

DEN ORTOGRAFISKA UTSKRIFTEN I TALKO

Manual för transkriberare

Inledning	4
1 Allmänna principer för den ortografiska utskriften	4
1.1 Ordgränser.....	5
1.2 Namn.....	6
1.3 Tal.....	7
1.4 Kortformer av ord.....	7
1.5 Skiljetecken: punkt, kolon, et-tecken.....	7
2 Innehållsord	8
2.1 Verb	8
2.1.1 <i>Tempus</i>	8
<i>Dubbelsupinum</i>	8
<i>Pseudosamordning</i>	9
2.1.2 <i>Hjälppverb</i>	9
<i>Måste</i>	9
<i>Böra</i>	10
<i>Varda</i>	10
<i>Skola</i>	10
<i>Löns</i>	10
2.1.3 <i>Sättas och Läggas</i>	10
<i>Sättas</i>	10
2.1.4 <i>Avledning</i> ar på <i>-as</i>	10
2.2 Substantiv	11
2.2.1 <i>Species</i>	11
<i>Formsammanfall i nÖB och sÖB</i>	11
<i>Artikelsynkopering</i>	12
2.2.2 <i>Genus</i>	12
<i>Tregenussystemet</i>	12
<i>Avsaknad av genus</i>	12
<i>Annat genus än standardsvenskan</i>	12

2.2.3 Numerus.....	13
2.3 Adjektiv.....	13
2.3.1 Adjektiv på -ig/-og/-ot.....	13
2.3.2 Genuskongruens.....	13
2.3.3 Komparering.....	13
3 Formord.....	14
3.1 Pronomen.....	14
3.1.1 Personliga pronomen.....	14
Generiskt pronomen.....	14
3.1.2 Demonstrativa pronomen.....	14
Sådan-här och sådan-där.....	15
3.1.3 Interrogativa pronomen / vilken.....	15
3.3 Adverb.....	15
3.3.1 Här.....	15
3.3.2 Där.....	16
3.3.3 Hit.....	16
3.3.4 Dit.....	16
3.3.5 Härifrån.....	16
3.3.7 Därtill.....	16
3.3.8 Inte.....	16
3.3.9 Också.....	16
3.3.10 Någorst och annars.....	17
3.3.11 Ord som slutar på -ledes.....	17
3.4 Prepositioner.....	17
3.5 Ord som slutar på -efter, -över.....	17
3.6 Subjunktioner och konjunktioner.....	17
3.6.1 Före och förrän.....	17
3.6.2 Eller.....	18
3.7 Infinitivmärkena att och till.....	18
3.8 Diskurspartiklar.....	18
3.9 Kraftuttryck.....	18
4 Ord som inte finns i SAOL.....	19
Ortografisk form för finska ord.....	19
Gränsdragningar.....	19

Talko – korpus över den talade svenskan i Finland.

4.1 Ordbokslänkning.....	20
Exemplet bulle.....	20
Böjningsformer med ordbokslänk.....	21
4.2 X-tagget.....	21
4.2.1 Slangord och inofficiella namnformer.....	21
4.2.2 Ord från andra språk.....	21
4.3 G-tagget.....	21
5 Translitteratorbrus	22
LITTERATURFÖRTECKNING	22

INLEDNING

Talspråkskorpusen Talko består av sökbara utskrifter som är lingvistiskt annoterade och kopplade till ljudfiler. Utskrifterna görs i två parallella versioner: en ljudenlig utskrift och en ortografisk utskrift.

Ljudskriften transkriberas helt manuellt i programmet Transcriber och används sedan som grund för den ortografiska utskriften. Det första utkastet till den ortografiska utskriften skapas av en *Dialekttranslitterator* som utvecklats av Tekstlaboratoriet vid universitetet i Oslo. Dialekttranslitteratorn gör en råversion som sedan kontrolleras och korrigeras manuellt.

Den ortografiska utskriften är väldigt central för korpusens sökbarhet eftersom den möjliggör sökningar på normaliserade former. Dessutom annoteras den ortografiska utskriften med hjälp av en automattaggare (TreeTagger) som tilldelar varje ord information om ordklass, genus, species, lemma etc. Det innebär att den ortografiska utskriften även är av stor betydelse för hur väl taggningen lyckas. En felaktig form i den ortografiska utskriften kommer sannolikt att leda till att ordklasstagningen också blir fel.

1 ALLMÄNNA PRINCIPER FÖR DEN ORTOGRAFISKA UTSKRIFTEN

Precis som den ljudenliga transkriptionen ska den ortografiska utskriften så noga som möjligt återge det som sägs på inspelningen. Det innebär att alla stakningar, upprepningar, pauser osv. som markerats i ljudskriften också ska finnas med i den ortografiska utskriften.

ja fikk en ny: (.) ny: sykkel ti ju- ju:lklapp *jag fick en ny (.) ny cykel till ju- julklapp*

Den ortografiska utskriften innehåller inga skiljetecken och har inte heller stor bokstav i början av repliker.

Den ortografiska utskriften är en normaliserad version av ljudskriften, en ord-för-ord överföring där den grammatiska informationen ur talspråket överförs till standardsvenskans böjningsformer. Som norm för att konstruera de normaliserade formerna används i första hand *Svenska Akademiens ordlista (SAOL)* men för ord som inte ingår där används *Ordbok över Finlands svenska folkmål (FO)* och *Finlandssvensk ordbok (FSOB)* som norm. Dialektord, slang, finlandismer och utländska ord förses även med en egen tagg, x. Dessutom används taggen g för s.k. substantiverade passiver. Se avsnitt 4.

På samma sätt som den ljudenliga utskriften innebär den ortografiska utskriften ofta en tolkning av det som sägs. Det är vanligen kontexten som avgör hur ordet translittereras. I exemplen nedan uttalar informanten *ta:la* på samma sätt men kontexten visar att den första repliken avser preteritum *talade* och den andra avser infinitiv *tala*:

sen fick vi en massa bekantta såm **ta:la** de dä språ:ke sen fick vi en massa bekanta som talade
det där språket

åm ja sku ha: sånnä:n talang att li:ksåm bara: (.) om jag skulle ha sådan-där talang att
ta:la å:t mej ett stelle å så:va liksom bara (.) tala åt mig ett ställe att
sova

1.1 ORDGRÄNSER

Antalet ord måste vara lika i ljudskriften och den ortografiska utskriften, så att de alltid löper helt parallellt. I exemplet nedan står den assimilerade formen *messej* 'med sig' hopskriven i ljudskriften men som separata ord i den ortografiska utskriften. Det skapas en förskjutning mellan utskrifterna som gör att den ortografiska formen *sig* kombineras med uttalet *hem* i ljudskriften. En sökning i Talko på sig skulle alltså visa ett uttal "hem"

Ortografisk utskrift:	<i>han</i>	<i>tog</i>	<i>den</i>	<i>med</i>	<i>sig</i>	<i>hem</i>
Ljudskrift:	<i>han</i>	<i>to:</i>	<i>den</i>	<i>messej</i>	<i>hem</i>	

Lösningen är att separera orden även i ljudskriften *me ssej*.

Ortografisk utskrift:	<i>han</i>	<i>tog</i>	<i>den</i>	<i>med</i>	<i>sig</i>	<i>hem</i>
Ljudskrift:	<i>han</i>	<i>to:</i>	<i>den</i>	<i>me</i>	<i>ssej</i>	<i>hem</i>

Med tanke på ordklasstagningen är det viktigt att ord som hör till två olika ordklasser inte skrivs ihop, varken i ljudskriften eller i den ortografiska utskriften, även om de uttalas assimilerat.

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>
na lete	något lite
me sse	med sig
he e	det är
me: n	med honom
na meir	något mera

Undantagsvis kan ord som är hopskrivna i ljudskriften och som kan tolkas som *ett* ord med *en* ordklass vara skrivna ihop också i den ortografiska utskriften:

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>	<u>Ordklass</u>
sånhä:n	sådan-här	adjektiv
sånhä:nt	sådant-här	
sånhä:na	sådana-här	
såndä:na	sådana-där	
vetja, vetdu	vet-jag, vet-du	diskurspartiklar
sidu, hödu	serdu, hördu	

1.2 NAMN

Namn skrivs vanligen enligt standardsvenska skrivregler, dvs. med inledande stor bokstav. Också namn som anonymiserats, t.ex. **Förnamn Efternamn**, skrivs med stor bokstav.

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>	<u>Typ av namn</u>
va:sa	Vasa	städer
svärje	Sverige	länder
huvustassbla:de	Hufvudstadsbladet	tidningar/tidskrifter
ra:diå ve:ga	Radio Vega	radio-/tv-kanaler
yle	Yle	radio-/tv-kanaler
svenn svensån	Förnamn Efternamn	anonymiserade namn
ike:a	Ikea	företagsnamn
fj:ra hå:	4H	organisation
diskavery tjänål	Discovery Channel	namn på andra språk
kåkkåla	Kokkola	namn på andra språk
krombi	Kronoby	dialektala namnformer
nyka:bi	Nykarleby	dialektala namnformer
årvas	Oravais	dialektala namnformer

Notera att kortformer som *Björkö* används i ortografiska utskriften ifall informanten använder den istället för *Björköby*.

Notera även att namn förekommer i bestämd form:

<i>bärgöjen</i>	<i>Bergön</i>
<i>lafåssby:ji</i>	<i>Lappforsbyn</i>

1.3 TAL

Årtal och sifferuttryck skrivs med bokstäver:

nittånsekstifem *nittonsextiofem*
sju:gi trettita:le *tjugo- trettioalet*
cåmmåå:r sekstify:ra *Commodore sextiofyra*

Uttalen *a:rtån* och *adårtån* normaliseras båda till *aderton*.

I vissa namn och ord ingår siffror och då kan siffror användas även i den ortografiska utskriften:

emmpietri:spelar *mp3-spelare*
emmte:ve:tre: *MTV3*
entjugofem *N25*

1.4 KORTFORMER AV ORD

De längre formerna av följande ord används fastän kortformer förekommer i ljudskriften:

Grov ljudskrift	Ortografisk utskrift
sta:n	staden
sen	sedan
da:n	dagen
nån	någon
nånting	någonting
mommi, mommo	mormor
famo	farmor

Notera att ifall informanten säger *ska* skrivs *ska* även i den ortografiska utskriften.

Notera även att följande kortformer används i ortografiska utskriften (utan x-tagga).

sent *cent* 'centimeter'
kvadrat *kvadrat* 'kvadratcentimeter'
milli *milli* 'millimeter'

1.5 SKILJETECKEN: PUNKT, KOLON, ET-TECKEN

Inga skiljetecken sätts ut mellan satser men punkter och kolon förekommer t.ex. i förkortningar:

fi:lmag *fil.mag.*
e:tre:an *E3:an*
emmi:s kursär *MI:s kurser*

Notera att om informanten säger punkt (t.ex. i en webbadress) skrivs det ut med bokstäver:

a:rla punkt se

Arla punkt se

Et-tecken förekommer t.ex. i namn:

såssåkåmm Soc&Kom

Observera att &-tecknet kan skrivas vanligt i Transcriber men vid redigering i Notepad++ måste det stå &

2 INNEHÅLLSORD

Innehållsorden, dvs. verb, substantiv och adjektiv, behandlas enligt olika principer, delvis på grund av det taggset som används vid ordklasstagningen. Taggaren anger t.ex. genus och species för substantiv vilket innebär att den normaliserade formen av ett substantiv måste ha samma genus och species som i ljudskriften för att taggningen ska bli rätt. Taggaren anger däremot inte vilken konjugation ett verb tillhör vilket innebär att taggningen blir rätt trots att verben normaliseras enligt standardsvenskt böjningsmönster.

2.1 VERB

Dialektala och talspråkliga verbformer normaliseras enligt standardsvenskt mönster.

2.1.1 TEMPUS

För första konjugationens verb kan uttalet i talspråket vara samma i alla tempusformer. I den ortografiska utskriften blir kontexten ofta avgörande för vilken tempusform som väljs.

Verb:	kasta	Ortografisk utskrift
Infinitiv:	<i>att kasta bollen i mål e svå:rt</i>	<i>kasta</i>
Presens:	<i>ja kasta(r) bollen</i>	<i>kastar</i>
Preteritum:	<i>ja kasta bollen ren</i>	<i>kastade</i>
Supinum:	<i>ja ha kasta fö länge sen ren</i>	<i>kastat</i>

Observera alltså att den ortografiska utskriften i fråga om **tempus** ska följa det som informanten säger. Den exakta **formen** som ifrågavarande tempus har i dialekten återges däremot inte utan följer standardsvenskans böjningsformer.

*vi ha no itt **havi** an så mangg ganggor* *vi har nog inte **haft** den så många gånger*
*vi **å:ka** metrå* *vi **åkte** metro*
*man **skrann** föståss på vintärn* *man **skrinnade** förstås på vintern*

DUBBELSUPINUM

När informanten använder en annan tempusform än vad som används i standardsvenskan ska den ortografiska utskriften återge den tempusform som informanten använder. Detta gäller bl.a. i fråga om dubbelsupiner:

*så ha man **kuna fari** på nån re:sa* *så har man **kunnat farit** på någon resa*

Detta är viktigt för att man vid sökning på uttal av infinitivformen *fara* i korpusen inte ska få träffar av typen *fari* som inte är en infinitivform. Kärnområdet för dubbelsupinum är södra Österbotten, men dubbelsupinum förekommer också bl.a. i Åboland. I fråga om första konjugationens verb är det inte alltid möjligt att veta om det handlar om en infinitiv eller en supinumform. I södra Österbotten är infinitiver ändå i allmänhet apokoperade, vilket gör att tolkningen som supinum är rimligast vid a-ändelse, i synnerhet om informanten i övrigt använder dubbelsupinum. (Läs om dubbelsupinum hos Ivars 2010:147ff och Leinonen 2015:104-105.)

PSEUDOSAMORDNING

Ett annat fall där talspråket kan ha från standardsvenskan avvikande tempus är s.k. pseudosamordningar. Den standardsvenska konstruktionen "han var ute och rodde" kan i finlandssvenskt talspråk motsvaras av "han va u:te å ro:", vilket normaliseras till "han var ute och ro". Det finns ändå fall där det inte är möjligt att avgöra om det senare verbet i en pseudosamordning är en infinitiv eller en finit form (se Leinonen 2015: 114-115). När det inte går att avgöra vilken tempusform det är fråga om görs normaliseringen enligt standardsvenska.

*sen hadd vi ju alti nångång at man for in ti å:bo å **simma***

*sedan hade vi ju alltid någongång att man for in till Åbo och **simmade***

Var även uppmärksam på translittereringen av "å". Endast infinitivfraser får "att", övriga fall får "och":

*så ha vi in ens brytt åss åm **å taiji beta:lt** så har vi inte ens brytt oss om **och tagit betalt***

2.1.2 HJÄLPVERB

MÅSTE

För hjälp verbet *måste* används de standardsvenska tempusformerna kompletterat med den finlandssvenska infinitivformen *måsta*.

infinitiv: *måsta*
presens: *måste*
preteritum: *måste*
supinum: *måst*

Grov ljudskrift

Ortografisk utskrift

ja måsta springa ti bussen igå:r
ja sku måsta
ja sku ha måsta

jag måste springa till bussen igår
jag skulle måsta
jag skulle ha måst

Notera att böjningsformer inte ska x-taggas, utan följer den standardsvenska böjningen (som i SAOL).

Talko – korpus över den talade svenskan i Finland.

BÖRA

Var uppmärksam på böjningen av *böra*, enligt SAOL:

infinitiv: *böra*
presens: *bör*
preteritum: *borde*
supinum: *bort*

ja sku bo:rda *jag skulle böra*
ja sku ha bo:rda *jag skulle ha bort*

VARDA

Varda böjs enligt SAOL med tillägg av supinumformen *vordit* från SAOB.

infinitiv: *varda*
presens: *varder*
preteritum: *vart*
supinum: *vordit*

SKOLA

Observera att i sÖB används presensformen *ska* också i betydelsen 'skulle'. Jfr Ivars 1988:133. I den ortografiska utskriften överförs uttalet *ska* för preteritum till *skulle*

he va oåre furi vi ska buri ga i skreftskåolan *det var året före vi skulle börja gå i skriftskolan*

LÖNS

Observera att uttalet *lö:ns* i presens passiv av *löna* får ortografiska formen *lönas*.

2.1.3 SÄTTAS OCH LÄGGAS

För deponens verbet *sättas* 'sätta sig' används följande konstrueras följande böjningsformer:

SÄTTAS

infinitiv: <i>sess</i>	<i>sättas</i> (+ lang=x)	<i>sätta sig</i>
presens: <i>sess(är)</i>	<i>sättas</i> (+ lang=x)	<i>sätter sig</i>
preteritum: <i>sesst</i>	<i>sattes</i> (+ lang=x)	<i>satte sig</i>
supinum: <i>sesstis</i>	<i>satts</i> (+ lang=x)	<i>satt sig</i>

Formerna konstruerade i samråd med FO/Caroline Sandström. Se även FO **läggas** och Huldén Verbböjningen II s. 239 (*sättas*); 249 (*läggas*).

2.1.4 AVLEDNINGAR PÅ -AS

Så kallade substantiverade passiva infinitiver normaliseras till former på *-as* i obestämd form och *-aset* i bestämd form. Förekommer i synnerhet i Österbotten men även i Åboland och Nyland. Jämför Ivars 1988:121, Huldén 1959: 255 f.

Talko – korpus över den talade svenskan i Finland.

<i>tjö:ras</i>	köras (+lang=g)
<i>te:veskådas</i>	tv-skådas (+lang=g)
<i>sji:dase</i>	skidaset (+lang=g)
<i>jaktase</i>	jaktaset (+lang=g) (+lang=x)

Kontrollera att bestämd/obestämd följer det som sägs!

Obs! Ett ord kan få både x- och g-taggar.

2.2 SUBSTANTIV

Huvudprincipen när det gäller standardiseringen av substantiv är att de behåller samma species, genus och numerus som i ljudskriften.

2.2.1 SPECIES

I dialekterna förekommer bestämd form i större utsträckning än i standardsvenska, t.ex. *koka gröten*. Använd samma species som informanten.

Grov ljudskrift	Ortografisk skrift
<i>vi ha: jo gambe <u>namne</u> jär å fo:rnfynde</i>	<i>vi har ju gamla <u>namnen</u> här och <u>fornfynden</u></i>
<i>vi fo:r no: mest ti lafåssby:ji</i>	<i>vi for nog mest till Lappforsbyn</i>
<i><u>tsöti</u> på <u>beena</u></i>	<i><u>köttet</u> på <u>benen</u></i>

FORMSAMMANFALL I NÖB OCH SÖB

Notera formsammanfall i bl.a. norra och södra Österbotten mellan bestämd och obestämd form plural. Alla förekomster där standardsvenskan har bestämd form blir bestämd form men eftersom dialekterna har utvidgad användning av den bestämda formen kan det vara skäl att använda bestämd form också i övriga fall. Ibland kan man med hjälp av kontexten sluta sig till om det är bestämd eller obestämd form, t.ex. ifall övriga substantiv i satsen har en entydig form. I exemplet nedan är *åsiktne* entydigt en bestämd form vilket gör att *fonde:ringga* också kan normaliseras till bestämd form

*no: hadd man jo myki å:siktne å **fonde:ringga** nog hade man ju mycket åsikterna och **funderingarna***

Även liknande exempel i samma inspelning kan visa det grammatiska system som talaren använder. I exemplet *haldi:ns bossa* kan *bossa* stå för antingen *bussar* eller *bussarna* men p.g.a. att bestämd form vanligen används efter s-genitiv i Österbotten (Ivars 2015:161) och motsvarande exempel med entydigt species finns i inspelningen (*visas **halln** Visas **hallen***) kan bestämd form väljas (*Haldins **bussarna***)

Ta stöd av Ivars beskrivning av dialekternas formsystem: norra Österbotten (Ivars 2015:96–99) och i södra Österbotten (Ivars 2015:140–142). Thereses Leinonen har skrivit om species i Jakobstad 17.10.2017 för Talko, separat.

ARTIKELSYNKOPERING

Var uppmärksam på att vissa ord kan ha samma uttalsform både i obestämd och bestämd form p.g.a. uttalslättnad (artikelsynkope). I dessa fall normaliseras orden till den standardsvenska formen.

Grov ljudskrift

de e tungt ti sti:ga åpp på mårån
ja pra:ta me hånåm i telefo:n
de såm tjenns fe:l i mun
två: gångär i må:n

Ortografisk utskrift

det är tungt till stiga upp på morgonen
jag pratade med honom i telefonen
det som känns fel i munnen
två gånger i månaden

2.2.2 GENUS

För substantiven finns dels fler genus i dialekterna, dels finns ord med annat genus än i standardsvenskan.

TREGENUSSYSTEMET

Tregenussystemet finns bevarat i många finlandssvenska dialekter. I den ortografiska utskriften har normaliseringen gjorts enlighet med svenskans tvågenussystem eftersom den taggupsättning som används har endast utrum och neutrum. Maskulina och feminina ord normaliseras alltså till utrum.

honde klåkkon den-där klockan
hande vejjin den-där vägen

AVSAKNAD AV GENUS

Dialekterna i Karleby och Nedervetil i nÖB har ett förenklat genussystem: alla substantiv har en-genus i obestämd och det-genus i bestämd form (jfr Ivars 2015:88):

ein stell en ställe
hedi bo:me det-där bommen

ANNAT GENUS ÄN STANDARD SVENSKAN

När informanterna använder substantiv med annat genus än motsvarande ord i standardsvenskan behålls talarens genus:

et fabri:k ett fabrik
en jä:lo en gärda + ordbokslänk

Se avsnitt 4.1 om länkning till ordböcker. I vissa fall har böjningsformer konstruerats eftersom motsvarande form inte finns i standardsvenska:

börjane/börjani/byrjani börjanet + ordbokslänk

Då taggas ordet korrekt som bestämd form neutrum. I fall där vi använt konstruerade former hittas de i korpusen genom en sökning på ordets standardsvenska form som lemma (i exemplet ovan lemma: början)

2.2.3 NUMERUS

I vissa fall i norra Österbotten saknas pluralformer efter räkneord. Dessa fall normaliseras enligt standardsvenskans mönster:

<i>mangg hest</i>	<i>många hästar</i>
<i>mangg gangg</i>	<i>många gånger</i>

Notera även att pluralformer som *bastun*, *hobbyn* och *kaverin* ska standardiseras till *bastur*, *hobbyer* och *kaverir*. Detta görs för att ordklasstagningen ska bli korrekt (annars taggas motsvarande ord till bestämd form singularis i stället för obestämd form pluralis).

de fans fle:ra bastun *det fanns flera bastur*

2.3 ADJEKTIV

Den generella principen är att adjektiven överförs till de standardsvenska böjningsformerna.

2.3.1 ADJEKTIV PÅ -IG/-OG/-OT

Adjektiv som slutar på något av suffixen -ig, -og eller -ot sammanförs alla till en form på -ig.

<i>sveitto</i>	<i>svettig</i>
<i>råjso</i>	<i>rojsig</i> (+ länk till FSOB)
<i>pöllo</i>	<i>pöllig</i> (+x-tag)
<i>gåråt</i>	<i>gorrigt</i> (+länk till FO)
<i>nålot</i>	<i>nåligt</i> (+länk till FO)

På så sätt hittas alla adjektiv med samma rot under samma standardiserade form trots att suffixen kan vara olika i olika dialekter. Se avsnitt 4.1 om länkar till ordböcker.

2.3.2 GENUSKONGRUENS

De attributiva adjektivens genuskongruens i maskulinum, femininum eller neutrum överförs till utrum och neutrum enligt standardsvenskt mönster. Jfr 2.2.2 Substantivets genus.

mask.	<i>en långär ve:g</i>	<i>en lång väg</i>
fem.	<i>grömt granna tall</i>	<i>grymt grann tall</i>
neutr.	<i>ny:ji namn</i>	<i>nytt namn</i>

Exemplen ur Ivars 2015.

2.3.3 KOMPARERING

Även de komparerade formerna överförs till standardsvenskans böjningsformer.

fi:n – fi:nan – fi:nast *fin – finare – finast*

Observera att formerna *gamlare* (komparativ) och *gamlast* (superlativ) förekommer som komparering av *gammal*. Dessa förses även med x-tagg, se avsnitt 4.2.

3 FORMORD

Till formorden hör småord i ordklasserna prepositioner, subjunktioner, pronomen och adverb. Formorden standardiseras i Talko vanligen enligt huvudprincipen ord-för-ord. Det innebär att de förekommer även i andra sammanhang än enligt standardsvenskt mönster.

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>
<i>ja titta på de frå:n te:ve:n</i>	<i>jag tittade på det från teven</i>
<i>ja bo:r hi:t i låvi:sa</i>	<i>jag bor hit i Lovisa</i>
<i>vart bo:r du</i>	<i>vart bor du</i>

3.1 PRONOMEN

Huvudregeln är att pronomenen normaliseras till standardsvenskans former.

3.1.1 PERSONLIGA PRONOMEN

De/dem och *han/honom, hon/henne* normaliseras utgående från om det är fråga om ett subjekt eller objekt:

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>
<i>ja så:g dåm</i>	<i>jag såg dem</i>
<i>dåm kåm hem</i>	<i>de kom hem</i>
<i>ja slo:g han hå:rt</i>	<i>jag slog honom hårt</i>

När *han* och *hon* används om inanimata substantiv standardiseras de till *den*, jfr punkt 2.2.2.

GENERISKT PRONOMEN

Den form som informanten säger används i den ortografiska utskriften, vanligen *man, han* eller *en*.

<i>ha, a, an, in, 'n</i>	<i>han</i>
--------------------------	------------

3.1.2 DEMONSTRATIVA PRONOMEN

Om *han-där/hon-där* används om saker eller personer ska det normaliseras till *den-där*.

<i>onde Förnamn</i>	<i>den-där Förnamn</i>
<i>onde: systärn såm e kva:r</i>	<i>den-där systemen som är kvar</i>
<i>han råopa jo hande såm tjö:rt hande bå:tin</i>	<i>han ropade ju den-där som körde den-där</i>
<i>båten</i>	

Även former som motsvarar FO:s **hissen** (främst nÖB och sÖB) och **dessen** normaliseras till *den-här* respektive *den-där*.

Talko – korpus över den talade svenskan i Finland.

nå **hison** vikon nu då nå **den-här** veckan nu då
he va int så myki vi:dske:pelse he:r på **tessn** traktn det var inte så mycket vidskepelse här på
den-här trakten

Alla varianter av de plurala formerna normaliseras till *de-här* och *de-där*:

dihē:, *dihä*:, *dömhi*:, *dendähä*:, *teije*, *täisje*, *tåmsje*, osv. *de-här*

dendi, *didä*:, *dåmde*, *teije*, *täide*, *tåmtä*:r, *tömte* osv. *de-där*

Notera

Hete/heti/hede i nÖb har genomgående standardiseras till *det-där*.

SÅDAN-HÄR OCH SÅDAN-DÄR

Det förekommer stor variation i talspråket, ofta fogas böjningssuffixet till *här* i stället för till *sådan*, t.ex. *sånhän*:t. I den ortografiska utskriften används standardsvenskans böjningsformer. Kontexten avgör ofta vilken form som ska väljas.

Sånhä:r, *sånhä*:n, *sånje*: *sådan-här*
sånthä:r, *sånhä*:nt *sådant-här*

Även former med *tocken* (främst i Österbotten) och *slik* (i östra Nyland) har överförts till *sådan-här* respektive *sådan-där*

tåje, *tålje*, *tålije*, *tåsse* *sådan-här*
toande, *tode*, *tokande*, *tokian*, *toko*, *sådan-där*
tokode:, *tådde*, *tålde*, *tåle*, *tålidennar*

slikantä:r *sådan-där*

3.1.3 INTERROGATIVA PRONOMEN / VILKEN

Former som motsvarar FO:s **holgen** och **hocken** normaliseras till *vilken*.

nåo kan ann finnas men na: veit int **hokan** *na: meir* *nog kan den finnas men någon vet inte*
vilken *något mera*

3.3 ADVERB

För adverb med väldigt stor formrikedom har en överföring till standardsvenskans former gjorts.

3.3.1 HÄR

För alla varianter av 'här' används den standardsvenska formen:

hije, *hijennan*, *hä*:r, *hänn*, *ije*:, *je*:r, *jenn*, *jennar*, *sjenn*, *sjennar* osv. *här*

Här motsvaras i FO av bl.a. **härna**, **härnan**, **härnar**; **hit-hithjär**, **hit-hjär**

Talko – korpus över den talade svenskan i Finland.

3.3.2 DÄR

För alla varianter av 'där' används den standardsvenska formen:

de:r, dä:r, te:r, tide, tite:, tä, tä:r, tänn, tännan osv *där*

Där motsvaras i FO av bl.a. **där, dit-där, -däran, -därna, -därnan, -därnandes, -därnar**

3.3.3 HIT

För alla varianter av 'hit' används den standardsvenska formen:

hi:, hi:d, hi:je, hi:tjennan, hitt osv. *hit*

Hit motsvaras i FO av bl.a. **hit, hit-hit-hjär, hit-hjär.**

3.3.4 DIT

För alla varianter av 'dit' används den standardsvenska formen:

di:, di:t, ti:, ti:d, ti:dde, ti:t, tid:enand, titi: *dit*

Dit motsvaras i FO av bl.a. **dit, ditdär, dit-därna**

3.3.5 HÄRIFRÅN

För alla varianter av 'häriifrån' används den standardsvenska formen:

herifrå:n, hi:jifrå:n, hitifrå:n, ja:n, ja:nfrå:n, ja:nje, jedan, jedanfron, sja:nfroån osv. *häriifrån*

Häriifrån motsvaras i FO av bl.a. **här-ifrån, hit-ifrån, hädan, hädan-från.**

3.3.6 DÄRIFRÅN

För alla varianter av 'däriifrån' används den standardsvenska formen:

derifrå:n, di:ifrå:n, däriifrå:n, ta:n, ta:nde, ta:nifrå:n, te:rifrån, tä:rifroån osv. *däriifrån*

Häriifrån motsvaras i FO av bl.a. **där-ifrån, dädan, dädan-från, dädan-ifrån.**

3.3.7 DÄRTILL

Uttalsformer som *tärt, tärtti, tertil* normaliseras precis som i FO till *därtill*.

3.3.8 INTE

Alla varianter av negationen *inte* normaliseras till den standardsvenska formen

snapsvi:sår sjungär vi inga hi: *snapsvisor sjunger vi inte hit*

Inte motsvaras i FO av **inga** och **inte**.

3.3.9 OCKSÅ

Alla varianter av 'också' normaliseras till den standardsvenska formen:

å, å: åkså *också*

Talko – korpus över den talade svenskan i Finland.

Också motsvaras i FO av **ock, också(n)**.

3.3.10 NÅGORST OCH ANNARS

någårs 'någonstans' normaliseras som i FO till *någorst* och förses med ordbokslänk.

anars 'annanstans' normaliseras som i FO till *annars*.

3.3.11 ORD SOM SLUTAR PÅ -LEDES

Ord som slutar på -ledes skrivs utan bindestreck. De ord som inte förekommer i SAOL men i FO får en ordbokslänk

<i>a:leis, a:deleis</i>	<i>annorledes</i>	finns i SAOL
<i>holeis, horosleis</i>	<i>huruledes</i>	finns i SAOL
<i>li:kaleis</i>	<i>likaledes</i>	finns i SAOL
<i>na:leis, nå:leis, na:rleis</i>	<i>någonledes</i> + länk till FO	
<i>samaleis</i>	<i>sammaledes</i>	finns i SAOL
<i>sålis, såleis, söis</i>	<i>således</i>	finns i SAOL

3.4 PREPOSITIONER

Lägg märke till skillnaden mellan *på* och *uppå* (uttalas vanligen *åpå*: eller *opa*). Notera särskilt *mårade uppå*, INTE ~~*mårade och på*~~.

Lägg också märke till skillnaden mellan *av* och *utav* (uttalas bl.a. *uta, ota, otta*).

3.5 ORD SOM SLUTAR PÅ -EFTER, -ÖVER

Ord som slutar på -efter eller -över skrivs utan bindestreck. De ord som inte förekommer i SAOL men i FO får en ordbokslänk.

<i>bakit, baket, bakett</i>	<i>bakefter</i>	finns i SAOL
<i>jedanett</i>	<i>hädanefter</i>	finns i SAOL
<i>järett</i>	<i>härefter</i>	finns i SAOL
<i>langsett</i>	<i>längsefter</i>	finns i SAOL
<i>mittett</i>	<i>mittefter</i> + länk till FO	
<i>nierett</i>	<i>neretter</i>	finns i SAOL
<i>nårrett</i>	<i>norrefter</i> + länk till FO	
<i>söderett</i>	<i>söderefter</i> +x-tag	

3.6 SUBJUNKTIONER OCH KONJUNKTIONER

3.6.1 FÖRE OCH FÖRRÄN

Var uppmärksam på skillnaden mellan *före* och *förrän* och normalisera till samma som informanten använder.

3.6.2 ELLER

Eller uttalat som *hellär* med ett initialt h- normaliseras till *eller*, jfr FO **eller**.

3.7 INFINITIVMÄRKENA ATT OCH TILL

Infinitivmärket *att* uttalas ofta *å*. Var uppmärksam på att den ortografiska formen ska vara *att* (inte *och*).

Även *till* förekommer som infinitivmärke

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>
<i>å såva e sjö:nt</i>	att sova är skönt
<i>ti såva e sjö:nt</i>	till sova är skönt

3.8 DISKURSPARTIKLAR

Uppbackningar normaliseras enligt följande:

<i>nä:ä, jå:å</i>	<i>nä-ä, ja-a</i>
<i>jå:, ja:, jå, ja</i>	<i>ja</i>
<i>jo:, jo</i>	<i>jo</i>
<i>jåo</i>	<i>ja</i>
<i>mm-m</i>	<i>mm-m</i>
<i>sjo (inandning)</i>	<i>jo</i>

3.9 KRAFTUTTRYCK

Kraftuttrycken varierar beroende på område och generation. Eftersom de förändras snabbt är det troligt att ordböcker inte heller täcker den här variationen. De finlandssvenska varianterna kan därtill se lite annorlunda ut än dem som förekommer i standardsvenskan. De här formerna är dock ofta härledda ur samma rot och får därför stå som informanterna använder dem. Former som inte finns i SAOL men i FO förses med ordbokslänk, övriga med x-tagg.

<i>sa:ken</i>	<i>saken</i> +FO
<i>sa:kran</i>	<i>sakran</i> +FO
<i>jestas</i>	<i>jestas</i> +länk till FO <i>jestas</i>
<i>jessus</i>	<i>jessus</i> + länk till FO <i>Jesus</i>
<i>herrejess</i>	<i>herrejess</i> +länk till FO <i>jess</i>
<i>vitsi</i>	<i>vitsi</i> +x-tagg

4 ORD SOM INTE FINNS I SAOL

För ord som inte finns i SAOL konstrueras en normaliserad form med stöd av följande ordböcker:

Finlandssvensk ordbok (FSOB) -

Ordbok över Finlands svenska folkmål (FO)

Svenska Akademiens ordbok (SAOB)

Lånord från andra språk skrivs enligt ursprungsspråket. Det är ibland svårt att avgöra om ett ord är ett tillfälligt lån eller ett integrerat ord som ska behandlas enligt svenska stavningsregler.

ORTOGRAFISK FORM FÖR FINSKA ORD

Många ord som vid första anblicken kan verka vara ett finskt lånord finns också i FO. FO konstruerar vanligen ortografisk form enligt svenskt mönster medan FSOB vanligen håller den finska stavningen. Om ordet finns i FSOB, välj den ortografiska formen i första hand, i andra hand en form ur FO.

Ordbok över Finlands svenska folkmål	Finlandssvensk ordbok	TALKO	
jutta s.	juttu	juttu	
pajka s.	-	paikka	
kejka s 2	keikka	keikka	
pajta s.	pajta/pajtu	paita	
nålog, -ot	-	nålot/nålogt	
-	mopo	mopo	

Observera att böjningsformerna ska skrivas ut enligt svenskt mönster, t.ex. *kaverir* (pl.), *kaverirna* (best. pl.) annars blir taggningen fel!

Alla ord som inte finns SAOL ska ha en lang-tag. Taggen kan antingen vara en ordbokslänk eller en x-tag.

GRÄNSDRAGNINGAR

Det är inte helt entydigt när en standardsvensk form ska väljas och när en normaliserad form som ligger närmare dialekten ska väljas. Till exempel för vissa formord, t.ex. adverbena *hit-här*, *dit-där*, med stor variation i dialekterna har den standardsvenska formen valts, se avsnitt 3.

Även för vissa substantiv väljs den standardsvenska formen i Talko, även om FO har flera olika uppslagsord: t.ex. hare (FO: hare, häre), gran (FO: gran, grän), björk (FO: björk, bjärk), broms (FO: broms, bröms). Ingen x-tagg eller ordbokslänk läggs ut eftersom den ortografiska formen finns i SAOL.

4.1 ORDBOKSLÄNKNING

I den ortografiska utskriften kan ett ord försees med direktlänkar till uppslagsord i FO eller FSOB. Ofta finns ett ord i båda ordböckerna och då kan två länkar läggas in.

Även sammansättningar kan få en ordbokslänk ifall ena leden finns i FSOB eller FO.

de e jettekiva *det är jettekiva (+länk till FO + FSOB)*

Observera att ord som finns i SAOL inte får ordbokslänk även om betydelsen är delvis annan, t.ex. *söka* i betydelsen 'hämta' eller *bro* i betydelsen 'brygga'. Många ord som man associerar med dialekterna finns i SAOL och ska inte ha länk, t.ex. *håksa*, *bakefter*, *alltjämt*.

Till FO länkas ord som förekommer enbart i dialekterna eller talspråket. Till FSOB länkas huvudsakligen finlandismer, även ord som förekommer i SAOL med markeringen finl., t.ex. vappen och skinksemla.

EXEMPLET BULLE

FO har uppslagsordet **bulle** (bolla, bulla) för svaga maskulinum-former och **bull**a (bollo, bolla) för svaga femininum-former. I FSOB finns uppslagsordet **bull**a, **bullor**.

>bollo, bollor, bollona blir *bull*a, *bullor*, *bullorna* i ortografiska utskriften
tjöttbollona > köttbullorna + FO-länk + FSOB-länk till köttbulla

fetti:sdasbollona > fettisdagsbullorna + FO-länk till fettisdagsbulla

>uttalen *bull*a/*bolla*, *bullar*/*bollar*, *bullana*/*bollana* får olika form i FO enligt genus
oja_yw02_sp

en vissan bolla > mask? > en viss bulle (>ingen länk)

men talare senare om *tsöttbollona* > fem > köttbullorna + FO-länk

kronoby_ow02: pakka jo veskon me saft å smörgå:san å bolla > bulle/bulla (bullarna?!)

"fasta uttryck": *baka bolla*, *kaffi å bolla*

bollabju:dasi > bullabjudaset + länk till FO, FSOB bulla + g-tagg

bollasji:vå > bullaskiva + länk till FO, FSOB bulla

BÖJNINGSFORMER MED ORDBOKSLÄNK

-*gamlare* och *gamlast* som kompareringsformer till *gammal* + länk till FO *gammal*

-*börjanet* + länk till FO *början*

4.2 X-TAGG

Övriga ord som inte finns i standardsvenskan märks ut med taggen x (lang=x) för att göras sökbara i korpusen.

4.2.1 SLANGORD OCH INOFFICIELLA NAMNFORMER

Inofficiella ortnamn och slangformer av ortnamn, föreningsnamn, tidningar o. dyl förses med x-tag:

<i>Hesa</i> + x-tag	Helsingfors
<i>Jeppis</i> + x-tag	Jakobstad
<i>Västis</i> + x-tag	tidningen Västra Nyland
<i>Kiffen</i> + x-tag	Kyrksläpps IF
<i>handels</i> + x-tag	handelsläroverket
<i>choko</i> + x-tag	'choklad'
<i>Medi</i> + x-tag	'medicinska fakulteten'
<i>Tyksen</i> + x-tag	Tyks, Turun yliopistollinen keskussairaala

4.2.2 ORD FRÅN ANDRA SPRÅK

yes + x-tag

alright + x-tag

En hel del engelska ord förekommer i SAOL och behöver inte ha x-tag: *showbusiness*, *action*. En del ord finns i Svensk ordbok (men inte i SAOL) och behöver inte heller få en tagg: *sorry*, *inside information*

I längre citat på andra språk förses varje ord med x-tag.

mitä kieltä toi puhuu > *mitä* +x-tag *kieltä* +x-tag *toi* +x-tag *puhuu* +x-tag

Notera att namn på andra språk inte får x-tag.

4.3 G-TAGG

S.k. substantiverade passiva infinitiver markeras med g-tag. Jfr Huldén 1959: 255 f. och Ivars 1988: 121. Observera skillnaden mellan bestämd och obestämd i den ortografiska formen.

<u>Grov ljudskrift</u>	<u>Ortografisk utskrift</u>	
<i>ringase</i>	ringaset + g-tag	'ringandet'
<i>tjö:ras</i>	köras + g-tag	'körande'
<i>messåmasfi:rased</i>	midsommarfiraset +g-tag	'midsommarfirandet'
<i>jaktase</i>	jaktaset + g-tag + FO-länk	'jagandet'

5 TRANSLITTERATORBRUS

Translitteratorn skapar den ortografiska utskriften utgående från enskilda ord men klarar inte av att ta kontexten i beaktande. Vid homonymer väljer translitteratorn den variant som är mest frekvent. Det är skäl att vara extra uppmärksam vid följande exempel:

nu> nog/nu å> och/också
va> vad/var ti:d > tid /dit
visst > visst /visste e >det/är
ha: > ha/har "skulle har" jább > jobb/jobba
tala > tala, talar, talade
tenkt > tänkt/tänkte

LITTERATURFÖRTECKNING

Huldén, Lars (1959), *Verbböjningen i Österbottens svenska folkmål II*. Studier i nordisk filologi 48. SLS: Helsingfors.

Ivars, Ann-Marie (1988): *Närpesdialekten på 1980-talet*. Studier i nordisk filologi 70. SLS: Helsingfors

Ivars, Ann-Marie (2010). *Sydösterbottnisk syntax*. Studier i nordisk filologi 84. SLS: Helsingfors.

Ivars, Ann-Marie (2015). *Dialekter och småstadsspråk*. Svenskan i Finland – i dag och i går I:1. SLS: Helsingfors.

Leinonen, Therese (2015). [Talet lever!](#) Studier i nordisk filologi 86. SLS: Helsingfors.