) och *sammansmälta teckensekvenser* (*), se Mesch & Wallin (2021).

Sammansättningar annoteras som en egen enhet i en gemensam cell och får en egen ID-glossa som förs in i Signbank i följande fall:

- De utgör tydligt en sammansmält enhet med en egen betydelse och uppfattas nästan som ett enda tecken.

BETYDA(GB)

AVSLUTA(R_vinkelrät)

I-DAG

FARA-HEM(B)

- Man urskiljer ännu delarna trots formförändringar men det handlar ändå om ett eget begrepp.

TECKENSPRÅK

TILL-SIST

BEHÖVA-INTE

DÖVFÖRENING

- De förknippas med teckenspråkig kontext och kultur eller är ofta förekommande i korpusen.

DÖVSKOLA

DAGHEM

ARBETSPLATS

SNÖGUBBE

- De utgör egennamn.

HUFVUDSTADSBLADET

KORSNÄS

STOCKHOLM

Observera att sammansättningar av typen FINNAS-INTE annoteras på både höger och vänster ID-rader trots att förleden FINNAS endast tecknas med ena handen:

ID_1_höger:	FINNAS-INTE
@_1_höger:	@neg
ID_1_vänster:	FINNAS-INTE
@_1_vänster:	@neg
ID_1_observationer:	@ssg
Översättning_1:	<i>finns-inte</i>

4.4 Språkkontakt (@sk)

Av naturliga skäl finns det mycket i FSTS som har likheter med SVK och STS eftersom dessa tre språk är besläktade. Det finlandssvenska teckenspråket har dessutom utvecklats under årens lopp via kontakt med dessa (och andra) språk. Som konstaterats i inledningen har finlandssvenska teckenspråkiga kompetens i flera olika språk och använder många språk i sin vardag. I den bakgrundsinformation som informanterna i korpusen delat med sig av framkommer det att förutom FSTS har många informanter kompetens också i SVK och STS. Utöver kunskaper i svenska har många också kunskaper i bland annat finska och engelska.

Språkkontakt med olika språk tar sig i korpusen uttryck på varierande sätt. Påverkan från svenskan syns i teckenbildningen bland annat som etablerade översättningslån (sammansättningar av typen FRISÖRLINJE, TILL-SIST, GLÖMMA-BORT och SE-UT), tecken som innehåller bokstaverade element

(se 4.5) och bokstaverings med handalfabetet (se 3.5). Kontakt med andra teckenspråk, så som till exempel STS eller SVK syns vid användning av tecken som finlandssvenska teckenspråkiga anser att inte hör hemma i FSTS eller som utgör (tillfällig) inlånning eller kodväxling. Indirekt språkkontakt med finska syns också i form av finskinlånade munrörelser via kontakten med SVK¹⁸.

Eftersom vi som byggt korpusen vid JYU inte har FSTS som modersmål har vi jobbat i nära samarbete med de finlandssvenska språkrådgivarna vid Finlands Dövas Förbund rf och med teckenspråkiga som sammankommit till korpusworkshoppar under korpusbyggets gång (8 sammankomster under åren 2022–2024). Målsättningen med annoteringen av kontaktfenomen har varit att följa den intuition om språkbruket som språkbrukarna har. Det har handlat om en balansgång mellan insikten om att FSTS varierar väldigt mycket på individnivå, och allting egentligen är "språkande" på ett kontinuum, till att försöka följa språkbrukarnas språkliga magkänsla från fall till fall. För att kunna markera @sk har man tillsammans diskuterat "gränser" mellan de olika teckenspråken och vad som kan anses etablerat i språket och vad som ännu känns nytt.

När det kommer till FSTS och dess "gränser" visade workshopdiskussionerna att de olika individernas språkbruk och språkkänsla för språkkontakt varierar. Att försöka skilja mellan vad som är tillfällig och vad som är etablerad språkkontakt i språkgemenskapen är ingen okomplicerad uppgift¹⁹. När handlar det om kodväxling, när är det språkförändring och när är det fråga om försvagade språkkunskaper och attrition, dvs. språklig förvittring hos språkbrukare av ett hotat språk (se Hoyer 2013)?

Med kategoritagggen @sk markeras endast sådan kontakt som kan anses som "inte helt etablerad" under den tid korpusen varit under uppbyggnad (tabell 5). En del av den språkkontakt som vi nu markerat kan sannolikt komma att bli en etablerad del av språket i framtiden. Korpusens videoinspelningar utgör en ögonblicksbild av språkbruket hos en begränsad mängd teckenspråkiga i en viss social kontext med välavgränsade samtalsämnen under en viss begränsad tid under åren 2015 och 2017. Språkbrukarnas språkliga intuition när det kommer till annotering av språkkontakt är däremot av senare datum och härstammar från tiden för korpusbygget 2021–2024.

Korpusannoteringarna är inte avsedda att styra språkbruket eller vara normerande. Men när det kommer till ett språk som FSTS (med begränsad forskning från förr) kommer man inte ifrån att detta systematiska klassificeringsarbete kan ha normativa följder för språket. Det här är något vi som annoterat är medvetna om och till en del inte går att undvika eftersom korpusens grundidé är att kategorisera och namnge delar i språket. Annoterarnas språkliga kompetens är också en faktor som påverkar annoteringsresultatet – för att kunna identifiera språkkontakt behövs förstås språkkunskaper i minst dessa tre olika teckenspråk och svenska och finska. Se annoteringsexempel i tabell 6.

Värt att poängtera är att den språkpåverkan som är en etablerad del av FSTS i dag alltså inte beaktas i annoteringen av språkkontakt. Etablerad språkkontakt från andra språk kan i korpusen dock sökas på många olika sätt, till exempel:

- via sökning på kategorin bokstaverings (_bokst)

¹⁸ Se också Huhtinen (2020) för en diskussion om språkkontakt och korpusannotering för FSTS.

¹⁹ Också Södergård & Leinonen (2017) beskriver annoteringsutmaningar med att avgöra skillnaden mellan tillfälliga och etablerade lån vid korpusbygget över den talade svenskan i Finland.

- via sökning med taggen @ssg hittas en del översättningslån
- med taggen @bt (se 4.5) hittas tecken som innehåller bokstaverade element.

Med taggen @sk markeras fyra kategorier av "icke-etablerad" språkkontakt:

1. GLOSSA_STS och GLOSSA_SVK markeras alltid med @sk på ID-glossans underordnade rad i ELAN. Dessa tecken anses till sin form och betydelse inte höra hemma i FSTS utan i STS eller SVK enligt respektive kategorimarkering. I Signbank tilläggs för deras del @sk i fältet för "Mer information".
2. Tecken där språkbrukarnas åsikter går isär när det gäller om tecknet är ett etablerat lån i FSTS eller bara en tillfällig kodväxling/inlån markeras endast med @sk på ID-glossans underordnade rad. Dessa har alltså inte någon kategorimarkering i själva ID-glossan. I korpusen har största delen av tecken av denna typ sitt ursprung i STS. Till denna grupp hör också sådana tecken där man i teckenbildningen kan se tydlig språkkontakt med SVK och därmed indirekt finskan. Om det handlar om en sammansättning markeras @sk på ID_observationer-raden. I Signbank tilläggs @sk under "Mer information".
3. Om ett tecken i korpusen används i den betydelse tecknet har i STS ges tecknet inte en egen ID-glossa (GLOSSA_STS), utan det annoteras med den ID-glossa som hänvisar till tecknet i FSTS. Därtill markeras @sk på ID-glossans underordnade rad. Både @sk och tecknets betydelse i STS markeras i Signbank.
4. Finskinlånade munrörelser annoteras med @sk på ID-glossans underordnade rad. På ID_observationer-raden skrivs den finskinlånade munrörelsen ut (hela eller delar av motsvarande finska ord) med markeringen [fi. mun:]. I oklara fall annoteras både @sk och [fi. mun:] endast på ID_observationer-raden och markeringarna följs av frågetecknet.

Tabell 5. Olika typer av språkkontakt som markerats i korpusen.

Typ av språkkontakt som markeras med @sk	EXEMPEL på @sk
1. Tecken ur andra teckenspråk	FJORTON_num_STS FOLK_STS TIETOKONE_SVK
2. Gränsfall: tillfällig kodväxling eller redan etablerat lån?	STS: • FÖR • OVAN • SLAPPNA-AV SVK: • PERSONLIG (@ssg MÄNNISKA+MÖTA)

	RÖR (i betydelsen 'på raken')
3. Tecken som används i den betydelse tecknet har i STS	ELEV i STS-betydelsen 'yrke' FÖRSTÅ(G) i STS-betydelsen 'ont' FATTA(55) i STS-betydelsen 'kultur'
4. Finskinlånade munrörelser	FINLAND [fi. mun: suomi] TECKEN [fi. mun: vi] AVSLUTA(R_ifrån-varandra) [fi. mun: lop]

Egennamn så som ortnamn eller namn på instanser eller aktörer i utlandet (HÄRNÖSAND, BORNHOLM, VIRGINSKA-SKOLAN) markeras **inte** med @sk i korpusen.

Tabell 6. Exempel på annoteringar med taggen @sk

ID_1_höger:	FJORTON_num_STS	FINLAND	ÄNNU-INTE
@_1_höger:	@sk	@sk	
ID_1_vänster:	FJORTON_num_STS		
@_1_vänster:	@sk		
ID_1_observationer:		[fi. mun: suomi]	@sk? [fi.mun: vie]?

4.5 Tecken som innehåller bokstaverade element (@bt)

Med kategoritaggen **@bt** markeras tecken som innehåller element från handalfabetet. I korpusen finns det olika typer av tecken taggade med @bt:

- Lexikaliserade bokstaveringar utgör en sammansmält helhet där den ursprungliga bokstaveringen till sin form ändrats eller reducerats²⁰.

OM JA HELSINGFORS(F) NU(H) HOBBY UT EU

- Initialtecken där handformen utgör en bokstav från handalfabetet.

TACK FIN(F) NEJ ONSDAG TE TUR NÄR(H)

- Övriga: till exempel sammansättningar som delvis innehåller element från handalfabetet.

I-DAG KRISTINASKOLAN

²⁰ Denna typ av tecken kallas *bokstaverade tecken* i STS, se Institutet för språk och folkminnen (2024).

4.6 Listboj/Listkonstruktion (@bojl)

Listbojar eller *listkonstruktioner* (se närmare Siltaloppi 2019, 2023) annoteras med kategoritaggen **@bojl**. Tecknarens båda händer utgör delar i en språklig enhet där oftast den icke-dominanta handen utgör målet för en pekning utförd av den dominanta handen (figur 9). I Signbank finns ID-glossorna PEK:tummen, PEK:pekfinger, PEK:långfinger, PEK:ringfinger och PEK:lillfinger som väljs vid annoteringen beroende på vilket finger som pekningen riktas mot. Den hand som är målet för pekningen annoteras med den ID-glossa som handformen motsvarar bland numeralerna (till exempel TRE_num eller FYRA_num). På den underordnade raden för denna ID-glossa markeras **@bojl**. Betydelsen av konstruktionen i den kontext där den förekommer framkommer på Översättning-radens.

Dominant hand (höger)

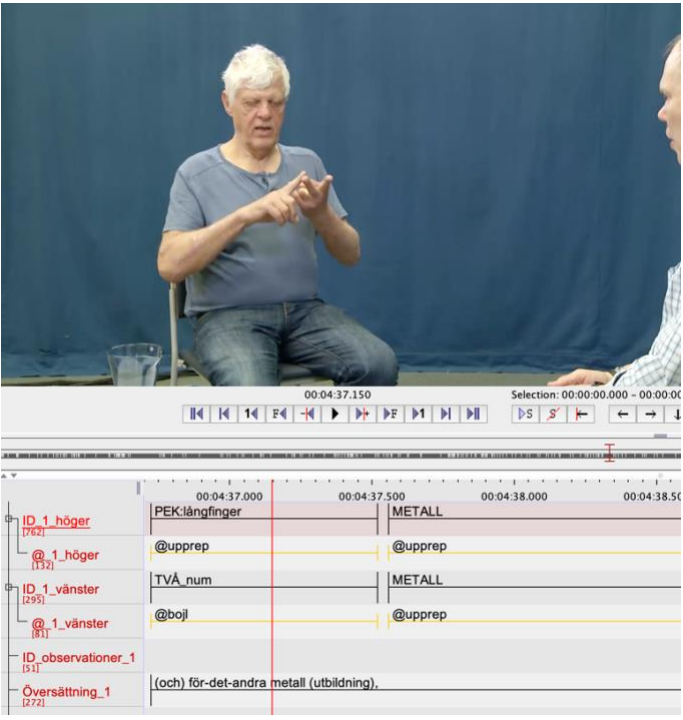
- vilket finger pekats på

Icke-dominant hand (vänster)

- hur många fingrar
- på underordnad rad @bojl

I bildexemplet:

ID_1_höger:	PEK:långfinger
@_1_höger:	@upprep
ID_1_vänster:	TVÅ_num
@_1_vänster:	@bojl
Översättning_1:	(och) för-det-andra metall (utbildning),



The figure illustrates the annotation of listboj/listkonstruktion. The top part is a video frame showing a man signing. The bottom part is the ELAN annotation interface, showing a timeline with various tracks. The tracks include ID, glosses, and translation. The glosses are PEK:långfinger, @upprep, TVÅ_num, and @bojl. The translation is (och) för-det-andra metall (utbildning).

Figur 9. Exempel på annotering av listboj/listkonstruktion.

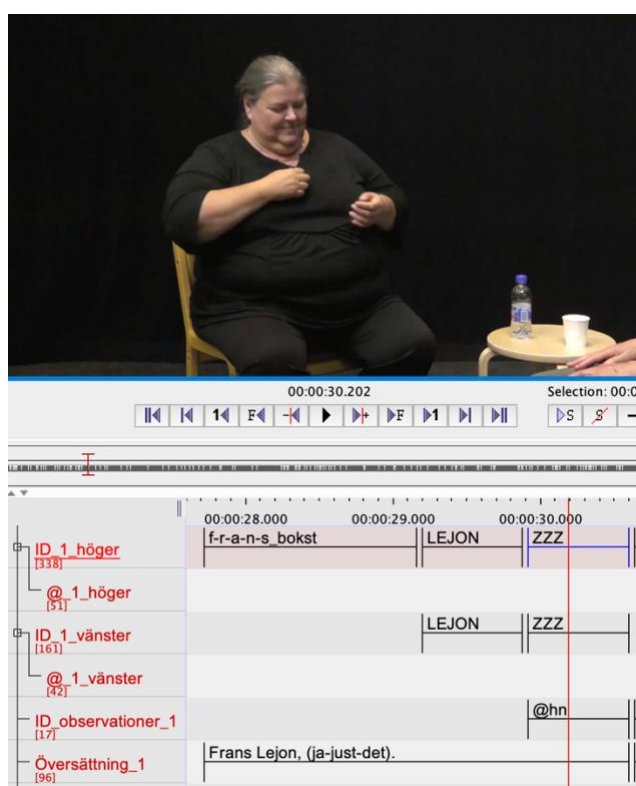
4.7 Huvudskakning (@hs) och huvudnickning (@hn)

Icke-manuella signaler, så som till exempel kroppens och huvudets rörelser, är betydelsebärande element i tecknade språk. Annoteringskonventionerna för STS (Mesch & Wallin 2021:43–44) innehåller instruktioner om hur huvudets aktivitet kan annoteras. Olika typer av huvudrörelser kan markeras detaljerat på en egen rad i ELAN. Också i SVK finns det studier i huvudrörelser och deras funktion (till exempel Puupponen 2012, 2019)²¹.

²¹ Se Puupponen m.fl (2016) för en jämförande studie i SVK och STS.

Även i korpusen i FSTS förekommer det huvudrörelser: en nickning med huvudet kan ha många funktioner, den används både i pågående taltur och som respons till den som tecknar för tillfället. Huvudskakning används åtminstone för att uttrycka nekande betydelse. Forskning i användning av huvudrörelser i FSTS behövs.

I vår grundannotering markeras huvudskakning vid negation på ID-glossans underordnade rad med @neg@hs. För övrigt annoteras huvudnickning (@hn) och huvudskakning (@hs) endast då tecknaren inte artikulerar något manuellt. Med andra ord annoteras nickning eller skakning på huvudet i de fall inga tecken samtidigt utförs med händerna. I dessa fall markeras ZZZ på ID-glossaraden (både ID_höger och ID_vänster) och @hn eller @hs på raden ID_observationer (figur 10). Värt att observera är att det här endast gäller för den tecknare som har ordet för tillfället. Den samtalsrespons som lyssnaren ger i form av huvudnickning annoteras inte i detta korpusens grundannoteringsskede. Huvudrörelserna (liksom andra icke-manuella signaler) beaktas däremot i översättningen med strävan att hitta ett motsvarande uttryck på svenska.



Figur 10. Exempel på annotering av huvudnickning @hn.

4.8 Oklart fall (@oklar)

I en del fall kan det vara svårt att uppfatta vad personen tecknar eller så kan tecknet till sin form eller betydelse vara oklart. I dessa fall markeras @oklar på den underordnade raden till ID-glossan. Som ID-glossa väljs det tecken som annoteraren antar att det handlar om. Ifall det inte är möjligt att identifiera tecknet överhuvudtaget annoteras ZZZ (finns i Signbank) på raden för ID-glossan. På ID_observationer-raden kan man skriva in en kommentar om möjliga eller alternativa tolkningar till tecknet.

4.9 Söker tecken (@st)

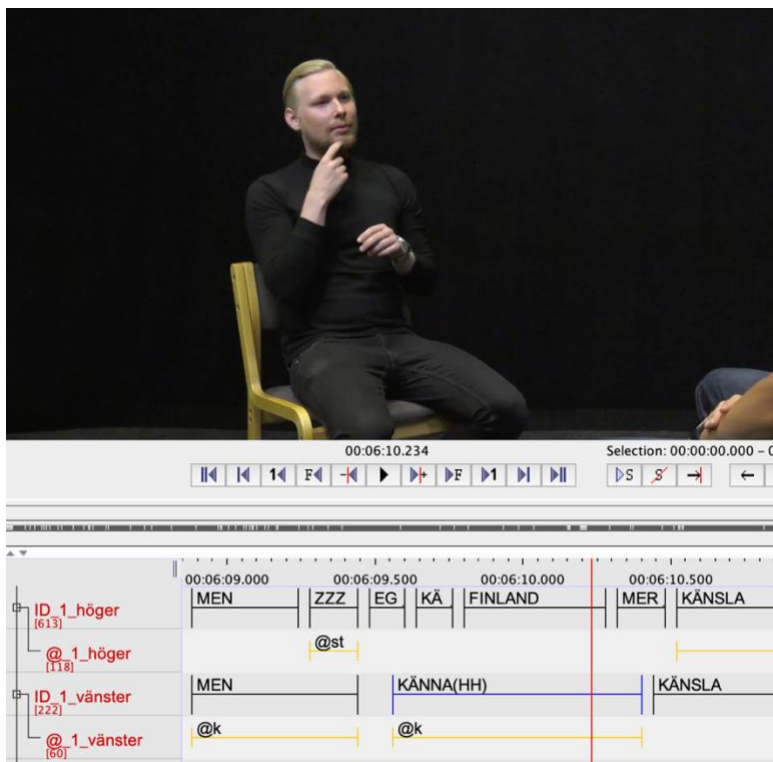
När tecknaren stannar upp i sitt tecknande och tvekar hur hen skall fortsätta annoteras det med ID-glossan ZZZ och @st på den underordnade raden. @st står för "söker tecken" och kan synas som en paus med funderande minspel och spelande fingrar.

4.10 Avbrott (@avbr)

Det förekommer att tecknaren avbryter sig och tecknet inte utförs i sin helhet. Dessa fall annoteras med ID-glossan för tecknet som blir avbrutet och @avbr markeras på den underordnade raden till ID-glossan. Eventuella kommentarer (till exempel om avbruten bokstavering) kan skrivas in på raden ID_observationer.

4.11 Handen blir "kvar" (@k)

Om ena handen hålls kvar längre än fem bildrutor efter att tecknet artikulerats, markeras det här med @k på ID-glossans underordnade rad (se exempel i figur 11). Att dessa fall markeras i korpusen innebär att fenomenet och vilka funktioner det har kan studeras närmare. Det kan finnas olika mer eller mindre betydelsebärande orsaker till att handen blir kvar.



Figur 11. Exempel på annotering av @k. Vänster hand är kvar i handformen från föregående tecken KÄNNA(HH) medan höger hand artikulerar FINLAND.

4.12 Raden *ID_observationer* och principer för att skriva in observationerna

De flesta @-kategoritagggar annoteras på ID-glossans underordnade rad. Undantagen är **@ssg** och **@hn** som båda annoteras på raden *ID_observationer*. På raden för *ID_observationer* markeras också huvudskakning **@hs**. Huvudskakning vid negation markeras däremot på ID-radens underordnade rad (se kapitel 4.7).

På raden *ID_observationer* markeras också oklara fall av språkkontakt för finskinlånade munrörelseers del med **@sk** följt av ett frågetecken, **@sk?**. På raden för *ID_observationer* markeras de munrörelser som annoteras som språkkontakt (se 4.4) med markeringen [**fi. mun:**]. Dessutom markeras här den allmänna kommentaren **icke-manuellt uttryck** då ID-glossan är ZZZ och inget tecken artikuleras med händerna.

På raden kan man skriva in också andra kommentarer vid behov – till exempel vid oklara fall eller då det finns möjlighet till en alternativ tolkning. Dessa kommentarer skrivs med liten begynnelsebokstav och utan punkt. Ifall det finns många markeringar eller observationer på raden särskiljs dessa med semikolon, exempelvis:

@sk?; eventuell påverkan från finskans "tällä hetkellä"

eller ÄLG @avbr; eller _at_ho

eller KOMMA(G); eller PEK:

5 ÖVERSÄTTNING

5.1 Allmänna översättningsprinciper

I korpusen skrivs översättningen till svenska in på raden **Översättning**, och eventuella observationer eller kommentarer skrivs in på raden **Översättning_observationer**. Korpusöversättningen är inte helt självständig utan är avsedd att användas tillsammans med de teckenspråkiga videofilmerna. Översättningsprinciperna utgår från praxisen utvecklat för arbetet med JYU finska teckenspråkskorpus (Pippuri 2015; Salonen m.fl. 2019a). Principerna är kompletterade och utvecklade för att passa översättning till svenska. Målet med översättningen är att den både ska utgå ifrån och avspegla teckenspråkets uttryckssätt. På det här sättet kan översättningen utgöra ett stöd för forskningen i FSTS. Översättningen möjliggör också mer omfattande (tematiska) sökningar i korpusen.

En bra korpusöversättning är strukturerad så att den följer de tecknade satsernas uppbyggnad och betydelse. Samtidigt ska översättningen återge helheten av det teckenspråkiga innehållet på en möjligast idiomatisk svenska med motsvarande fraser, satser och meningar. Det här innebär att översättningen ofta kan stå närmare talat språk än en traditionell översatt skriven text. I översättningen ska bland annat avbrott, pauser och upprepningar synas, på motsvarande sätt som i en transkriberad text av ett samtal. Tvekan (*öh/eh, hur-var-det-nu, alltså*), återkoppling eller respons från lyssnare till tecknare (*mm, hm, just-det*) skall också komma fram i översättningen.

Det är skäl att komma ihåg att man på raden för ID-glossorna endast annoterar den information som händerna förmedlar, det vill säga manuella tecken. Därför ska övrig information, så som icke-manuella signaler (ansiktets mimik, huvud- och kroppsrörelser) synas i översättningen. Ibland kan till exempel frågor eller negation uttryckas endast på icke-manuell väg och syns då i översättningen, men inte på raden för ID-glossan. Ansiktsuttrycket, så som höjda ögonbryn och uppspärade ögon kan förmedla stark betoning (*menar-du-det?, verkligen?, det-var-värst!*). Ibland kan det räcka med ett utropstecken för att markera samtalstonen.

Översättaren ska sträva efter att dela upp teckenspråkstexten i lämpliga delar genom att följa teckenspråkets satsstruktur. Typiskt för det teckenspråkiga uttrycket är långa kedjor med många verbala tecken. Gränsmarkörer för satsbyte i kedjelika yttranden kan i tecknade språk vara icke-manuella markörer så som ögonkontakt, markering med blicken, huvudets och överkroppens rörelser och vridning, eller minimala pauser i yttrandet. Långa yttranden kan i den svenska översättningen med fördel kapas av till kortare sekvenser i form av huvudsatser och bisatser avskilda med passande skiljetecken (, ; -), eller till flera separata meningar. Översättaren får ofta förlita sig på sin språkliga intuition för att göra logiska översättningshelheter.

För att översättningen skall motsvara det ibland väldigt komprimerade teckenspråkiga uttrycket består den svenska översättningen ofta av flera ord än det finns manuella tecken. De ord som i översättningen inte har en explicit motsvarighet på teckennivå skall sättas inom parentes eller kombineras med bindestreck. Det här görs för att översättningen skall spegla det teckenspråkiga manuella uttrycket så bra som möjligt. Orden i parentes eller kombinerade med bindestreck

behövs för att förmedla betydelsen så exakt som möjligt och för att översättningen skall fungera på svenska.

5.2 Parenteser och hakparenteser

- Parenteser används för att i översättningen visa vilka ord som inte har en direkt motsvarighet till manuella tecken i det teckenspråkiga uttrycket.
EGEN:min NAMN l-e-n-a_bokst => *Mitt namn (är) Lena.*
 - Svenskans artiklar (*en, det, de*), konjunktioner (*och, eller, men*), subjunktioner (*så, att, som*) och prepositioner (*på, i, till, för, av*), adverb och partiklar (*ju, visst, väl, nog*) är ord som i översättningen ofta skall sättas inom parentes eftersom de inte alltid har en direkt manuell motsvarighet i teckenspråkssatsen.
 - Om det finns något som uttrycks endast via kroppsrörelse eller mimik så skall svenskans motsvarande adverb eller adjektiv sättas inom parentes.
(*Nedstämd*) *gick-(han)-därifrån.*
Jag såg-på (honom) (och tänkte)...
 - I teckenspråkssatsen behöver det inte synas explicit vem det är som utför handlingen utan detta kan framkomma av kontexten. Då måste ett subjekt inom parentes tillföras den svenska översättningen.
BÄRA _at_gh => (*Mannen*) *tar-bagen (och) öppnar-bagageluckan,*
- Det samma gäller då det är en pekning (eller blickriktning) mot en känd referent i teckenutrymmet som skall översättas.
PEK: FÖRSÖKA(VcVc) KLÄDER(2S) => *Snögubben provar kläder.*
- Om många ord efter varandra sätts inom parentes av samma orsak, kan orden sättas inom samma parentes.
(*Sen för det*) *andra,*
 - Förklarande tillägg av slaget [*tecken visas*] eller [*personen visar*] sätts inom hakparentes.
PEK:jag l-e-n-a_bokst => *Jag (heter) Lena.*
PEK:jag persontecken => *Jag (är) Lena [persontecken visas]*

Hakparenteserna används främst för läsbarhetens skull så att man lättare kan skilja åt dessa utomtextuella hänvisningar från översättningens vanliga parenteser. Kommentarer om munrörelser och val av handalfabet markeras också med hakparenteser i korpusen.

5.3 Bindestreck

- Bindestreck skall användas om man i den svenska översättningen måste använda sig av flera än ett ord för att uttrycka information som framkommer i ett enda tecken (*fara-hem,*

gå-av, sätta-på). Ifall bindestreck inte används i detta fall, kan man felaktigt tro att det handlar om två olika tecken.

- Bindestreck används ofta i översättning av sådana teckenspråkiga uttryck som kodar in mycken information eller händelser som följer på varandra. Till exempel det avbildande tecknet:

ID_1_höger: _at_gh
 @_1_höger: @upprep@k
ID_1_vänster: _at_gh
 @_1_vänster: @upprep
Översättning_1: *(då)-jag-bläddrar-och-tittar-i-tidningar*

Om bindestrecken utelämnas från översättningen blir tolkningen felaktigt att det handlar om en teckenspråkig fras bestående av många tecken efter varandra.

- Händelseförlopp som beskrivs via samtidig artikulation med båda händerna skall markeras med bindestreck. Exempelvis då all informationen i det teckenspråkiga uttrycket uttrycks simultant med hjälp av enda avbildande tecken som representerar olika parter.

bilen-kolliderar-med-cykeln

- Tecken som kan modifieras enligt riktning och placering i teckenrummet (till exempel riktade tecken) skall översättas med bindestreck för att det skall synas att det handlar om ett enda tecken.

ID_1_höger: KASTA(5)
Översättning_1: *(jag) skickar-åt (honom)*

ID_1_höger: HATT(2-handstecken) _at_gh
ID_1_vänster: HATT(2-handstecken) _at_gh
Översättning: *(Så tar pojken) hatten (och) sätter-den-på (snögubben).*

- När man delar in ett långt yttrande i mindre delar kan man förutom kommatecken i översättningen använda bindestreck. Det sker på motsvarande sätt som man i svensk text använder sig av tankstreck (pratminus) runt ett parentetiskt tillägg eller en kort bisats med sidoinformation.

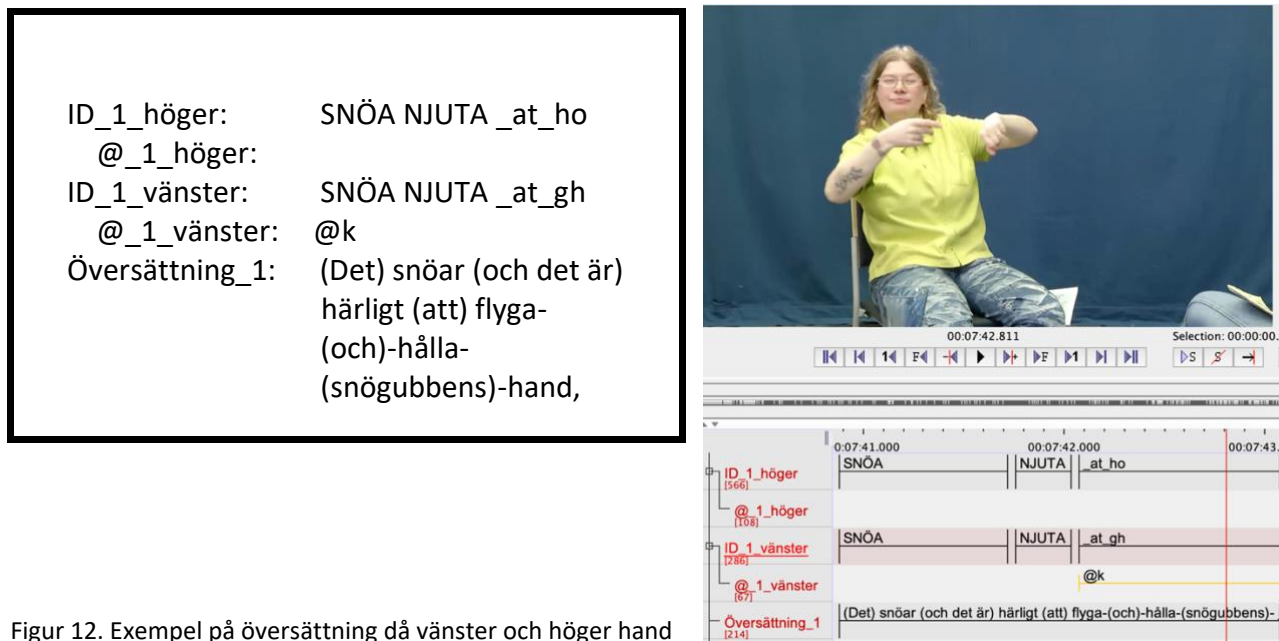
(det) tror (jag) - (på) något (sätt) bara - (jag-vet-inte).

- För att få svenskan att fungera är man ibland tvungen att kompromissa gällande kravet att följa det tecknade språkets teckenföljd²².

ID_1_höger: PEK:jag FARA-IVÄG FARA(B)
Översättning_1: *(då) stack jag iväg,*

²² I svenskan blir det omvänd ordföljd om bisatsen kommer först. Jämför ordföljden: Jag hade inget persontecken när jag var liten. <=> Förut när jag var liten, hade jag inget persontecken.

Också när händerna artikulerar olika avbildande tecken simultant måste översättningarna (innehållande bindestreck) ändå av texttekniska skäl komma efter varandra (figur 12).



Figur 12. Exempel på översättning då vänster och höger hand är annoterade med simultant olika (avbildande) tecken.

- Det är skäl att minnas att bindestrecken påverkar de sökfunktioner i ELAN som riktas mot översättningarna. Om man vill söka flerordsuttryck på svenska som motsvaras av ett tecken måste bindestrecken beaktas vid sökningen (då-och-då, på-grund-av).

5.4 Val av tempus

FSTS skiljer sig från svenskan bland annat när det kommer till hur man uttrycker tid. I svenskan syns tidsbestämningen i verbets tempusböjning, medan FSTS uttrycker tid på annat sätt, bland annat med hjälp av tidslinjer i teckenrummet (Wenman 2022). Om det av teckenspråkstexten tydligt framgår att det handlar om förfluten tid eller framtid skall detta synas i översättningen. Förutom tidslinjer kan det finnas tecken så som I-GÅR, FÖRUT(B), REDAN, I-MORGON eller så kan tempus framkomma av kontexten. Om det inte finns något som definierar tidsformen skall man i översättningen välja presens.

5.5 Om citering och användning av citattecken

I korpusen används citattecken i översättningen i följande olika fall:

- Vid direkt anföring sätts det återberättande inom citattecken.
Chefen (säger): "Kom-in",

- När man översätter referens i konstruerad handling (eng. constructed action, på svenska också konstruerat agerande, kallades tidigare "rollbyte") kan man inom parentes tillägga information om vem det är som står bakom det sagda, tänkta, känslan, attityden eller handlingen. På det här sättet blir översättningen tydligare. Citatet föregås då av ett kolon som följer på parentesen:

(Pojken säger): "Då-var-det-klart!" (men) mamma (ser det och) tar-sig-för-pannan!

- Då tecknaren använder sig av tecknet CITATTECKEN översätts det med lämpligt uttryck på svenska (till exempel *liksom*, *så-kallad*) förutom med eventuella citattecken runt rätt ord.

(då yrkesskolan) liksom så-att-säga avslutades.

(i) så-kallad "statlig nivå",

- Citattecken används som motsvarighet till olika slag av markeringar med *kursivering*, **fet stil**, 'enkla citattecken' eller "dubbla citattecken". Denna typ av markering används vanligtvis i löpande text för att markera betoningar, titlar, termer, betydelser eller namn (på ett annat språk), eller uttryck som metaforer, osv:

*Förut (var) termen "banmästare",
..eh (som) yngre (än) "(den) andra",*

5.6 Översättning av ofullständigt tecknande

Paus eller avbrott i tecknandet uttrycks i översättningen med tre punkter (...) eller tre streck (---). Om en del av tecknet redan artikulerats kan man uttrycka avbrottet med att skriva ut början av ordet (arbetspla...).

5.7 Bokstavering med handalfabet och användning av versaler i översättningen

- Bokstavering med handalfabetet annoteras på raden för ID-glossan med bindestreck (l-e-n-a_bokst) medan bokstaveringen skrivs ut utan bindestreck i översättningen, alltså *Lena*.
- Att bokstaveringen sker med gamla handalfabetet, nya handalfabetet, handalfabetet för STS eller med en blandning av dessa behöver inte markeras i översättningen. Eventuella kommentarer om användningen av olika handalfabet skrivs in på ID_observationer-raden.
- VERSALER (stora bokstäver) skall inte användas i översättningen. Detta för att sökfunktionen med versaler i ELAN är reserverad för raden för ID-glossan. Om man skall översätta att ett visst tecken diskuteras översätts tecknet med gemener (små bokstäver) och man kan vid behov sätta ett förtydligande inom hakparentes, av typen [*tecknet visas*].

5.8 Sifferuttryck i översättningen

I översättningen skrivs siffror och sifferuttryck under tolv (12) med bokstäver och högre tal med siffror.

om-fem-år elva 20 år sedan 1953

Undantag är tekniska och andra uttryck där det är logiskt att skriva också siffror under tolv med siffror. Till exempel:

klockan 10.30 programversion 2.0

5.9 Kommentar på raden **Översättning_observationer**

På raden **Översättning_observationer** kan man skriva in observationer och kommentarer som anknyter sig till översättningen. Det kan gälla att motivera val vid översättningen, det kan handla om översättningsförtydliganden på ett mer allmänt plan eller vara frågor av olika slag, exempelvis:

- Om översättaren har behov av att för översättningens del kommentera språkkontakt (@sk) mer detaljerat – till exempel munrörelser, teckenval eller val av handalfabet – kan hen skriva in en kommentar här.
- Om översättaren måste välja ett annat ord på svenska än motsvarigheten till tecknet för att det skall bli idiomatisk svenska kan man skriva in en observation.

HÖRA(G) LÄKARE => öronläkare.

- Översättaren har ett behov att motivera sin översättningslösning då det finns flera alternativ. Hen väljer det uttryck som passar bäst i kontexten och skriver in en kommentar om ordvalet på denna rad.

6 DEN LEXIKALA DATABASEN SIGNBANK

Suomen Signbank
Finlands Signbank
Finnish Signbank

fi sv en Logga in Skapa användarnamn

Startsida Om Tecken Sök glossa Sök

Sök tecken [Hjälp](#)

Visa glossor:

_	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	R	S	T	U
V	W	Y	Å	Ä	Ö															

Alla teckenspråk finlandssvenskt teckenspråk finskt teckenspråk

Sök glossa

Sök glossa

Sök översättningsmotsvarighet

Sök översättningsmotsvarighet

Du kan avgränsa sökningen till de här lexikonerna

☐ SVK_JYU-korpus ☐ SVK_Kipo-korpus ☒ FSTS_JYU-korpus [Lexikon ?](#)

Sök

1 2 3 4 5 6 7 8 9 10 ... 34 »

AKTIE FSTS_JYU-korpus

ALARM FSTS_JYU-korpus

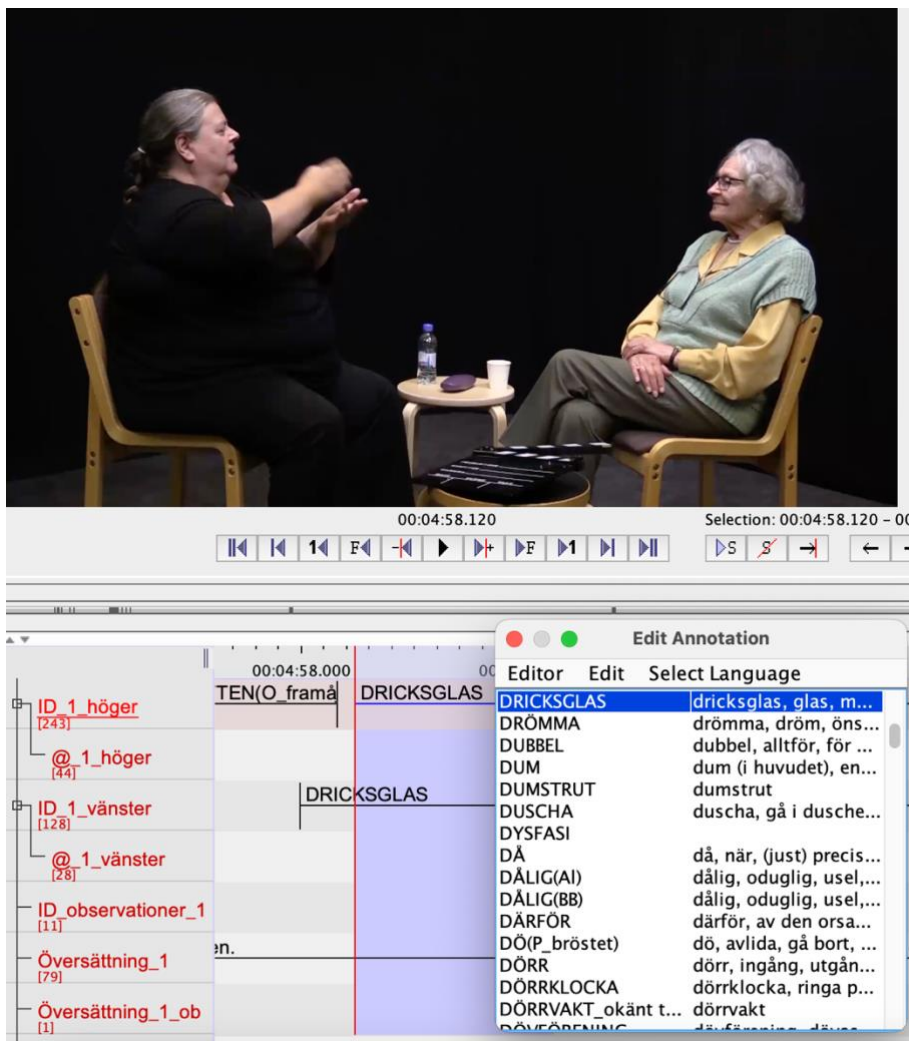
Figur 13. Databasen Signbank finns på webbadressen <https://signbank.csc.fi/>

6.1 Vad är Signbank och hur fungerar den?

Då man annoterar korpusen använder man förutom ELAN-programmet också den lexikala databasen *Finlands Signbank* (figur 13). I Signbank finns huvudsakligen lexikala tecken från det annoterade videomaterialet samlade i bokstavsordning enligt ID-glossans namn. Förutom videofilmer av tecknens grundformer finns där översättningsmotsvarigheter till svenska och annan information om tecknen.

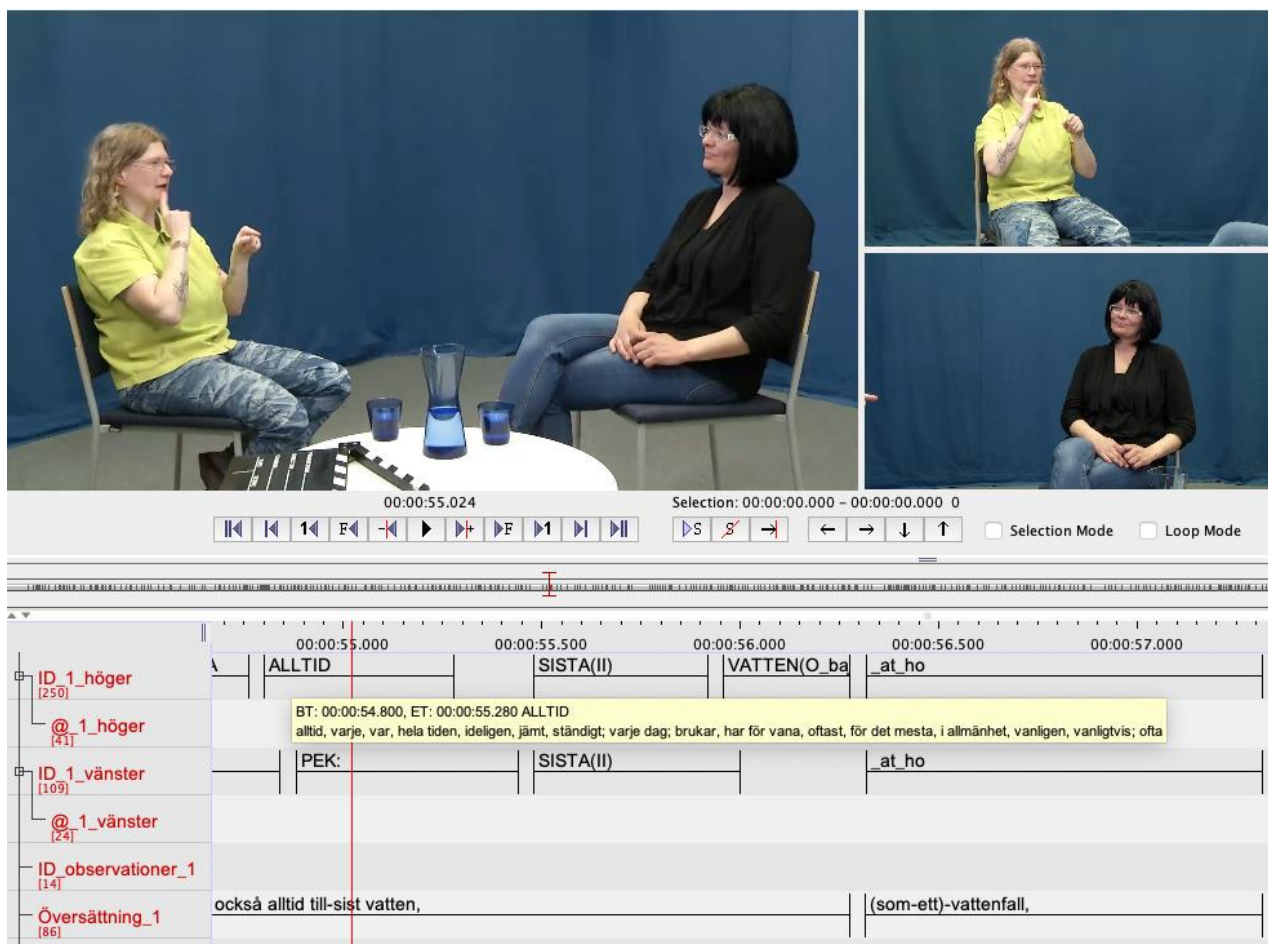
Databasen Signbank är kopplad till ELAN över nätet (ECV, *External Controlled Vocabulary*). Det innebär att en lista med ID-glossor från Signbank kan hämtas fram i ELAN under annoteringens gång. Då man annoterar kan man välja ID-glossan direkt ur Signbank-fönstret (figur 14) om glossan redan finns insatt där. Också översättningsmotsvarigheterna till svenska är synliga i fönstret, och

då man söker i listan görs sökningen både bland ID-glossor och översättningsmotsvarigheter. Det här gör annoteringsprocessen snabbare och betydligt säkrare, eftersom man undviker stavfel som kan uppkomma då man skriver in ID-glossan manuellt. Dessutom medför Signbank en stor fördel när man skall uppdatera eller ändra i annoteringarna. Alla ändringar som görs i Signbank uppdateras automatiskt också i de annoteringsceller i ELAN som är länkade med ECV-kopplingen till databasen. Ifall kontakten mellan Signbank och ELAN inte fungerar friktionsfritt – om till exempel Signbanks ID-glossor inte uppdateras i ELAN – lönar det sig att kontrollera att svenska är valt i språkinställningarna (Tier > Change Tier Attributes > Content Language).



Figur 14. ELAN och Signbank är i kontakt med varandra via internet. På bilden syns ett öppnat fönster i ELAN där kontakten till Signbank syns. Till vänster på listan finns ID-glossor och till höger översättningsmotsvarigheter.

När man för kursorn ovanpå en ID-glossa i en annoteringscell i ELAN öppnas ett fönster där glossans översättningsmotsvarigheter syns om glossan finns publicerad i Signbank (se exempel i figur 15). Det här hjälper användaren av korpusen att se tecknets olika betydelser. Den betydelse som tecknet har i just denna kontext framkommer på raden *Översättning* i ELAN.



Figur 15. När man exempelvis för kursorn i ELAN på ID-glossan ALLTID framträder de svenska översättningsmotsvarigheterna som finns insatta i Signbank.

Signbank är utformad huvudsakligen som korpusens lexikala databas, vilket betyder att den fungerar som ett verktyg för annotering av teckenspråkstexter. Signbank är inte en ordbok, trots att den kan ha likheter med en nätordbok. Den utgör i stället mer en lista på ID-glossor av (främst) grundformer av de tecken som finns dokumenterade i korpusens videoinspelningar. Det handlar alltså om språket som användes då korpusens 12 informanter parvis diskuterade kring vissa teman. Signbank är därmed inte en uttömmande ordbok innehållande alla olika tecken och deras former som finns i FSTS. Eftersom korpusen också innehåller tecken från STS och SVK (till exempel STAD_STS och LOPPUUN-ASTI_SVK) finns dessa medtagna i Signbank, trots att de inte hör till FSTS.

I Signbank är det tecknets formegenskaper som i första hand avgör om olika tecken kombinerats under samma ID-glossa (se 3.1, figur 5). Det här innebär att exempelvis tecken som betyder 'bekvämt' eller 'automatisk' båda står under ID-glossan BEKVÄM. Det finns dock en del undantag, till exempel under ID-glossan HÄNDERNA-IHOP_gest finns många till sin form olika tecken samlade. Framtida studier klargör förhoppningsvis gesters funktion och betydelser i FSTS. Eftersom endast knappt hälften av alla korpusens ID-glossor publicerades i Signbank före utgången av 2024 återstår ännu arbete²³. Orsaken till att arbetet med Signbank inte är slutfört är att det

²³ De ID-glossor som publiceras i detta första skede avspeglar annoteringsprocessen: tecken som eventuellt är mer frekventa i korpusen men namngavs med en ID-glossa först i slutet av annoteringsskedet finns inte med i den första delen av publiceringen.

saknats forskning från förr om alla tecknens grundformer och betydelser. Grundlig genomgång av varje tecken är en långsam analysprocess. Analysen kräver organiserat samarbete med finlandssvenska teckenspråkiga personer som har det metaspråkliga kunnandet om tecknens former och betydelser som behövs. An efter att man annoterar mer korpusmaterial i framtiden måste databasen också ständigt kompletteras och uppdateras.

Förutom finlandssvenskt teckenspråk finns det i Signbank också två lexikon i finskt teckenspråk: Finlands Dövas Förbund rf:s SVK_Kipo-korpus och teckenspråkscentrets SVK_JYU-korpus. Utvecklingsarbetet av Signbank har skett i samarbete mellan förbundet och teckenspråkscentret. Finlands Signbanks uppbyggnad baserar sig på på Signbank-databaser för Auslan och nederländskt teckenspråk. Källkoderna för alla tre databaser finns tillgängliga på GitHub (<https://github.com/Signbank>).

6.2 Vilken information finns för ID-glossorna i Signbank?

6.2.1 Videofilmer och namnet på videon

Videofilmer

I videorna visas olika former av samma tecken som är sammanförda under en och samma ID-glossa. Till exempel står videofilmerna PAPPA(G), PAPPA(Ax) och PAPPA(A) alla samlade under ID-glossan PAPPA. Filmer finns också för de olika lexikala tecken som är samlade under samma ID-glossa – till exempel finns filmer på tecken som betyder 'försvinna', 'borta' och 'vår' under ID-glossan BORTA. I videon presenteras främst grundformsvarianter. Det kan finnas videofilmer av både singulära och plurala former (ÖGA(2S), ÖGA(1S) och ÖGON), och också olika former av riktade tecken kan finnas medtagna (GE(1-handstecken), GE(2-handstecken) och GE-ÅT-MIG).

I korpusen förekommer tecknet ofta i modifierad form och det har inte alltid varit möjligt att avgöra vilken av varianterna som är tecknets grundform. Därför kan till exempel både enhandsvarianter och tvåhandsvarianter av ett tecken finnas dokumenterade i Signbank, trots att endast ena formen hittats i korpusen. För en del teckens del innehåller korpusen många varierande former och om inte grundformerna lätt kunnat identifieras har vi valt att spela in endast några av korpusformerna i Signbank. Genom att söka på ID-glossan i korpusen kan man se de olika varierande formerna och kontexten de används i.

Vi har strävat efter att spela in grundformerna i neutral form. Munrörelserna baserar sig på de två teckenaktörernas intuition då grundformerna spelades in, och kan variera. Det kan finnas alternativa munrörelser i korpusen som inte finns medtagna i Signbank. Tecknets betydelse kan ibland ändras om munrörelsen är en annan, till exempel BONDE får betydelsen 'Pedersöre' om det tecknas med en svenskinlånad munrörelse. Alla översättningsmotsvarigheter finns inte representerade i videofilmerna. Videofilmernas ordningsföljd korrelerar inte med variantens frekvens i korpusen.

Namn på videon

Namnet på videofilmerna kan också bidra med information. Observera att videons namn dock inte är sökbart i Signbank. Om det finns betydelseskillnad mellan olika varianter kan det här synas i videons namn. Till exempel har ID-glossan FÄNGELSE tre videofilmer: FÄNGELSE, FAST och FÅNGE. Namnen på videofilmerna beskriver ibland formen på de olika varianterna, till exempel GRÄL(2S) har de tre varianterna GRÄL(R_rak), GRÄL(R_cirkulär) och GRÄL(R_av-och-an). Ibland är videonamnen åtskilda med endast siffror, så som i BORTSKÄMD_1 och BORTSKÄMD_2 i figur 16.

BORTSKÄMD

FSTS_JYU-korpus

Översättningsmotsvarigheter

svenska:

bortskämd, skämma bort, bortskämdhet, van med bekvämligheter

Mer information

@ssg

BORTSKÄMD_1



BORTSKÄMD_2



Relationer till andra glossor

Relationer från denna glossa

Inga relationer.

Relationer till denna glossa

Inga omvända relationer.

Webbadresser

Figur 16. Skärmdump av ID-glossan BORTSKÄMD i Signbank.

6.2.2 Översättningsmotsvarigheter (svenska)

Tecknen som ID-glossan representerar kan ha flera betydelser och i olika kontexter motsvaras av olika svenska översättningsmotsvarigheter. De mest centrala betydelserna finns medtagna, och de kompletteras under annoteringens gång. Det här sker när översättaren hittat tecknet i en betydelse vars svenska motsvarighet saknas i Signbank.

Ibland förknippas en del betydelser mer med någon av videovarianterna än de andra, men det här är inte alltid klart utskrivet i Signbank. Tecknet kan förekomma i sammansättningar, och då uttrycks det här med bindestreck framför eller efter den svenska översättningsmotsvarigheten (-bank och bank-).

6.2.3 Mer information

Under "Mer information" kan det finnas mer detaljerade uppgifter om tecknet eller tecknen. Se koderna och deras betydelse i tabell 7.

Tabell 7. Koder för "Mer information"

Koden i fältet "Mer information"	Koden står för
@ssg	sammansättning
@neg	negation
@bt	tecken som innehåller bokstaverade element
@sk	språkkontakt – också språket i fråga kan markeras, t.ex. STS eller SVK
@bojl	listboj/listkonstruktion
@numg	icke-inkorporerade numeraler, dvs. grundtal
@numi	Inkorporerade numeraler
@numår	årtal
@eno	egennamn: ortnamn (städer, platser, berg, sjöar, osv.)
@enf	egennamn: företag (organisationer, skolor, instanser, osv.)
@enp	egennamn: persontecken
@gest	gestikulering och gester

Även om informationen endast gäller en av varianterna sätts koden ut. Det kan alltså stå @ssg också om bara en av teckenvarianterna är en sammansättning. Också övriga allmänna kommentarer kan skrivas in i det här fältet.

6.2.4 Relationer till andra glossor

I fältet "Relationer till andra glossor" finns länkar till andra ID-glossor för att man bland annat skall kunna:

- jämföra likheten i betydelse hos likartade former
FÅ(BB) ⇔ GE(BB)
LÄSA(B) ⇔ SKOLA(BB)
RÄTT ⇔ STÄMMA
- jämföra synonymer då betydelsen inte syns i ID-glossans namn
RÄTT ('i mitten') ⇔ MITTEN(BB)
- jämföra enkla och sammansatta former
BJÖRN ⇔ BJÖRNEBORG
- kolla upp motsvarande tecken med samma eller liknande form och betydelse i SVK_JYU-korpusen.
NY ⇔ UUSI
PAPPA ⇔ ISÄ

6.2.5 Webbadresser

I Signbank är det möjligt att länka ID-glossan till webbordböcker i teckenspråk. Det här gör det möjligt för användaren att jämföra tecken i olika webbpublikationer. Webblänkarna finns i fältet "Webbadresser". Förutom webblänken finns ett omnämnande om vilket teckenspråk ordboken representerar. Det finns främst länkar till ordböckerna *Suvi* och *Signwiki* i FSTS och ordboken *Teckenspråkslexikon* i STS. Det finns för tillfället länkar endast i ett begränsat antal ID-glossor, men arbetet kan på ett enkelt sätt fortsättas när de finns resurser för det.

KÄLLOR

- Andersson-Koski, M. (2024). "Jag balanserar liksom mitt emellan." Uppfattningar om språkliga och kulturella gränser bland fyra finlandssvenska teckenspråkiga. I D. Jansson, I. Melander, G. Westberg & D. Y. Falk (Red.), *Svenskans beskrivning 38: Förhandlingar vid trettioåttonde sammankomsten. Örebro 4–6 maj 2022, del II*. Örebro universitet.
<http://www.diva-portal.org/smash/get/diva2:1854426/FULLTEXT01.pdf> [1.12.2024]
- ELAN [dataprogram]. (2024). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/tla/elan> [1.12.2024]
- Finlands Dövas Förbund rf & Finlandssvenska teckenspråkiga rf. (2024). Revitaliseringsprogram för finlandssvenskt teckenspråk. Opublicerat utkast.
- Finlandssvenska teckenspråkiga rf & Finlands Dövas Förbund rf. (2017). *Språkplan för finlandssvenskt teckenspråk*. Version 7.0.
https://www.dova.fi/wp-content/uploads/2020/05/Spra%CC%8Akplan_7.0_2017.pdf [1.12.2024]
- Hanke, T. & Fenlon, J. (2022). Creating Corpora: Data Collection. I J. Fenlon & J. Hochgesang (Red.), *Signed Language Corpora* (s. 18–45). Gallaudet University Press.
- Heimonen, O. (2024). *Suomenruotsalaisen viittomakielen viittomatyyppijakauma*. Magisteruppsats. Suomalainen viittomakieli, Institutionen för språk- och kommunikationsvetenskap, Jyväskylä universitet. <http://urn.fi/URN:NBN:fi:juu-202406244958>
- Hochgesang, J. & Palfreyman N. (2022). Signed Language Corpora and the Ethics of Working With Sign Language Communities. I J. Fenlon & J. Hochgesang (Red.), *Signed Language Corpora* (s. 158–195). Gallaudet University Press.
- Hodge, G. & Crasborn, O. (2022) Good Practice in Annotation. I J. Fenlon & J. Hochgesang (Red.), *Signed Language Corpora* (s. 46–89). Gallaudet University Press.
- Hoyer, K. (2000). *Variation i teckenspråk – en studie i släktskapsterminologi i teckenspråk hos finlandssvenska döva*. Avhandling pro gradu. Allmän lingvistik. Helsingfors universitet. [Publikationsserie L 4/2000, Finlands Dövas Förbund r.f.]
- Hoyer, K. (2004). The sociolinguistic situation of Finland-Swedish Deaf people and their language, Finland-Swedish Sign Language. I M. Van Herrevege & M. Vermeerbergen (Red.), *To the lexicon and beyond: Sociolinguistics in European Deaf communities*. (s. 3–23). Gallaudet University Press.
- Hoyer, K. (2005). "Vi kallade dem Borgåtecken". Det finlandssvenska teckenspråket i går och i dag. I J.-O. Östman (Red.), *FinSSL – Finlandssvenskt teckenspråk*, (s. 21–80). [Teckenspråksstudier 1] Nordica, Institutionen för nordiska språk och nordisk litteratur, Helsingfors universitet.

- Hoyer, K. (2012). Dokumentation och beskrivning som språkplanering – perspektiv från arbete med tre tecknade minoritetsspråk. [Nordica Helsingiensia 29] Doktorsavhandling. Nordica. Finska, finskugriska och nordiska institutionen. Helsingfors universitet. <http://urn.fi/URN:ISBN:978-952-10-7612-1>
- Hoyer, K. (2013). Suomenruotsalainen viittomakieli – attritiosta ahdinkoon. I: K. Granqvist & P. Rainò (Red.), *Rapautuva kieli. Kirjoituksia vähemmistökielten kulumisesta ja kadosta*. (s. 232–251). Suomalaisen Kirjallisuuden Seura.
- Hoyer, K. & Kronlund-Saarikoski, K. (2002). *Se vårt språk! Näe kielemme!* Finlandssvenskt teckenspråk, 38 ordboksartiklar. Suomenruotsalainen viittomakieli, 38 sanakirja-artikkelia. Helsingfors: Finlands Dövas Förbund r.f. och Forskningscentralen för de inhemska språken.
- Huhtinen, H. (2020). *Mikä ihmeen korpusannotointi? Taustatietoa suomenruotsalaisen viittomakielen annotoinnin tueksi*. Humak, opinnäytetyö Tulkki AMK. <https://urn.fi/URN:NBN:fi:amk-2020112724713>
- Institutet för språk och folkminnen (2024). *Bokstaverade tecken*. <https://www.isof.se/svenskt-teckensprak/laromedel-och-utbildningsmaterial/det-svenska-teckensprakets-grammatik/teckenbildning/lexikala-tecken/bildningssatt-for-lexikala-tecken/inlan-fran-andra-sprak/bokstaverade-tecken>
[1.12.2024]
- Jantunen, T. (2013). Signs and transitions: Do they differ phonetically and does it matter? *Sign Language Studies* 13:2, (s. 211–237). <https://www.jstor.org/stable/26190844>
- Jantunen, T. (2015). How long is the sign? *Linguistics* 53:1, (s. 93–124). <https://doi.org/10.1515/ling-2014-0032>
- Johnston, T. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. I *International Journal of Corpus Linguistics* 15:1. (s. 106–131). John Benjamins Publishing Company. <https://doi.org/10.1075/ijcl.15.1.05joh>
- Johnston, T. (2019). *Auslan Corpus Annotation Guidelines*. Centre for Language Sciences, Department of Linguistics, Macquarie University, Sydney, Australia. https://www.academia.edu/40088269/Auslan_Corpus_Annotation_Guidelines_August_2019_version [1.12.2024]
- Johnston, T. & Schembri, A. (2013). Corpus Analysis of Sign Languages. I C. Chapelle (Red.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal0252>
- Keränen, J., Syrjälä, H., Salonen, J., & Takkinen, R. (2016). The Usability of the Annotation. I E. Efthimiou, S.-E. Fotinea, T. Hanke, J. Hochgesang, J. Kristoffersen, & J. Mesch (Red.), *Workshop Proceedings : 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining / Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (s. 111–116). European Language Resources Association (ELRA). <http://www.lrec->

conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-SignLanguage_Proceedings.pdf
[1.12.2024]

Keränen, J., Nieminen, H., Salonen, J., Green, O., Pippuri, O., Wainio, T. & Takkinen, R. (2017). *CFINSL:n ELAN-ohjeet*. Versio 0.1 https://www.jyu.fi/sites/default/files/2024-02/cfinsl_elanohjeet_2018_1versio.pdf [1.12.2024]

Lyxell, T. (2009). *Korpusarbete i teckenspråk. Förstudie om hur korpusarbete av teckenspråk ska bedrivas på Språkrådet*. Institutet för språk och folkminnen, Språkrådet.
<http://isof.diva-portal.org/smash/get/diva2:1560039/FULLTEXT01.pdf> [1.12.2024]

Mesch, J. (2023). Creating a multifaceted corpus of Swedish Sign Language. Visual, tactile, and L2 signing. I E. Wehrmeyer. (Red.), *Advances in Sign Language Corpus Linguistics*. (s. 242–261). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.108>

Mesch, J. & Wallin, L. (2021). *Annoteringskonventioner för teckenspråkstexter. Version 8*. Institutionen för lingvistik, Teckenspråk. Stockholms universitet. <https://su.diva-portal.org/smash/get/diva2:1556259/FULLTEXT01.pdf> [1.12.2024]

Mesch, J. & Eir Cortes, E. (2021). *Bruksanvisning för annotering av teckenspråkstexter i ELAN. Version 4*. Institutionen för lingvistik, Avdelningen för teckenspråk. Stockholms universitet. <https://www.diva-portal.org/smash/get/diva2:1621750/FULLTEXT01.pdf> [1.12.2024]

Mesch, J., Eir Cortes E., Björkstrand T., Kankkonen, N. R., Bäckström, J. & Hansson, P. (2023). Teckenspråkslexikografi – utmaningar i en annan modalitet. I L. Holmer, G. Horn, H. Landqvist, P. Nilsson, E. Nordgren, E. Sköldberg (Red.), *Nordiska studier i lexikografi 16. Rapport från 16:e konferensen om lexikografi i Norden. Skrifter utgivna av Nordiska föreningen för lexikografi, Skrift nr 17*. (s. 225–240). <https://hdl.handle.net/2077/79763>

Nordlund, S. & Savolainen L. (2017). *Kipo-korpuksen annotointikäytänteet*. Opublicerad.

Pippuri, O. (2015). *Annotointi- ja käännöstyö ProGram-projektissa*. Avhandling för utbildningsprogrammet för teckenspråkstolk (240 sp). Humak. <https://urn.fi/URN:NBN:fi:amk-201505127633>

Puupponen, A. (2012). *Horisontaaliset ja vertikaaliset päänliikkeet suomalaisessa viittomakielessä*. Avhandling Pro gradu. Kielten laitos, suomalainen viittomakieli. Jyväskylän universitet. <http://urn.fi/URN:NBN:fi:ju-201207242120>

Puupponen, A. (2019). *Understanding nonmanuality – A study on the actions of the head and body in Finnish Sign Language*. Doktorsavhandling. Jyväskylän universitet. <http://urn.fi/URN:ISBN:978-951-39-7761-0>

Puupponen, A., Jantunen, T., & Mesch, J. (2016). The alignment of head nods with syntactic units in Finnish Sign Language and Swedish Sign Language. I J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Red.), *Speech Prosody 2016 : Proceedings of the 8th International Conference on*

Speech Prosody, Boston University, USA, 31 May - 3 June 2016 (s. 168–172). International Speech Communication Association. Speech prosody. <https://doi.org/10.21437/SpeechProsody.2016-35>

Safar, J. & Webster, J. (2014). *Cataloguing endangered sign languages at iSLanDS* https://islandscentre.files.wordpress.com/2014/08/report-endangered-sls_070814.pdf [1.12.2024]

Salonen, J., Wainio, T., Kronqvist, A., & Keränen, J. (2019a). Suomen viittomakielten korpusprojektin (CFINSL) annotointiohjeet. Kieli- ja viestintätieteiden laitos, viittomakielen keskus. Jyväskylän yliopisto. <https://www.jyu.fi/fi/file/suomen-viittomakielten-korpusprojektin-cfin-sl-annotointiohjeet> [1.12.2024]

Salonen, J., Puupponen, A., Takkinen, R., & Jantunen, T. (2019b). Suomen viittomakielten korpusta rakentamassa. I J. H. Jantunen, S. Brunni, N. Kunnas, S. Palviainen, & K. Västi (Red.), *Proceedings of the Research data and humanities (RDHum) 2019 conference : data, methods and tools* (s. 83–98). Uleåborgs universitet. Studia Humaniora Ouluensia, 17. <http://urn.fi/urn:isbn:9789526223216>

Salonen, J., Andersson-Koski, M., Hoyer, K., & Jantunen, T. (2022). Building the Corpus of Finland-Swedish Sign Language : Acknowledging the Language History and Future Revitalization. I J. H. Jantunen, J. Kalja-Voima, M. Laukkarinen, A. Puupponen, M. Salonen, T. Saesma, J. Tarvainen, & S. Ylönen (Red.), *Diversity of Methods and Materials in Digital Human Sciences: Proceedings of the Digital Research Data and Human Sciences DRDHum Conference 2022, December 1-3, Jyväskylä, Finland* (s. 187–199). Jyväskylä universitet. <http://urn.fi/URN:ISBN:978-951-39-9450-1>

Signwiki – wikiordbok i finlandssvenskt teckenspråk. Finlands Dövas Förbund. <https://finssl.signwiki.org/> [1.12.2024]

Siltaloppi, S. (2016). *Den andra handen. Den icke-dominanta handens funktioner i finlandssvenskt teckenspråk*. Avhandling pro gradu. Finska, finskugriska och nordiska institutionen. Helsingfors universitet. https://helda.helsinki.fi/bitstream/handle/10138/161842/Siltaloppi_Satu_Progradu_2016.pdf

Siltaloppi, S. (2019). Finlandssvenskt teckenspråk. Listkonstruktion, koherens och konstruktionsgrammatik. I M. Bianchi, D. Håkansson, B. Melander, L. Pfister, M. Westman & C. Östman (Red.), *Svenskans beskrivning 36. Förhandlingar vid trettiosjätte sammankomsten. Uppsala 25–27 oktober 2017*. Uppsala: Institutionen för nordiska språk vid Uppsala universitet. (s. 281–293). <http://uu.diva-portal.org/smash/get/diva2:1313414/FULLTEXT01.pdf> [1.12.2024]

Siltaloppi, S. (2023). *List construction in Finland-Swedish Sign Language*. Doktorsavhandling. Humanistiska fakulteten, Helsingfors universitet. <http://urn.fi/URN:ISBN:978-951-51-8955-4>

Språkbanken. FIN-CLARIN. <https://www.kielipankki.fi/sprakbanken/> [1.12.2024]

STS-korpus [webbaserat korpusverktyg för teckenspråksmaterial] (2023). <https://teckenspraskorpus.su.se/> [1.12.2024]

Suvi – nätordbok i de finländska teckenspråken. Finlands Dövas Förbund.

<https://suvi.teckensprak.net/> [1.12.2024]

Södergård, L. & Leinonen, T. (2017). Talko – korpus över den talade svenskan i Finland: Korpusbygge i teori och praktik. I J.-O. Östman, C. Sandström, P. Gustavsson & L. Södergård (Red.), *Ideologi, identitet, intervention. Nordisk dialektologi 10* (s. 331–340). Nordica vid Finska, finskugriska och nordiska institutionen, Helsingfors universitet.

<https://www.doria.fi/bitstream/handle/10024/134741/sodergard-leinonen2017.pdf> [1.12.2024]

Wehrmeyer, E. (2023). Sign language corpus linguistics. Taking stock of an emerging discipline. I E. Wehrmeyer (Red.), *Advances in Sign Language Corpus Linguistics*. (s. 1–27). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.108>

Wenman, L. (2022). *Tidslinje och tecknets rörelse i finlandssvenskt teckenspråk. En korpusbaserad studie*. Examensarbete, Teckenspråkslingvistik. Institutionen för lingvistik, Stockholms universitet.

<https://www.diva-portal.org/smash/get/diva2:1670702/FULLTEXT01.pdf>