

# Transkripsjonsrettleiing for LIA - samisk

1.12.2016

Av Kristin Hagen, Live Håberg, Arnstein Johnskareng, Eirik Olsen og Åshild Søfteland <sup>1</sup>

## Innhald

0 Om rettleiinga .....	2
1 Bakgrunn .....	3
1.1 Gode råd for transkripsjon .....	3
1.2 Filnamn .....	3
1.3 Namn på informantar .....	4
1.4 Loggføring .....	4
2 Transkripsjon og korrektur i ELAN .....	5
2.1 Transkripsjonsprogrammet ELAN .....	5
2.2 Starte ein ny transkripsjon og definere talarar .....	5
2.3 Halde fram med ein påbegynt transkripsjon .....	6
2.4 Segmentering .....	6
2.4.1 Segmentering i Segmentation Mode .....	6
2.4.2 Nyttige snarvegar, Segmentation Mode .....	8
2.5 Transkripsjon.....	8
2.5.1 Transkripsjon (og korrekturlesing) i Transcription Mode .....	8
2.5.2 Nyttige snarvegar, Transcription Mode.....	11
2.6 Korrekturlesing .....	11
2.6.1 Korrekturlesing i Annotation Mode .....	11
2.6.2 Nyttige snarvegar, Annotation Mode .....	12
3 Transkripsjonsreglar.....	13
3.1 Når reglane ikkje strekk til.....	13
3.2 Inndeling av talen i turar og segment.....	13
3.2.1 Nokre grunnbegrep: Tur, replikk og ikkje-replikk .....	13
3.2.2 Inndeling av turar i segment.....	14
3.3 Taggar til markering av ekstrainformasjon og ikkje-språklege lydar .....	15
3.3.1 Generelle prinsipp for taggar .....	15
3.3.2 Liste over avhengige og uavhengige taggar med tyding.....	17
3.3.3 Sensitive personopplysningar og annan informasjon som ikkje skal med.....	17
3.4 Språkform.....	19
3.4.1 Namn.....	19
3.4.2 Ikkje-språklege lydar.....	20
3.4.3 Avbrot, pausar og uklare passasjar .....	20
3.4.3.1 Avbrotne ord.....	20
3.4.3.2 Uavslutta ytringar.....	20
3.4.3.3 Pausar .....	21
3.4.3.4 Passasjar som er vanskelege å forstå.....	22

<sup>1</sup> Transkripsjonsrettleiinga er eit resultat av samarbeid mellom mange tilsette i fleire prosjekt.

LIA/Norsk i Amerika: Tor Erik Jenstad, Andre Kåsen.

ScanDiaSyn/Nordiasyn: Janne Bondi Johannessen, Signe Laake, og Øystein Vangsnes samt Sine Bjordal, Ingebjørg Før Gjermundsen, Astrid Grov, Margit Ims, Kari Kinn, Ann Helen Lea, Arne Martinus Lindstad, Alexander Lykke, Stig Rognes, Karine Stjernholm. Dialektendringsprosessar: Bente Selback

3.4.4 Stor bokstav og teiknsetjing .....	22
3.4.4.1 Stor bokstav.....	22
3.4.4.2 Spørjeteikn .....	23
3.4.4.3 Hermeteikn.....	23
4 Korrektur.....	23
4.1 Lese korrektur .....	23
5 Oversikt over hurtigkommandoar, eventar og ikkje-språklege lydar.....	24
5.1 Oversikt over hurtigkommandoar .....	24
5.2 Oversikt over taggar.....	25
5.3 Liste over interjeksjonar.....	26
6 Referansar .....	27
7 Vedlegg.....	28
7.1 Når du begynner å jobbe på ei ny maskin.....	28
7.2 Hugseliste for transkripsjon .....	29
7.3 Hugseliste for korrektur .....	30

## 0 Om rettleiinga

Denne rettleiinga bygger på transkripsjonsrettleiingar frå prosjekta NoTa-Oslo (UiO), Scandiasyn/Nordiasyn (UiO) og Dialektendringsprosessar (UiB).

I avsnitt 1 finn du ein generell introduksjon om transkripsjon. Avsnitt 2 er meir detaljert og praktisk, og gir gode råd om transkripsjon i tillegg til ei kort skildring av korleis du bruker transkripsjonsprogrammet *Elan*. Avsnitt 2.6.2 tek for seg dei konkrete transkripsjonsreglane. Dette avsnittet bør du ha lett tilgjengeleg medan du transkriberer. Avsnitt 0 handlar om korrektur. I avsnitt 5 finn du ei oversikt over hurtigkommandoar og taggar, og avsnitt 6 inneheld referansar. I avsnitt 7 finn du nyttige vedlegg.

# 1 Bakgrunn

For at ein skal kunne søkje i eit talemålsmateriale, må materialet transkriberast. Det er mange måtar å transkribere talespråk på, men ingen av dei er perfekte. Ein *lydnær* transkripsjon gir det beste og mest detaljerte biletet av kva informanten seier, spesielt om ein bruker IPA eller ei anna lydskrift. Men ein slik transkripsjon er tidkrevjande. Ein lydnær transkripsjon kan også vere tung å lese og vil dessutan vere vanskeleg å søkje i fordi ein må vite på førehand kva for ordformer som finst.

Ein *ortografisk* transkripsjon er raskare, lett å lese og gir eit resultat som er enkelt å søkje i. Men sidan mange nyansar kan gå tapt, både når det gjeld fonologi og morfologi, er det heller inga ideell løysing. Moderne talespråkskorpus har rett nok lenka saman transkripsjon og lydfil slik at ein enkelt kan lytte seg igjennom dei eksempla ein ønskjer. Det kan likevel vere vanskeleg å finne fram til mange fenomen, som for eksempel dativbøying som vil gå tapt ved ein ortografisk transkripsjon, og det kan bli mange eksempel å lytte seg igjennom.

NoTa-Oslo-prosjektet og Dialektendringsprosessar-prosjektet valde begge ortografisk transkripsjon, det første bokmål og det andre nynorsk. Scandiasyn-prosjektet valde ein talemålsnær transkripsjon med norsk alfabet og deretter ei halvautomatisk omsetjing til bokmål.

I LIA-prosjektet skal vi bruke metoden frå Scandiasyn med lydnær transkripsjon først for å få fram dei viktigaste nyansane i dialektopptaka. Deretter skal vi nytte transliteratoren frå Scandiasyn-prosjektet for å omsetje til ortografisk transkripsjon, denne gongen til nynorsk. Vi skal framleis behalde den originale syntaksen og dialektord og andre ord som ikkje let seg omsetje.

Tidlegare prosjekt har nytta transkripsjonsprogramma Transcriber og Praat. I LIA satsar vi på eit nytt program: *ELAN*.

## 1.1 Gode råd for transkripsjon

Det er viktig å finne fram til gode rutinar og vanar for transkripsjon. Målet er at arbeidet skal gå både raskt og så feilfritt som mogleg. Det er ikkje sikkert at det som fungerer for éin transkribør, fungerer for alle, men under finn du likevel nokre gode råd du bør merke deg:

- Ta hyppige, korte pausar der du fokuserer auga på noko anna og strekkjer på beina.
- Arbeid aldri meir enn fire timar av gongen med transkripsjon.
- Lær deg flest mogleg hurtigkommandoar. Det sparar tid og ikkje minst hendene dine.
- Fokuser på det du høyrer, det er dessverre fort gjort å skrive det ein ventar i staden for det som faktisk blir sagt.

## 1.2 Filnamn

Lydfilene har namn etter dette mønsteret:

stadnamn \_ universitetsnamn forkorta \_ eit nummer (som ikkje er informantnummer. Filene frå same stad får nummer fortløpande).

Eksempel: *valdres\_uio\_01*, *valdres\_uio\_02*

Transkripsjonsfilene får namn etter lydfila. Legg dine eige initialar bakarst i namnet slik: *\_kh*.

Eksempel: Eksempel: *valdres\_uio\_01*, *valdres\_uio\_02\_kh*

### 1.3 Namn på informantar

Informantane får også namn etter lydfila. Altså først stadnamn, deretter forkorta universitetsnamn og til slutt nummer på fila og nummer på informant:

Eksempel:

*valdres\_uio\_0101*,

*valdres\_uio\_0102*,

*valdres\_uio\_0103*

*valdres\_uio\_0201*

*valdres\_uio\_0202*

Intervjuaren kallar vi *int*. Dersom det er fleire intervjuarar, kallar vi dei *int1*, *int2* osv.

### 1.4 Loggføring

Kvar transkribør skal ha ei loggfil der arbeidet for kvar arbeidsøkt er loggført. Loggen skal ligge lagra i mappa med transkribøren sitt namn. Det er viktig at alle fører nøyaktig logg. Loggen skal innehalde dato og kor mange timar du har jobba. I tillegg skal du føre nøyaktig informasjon om kva fil/filer du har jobba med i den aktuelle økta. Skriv opp tidskoden du starta på, og tidskoden for der du var då du avslutta økta. Det er også viktig at du opplyser om kva type arbeid du har gjort (transkribering eller korrekturlesing). Dersom det er noko spesielt med fila du jobbar med (lydkvalitet, spesielle problem med informanten e.l.), skriv du også opp dette. Noter også tid som har gått med til møteverksemd eller liknande. Her er eit eksempel på ei loggfil:

Logg for xx			
DATO	TIMER	ARBEID	ANNET/KOMMENTAR
x.x.xxxx	4	skreia_03-04.kh.trs – 2.40 – 6.10	Transkribert
x.x.xxxx	4	skreia_02-sl_ amg_kh.trs skreia_03-04.kh.trs – 6.10 – 7.20	Korrekturlest – tok lang tid, mye mumling Transkribert – mange spørsmål – mye e/æ

Når alle transkribørane fører ein slik logg, vil ein med jamne mellomrom kunne danne seg eit bilete av framgangen i prosjektet generelt. Det gjer det også mogleg å komme med tilbakemeldingar til kvar enkelt transkribør. Til dømes kan det vere aktuelt at ein som transkriberer svært raskt, men slurvar ein del, kan sakke farten litt. På same måte vil ein som bruker lang tid, og har svært få feil, kanskje kunne arbeide raskare og «slurve» litt meir, sidan alle filene uansett skal korrekturlesast. Dette er først og fremst meint som hjelp til transkribørane, ikkje som overvaking og styring.

Transkribørane bør også ha sin eigen problemlogg der ein noterer alle spørsmål knytt til problem

med transkripsjonen. Når ein støyter på eit problem, vil det av og til ikkje vere nokon rundt ein kan spørje, eller ein vil kanskje bli einige om at problemet må diskuterast på neste transkripsjonsmøte. I problemloggen skriv du opp kva fil det gjeld, den nøyaktige tidskoden (slik at ein lett kan finne att staden) og kva problemet gjeld. Loggfilene og problemfilene har transkribørane sine initialar som fornamn og *logg* og *prob* som etternamn.

Hugs også å oppdatere dette regnearket i Google Docs:

[https://docs.google.com/spreadsheets/d/1vIqvjpwl\\_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846](https://docs.google.com/spreadsheets/d/1vIqvjpwl_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846)

## 2 Transkripsjon og korrektur i ELAN

### 2.1 Transkripsjonsprogrammet ELAN

*ELAN* er eit gratis transkripsjonsprogram frå Max Planck Institute og The Language Archive i Nederland. Vil du lese meir om det, kan du gå inn på nettsida: <http://tla.mpi.nl/tools/tla-tools/elan/>

I dette avsnittet skildrar vi kort korleis programmet fungerer. På nettsida finn du også ein manual for programmet.

Når du skal begynne å jobbe med ELAN på ei ny maskin, er det nokre innstillingar som må ordnast. Du finn eit vedlegg med manual for dette i 7.1.

ELAN har tre ulike modus å arbeide i:

- Segmentation Mode
- Transcription Mode
- Annotation Mode

I Segmentation Mode kan transkripsjonen delast opp i tidsbolkar (les meir om dette i kapittel 2.4). Når segmenteringa er unnagjort, kan du gjere sjølve transkripsjonen i Transcription Mode (sjå kapittel 2.5). I Annotation Mode er det mogleg å både segmentere og transkribere, men vi vil likevel rå til å berre bruke denne modusen til å sjå over eller korrekturlese transkripsjonen og korrekturlesinga.

Kvar transkribør må sjølvsgt velje sin eigen måte, men det vil sannsynlegvis lønne seg å veksle mellom periodar med segmentering og periodar med transkripsjon for å få litt variasjon i arbeidet.

### 2.2 Starte ein ny transkripsjon og definere talarar

I 7.2 finn du ei hugseliste som kan vere nyttig å sjå på når du skal starte på ein ny transkripsjon.

Flytt først lydfila du skal jobbe med, inn i di eiga mappe. Finn fila i dokumentet «LIA-filer (fane: «Samisk»):

[https://docs.google.com/spreadsheets/d/1vIqvjpwl\\_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1200260466](https://docs.google.com/spreadsheets/d/1vIqvjpwl_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1200260466)

Skriv namnet ditt i transkribørfeltet. Når du er ferdig med fila, kryssar du av for «Ferdig tanskribert».

Opne ELAN.

- Vel **File** → **New**
- I vindauget til venstre vel du riktig lydfil
- Kryss av for **Template** og vel fila *template-lia.etf* (nota\LIA\Transkripsjon) i vindauget til venstre
- Set namn på talarane i lydfila. Kvar talar skal ha sin **tier** (lag) under bølgeforma i **Annotation mode** (som er førehandsvalt modus i opningsvindauget til ELAN). Rekkjefølgja på talarane har ingen ting å seie. Du set namn på talarane slik:
  - Vel **Tier** → **Change Tier Attributes**.  
I dialogboksen skal **Tier Name** og **Participant** vere det same og ha same namn som i filnamnet. I *template\_lia* er *talari* og *intervjuari* førehandsdefinert, men namna skal endrast. I tillegg skal du sjekke:  
**Annotator:** sett initialane dine her.  
**Parent Tier** skal vere *none*.  
**Lingustic Type** er *Utterance Type*.  
**Default language** kan stå som *System default*.
  - Treng du å definere fleire talarar, klikk på **Add** og fyll ut dialogboksen på same måten. Du kan definere fleire talarar når som helst under transkripsjonen.

## 2.3 Halde fram med ein påbegynt transkripsjon

Dobbelklikk på transkripsjonsfila du vil jobbe med, og du vil automatisk få opp både transkripsjon og lyd-/videofil. Du kan også opne programmet ELAN og deretter velje **Open** og rett transkripsjonsfil. Lyd-/videofila kjem opp automatisk.

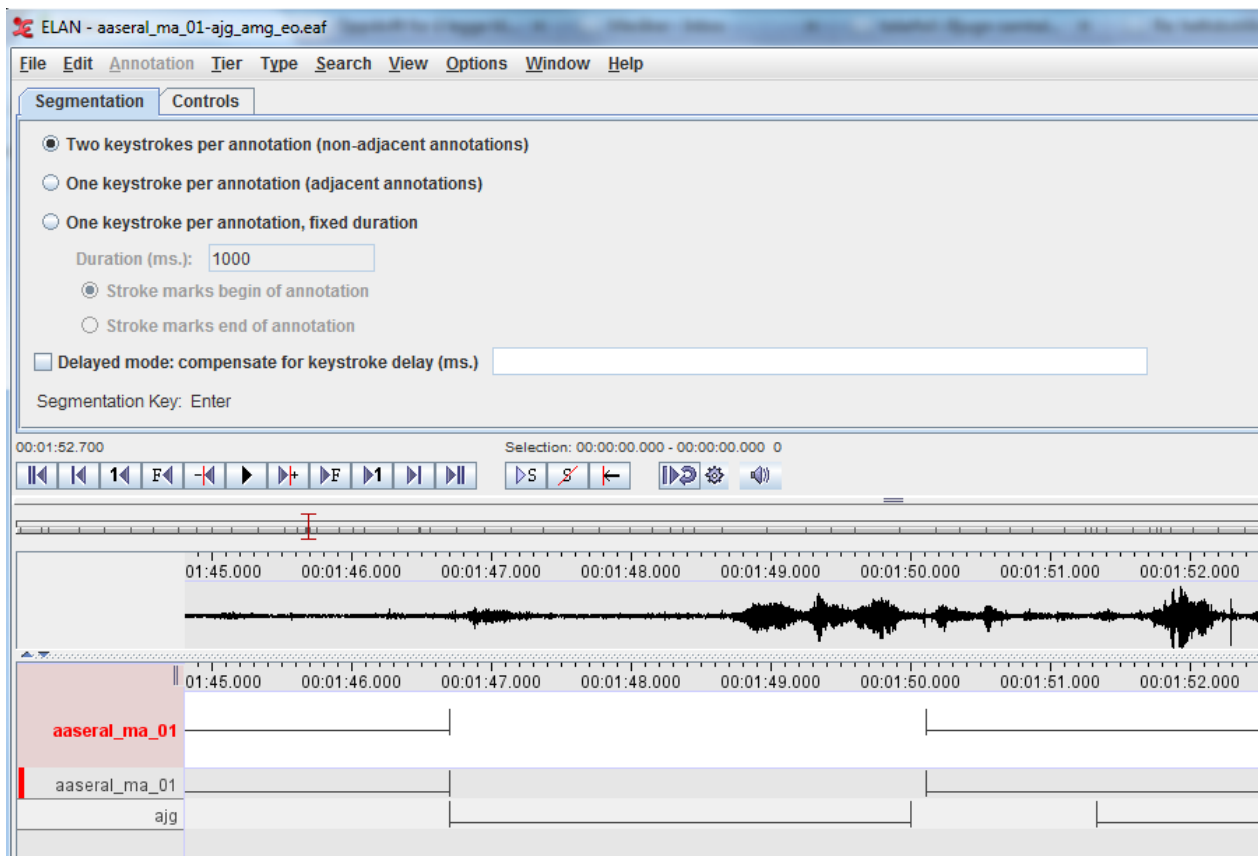
## 2.4 Segmentering

I segmenteringsdelen skal lyden delast opp i mindre bitar. For kvar gong vi startar eller avsluttar eit segment, blir det sett inn ein tidskode i transkripsjonen. Det er desse tidskodane som gjer at transkripsjonen blir lenka saman med lydfila. Du kan lese om turar, segment, replikkar og ikkje-replikkar i kapittel 3.2.

### 2.4.1 Segmentering i Segmentation Mode

Vel **Options** → **Segmentation Mode**.

Du er no inne i ein modus der det ikkje er mogleg å transkribere, men der det går raskt å segmentere, det vil seie å dele opp lydfila i tidssegment for kvar talar.

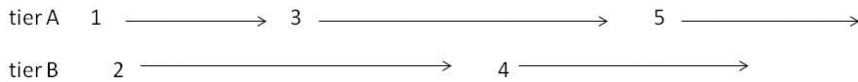


Du speler av lyden ved å trykke på play-knappen på spelaren eller enklare, trykke Ctrl+Mellomrom. Du stoppar lyden på same måten. Dersom du har markert eit område i bølgeforma, nyttar du Shift+Mellomrom for å spele av akkurat det området.

Du kan veksle mellom aktive *tiers* med pil opp og pil ned. Den aktive tieren kan vi dele inn i segment (*annotations*) ved hjelp av markøren (*crosshair*) og *enter*-tasten. Set markøren der segmentet skal starte i spektogrammet og marker starten på eit segment med *enter*-tasten. Set så markøren der segmentet skal slutte, og marker slutten med eit nytt trykk på *enter*-tasten. Og vips – du har laga eit segment mellom dei to markerte stadene. Dersom du transkriberer ein talar som snakkar lenge og ønskjer å dele opp i fleire segment utan pause mellom, trykk *enter* to gonger i slutten av eit segment, slik at slutten på førre segment automatisk blir markert som byrjinga på neste.

Du kan flytte markøren i spektogrammet med musa eller ved hjelp av snarvegar (sjå under). Segment kan flyttast og finjusterast ved å klikke og dra, eller eventuelt slåast saman eller delast opp ved hjelp av snarvegar.

I ein samtale eller eit intervju med to deltakarar, kan det vere greitt å konsentrere seg om å segmentere ein tier (og altså ein talar) om gongen: Først kan du dele inn segmenta til talar A heilt til talar B tek over turen. Gå så tilbake til starten av turen og segmenter ikkje-replikkane til talar B. Fortset deretter med talar B heilt til talar A tek tilbake turen. Gå så tilbake og segmenter ikkje-replikkane til talar A og fortset slik, delvis overlappende heile vegen, jf. illustrasjonen under, der rekkjefølgja på segmenteringa er nummerert.



Storleiken på spektogrammet kan du endre ved hjelp av den vesle knappen på lina nedst til høgre på sida.

## 2.4.2 Nyttige snarvegar, Segmentation Mode

Merk at nokre av snarvegane er ulike på PC og Mac (sidan funksjonstastane er ulike).

Funksjon	PC og Mac
Slå saman (utheva segment) med neste segment	Ctrl+A
Slå saman (utheva segment) med førre segment	Ctrl+B
Del opp (utheva segment)	Ctrl+Enter <sup>2</sup>
Gjer tieren over aktiv	Pil opp
Gjer tieren under aktiv	Pil ned
Spel/Pause	Ctrl+Mellomrom
Spel markert område	Shift+Mellomrom

Bruk musa til å flytte markøren, eller bruk snarvegane:

Funksjon	PC og Mac
Flytt markøren eit sekund til venstre	Shift+Pil venstre
Flytt markøren eit sekund til høgre	Shift+Pil høgre

Ctrl/Cmd+Shift+Piltastar kan nyttast for å flytte markøren eit lite hakk til høgre eller venstre.

## 2.5 Transkripsjon

I Transcription Mode går det raskt å transkribere den delen av opptaket som allereie er segmentert i Segmentation Mode. Du kan ikkje redigere sjølve segmenteringa i denne delen av programmet.

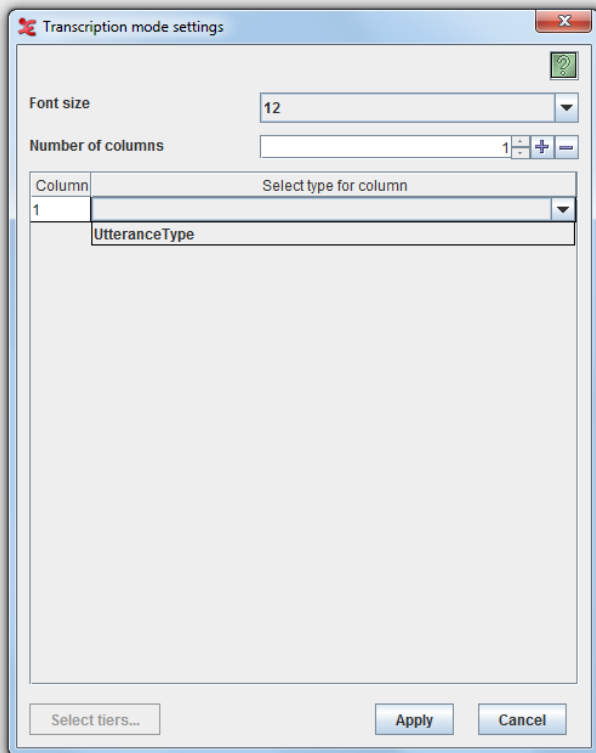
### 2.5.1 Transkripsjon (og korrekturlesing) i Transcription Mode

Vel **Options** → **Transcription Mode**.

Når du bytter til Transcription Mode første gong, må du konfigurere oppsettet. Talet på *columns* skal vere 1, og du skal velje *Utterance type*.

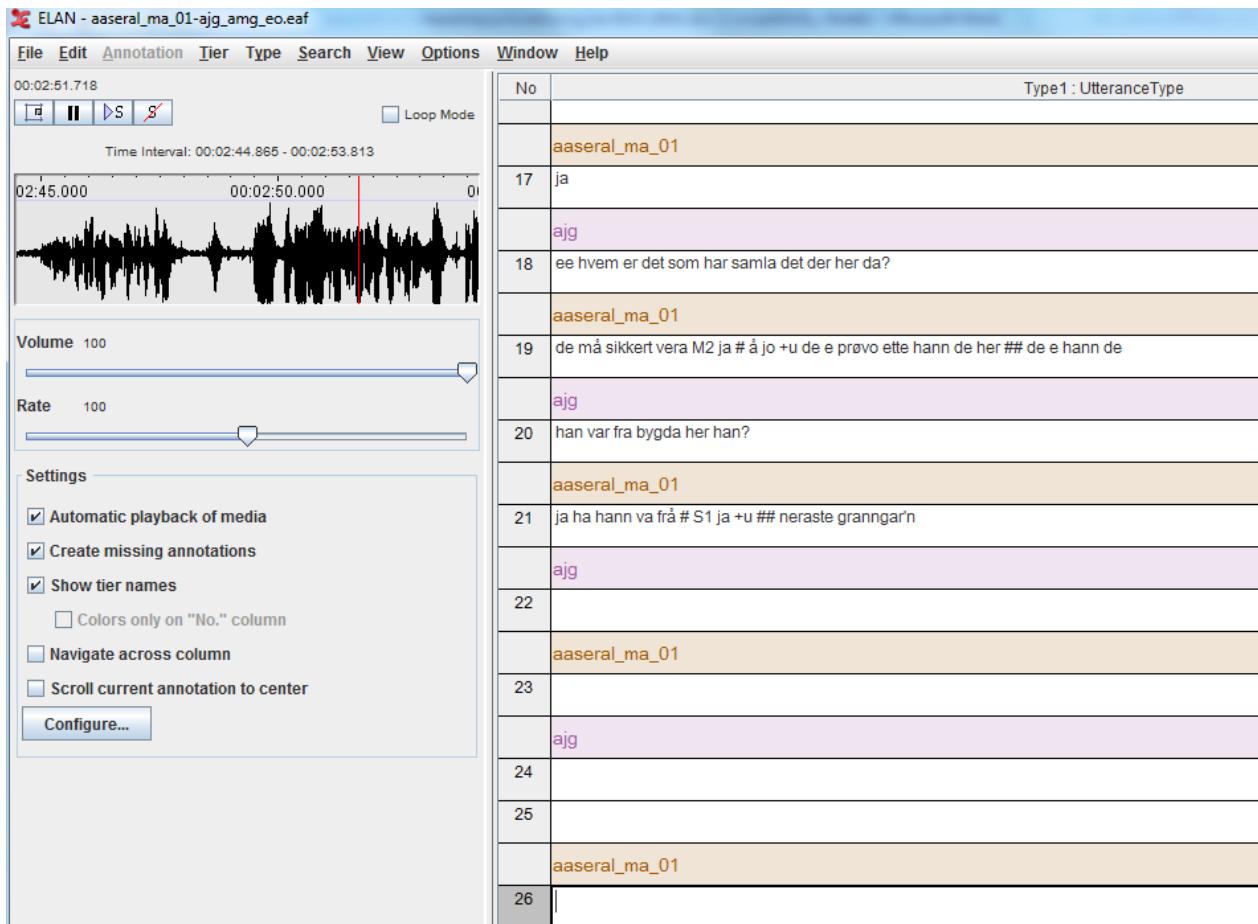
<sup>2</sup> Merk at denne snarvegen ikkje er førehandsdefinert. I kapittel 7.1 kan du lese om korleis du definerer egne snarvegar.





Under *Select tiers* treng du vanlegvis ikkje å gjere noko. Her skal det vere haka av for alle talarane. Vel gjerne 12 punkt som passande skriftstørrelse.

Du får no ei visning der segmenta du laga i Segmentation Mode er plasserte under kvarandre med namnet på dei ulike tierane i ulike fargar.



Det er fleire ulike funksjonar for å spele av lyden:

- Du kan klikke i det kvite feltet der transkripsjonen skal stå
- Tab fungerer som ein play/pause-knapp
- Shift+Tab speler av lyden frå starten av segmentet
- Enter går til neste felt og speler av lyden
- Alt+Pil opp/ned skiftar felt og speler av lyden

Hakar du av for *Loop Mode* (øvtst til høgre over spektogrammet), vert lyden i segmentet spelt om att og om att til du stoppar han med Tab.

Hastigheita på opptaket kan du endre ved å justere *Rate* (lina under *Volume*). Som regel bør *Rate* vere 100, men om noko er svært utydeleg eller talaren snakkar svært fort, kan du forsøke å setje ned hastigheita noko.

Elles er det ikkje så mykje meir å seie: Du høyrer, transkriberer og trykkar på Enter. Spektogrammet i vinduet til venstre kan vere til hjelp. Storleiken på spektogrammet kan du endre ved å dra i høgre kanten. Om du vil endre på segmenteringa, må du gå tilbake til Segmentation Mode, eventuelt til Annotation Mode (jf. avsnitt 2.6).

## 2.5.2 Nyttige snarvegar, Transcription Mode

Funksjon	PC og Mac
Hopp ned til neste segment	Enter/Alt+Pil ned
Hopp opp til forrige segment	Alt+Pil opp
Spel/Pause	Tab
Spel segmentet om igjen fra start	Shift+Tab

Dei same snarvegane kan brukast til å flytte markøren som i Segmentation Mode (sjå avsnitt 2.4.2).

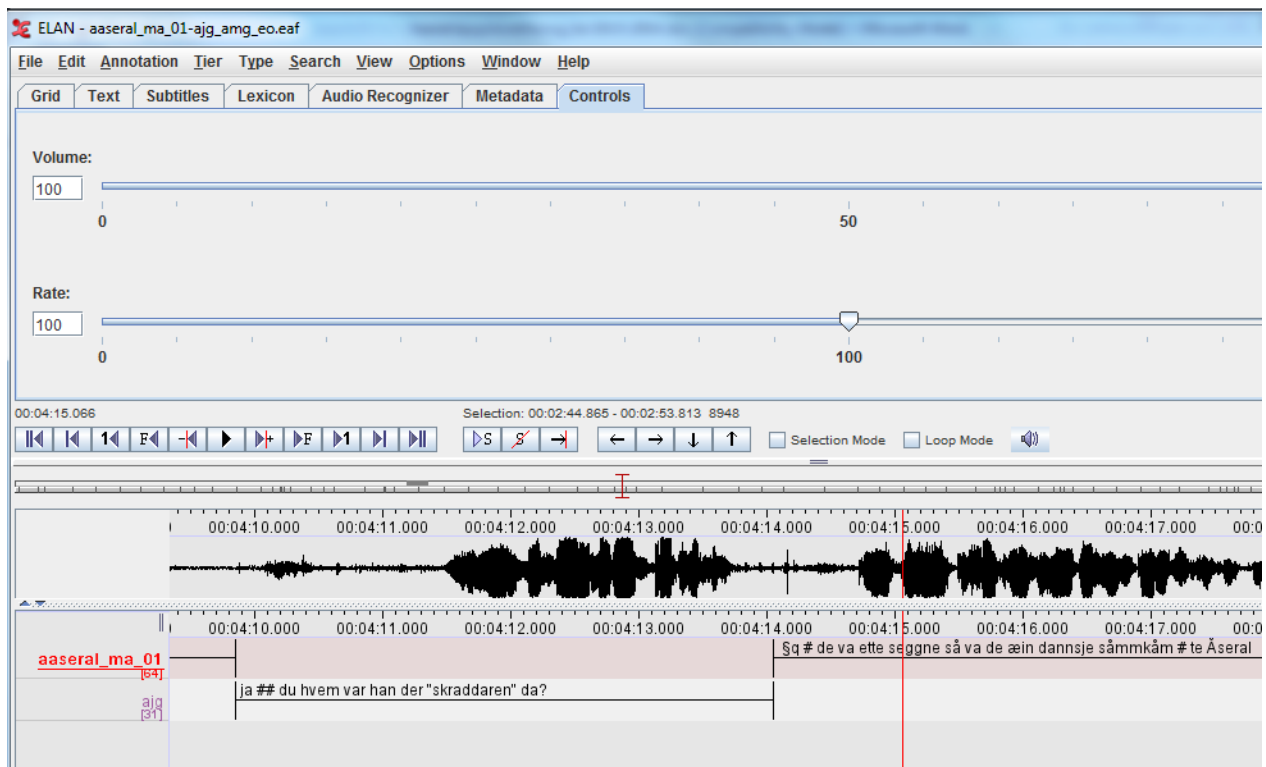
## 2.6 Korrekturlesing

Transcription Mode eignar seg godt til korrekturlesing, men du kan også gå over i Annotation Mode der du får god oversikt over både segmentering og transkripsjon. Det står meir om korrekturlesing i avsnitt 4.

### 2.6.1 Korrekturlesing i Annotation Mode

Vel **Options** → **Annotation Mode**.

I Annotation Mode kan du rette både på transkripsjon og segmentering. Det ser likevel ut til at det er raskast å bruke Segmentation og Transcription Mode kvar for seg til den første transkriberinga (og kanskje til korrekturen også) og bruke Annotation Mode berre til å sjå over.



Dobbelklikk på eit segment (annotation) for å gå inn og rette i transkripsjonen. Ved å bruke Alt+M får du transkripsjonen opp i eit eige vindauge og kan redigere han der. For at endringane skal bli lagra i dette vindauget, må du trykke Ctrl+Enter

Segmenteringa kan endrast ved å klikke og dra medan du held inne halde inne alt-knappen.

## 2.6.2 Nyttige snarveger, Annotation Mode

Merk også at nokre av snarvegane er ulike på PC og Mac (sidan funksjonstastane er ulike).

Funksjon	PC og Mac
Spel/Pause	Ctrl+Mellomrom
Spel markert område	Shift+Mellomrom
Ny annotation her (markert område)	Alt+N
Slett annotation	Alt+D
Gå til annotation over	Ctrl+pil opp
Gå til annotation under	Ctrl+pil ned
Gå til annotation til høgre og rediger	Ctrl+Alt+pil høgre
Gå til annotation til venstre og rediger	Ctrl+Alt+pil venstre
Flytt venstre segmenteringsgrensa eit hakk til venstre	Ctrl+J
Flytt venstre segmenteringsgrensa eit hakk til høgre	Ctrl+U
Flytt høgre segmenteringsgrensa eit hakk til venstre	Ctrl+Shift+J
Flytt høgre segmenteringsgrensa eit hakk til høgre	Ctrl+Shift+U

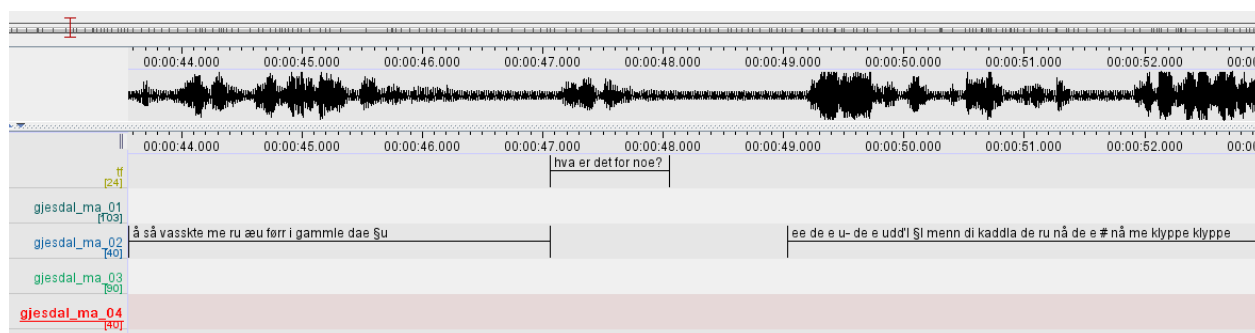
## 3 Transkripsjonsreglar

### 3.1 Når reglane ikkje strekk til

Sjølv om dei utarbeidde reglane er detaljerte, kjem ein likevel bort i tilfelle der reglane ikkje strekk til, anten fordi det ikkje finst nokon regel for dette, eller fordi tilfellet er eit grensetilfelle der vi er usikre på kva vi skal gjere. Er du usikker, skriv ned problemet i loggfila (sjå 1.4) og ta det opp på neste transkripsjonsmøte. Loggfør heller eitt problem for mykje enn eitt for lite. Det er viktig at vi blir einige og løyser vanskelege tilfelle på same måte.

### 3.2 Inndeling av talen i turar og segment

Når vi transkriberer, må talestraumen delast opp på ein meiningsfull måte. Resultatet kan sjå ut som replikkvekslinga i eit filmmanus. Det er fleire måtar å gjere det på. I transkripsjonsprogram som ELAN, er lyden representert ved ei bølgeform og dei ulike talarane har kvar sin *tier* (lag) der vi skriv replikkane. Inndelinga kan derfor bli ganske enkel, som til dømes her (i Annotation Mode):



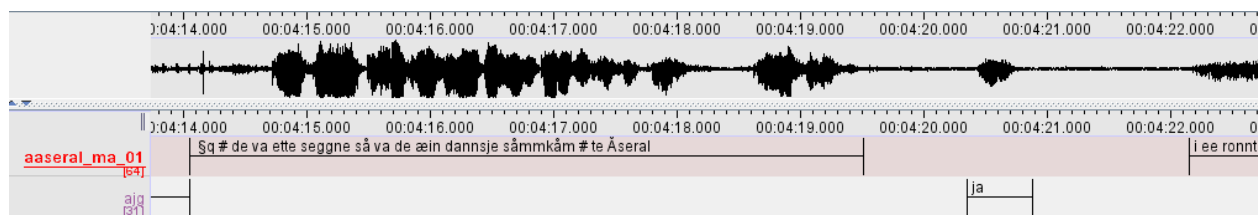
Vi må likevel ha nokre reglar for korleis dette skal gjerast, og vi startar med å forklare nokre grunnbegrep.

#### 3.2.1 Nokre grunnbegrep: Tur, replikk og ikkje-replikk

Omgrepet *tur* blir ofte brukt for ein sekvens der ein person har ordet. Dette omgrepet vil vi også bruke i denne rettleiinga saman med omgrepa *replikk*, *ikkje-replikk* og *segment*. Framstillinga er inspirert av Linell og Gustavsson (1987).

Linell og Gustavsson definerer *tur* som ein periode då ein person har ordet. Å ha ordet inneber at ein disponerer talerommet og har rett og eventuell plikt til å ytre seg. Ein kan tenke seg samtalen som ein situasjon der samtalepartnarane har ein stafettpinne dei sender fram og tilbake mellom seg. Samtalepartnaren som har pinnen, har også turen. Det skal litt til før ein samtalepartnar gir frå seg pinnen – den andre samtalepartnaren kan kanskje komme med små utrop eller kommentarar til det den første seier, men stafettpinnen vil framleis vere hos førstemann. Vi kan også sjå for oss at ein samtalepartnar avsluttar sin tur og legg stafettpinnen på bordet utan at den andre tar han opp. Så kjem kanskje ein lang pause før den første tar pinnen opp igjen og byrjar å snakke om noko anna. Ein slik sekvens kan reknast for to turar, sjølv om det er den same informanten som snakkar.

Turskifte oppstår altså når talar A ved pause eller intonasjon signaliserer at det er B sin tur, og B tar ordet. Det er også turskifte dersom B greier å overta turen ved å snakke så høgt eller overtydande at A til slutt sluttar å snakke.



Ein *replikk* er ei ytring som blir sagt av den som har turen. Ein *ikkje-replikk* blir sagt av den som ikkje har turen. Ein *ikkje-replikk* påverkar verken turhavaren sin tale i innhaldsmessig retning eller turhavaren sin tale med omsyn til aspekt på emne. Ein *ikkje-replikk* kan vere av to typar:

- *Feedback*: Talar B signaliserer at ho er einig, forstår og at A skal halde fram. Typisk feedback er ytringar som *ja*, *mm*, *oi*, latter osv.
- *Mislukka avbrottsforsøk*: B prøver å avbryte A, men mislukkast slik at A beheld turen. (Dersom B lukkast og "vinn turen", er det ein replikk.)

Andre kjenneteikn på *ikkje-replikkar*:

- ikkje svar på spørsmål (svar på spørsmål er *replikk*)
- kort latenstid (pause) eller overlapp mellom A og B sine ytringar
- svak lydstyrke eller lite omfang
- ofte omformuleringar eller korte spørsmål

I somme transkriberingsprogram er det ulike funksjonar for å markere turar og ikkje-replikkar. Slik er det ikkje i ELAN, der all tale blir markert på same måte. Som regel vil vi i ELAN markere ein tur med eit samanhengande replikkfelt i tieren for talaren under bølgefeltet. Om turen eller replikken er lang, kan han delast opp i fleire segment, og det kan gjerne vere mellomrom mellom desse segmenta om talaran tar korte pausar. Sjå også kapittel 3.2.2 og 3.4.3.3 nedanfor. Merk likevel at eitt segment aldri skal innehalde to turar. Turskifte skal alltid markerast med nytt segment.



Ikkje-replikkar markerer vi som oftast i ELAN som korte segment på tieren til den aktuelle talaren akkurat der ikkje-replikken vert yttra. Stille mellom ikkje-replikkane skal ikkje vere ein del av segmentet.



At samtalepartnarane snakkar i munnen på kvarandre, er svært vanleg i naturleg samtale. Ofte er det slik at B kjem med overlappende feedback samtidig med at A snakkar i sin tur. Overlappende tale kan også vere mislukka eller vellukka forsøk frå B si side på å ta turen. I ELAN markerer vi overlappende tale ganske enkelt ved at replikkfeltet på tierane til talarane overlappar, jamfør biletet ovanfor.



### 3.2.2 Inndeling av turar i segment



Ein *tur* (eller *replikk*) slik han er definert i førre kapittel, kan vere anten ganske kort eller temmeleg lang. I eit transkribert talespråkskorpus vil vi helst verken ha for korte eller for lange replikkar. For det første ønskjer vi at transkripsjonen skal vere delt opp i einingar som er meningsfulle å

ordklassetage med eit automatisk annoteringsverktøy. Det vil seie at vi ønskjer einingar som liknar mest mogleg på setningar – sjølv om vi ikkje har definerte setningar i talemål slik vi finn dei i skriftmål. For det andre ønskjer vi einingar som gir oss meningsfulle resultat når vi søker i transkripsjonane via eit søkjeverktøy som for eksempel *Glossa*, sjå biletet nedanfor. Glossa hentar ei og ei eining der søkjeordet inngår. Er eininga for lang, blir det mykje tekst å lese og høyre gjennom, er ho for kort, blir det liten kontekst.

  aal\_ma\_01 mått slik måtte me stå ennt'n de va # reggnver ell de va # sosjinnver å reggnver va re nusstre å va re sosjinn # so villde nesst'n kryture springge ifrå åss fårr # då sjena ræ  
[translate]

  aasnes\_ma\_02 så vi hadde nåe å sto åss i å tæll vi komm tæll a bæssmor sætærn  
[translate]

  alvdal\_02uk åss åss'n da lissåmm # ell sånn eksempel ...  
[translate]

  alvdal\_02uk næi # så re e korrt t stort sett korte avstanner mellomm åss alle hær

Linell og Gustavsson (1987) bruker omgrepet *meningseining* for den minste eininga ein samtale skal delast opp i. Vi har valt å kalle meningseiningane for *segment*. Eit segment er ei formelt og/eller semantisk avgrensa eining som utgjer ei ytring eller ein del av ei ytring. Vi kan bruke følgjande kriterium for dele replikkane inn i segment:

- Intonasjonskriterium (fallande, stigande og fortsettande tonefall)
- Avslutning av grammatisk eining (heilsetning, delsetning)

Kor langt eit segment bør vere, vil variere noko. Ein kan tenke seg eit ”idealsegment”, som varer i underkant av 10 sekund og inneheld ei passe lang meningseining. Dette vil sjølvstøtt ofte vere umogleg å få til, ettersom informantane snakkar med ulik fart. Likevel kan det vere nyttig å ha ”idealsegmentet” i bakhovudet som ein peikepinn.

### 3.3 Taggar til markering av ekstrainformasjon og ikkje-språklege lydar

Ein normal samtale inneheld ofte mange ikkje-språklige element som snufsing, hosting, latter osv. Samtalen kan også innehalde informasjon som ikkje bør transkriberast (som innleiande kommentarar eller sensitiv informasjon), og enkelte gonger kan det vere nødvendig å markere eitt eller fleire ord som utydelege eller uforståelege. ELAN har (i motsetnad til t.d. Transcriber) ikkje støtte for taggar eller hurtigkommandoar inne i sjølve transkripsjonen, og vi har difor laga eit eige system som bør vere raskt og enkelt å bruke for å tagge teksten fortløpande.

#### 3.3.1 Generelle prinsipp for taggar

- + Plussteiknet og bokstaven etter karakteriserer ordet eller ordgruppa som kjem etter. Dette kallar vi *avhengige* taggar.  
Døme:

+u han

→ transkribøren meiner «han» er uklart

- ( ) Vi set vanlege parentesar rundt ei ordgruppe som skal karakteriserast med ein +-tagg. Det skal ikkje vere mellomrom mellom taggen og parentesen.

Døme:

```
+u(han er)
```

→ transkribøren meiner «han er» er uklart

- { } Kommenterar markerer vi med krøllparentesar (AltGr+7 og AltGr+0), sjå også avsnitt 3.3.3.

Døme:

```
{gaskkalduvvan}  
{álggahan kommentárat}  
{sensitiiva persondieđut}  
{čoalkkuha beavddi}  
{informánttas leat bamsemumsat dievva njálmmi}
```

- % Uavhengige taggar markerer vi med %, og dei representerer ein sjølvstendig sekvens i talestraumen som ikkje kan uttrykkjast med skrift, for eksempel latter, hosting osv.

Døme:

```
det var moro %l
```

→ %l tyder at talaren ler etter at han har sagt «de va morro»

Fleire avhengige taggar kan stå til same ordet eller same uttrykket innanfor parentesar, rekkjefølgja på taggane er uvesentleg.

Døme:

```
+l +u(bla bla)
```

→ det som er inne i parentesen er utydeleg og blir sagt leande

Parentesuttrykk innafor parentesar er ikkje mogleg. Dersom fleire ord innafor ein parentes skal merkast, må kvart ord merkast separat (a), eller eventuelt må heile uttrykket merkast som (b).

```
(a) +l (bla +u bla +u bla)
```

```
(b) +l +u(bla bla bla)
```



### 3.3.2 Liste over avhengige og uavhengige taggar med tyding

<i>Avhengige</i>		<i>Uavhengige</i>	
+x			dialektord som ikke står i ordboka, usikre ord/ord det er tvil om
+h			fremmedspråklige ord tilpasset samisk (heivehuvvon)
+n			nordiske lån, og også lån fra engelsk og andre fremmespråk
+u	uklart	%u	uforståeleg
+l	leande	%l	latter
+g	gjespande	%g	gjesp
+v	kviskrande	%v	kviskring
+y	syngande	%y	Sang
		%k	kremting
+s	sukkande/stønnande	%s	stønn/sukk
		%q	(meningsbærande) klikkelyd
		%o	lydmalande ord (onomatopoetikon)
		%p	plystrande

Merk at vi kan bruke dei same taggane for avhengige og uavhengige taggar med ulikt symbol føre. I (a) blir heile ytringa blir sagt gjespande, mens i (b) kjem gjespen som ei enkeltstående hending etter at ytringa er sagt:

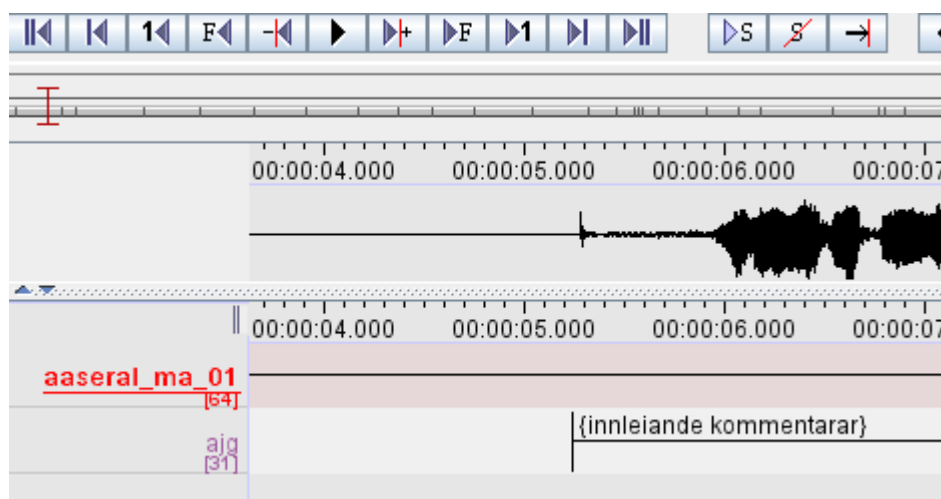
(a) +g(jeg er trøtt)

(b) jeg er trøtt %g

### 3.3.3 Sensitive personopplysningar og annan informasjon som ikkje skal med

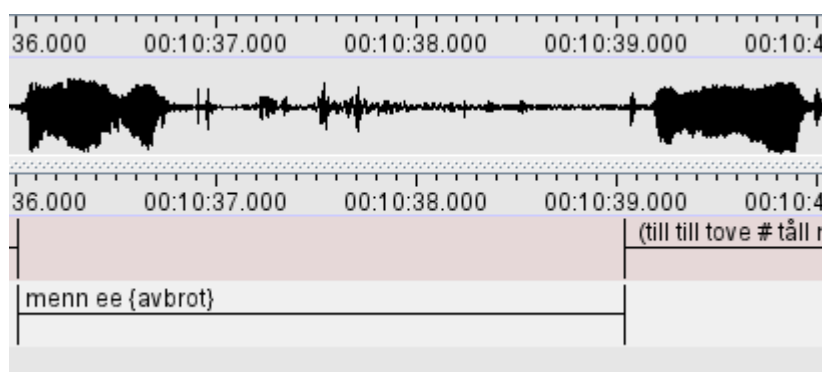
Følgjande skal ikkje transkriberast:

- Eventuelt småprat om opptaket/opptakssituasjonen før samtalen eller intervjuet er i gang. Dette blir berre transkribert med *{álggahan kommentárat}* (pá intervjuaren sin tier):



Vi bruker også *{loahpahan kommentárat}* p  same m te, dersom opptaket held fram etter at samtalen eller intervjuet er avslutta.

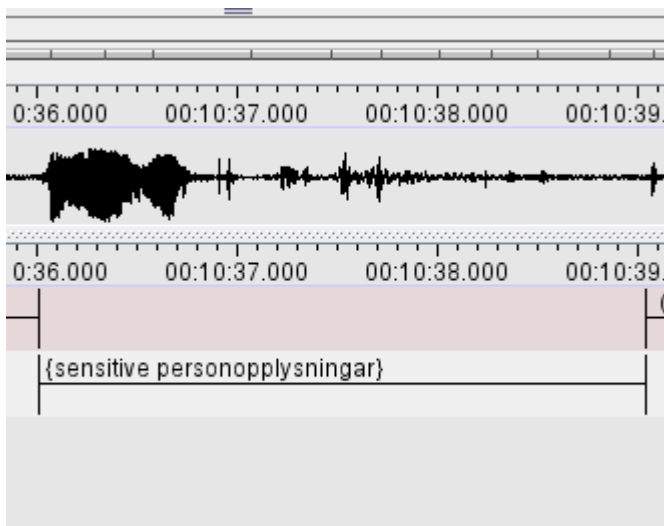
- Avbrot i samtalen som for eksempel kjem av at mikrofonen fell av, at nokon kjem inn i opptaksrommet og avbryt samtalen osv. Slike avbrot er stort sett uinteressante   ha med i transkripsjonen. I staden kan vi markere med ein *{gaskkalduvvan}*-kommentar:



Avbrot, innleiande og avsluttande kommentarar skal st  i tieren til intervjuaren. Om det ikkje er ein intervjuar i opptaket, set vi berre desse kommentarane i det  vste tieren i fila.

- Sekvensar som inneheld sensitive personopplysningar. Dette er if lge NSD (Norsk samfunnsvitenskapelig datatjeneste) f lgjande: *opplysningar om rasemessig eller etnisk bakgrunn, politisk, filosofisk eller religi s oppfatning, at ein person har vore mistenkt, sikta, tiltalt eller d mt for ei straffbar handling, helseforhold, seksuelle forhold, medlemskap i fagforeiningar.*

Sj lv om opptaka er gamle, skal vi vere varsame med   skrive inn sensitive personopplysningar. Dersom opptaket inneheld slike sekvensar, skal dei alt  ikkje transkriberast. Vi set heller kommentaren *{sensitiiva persondie ut}* i tieren til den aktuelle talaren:



- Namn på vener, familiemedlemmer, kollegaer osv. Vi har egne reglar for korleis vi skriv namn, sjå 3.4.1.

### 3.4 Språkform

I LIA har vi valt ortografisk transkripsjon av det samiske materialet. Sjå den samiske rettleiinga.

#### 3.4.1 Namn

Vi skriv vanlegvis ikkje inn namn på personar. I staden skriv vi kodar for kvart enkelt namn, *F* for jentenamn (fornamn eller fornamn pluss etternamn), *M* for gutenamn (fornamn eller fornamn pluss etternamn), og *E* for etternamn (utan fornamn) og namn vi er usikre på om tilhøyrrer gut eller jente. Namna får fortløpande nummer etter kvart som dei dukkar opp i teksten:

**Informanten seier:**

**jeg så at Kari Hansen ga Nina dokumentene til Nilsen**

**Vi skriv:**

**jeg så at F1 ga F2 dokumentene til E1**

Namn på offentlege personar treng vi som regel ikkje anonymisere.

Stadnamn, fotballag, namn på kjæledyr og organisasjonsnamn m.m. skriv ein normalt inn dersom ein ikkje vurderer at det er sensitive personopplysningar involvert, sjå 3.3.3. Bør slike namn anonymiserast, skriv vi *N1*, *N2* osv: "eg bur i *N1*", "jeg går på *N1* skole".

Namn på songar, filmar og bøker skriv vi normalt med stor bokstav. Namn på verk skriv vi i hermeteikn dersom det består av fleire ord. Namn på leikar som *sisten*, skriv vi med liten bokstav. Dersom slike namn består av meir enn eitt ord, set vi også hermeteikn rundt for å synleggjere at orda

høyrrer saman: ”boksen går”, ”slå på ring”, ”jeppe pinne”.

Dersom ein av samtalepartnarane bruker namnet på den andre samtalepartnaren, eller sitt eige namn (t.d. snakkar om seg sjølv i tredje person), bruker vi informantkodane (*gausdal\_uio\_0101* osv.) i staden for dei vanlege *F* og *M*. På same måte bruker vi intervjuarkoden om intervjuaren blir nemnd ved namn.

g

### 3.4.2 Ikkje-språklege lydar

Mange talarar lagar mange ikkje-språklege lydar når dei snakkar, og det er lett å henge seg opp i dei og bruke mykje tid på å transkribere dei nøyaktig. Lydane vi er mest interesserte i, er dei som potensielt gir meining. Latter skapar nesten alltid meining, men eit ukontrollert host gjer det ikkje. Likevel kan eit host eller krent til dømes vere ei type nøling som det er viktig å få med i transkripsjonen, men er ein informant forkjølt, er det ikkje nødvendig å ta med kvart host eller snufs. Utfordringa blir altså å få med ekstrainformasjon som har noko å seie for meininga i ytringa. Vi taggar berre dei ikkje-språklege lydane vi har definert taggar for i denne rettleiinga, sjå oversikta i avsnitt 3.3.2.

Døme: ”+g(*mun lean jo nu váiban*)”, ”%l *dat lei somá*”

### 3.4.3 Avbrot, pausar og uklare passasjar

#### 3.4.3.1 Avbrotne ord

Når eit ord blir avbrote, skal dei delane som er uttala, likevel transkriberast. Vi markerer at ordet blir avbrote ved hjelp av ein bindestrek (-), som vi skriv utan mellomrom rett etter den siste bokstaven, til dømes *hann viss-* eller *hann s-*.

Dersom ein person stammar eller nøler inne i ordet (*inne-ba-bane*, *inne- m -bane*), skriv vi det på vanleg måte: *innebane*. Ein anna variant av dette er når første delen av ei samansetting blir gjentekte fleire gonger, kanskje ikkje som avbrot, men som forsterking. Dette skriv vi slik: *kjemmpe-kjemmpe-kjemmpedæili*. Dersom personen derimot startar på eit ord og rettar seg sjølv inne i ordet, skriv vi det slik: *nitt'nsekks- søttiátte*.

Når det dreier seg om nøling eller stamming i starten av eit ord, markerer vi dette dersom det dreier seg om meiningsberande informasjon, men dersom informanten t.d. stammar mykje, markerer vi det ikkje (*hu- husska*, *frædd- fræddan*). Det kan vere vanskeleg å transkribere den eksakte lydkombinasjonen i slik nøling. Enkelte gonger vil nøling i starten av ord difor markerast med nølelyden *ee* eller med taggen %s for stønning, andre gonger kan ein velje å markere nølinga med taggen %u.

#### 3.4.3.2 Uavslutta ytringar

Dersom ei ytring ikkje blir skikkelig avslutta, og ein annen taler tek over, for eksempel ved avbrot, skal den avbrotne ytringa avsluttast med tre punktum med mellomrom føre. Dette gjeld også når ein talar prøver å overta turen, men blir avbroten og mislukkast. Somme tider må vi bruke både bindestrek for avbrot og tre punktum:

høres ut som sånn her ...  
høres ut som sånn he- ...

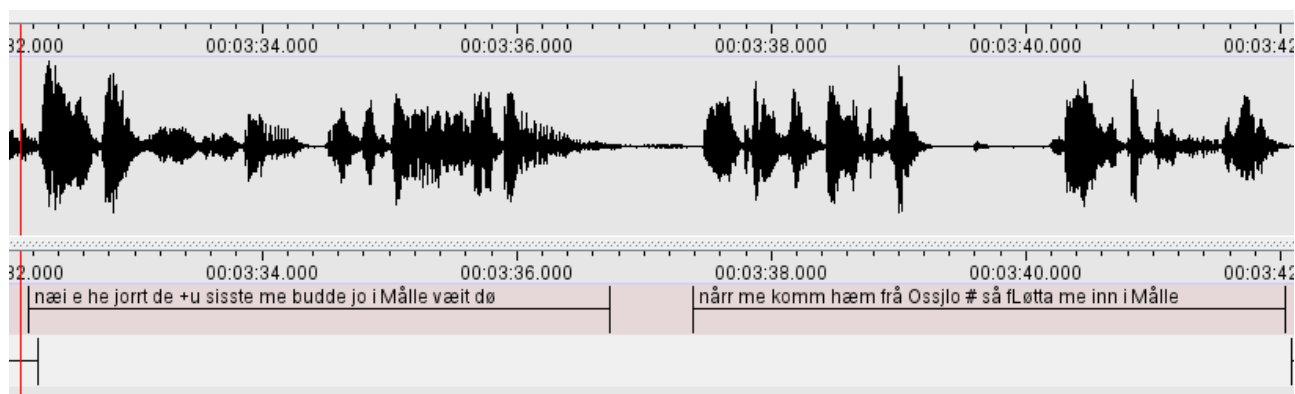
Dersom ein talar avbryt seg sjølv og held fram med ei anna ytring, skal vi ikkje skrive tre punktum. Skriv den påbegynte ytringa og hald fortløpande fram med den neste:

du har du har de er ikke sånn

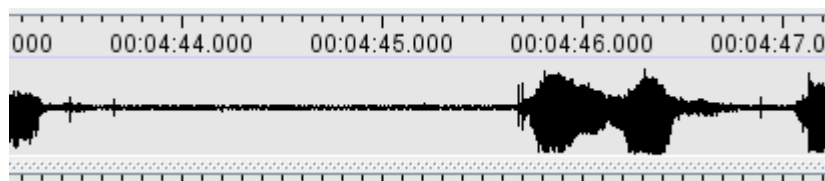
Dersom det er pausar i ytringa, sjå neste avsnitt.

### 3.4.3.3 Pausar

Pausar kan opptre innanfor eller mellom segment (3.2.2). Innanfor eitt segment blir korte pausar markerte med #, og mellomlange pausar med ##. Pausar mellom segment blir markert med avstand/opphald mellom segmenta. I eksempelet under ser vi to pausar, ein mellom to segment og ein innanfor eit segment:



Ein mellomlang pause skal vere så lang at det i praksis ville vore tid for ei knappenål å falle i golvet eller lengre (frå ca 1,5 sekund). Ein kort pause er kortare. Det kan vere vanskeleg å avgjere når ein skal markere korte pausar. Ein hugseregel kan vere at det skal vere eit markert opphald utan lyd. Ofte vil ein pause kunne sjåast som ein rett strek på i bølgeforma. Det er likevel verdt å merke seg at ein rett strek ikkje alltid vil vere ein pause. Enkelte konsonantlydar vil ikkje gi utslag i bølgeforma. Nedanfor er først ein lang og så ein kort pause:



eg ## kvila # mang ein god dag # å la han

Ekstra lange pausar er ikkje så vanleg, men vi kan definere ein ekstra lang pause som ein pause som er så lang at samtalepartnerane ville synest det begynte å bli pinlig stille. Slike lange pausar ser vi i ELAN oftast mellom segment, og blir altså markert med opphald mellom segment snarare enn #-ar. Den mellomlange pausen er den med størst variasjon, og kan vare ganske kort eller ganske lenge.

Dersom informanten pustar, vil vi av og til sjå dette som lyd i bølgeforma. Dersom det ikkje dreier seg om ikkje-språklege lydar som **latter** eller **hosting**, vil vi markere pusting som pausar (vi bruker ikkje nokon tagg for **pusting**). Ulike typar nøling som *ee* osb. markerer vi med ord frå lista i 5.3. Hosting, snufsing og klikkelydar markerer vi med tagg (sjå 3.3.2) når dei skapar meining i samtalen. Elles markerer vi dei ikkje. At informanten drar ut vokalane, som i *såååå de ææææ sjlik* markerer vi ikkje, heller ikkje andre former for pausar og nøling inne i ord som ved *inne-ba-bane* (sjå 3.4.3.1).

#### 3.4.3.4 Passasjar som er vanskelege å forstå

*Uforståeleg*: Skulle det vere lydkombinasjonar som du ikkje greier å forstå, let du vere å transkribere akkurat dette, og set inn ein *uforståeleg*-tagg i staden:

```
men takk for at du %u
```

Det er viktig at det du transkriberer, gir meining. Dersom du trur du høyrer alle lydane rett, men ikkje greier å forstå kva som blir sagt, skal du altså setje inn ein *uforståeleg*-tagg i staden for å prøve å transkribere det. Det vil føre til problem i t.d. omsetjinga til nynorsk dersom transkripsjonen inneheld passasjar som ikkje gir meining.

*Uklart*: Skulle det vere passasjar som du trur du forstår, men som du ikkje er sikker på, set du inn ein *uklart*-tagg:

```
han sa at +u(han visste) at de skulle komme +u
```

Dersom berre første delen av eit samansett ord er utydeleg, skriv du anten at heile ordet er uforståeleg eller du gjettar kva første delen skal vere og sett utydelegtagg på heile ordet.

#### 3.4.4 Stor bokstav og teiknsetjing

Vi bruker stor bokstav i namn. Vi bruker spørjeteikn i spørsmål og hermeteikn i mellom anna sitat. Semikolon, kolon, komma, utropsteikn og punktum bruker vi ikkje. Under har vi lista opp reglane for når stor bokstav, spørjeteikn og hermeteikn skal brukast. Det er viktig at vi ikkje bruker teiknsetjing til anna enn det som er nemnt i dette avsnittet.

##### 3.4.4.1 Stor bokstav

Stor forbokstav bruker vi ikkje i starten av ei setning, men i namn. Grensa mellom kva som er namn og kva som er vanleg substantiv er flytande. Namn på snop kan til dømes vere problematisk! Dersom ordet står i ordboka i tydinga snop (*seigmann*), skriv vi det med liten bokstav. Viss ikkje, skriv vi produktnamnet med stor bokstav (*Hobby*).

Når namn er bøygde, skriver vi dei framleis med stor bokstav og utan bindestrek.

```
gi meg en Seigmann  
nei men du kan få en Hobby hvis du gir meg Nokiaen min
```

### 3.4.4.2 Spørjeteikn

Spørjeteikn bruker vi som vanleg ved spørsmål, også dersom spørsmålet blir avbrote. (Skriv ... ?) Ved spørsmål ein ikkje ventar svar på (*de ær sjlik ikke sannt de bare blir sånn*), kan spørjeteikn droppast.

NB: Vi bruker mellomrom mellom spørjeteiknet og ordet som kjem føre!

### 3.4.4.3 Hermeteikn

Sitat set vi i hermeteikn. Dette gjeld også hypotetiske sitat:

da kjørte jeg den "jeg ? hæ ?" da juger jeg

Dersom ein informant snakkar metaspråkleg om uttalen av eit ord, eller om han les noko høgt, bruker vi også hermeteikn:

jeg sier ikke "snø" jeg sier "sne"

Merk at vi ikkje bruker hermeteikn dersom det dreier seg om vanlege forkortingar, produktnamn eller firmanamn med forkortingar (DVD blir devede A/S blir aess, NHH blir Ennhåå, AG3 blir agetre).

## 4 Korrektur

### 4.1 Lese korrektur

Alle filene skal korrekturlesast av ein annan transkribør enn den som opphavleg transkriberte fila.

Når du les korrektur, spelar du av korte sekvensar om gongen. Følg godt med, og spel uklare passasjar om att. Finn du feil i transkripsjonen, rettar du dei opp. Er du i tvil om noko, noterer du fila, tidskoden og problemet i problemlaggen (sjå 1.4).

## 5 Oversikt over hurtigkommandoar, eventar og ikkje-språklege lydar

### 5.1 Oversikt over hurtigkommandoar

#### Segmentation Mode

Funksjon	PC og Mac
Slå saman (utheva segment) med neste segment	Ctrl+A
Slå saman (utheva segment) med førre segment	Ctrl+B
Del opp (utheva segment)	Ctrl+Enter <sup>3</sup>
Gjer tieren over aktiv	Pil opp
Gjer tieren under aktiv	Pil ned
Spel/Pause	Ctrl+Mellomrom
Spel markert område	Shift+Mellomrom

#### Transcription Mode

Funksjon	PC og Mac
Hopp ned til neste segment	Enter/Alt+Pil ned
Hopp opp til forrige segment	Alt+Pil opp
Spel/Pause	Tab
Spel segmentet om igjen fra start	Shift+Tab

#### Annotation Mode

Funksjon	PC og Mac
Spel/Pause	Ctrl+Mellomrom
Spel markert område	Shift+Mellomrom
Ny annotation her	Alt+N
Slett annotation	Alt+D
Gå til annotation over	Ctrl+pil opp
Gå til annotation under	Ctrl+pil ned
Gå til annotation til høgre og rediger	Ctrl+Alt+pil høgre
Gå til annotation til venstre og rediger	Ctrl+Alt+pil venstre
Flytt venstre segmenteringsgrensa eit hakk til venstre	Ctrl+J
Flytt venstre segmenteringsgrensa eit hakk til høgre	Ctrl+U
Flytt høgre segmenteringsgrensa eit hakk til venstre	Ctrl+Shift+J
Flytt høgre segmenteringsgrensa eit hakk til høgre	Ctrl+Shift+U

<sup>3</sup> Merk at denne snarvegen ikkje er førehandsdefinert. I kapittel 7.1 kan du lese om korleis du definerer egne snarvegar.



## Generelt: Flytting av markøren

Funksjon	PC	Mac
Flytt markøren eit sekund til venstre	Shift+Pil venstre	Shift+Pil venstre
Flytt markøren eit lite hakk til venstre	Ctrl+Shift+Pil venstre	Cmd+Shift+Pil venstre
Flytt markøren eit lite hakk til høgre	Ctrl+Shift+Pil høgre	Cmd+Shift+Pil høyre
Flytt markøren eit sekund til høgre	Shift+Pil høgre	Shift+Pil høyre

## 5.2 Oversikt over taggar

<i>Avhengige</i>		<i>Uavhengige</i>	
+x			dialektord som ikke står i ordboka, usikre ord/ord det er tvil om
+h			fremmedspråklige ord tilpasset samisk (heivehuvvon)
+n			nordiske lån, og også lån fra engelsk og andre fremmespråk
+u	uklart	%u	uforståeleg
+l	leande	%l	latter
+g	gjespande	%g	gjesp
+v	kviskrande	%v	kviskring
+y	syngande	%y	Sang
		%k	kremting
+s	sukkande/stønnande	%s	stønn/sukk
		%q	(meningsbærande) klikkelyd
		%o	lydmalande ord (onomatopoetikon)
		%p	plystrande

### 5.3 Liste over interjeksjonar

Interjeksjonar vi ikkje endrar stavemåten på:

<i>ee</i>	(nøling - uansett lengde på <i>ee</i> )
<i>eh</i>	(avstandsindikerande)
<i>ehe</i>	("Eg forstår" - to stavingar)
<i>em</i>	(nøling)
<i>eh</i>	(imponert)
<i>hm</i>	(spørjande, undrande. I Nynorskordboka i tydinga kremting)
<i>m</i>	(nøling, ta til etterretning, nam)
<i>m-m</i>	(benektande)
<i>mhm</i>	("Eg forstår" - to stavingar)
<i>mm</i>	(bekreftande - to stavingar)

Interjeksjonar vi kan tilpasse informanten sin uttale:

<i>aha</i>	(overraska) Nynorskordboka
<i>jaha</i>	(forsterkande "ja") Nynorskordboka
<i>nja</i>	(tvilande) Nynorskordboka
<i>ops</i>	(oi då)
<i>u</i>	(imponert)

Samiske:

<i>áá</i>	(konstaterande - to stavingar)
<i>oh</i>	(utrop)
<i>o já</i>	(overraska)
<i>o-o</i>	("oj")
<i>gea (h)</i>	(forundring)
<i>há</i>	(spørjande)
<i>hei, he-hei</i>	(forundring, ergrelse)
<i>na</i>	(tvil - uansett lengde på <i>a-en</i> , bekræftende)
<i>naa</i>	(bekræftende - to stavingar)

## 6 Referansar

Hagen, Kristin. 2005. Transkripsjonsveiledning for NoTa-Oslo. Versjon 20.12.2005:  
<http://www.tekstlab.uio.no/nota/oslo/transkripsjon/NoTa-transkripsjonsveil7.pdf>

Hanssen, E. et al. 1978. *Oslomål. Prosjektbeskrivelse og syntaktisk analyse av oslomål med henblikk på sosiale skilnader. Talemålsundersøkelsen i Oslo (TAUS)*. Hovedrapport. Oslo, Novus forlag.

Linell, Per og Gustavsson, Lennart. 1987. *Initiativ och respons. Om dialogens dynamikk, dominans og koherens*. SIC 15, Universitetet i Linköping.

Papazian, Eric og Helleland, Botolv. 2005. *Norsk talemål. Lokal og sosial variasjon*. Høyskoleforlaget.

## 7 Vedlegg

### 7.1 Når du begynner å jobbe på ei ny maskin

**Montere nota-området** (der alle transkribørmapper, lydfiler og slikt ligg)

- Klikk på "Map network drive" i Computer.
- skriv in \\hf-tekstlab\nota i vindaugget som dukkar opp, og hak av for "reconnect at logon".
- Klikk "Finish".

**Lage snarveg til ELAN på "taskbaren"**

- Gå til Computer -> Local Disk (C:) -> Program Files og opne mapa som heiter ELAN (vel den nyaste versjonen om det er fleire å velje i).
- Dra fila som heiter ELAN.exe til taskbaren.

**Fjern skjermbildesnuing i Windows:**

Ctrl + Alt + F12.

Vel advanced mode --> options and support --> ikkje kryss av for Hot Key Functionality

**Dersom du opplever at språket endrar seg frå norsk til engelsk (eller andre språk) medan du bruker ELAN:**

Gå til kontrollpanelet. Klikk på «Change keyboards or other input methods». Klikk på «Change keyboards». Under «General» kan du fjerne dei språka du ikkje treng. Dersom berre norsk blir ståande att, blir problemet med språkskifte borte. Du kan òg gå til «Advanced Key settings» og fjerne hurtigtasten (Left Alt+Shift) som endrar språket.

**Definere egne snarvegar:**

I ELAN er det mogleg å definere egne tastatursnarvegar for gitte funksjonar. Går du inn på Edit -> Preferences -> Edit shortcuts, får du opp ei liste med snarvegane du kan velje egne hurtigtastar til. Her kan du setje Ctrl + Enter til å dele opp utheva segment (Split annotation), slik som det står i lista over snarvegane i transkripsjonsretteiinga.

## 7.2 Hugseliste for transkripsjon

1. Finn ei lydfil i mappa «Lydfiler».
2. Flytt lydfila til di eiga mappe. Namnet på lydfila skal ikkje endrast!
3. Start transkripsjonen etter oppskrifta i kapittel 2.2. Det er lurt å lagre med det same du startar å transkribere. Transkripsjonsfila får namn etter lydfila, med dine eigne initialar bakerst: *brandbu\_uio\_04\_lh*.
4. Gå inn på «LIA-filer» i Google Docs:  
[https://docs.google.com/spreadsheets/d/1vIqvjpw1\\_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846](https://docs.google.com/spreadsheets/d/1vIqvjpw1_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846)  
(Du finn òg ein snarveg til denne fila i mappa «Transkripsjon».) Finn det rette filnamnet (kvart fylke har si fane) og skriv namnet ditt i den tilhøyrande kolonnen "Transkribør". Fyll inn informantkodar etterkvart som du transkriberer (mange filer har berre éin informant, medan ein del har fleire).
5. Når du er ferdig med fila, set du ein X i kolonnen «Ferdig transkribert».
6. Flytt filene (både lydfil og transkripsjonssfil) til mappa «Til korrektur».

### 7.3 Hugseliste for korrektur

1. Finn ei fil du vil korrekturlese i mappa "Til korrektur".
2. Flytt fila og alle tilknytte filer (lydfil, txt-fil) over i di eiga mappe.
3. Legg til dine eigne initialar bakarst i filnamnet på transkripsjonsfila:

*brandbu\_uio\_03\_lh\_et*

Det er berre namnet på transkripsjonsfila du skal endre. Lydfiler skal aldri endrast!

4. Når du er ferdig med fila, legg henne (og alle tilknytte filer) i mappa «Korrekturlest» (felles mappe for alle transkribørane).
5. Gå inn på Gå inn på «LIA-filer» i Google Docs:  
[https://docs.google.com/spreadsheets/d/1vIqvjpw1\\_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846](https://docs.google.com/spreadsheets/d/1vIqvjpw1_1kj4tSNTZMo12MHzSmaUVq4qVim4TfiCm4/edit#gid=1993453846)

(Du finn òg ein snarveg til denne fila i mappa «Transkripsjon».) Finn det rette filnamnet (kvart fylke har si fane) og set ein X i kolonnen «Ferdig korrekturlese».