

Transkripsjonsrettleiing for ScanDiaSyn

31.03.2009

Av Janne Bondi Johannessen, Kristin Hagen, Live Håberg, Signe Laake, Åshild Søfteland og Øystein Vangsnes ¹

Innhald

0. Om rettleiinga	2
1. Bakgrunn	3
2. Praktisk rettleiing for transkripsjon.....	4
2.1 Gode råd for transkripsjon	4
2.2 Filorganisering	4
2.3 Transkriberingsprogrammet Transcriber.....	5
2.3.1 Starte på ein ny transkripsjon.....	6
2.3.1.1 Vel ei lydfil.....	6
2.3.1.2 Lag report.....	6
2.3.1.3 Definer talarane.....	7
2.3.2 Halde fram med ein påbegynt transkripsjon	7
2.3.3 Tidskoding og segment	7
2.3.4 Avspeling av lydfile.....	7
3. Transkripsjonsreglar.....	8
3.1 Når reglane ikkje strekk til.....	8
3.2 Utval og sensitive personopplysningar	8
3.3 Markering av ekstrainformasjon	9
3.3.1 Instantaneous event, apply to previous word, start/end of event	9
3.3.2 Ekstrainformasjon for intervjuaren	10
3.4 Språkform.....	10
3.4.1 Lydnær transkripsjon.....	10
3.4.1.1 Generelt.....	10
3.4.1.2 Fonologi	10
3.4.1.3 Kvantitet.....	11
3.4.1.4. Konsonantgrupper og diftongar	12
3.4.1.5 Stavingsberande konsonant.....	13
3.4.1.6 Diverse	13
3.4.1.7 Loggføring/informasjonsflyt.....	13
3.4.1.8 Standardortografi	13
3.4.2 Tal.....	14
3.4.3 Namn.....	14
3.4.4 Ikkje-språklege lydar.....	15
3.4.5 Sitat og ord som er stava eller lesne.....	16

¹ Transkripsjonsrettleiinga er eit resultat av samarbeid mellom mange tilsette på ScanDiaSyn gjennom transkripsjonsmøte med spørsmål, diskusjonar og innspel. Ein stor takk til Sine Bjordal, Ingebjørg Før Gjermundsen, Astrid Grov, Margit Ims, Kari Kinn, Ann Helen Lea, Arne Martinus Lindstad, Alexander Lykke, Eirik Olsen, Stig Rognes og Karine Stjernholm.

3.4.5.1	Sitat	16
3.4.5.2	Ord som er stava eller lesne	16
3.4.6	Avbrot, pausar og uklare passasjar	16
3.4.6.1	Avbrotne ord.....	16
3.4.6.2	Uavslutta ytringar.....	17
3.4.6.3	Pausar	17
3.4.6.4	Passasjar som er vanskelege å forstå.....	18
3.4.7	Teiknsetjing og stor bokstav.....	18
3.5	Tur, segment og meiningseining	19
3.5.1	Tur	19
3.5.1.1	Replikk og ikkje-replikk	19
3.5.1.2	Tur og turskifte.....	20
3.5.2	Meiningseining og segment	21
3.5.3	Innskot av ikkje-replikkar, overlapping og delte meiningseiningar.....	22
3.5.3.1	Innskot av ikkje-replikkar	22
3.5.3.2	Ikkje-replikk eller del av replikk?.....	23
3.5.3.3	Overlapping.....	23
3.5.3.4	Delte meiningseiningar	24
3.5.3.5	Skjulte turar.....	25
4.	Korrektur.....	26
4.1	Organisering av filene	26
4.2	Lese korrektur	26
5.	Oversikt over hurtigkommandoar, eventar og ikkje-språklege lydar.....	27
5.1	Hurtigkommandoar	27
5.2	Skjema over eventar.....	28
5.3	Liste over interjeksjonar.....	30
6.	Referansar	31
7.	Vedlegg.....	32
7.1	Når du begynner å jobbe på ei ny maskin	32
7.2	Hugseliste – når ein startar på ein ny transkripsjon:	33
7.3	Hugseliste – når ei fil er ferdig korrekturlese:	34

0. Om rettleiinga

Denne rettleiinga bygger i stor grad på transkripsjonsrettleiinga for NoTa (sjå <http://www.tekstlab.uio.no/nota/oslo/index.html>). Det tyder mellom anna at eksempel frå *Transcriber* i stor grad er henta direkte frå NoTa-rettleiinga, og dermed inneheld mange eksempel med ortografisk transkripsjon der vi i ScanDiaSyn vil transkribere lydnært. Prinsippa for turtaking og liknande er dei same for NoTa og ScanDiaSyn, og det er dette eksempla skal illustrere. For spørsmål om transkripsjonsnormalen må ein slå opp i avsnittet om språkform (3.4).

I avsnitt 1 finn du ein generell introduksjon om transkripsjon. Avsnitt 2 er meir detaljert og praktisk, og gir gode råd om transkripsjon samt ei kort skildring av korleis du bruker transkripsjonsprogrammet *Transcriber*. Avsnitt 3 tek for seg dei konkrete transkripsjonsreglane. Dette avsnittet bør du ha lett tilgjengeleg medan du transkriberer. Avsnitt 4 handlar om korrektur. I avsnitt 5 finn du ei oversikt over hurtigkommandoar og eventar (hendingar), og avsnitt 6 inneheld referansar. I avsnitt 7 finn du nyttige vedlegg.

1. Bakgrunn

For at ein skal kunne søke i et talespråkskorpus, må materialet transkriberast. Det er mange måtar å transkribere eit talespråk på, men ingen av dei er perfekte. Sidan ScanDiaSyn dreier seg om dialektopptak, er det mest naturleg for oss å velje ein lydnær transkripsjon, der ein tek omsyn til uttalen til informantane. Slik vil vi best kunne få fram variasjonane mellom dei ulike dialektane. Ein slik transkripsjon kan likevel også skape vanskar. Vi har bestemt oss for å bruke det norske alfabetet, ikkje IPA eller anna lydskrift. Dette vil gjere transkripsjonen mindre nyansert, men meir leseleg. I tillegg er det mykje arbeid med å utvikle ein transkripsjonsnormal som fungerer. Vi må også ta høgde for at akkurat dei fenomenane som ein bestemt fonolog er interessert i, likevel ikkje er markerte.

Dersom ein bruker berre ein lydnær transkripsjon, vil dette også gjere korpuset vanskeleg å søke i. Ein reint ortografisk transkripsjon vil løyse slike problem. Vi har difor valt å ha ortografisk transkripsjon i tillegg til lydnær transkripsjon i ScanDiaSyn. Vi transkriberer først lydnært, og denne transkripsjonen blir så omsett ved hjelp av eit halvautomatisk transliterator. Omsetjinga vil ta utgangspunkt i bokmålsnormalen, men vi vil framleis behalde den originale syntaksen og dialektord som ikkje let seg omsetje.

2. Praktisk rettleiing for transkripsjon

2.1 Gode råd for transkripsjon

Det er viktig å finne fram til gode rutinar og vanar for transkribering. 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 strekker på beina.
- Arbeid aldri meir enn fire timar av gongen med transkripsjon.
- Lær deg flest mulig hurtigkommandoar. Det sparar tid og ikkje minst hendene dine.
- Høyr på ca 30 sekund av lydfila først for å få ein oversikt over samtalen. Ta deretter for deg 10 og 10 sekund om gongen og transkriber.
- Fokuser på det du høyrer, det er dessverre fort gjort å skrive det ein ventar i staden for det som faktisk blir sagt.

2.2 Filorganisering

Lydfilene får namn etter målepunkta (opptaksstaden). På kvart målepunkt vil det normalt vere fire informantar, ein mannleg og ein kvinneleg informant under 30 år, og ein mannleg og ein kvinneleg informant over 50 år. Vi bruker det same systemet for å gi namn til lydfilene overalt: Den unge mannen får kode *stadnamn_01um*, den unge kvinna *stadnamn_02uk*. På same måte får den eldre mannen kode *stadnamn_03gm*, og kvinna *stadnamn_04gk*. Intervjuaren er representert ved initialar (t.d. amg). Bindestrek i filnamnet indikerer viser kven som er samtalepartnarane i opptaket.

NB! Merk at *æ*, *ø* og *å* blir erstatta med *ae*, *oe* og *aa* i stadsnamnet, men *a*, *e* og *a* i initialane til intervjuaren.

Med intervju og samtalar får vi til saman seks opptak frå kvart målepunkt. Lydfilene frå t.d. Rollag vil då få desse namna:

```
rollag_01um-amg.wav  
rollag_02uk-amg.wav  
rollag_01um-02uk.wav  
rollag_03gm-amg.wav  
rollag_04gk-sl.wav  
rollag_03gm-04gk.wav
```

Lydfilene ligg på ein eigen server – *scandiasyn on 'Samba 3.0.10-uio (hf-lager)'*. Her går du inn i *Scandiasyn – Lydfiler*. I denne mappa har alle transkribørane si eiga mappe. Når du transkriberer, legg du lydfila inn i di eiga mappe. Du skal ikkje flytte fila igjen når du er ferdig. Sidan det samlar seg opp ein del lydfiler i mappa di, kan det etter kvart vere nyttig å organisere filene for å unngå rot. Du kan t.d. lage ei eiga undermappe for filer du er ferdig med å transkribere, og ei for filer som er under arbeid.

Det er viktig at du ikkje skiftar namn på lydfilene, slik som vi gjer med transkripsjonsfilene. Finn du ut at ei lydfil har feil i namnet (feil intervjuar, feilstava stadnamn e.l.), gir du beskjed om dette til

nokon ansvarlege.

Transkripsjonsfilene skal lagrast på ein anna server: *nota on 'Samba 3.0.10-uio (foni)'*. Her går du inn i mappa *Scandiasyn* og lagrar fila i mappa med ditt eige namn. Når du er ferdig å transkribere, skal du flytte transkripsjonsfila til mappa *Til_korrektur* og samstundes slette henne frå di eiga mappe (sjå meir om korrekturlesing og filorganisering i [4.1](#)).

Alle som transkriberer eller les korrektur på ei fil, skal legge til sine eigne initialar etter ein understrek i slutten av filnamnet. Dette kan du lese meir om i [2.3.1.1](#) og [4.1](#).

Kvar transkribør skal ha ei loggfil der arbeidet for kvar arbeidsøkt er loggført. Loggen skal ligge lagra i mappa med transkribøren sitt namn. Malen ligg i mappa *Scandiasyn*, og kan kopierast herfrå. 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.), oppgir du også dette. Skriv også opp tida 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	FIL – FRA TID – TIL TID 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 mykje, kan sakke farten litt. På same måte vil ein som bruker lang tid, og har svært få feil, kanskje kunne arbeide raskare og «slurve» meir, sidan alle filene uansett skal korrekturlesast. Dette er først og fremst meint som hjelp til transkribørane, ikkje overvaking og styring.

Transkribørane bør også ha sin eigen problemlogg der ein noterer alle spørsmål tilknytt problem med transkripsjonen. Når du 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 diskuteras på neste transkripsjonsmøte. I problemlaggen skriv du opp kva fil det gjeld, den nøyaktige tidskoden (slik at ein lett kan finne att staden) og kva problemet gjeld. Loggfilene og problemlaggen har transkribørane sine initialar som fornamn og *logg* og *prob* som etternamn.

2.3 Transkriberingsprogrammet *Transcriber*

Transcriber er eit gratis fransk transkriberingsprogram. Vil du lese meir om det, kan du gå inn på nettsida: <http://trans.sourceforge.net/en/presentation.php>

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 Transcriber på ei ny maskin, er det ein del innstillingar som må ordnast. Du finn eit vedlegg med manual for dette i [7.1](#).

2.3.1 Starte på ein ny transkripsjon

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

2.3.1.1 Vel ei lydfil

Gå inn i *lydfiler*-mappa under *scandiasyn (hf-lager) – Scandiasyn – Lydfiler*. Filene som ikkje ligg i dei namngitte mappene, er ledige. Fila du skal transkribere, legg du i mappa med ditt eige namn. Deretter går du inn i *Transcriber*. Med **<Ctrl-n>** eller **New trans** under **File** opnar du ei ny fil for transkripsjon. Du blir no beden om å velje ei lydfil. Gå til mappa di i *lydfiler*-mappa, og vel den fila du nettopp har lagt her. Lagre fila med ein gong i mappa med ditt namn i *Scandiasyn* ved å bruke **<Ctrl-s>**. Transkripsjonsfila vil automatisk få det same namnet som lydfila, men med **.trs** som etternamn. Du skal legge til dine eigne initialar før **.trs**, etter ein understrek (sjå meir om filorganisering i [2.2](#)).

Namn på lydfil:

```
gausdal_03gm-04gk.wav
```

Namn på transkripsjonsfil:

```
gausdal_03gm-04gk_1h.trs
```

Når du har lagra fila, vel **Edit episode attributes** under **File**. Her skriv du inn namnet ditt (for- og etternamn) under **Transcribers name**.

2.3.1.2 Lag report

Klikk på den raude firkanten der det står *report*. Vel *New Topic* og skriv inn det som passar best av følgjande karakteristikkar:

```
Intervju  
Samtale – familie  
Samtale – venner  
Samtale – bekjente  
Samtale – ukjente
```

Trykk deretter *OK*. Er du usikker på kva slags forhold samtalepartnarane har til kvarandre, kan du gå inn i *Scandiasyn – Informanter* og gå inn i mappa for det aktuelle målepunktet. Under *Skjema B* opnar du skjemaet til ein av dei aktuelle informantane. Her kan du lese kva relasjon vedkommande har til samtalepartnaren sin.

2.3.1.3 Definer talarane

Under **Segmentation** skal du velje **Edit turn attributes**. Du kan også klikke på den grønne firkanten der det står *no speaker*. Vel **Create speaker**. Skriv inn koden til informanten. Koden skal vere den same som i lydfila. Dersom lydfilnamnet er *gausdal_03gm-04gk*, får samtalepartnerane kodane *gausdal_03gm* og *gausdal_04gk*. Dersom det er eit intervju, bruker vi initialane som namn på intervjuaren. Dei andre moglege vala i skjemaet let du vere umarkerte.

Kvar gong ein av informantane tar ordet, trykk **<Ctrl-t>** og vel informant. Vi har valt å bruke omgrepet *tur* for ein slik sekvens der ein informant har ordet (sjå [3.5](#)).

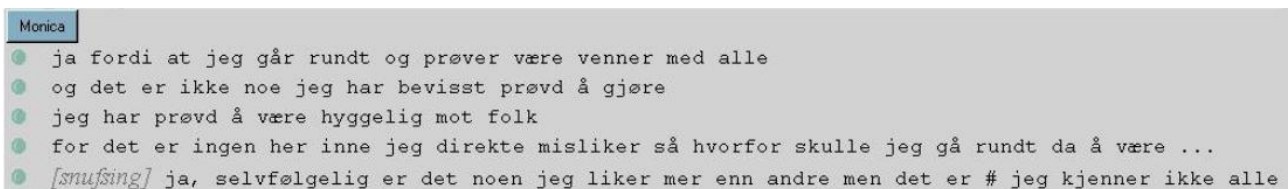
Du kan endre namnet på ein speaker ved å velje **Modify speaker**. Dersom du ved eit uhell lagar ein speaker du ikkje treng, kan du slette denne ved å velje *Edit – Speakers – Remove unused speakers*.

2.3.2 Halde fram med ein påbegynt transkripsjon

Med **<Ctrl-o>** eller **Open trans** under **File** opnar du ei påbegynt transkripsjonsfil. Du kan også gå inn i transkripsjonsmappa di og dobbeltklikke på den fila du vil halde fram med. Når transkripsjonen er opna, vil du bli beden om å opne den tilhøyrande lydfila.

2.3.3 Tidskoding og segment

Ved å trykke **Enter**, blir det automatisk sett inn ein tidskode i dokumentet slik at lydfila blir forbunde med transkripsjonen. Dette blir markert med ein grøn runding. Teksten mellom to tidskodar kallar vi *segment*, sjå [3.5.2](#). Under ser vi ein tur som er delt opp i fem segment:

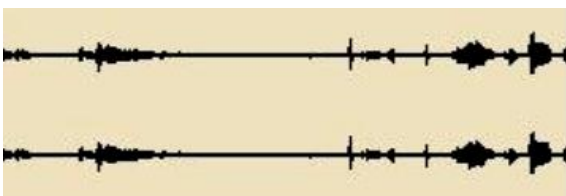


```
Monica
● ja fordi at jeg går rundt og prøver være venner med alle
● og det er ikke noe jeg har bevisst prøvd å gjøre
● jeg har prøvd å være hyggelig mot folk
● for det er ingen her inne jeg direkte misliker så hvorfor skulle jeg gå rundt da å være ...
● [snufsing] ja, selvfølgelig er det noen jeg liker mer enn andre men det er # jeg kjenner ikke alle
```

2.3.4 Avspeling av lydfila

Nedst i Transcriber-vindauget ser du eit spektrogram. Med **Tab** startar og stoppar du avspeling, og kan sjå korleis ein raud strek flyttar seg bortover i spektrogrammet. Spektrogrammet er til hjelp når du skal setje inn ein tidskode.

Dersom du trykker **Shift-Tab**, speler du av det segmentet markøren står på. Ved **Tab** startar avspeling på starten av det segmentet markøren står på, og fortset til han blir stoppa. Du kan også markere eit område i spektrogrammet og høyre på akkurat dette området ved å trykke **Tab**. Under ser du eit lite utsnitt av eit spektrogram:



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, loggfør problemet i loggfila (sjå [2.2](#)) 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 Utval og sensitive personopplysningar

Følgjande skal ikkje med i transkripsjonen:

- Eventuelt småprat om opptaket/opptakssituasjonen før samtalen eller intervjuet er i gang. Skriv ein *innledende bemerkninger*-kommentar <Alt-i>:



Vi bruker også *avsluttende bemerkninger* på same måte, dersom opptaket fortset etter at samtalen eller intervjuet er avslutta.

- Avbrot i samtalen som for eksempel skuldast at mikrofonen fell av, at nokon kjem inn i opptaksrommet og avbryt samtalen osv. Dersom dette er krevjande å transkribere, eller heilt uinteressant å ha med i transkripsjonen, kan vi markere eit avbrot. Skriv *avbrudd*-kommentar <Alt-a> og vel **no speaker** :



- Namn på vener, familiemedlemmer, kollegaer osv. Vi har egne reglar for korleis vi skriv namn, sjå [3.4.3](#).

- Sekvensar som inneheld sensitive personopplysningar. Dette er ifølge NSD (Norsk samfunnsvitenskapelig datatjeneste) følgjande: *opplysningar om rasemessig eller etnisk bakgrunn, politisk, filosofisk eller religiøs oppfatning, at ein person har vore mistenkt, sikta, tiltalt eller dømt for ei straffbar handling, helseforhold, seksuelle forhold, medlemskap i fagforeiningar.*

Dersom opptaket inneheld slike sekvensar, skal det merkast med ein kommentar (<Alt-s>) som fortel at sensitive opplysningar er tekne ut av transkripsjonen:



Informantane skal ha fått beskjed om at dei ikkje skal snakke om forhold som karakteriserast som sensitive personopplysningar, eller i negative ordelag om tredjeperson. Vi får derfor håpe at slikt ikkje dukkar opp så ofte i lydfilene.

3.3 Markering av ekstrainformasjon

All slags ekstrainformasjon som skal markerast i transkripsjonane, må vi gi som *event* (hending). **Insert event** finn du under **Edit**-menyen. Vi skal bruke fire typar eventar:

- **noise** til lydar som hosting, krenting, sukking og klikkelyd. Hurtigkommando <**Ctrl-d**>. To av elementa på denne lista – *pause* og *lattermild* – skal vi ikkje bruke. Dei andre kan vi bruke, men med måte. Hovudregelen er at lydar som er ein del av ei ytring, skal markerast.
- **pronounce** til å markere ei meiningseining (me) som går over fleire segment (sjå [3.5.3.4](#)), eller til å markere når noko t.d. er uklart uttala eller stava i staden for å bli uttalt på vanleg måte. Hurtigkommando <**Ctrl-r**>
- **language** til å merke ord som ikkje står i Bokmålsordboka, t.d. engelske ord. Denne typen event bruker vi berre når vi transkriberer intervjuaren. Hurtigkommando <**Ctrl-l**>
- **comment** til kommentarar som *sensitive personopplysningar* (sjå [3.2](#)) eller *uforståelig* (sjå [3.4.6.4](#))

For mange eventar, mellom anna for alle kommentarane, har vi egne hurtigkommandoar. Ei oversikt over alle eventar og hurtigkommandoar finst i kapittel [5](#).

Er ein informant forkjølt, er det ikkje nødvendig å ta med kvart host eller snufs. Utfordringa blir å ta med ekstrainformasjon som har noko å seie for meininga med ytringa. Eitt host eller krent er til dømes ofte ei type nøling som det er viktig å få med i transkripsjonen.

3.3.1 Instantaneous event, apply to previous word, start/end of event

Det skal alltid vere mellomrom før og etter eventane. Dersom eventen skildrar eitt særskilt ord, skal eventen stå etter ordet og markerast med ”apply to previous word”. Når ein event har eit plussteikn framfor seg, indikerer det at eventen er markert på denne måten. Merk at vi ikkje bruker ”apply to next word”.

Eventar som hosting, latter og liknande skal setjast inn der dei finn stad i samanhengen – først, sist eller midt i ei ytring – og markerast som «Instantaneous event». Dette betyr at dei ikkje står til eitt eller fleire ord, men er sjølvstendige delar av ytringa.

Nokre eventar kan gå over fleire ord – eller over fleire tidskodar. Sett inn ein event før det første ordet ved å velje eventtype og den eventen du vil ha (til dømes *leende*). Marker i tillegg **Start of event** i eventboksen. Etter det siste ordet går du igjen inn i eventboksen og vel same event, men marker **End of event** før du trykker **OK**. Du vil då sjå at taggane har fått bindestrekar som viser at dei høyrer saman (sjå under). Du kan også markere teksten det gjeld, bruke hurtigkommando **<ctrl-d>**, og velje den aktuelle eventen. Då vil **Start of event** og **End of event** bli markert automatisk. Legg merke til at taggane vil legge seg heilt inntil orda dei ramar inn, slik at du må legge inn mellomrom.

[leende-] det var morsomt [-leende]

3.3.2 Ekstrainformasjon for intervjuaren

Intervjuaren skal transkriberast ortografisk (sjå [3.4.1.8](#)). Når intervjuaren bruker ord som ikkje står i Bokmålsordboka, må vi bruke nokre eventar for intervjuaren som vi aldri treng å bruke i dialekttranskripsjonen.

Av og til kan det hende at intervjuaren bruker dialektord som ikkje utan vidare let seg omsetje til standard bokmål. Eit eksempel er *flingre*, som kan tyde *liten bit* eller *flis*. I slike tilfelle vil det vere meningslaust å skulle avgjere kva ord ein skal setje inn i transkripsjonen. Vi skriv difor *flingre*, men set også inn eventen **language** (<Ctrl-l>) og vel **X**, eller bruker hurtigkommandoen <Alt-x>. Vi bruker også denne eventen ved engelske ord. Hugs å sjekke først om ordet står i Bokmålsordboka!

3.4 Språkform

3.4.1 Lydnær transkripsjon

3.4.1.1 Generelt

- Vi skriv det dei seier – så lydnært som det er muleg med vanlege bokstavar.
- Ved tvil: Vi bruker trykksterk variant. (Døme: *mæg* eller *mæ?* Vel *mæg*.)
- Ved tvil: Vi bruker 'dialektnorma'. (Døme: *dæmm* eller *demmm?* Vel *dæmm*.)
- Ved tvil der ein ikkje får hjelp av punkta over: Vi bruker standarduttale etter ordbøkene.
- Vi held ordgrensene etter ordbøkene for å kunne omsette direkte til ortografisk transkripsjon
Døme: 'Jeg har fått den' /jæ ha 'fåtn/ blir notert *jæ ha fått n*.
Døme: 'Jeg har ikke penger' /jæ 'hake 'peŋer/ blir notert *jæ ha kke pennger*.
- Vi markerer heller ikkje assimilasjon over ordgrenser: /væsjåsnill/ skrivast: *vær så snill*, jf. Papazian og Helleland (2005:21): "Det gjer teksten meir lettlesn og skulle ikkje føre til noka uklarheit med omsyn til uttalen, ettersom denne assimilasjonen over ordformgrensene er ein fast regel i mål som har retrofleksar." Det vil vere ekstra vanskeleg å lese transkripsjonen om ein ikkje kjenner igjen starten av ord. Dessutan vil det med dette bli lettare for dialekt-omsetjaren å kjenne igjen ord og analysere riktig.

3.4.1.2 Fonologi

- *å* står for /o/
 - *o* står for /u/
 - *u* står for /µ/
 - *ao* står for /au/
 - *æu* står for /æµ/
 - *æi* står for /æi/
 - *ei* står for /ei/ (/ej/)
 - *ai* står for /ai/ (/aj/)
 - *âi* står for /âi/ (/âj/)
 - *øy* står for /øy/ (/øj/)
- I t.d. Vest-Telemark finn vi ein mellomting mellom /ø/ og /o/, /ø/, som har utvikla seg frå norrøn

kort /o/. Vi skriv *ø* for denne lyden.

- I enkelte dialektar (t.d. i Sogn) har ein ein mellomting mellom /u/ og /o/. Vi skriv *u* for denne lyden.

- /j/ blir skrive *j*.

Døme: *jore, jikk, jeit* (for gjorde, gikk, geit).

- /k/ blir skrive *k*.

Døme: *vagt > vakt, lags > lakks*

- sje-lyden /ʃ/ blir skrive *sj*, også når det er snakk om assimiliasjon av *r* og *s*.

Døme: *sjønne, sji, Lasj, sjlå* (for skjønne, ski, Lars, slå).

- kje-lyden /ç/ blir skrive *kj*.

Døme: *kjino, kjysse, takje, kjue* (for kino, kysse, taket, tjue).

- kje-lyden i enkelte Vestlands-dialektar, /tʃ/, blir skrive *tj*, jf. Papazian og Helleland (2005).

Døme: *tjino, tjysse, tjue*

- Tjukk l (for l og rð) blir skrive L. (Tjukk l har ikkje same plassen i lydssystemet som tynn l; dialektar som har tjukk l har også tynn l, og nokre har tjukk l for både l og rð, andre berre for l)

Døme: *æLLj, sjæL, soLa, gaL, joLe* (for elg, sjølv, sola, gard, jordet)

- Retrofleks t, d og n blir skrive *rt, rd* og *rn*.

Døme: *snart, færdi, værn* (for snart, ferdig, vern).

- η-lyden (nasal velar) blir skrive *(n)ng*.

NB: tegn > *tenngn*, ligning > *linngning*, Signe > *Sinngne*. NB: tango > *tannggo*

- η+k (nk i skrift) skrivast *(n)ngk*.

Døme: *tanngke, lenngke, vinngk*. (Men: tenkte blir oftast *tenngte* og tenkje *tennkje*.)

- Vi skil ikkje mellom skarre-r og rulle-r. (Dei har same plassering i lydssystemet, ingen norske dialektar har begge lydar.)

Skarre-r på Sørlandet, t.d. i Kristiansand, er ofte vanskeleg å høyre, særleg i utlyd. Vi skriv lyden berre dersom vi høyrer han.

- Palatalisering blir ikkje markert

- *Sj, kj, rt/rd/rn* og *ng* kan altså stå for enten to lydar uttalt kvar for seg eller ein assimilert lyd. Vi vel å ikkje innføre andre bokstavteikn enn i eit vanleg alfabet, blant anna fordi dette ville blitt problematisk for dialektomsettaren og for ein del nettlesarar. Dessutan ville det vore vanskeleg å bruke til dømes IPA-teikn i transkripsjonsprogrammet Transcriber.

Når vi har valt å bruke stor bokstav for tjukk l, L, er det mykje fordi dette har ein viss tradisjon. Stor L for tjukk l er mellom anna brukt i TAUS, der med eit plussteikn i tillegg: L+ (jf. Hanssen et al. 1978:24).

Når vi vel å ikkje innføre store bokstavar for retrofleksar, sje-/kje-lyden, η-lyden eller anna, er det fordi det ikkje finst tradisjonar for dette, og det dermed ville vore vanskeleg å lese transkripsjonen utan å sette seg svært nøye inn i transkripsjonsretteiinga på førehand.

- Engelsk /ɹ/ skriv vi *r*. Døme: *Braien* (for Brian)

- /w/ skriv vi *v*. Døme: *Vesstlaif* (for Westlife)

3.4.1.3 Kvantitet

Vokalkvantiteten i stavingar med hovudtrykk blir markert slik:

- Kort vokal + lang konsonant markerast med dobbel konsonant. Døme: *håppe* (hoppe)

- Lang vokal + kort konsonant markerast med enkel konsonant. Døme: *håpe* (håpe)

- Bitrykk i samansetningar blir markert tilsvarende: *ryggsekk*, *teveprogramm*, *Trånnjæmm*.
- Preposisjonar, adverb og determinativ reknast også som leksikalske ord i samansetningar: *åppvåkkest* (opp-), *nelakkt* (ned-), *ommgang* (om-), *fårrnavn* (for-), *inngentinn* (ingen-).
- Når to like konsonantar møtast, skriv vi det vi hører: *utelling* (~~uttelling~~), *boklubb* (~~bokklubb~~)
- Grammatiske småord m.m. som har kort konsonant i ordbøkene, skriv vi med dobbel konsonant viss hovudstavinga har kort vokal i trykksterk stilling.
Døme: *hos* > *hoss*, *han* > *hann*, *men* > *menn*, *vil* > *vill*, *at* > *att*, *tom* > *tomm*.
Skriv likevel berre dei lydane du høyrer. Ofte må transkripsjonen heller bli slik: *oss*, *ann*, *me*, *vi*.
- Stavingar som er bøyingsendingar er alltid korte og kan aldri ha trykk. Vi noterer dei med enkel konsonant – sjølv om vokalen er kort:
/²bå:tan/ > *båtan* (~~båtann~~), */²vi:ker/* > *viker* (~~vikerer~~), */²jentun/* > *jenntun* (~~jenntunn~~)
/²vide're/ > *videre* (~~viderre~~), */²ein'aste/* > *einaste* (~~einaste~~), */²spen'ans/* > *spennans* (~~spennanns~~)
jf. Papazian og Helland: "Lengd-motsetninga står ved lag også i stavingar utan hovudtrykk, i *avleiingar* og *samansetjingar*." (2005:18). Dersom ein skal lese transkripsjonen rett, må ein altså vite at bøyingsendingar i norsk alltid er korte stavingar, men vi meiner det ville vore svært mykje vanskelegare å lese teksten med det motsette valet; det ville ofte trekke fokus vekk frå hovudstavinga mot ei trykklett endestaving, slik at det kunne sjå ut som om det var endestavinga som hadde trykk: *rutebåtann*, *Bjørnskogenn*. Vi skriv *rutebåtan*, *Bjørnskogen*.
- Kvantitet blir normalt heller ikkje markert i avleiingsendingar:
fissking (~~fisskingg~~), *sjåpping* (~~sjåppingg~~), *snakkes* (~~snakkes~~), *trålldåm* (~~trålldåmm~~)
Men i t.d. bestemt form av ord på -dom, skriv vi likevel dobbel konsonant: *trålldåmmen*.
- Når bitrykket er spesielt sterkt, t.d. i ord med trykk på fyrste staving og tonem 2, må vi ofte rekne bitrykkstavinga som så trykksterk at kvantiteten bør markerast også her:
gymnastikk (~~gymnastik~~), *mussikk* (~~mussik~~), *butikk* (~~butikk~~), *spesiell* (~~spesiel~~).
Noko tilsvarende gjeld også for visse uttalevariantar av ord som *akkuratt*, *nettåpp*.
- Nokre få dialektar kan ha korte stavingar (VK) i trykksterk stilling (t.d. Nord-Gudbrandsdalen), andre kan ha overlange stavingar (V:K:) (t.d. Setesdal). Vi noterer likevel alle stavingar (utanom bøyingsendingar) som normallange, med utgangspunkt i vokallengda:
I korte stavingar er vokalen kort, og vi markerer dette med dobbel konsonant.
I overlange stavingar er vokalen lang, og vi markerer dette med enkel konsonant.

3.4.1.4. Konsonantgrupper og diftongar

- Vi skil mellom lang og kort konsonant i konsonantgrupper:
Døme: */spi:ste/* > *spiste*, */spiste/* > *spisste*, */på:ske/* > *påske*, */påske/* > *påsske*
>> I praksis blir dette slik at vi fyrst noterer den fyrste lyden i konsonantgruppa: er vokalen framfor kort, blir denne markert lang. Deretter noterer vi den neste lyden i konsonantgruppa.
- Også 'falske' konsonantgrupper – enkeltlydar som blir skrivne med to konsonantar i vår transkripsjon – har slike korte og lange variantar: *rn/rrn*, *rt/rrt*, *rd/rrd*, *sj/ssj*, *kj/kkj*, *ng/nng*.
Døme: *mårn*, *harrt*, *Fårrd*, *væssj*, *mykkje*, *sannng* – alle med korte vokalar.
- Diftongar blir behandla som monoftongar – dei kan vere både lange og korte:
/greit/ skrivast *greit* og */gr^{ei}t/* skrivast *greitt*. */ein/* skrivast *ein* og */eⁱn/* skrivast *einn*.

3.4.1.5 Stavingsberande konsonant

- Apostrof ' markerer stavingsberande konsonantar. (som i TAUS, jf. Hanssen et al. 1978:26)
Døme: guten > *gut'n*, tretten > *trett'n*, mannen > *mann'n*, dalen > *dal'n*, verden > *værrd'n*
Men: *brodærn*, *vintærn*, *dåmmarn*. Og /nån/ > *nånn*.
- I mange tilfelle blir det sett inn ein ny vokal slik at konsonanten ikkje ber stavinga aleine. Da skriv vi denne vokalen – ikkje apostrof:
Døme: skolen > *skoorn*, baren > *baarn*, gæren > *gæærn*,
Det kan vere vanskeleg å høre kva kvalitet denne vokalen har. Ved tvil skriv vi den same to gonger – som i døma over.
- Apostrof kan i visse tilfelle også brukast for å markere stavingsberande vokal: *ne'i* (nedi), først og fremst i tilfelle der ein vil unngå at vokalane blir oppfatta som ein diftong. I tilfelle som *kuer* treng ein ikkje å bruke apostrof, sidan *ue* ikkje er ein vanleg diftong.

3.4.1.6 Diverse

- Når *ikkje* og *ikke* er **klitikon** skriv vi *kje* og *kke*.
Døme: /vi:kje/ = *vi kje*, /vike/ = *vi kke*.
 - Der assimilasjon over ordgrenser *ikkje* er regelbunde, markerer vi han likevel:
/ha:rem/ (har dem) → *ha rem*. /skarul/ (skal du) → *ska ru*. /må:Lu/ (må du) → *må Lu*.
 - Der assimilasjon mellom stamme og bøyingsending *er* regelbunde, markerer vi det ikkje:
/¹værdʳn/ (verden) → *værrd'n*, /¹fartʳn/ (farten) → *farrt'n*
- Har ein retrofleks *d / t*, vil ein automatisk få retrofleks *n* rett etter. Derfor kan vi la vere å markere det for å gjere transkripsjonen litt meir lettlesen.
- Nølelyden *e* (i alle vokalvariantar) skrivast *ee*. I tillegg bruker vi nølelydane *em* og *m*. (sjå liste over interjeksjonar i 5.3)
 - Når ein informant seier /vetu/ eller /ve:tu/ (vet du), må vi skrive det i to ord på grunn av omsetjaren. Ein kan seie at d-en ikkje er høyrbar, men vi ser likevel ein assimilasjon her, ved at t-en blir uaspirert på grunn av den etterfølgande d-en. Vi skriv difor *vett du* eller *vet du*.
 - I samansette ord der det første ordet sluttar på den same bokstaven som det neste startar på (bokklubb, skoggård), vil ein vanlegvis berre høyre ein av dei to konsonantane, og vi skriv då t.d. *boklubb* i transkripsjonen. Dersom ein likevel kan høyre begge konsonantane, skriv ein t.d. *skoggård*, sjølv om vokalen framleis er lang.
 - Som nemnt tidlegare, blir endingsvokalen føre skarre-r i mange sørlandsdialektar assimilert, ofte til ein /å/, men det er ikkje alltid eintydig kva vokalkvalitet det er snakk om her. Hovudregelen er at vi vurderer kvar enkelt informant, og ved tvil skriv vokalen vi oftast høyrer.

3.4.1.7 Loggføring/informasjonsflyt

- Vi kan ved behov lage ei kort liste over dei viktigaste 'dialektnormene' for dialektar transkribørane oppfatar som særskilt vanskelege å transkribere.
Døme: Pronomen, spørjeord, bøyingsmorfologi, spesielle fonologiske trekk
- Vi kan også lage ein felles problemlogg for dialekten, og noterer her når eit problem er løyst / eit spesifikt val er tatt.

3.4.1.8 Standardortografi

- Intervjuaren skal transkriberast ortografisk etter bokmålsordboka.
- Alle transkripsjonane skal omsettast til standardortografi med ein dialektomsettar. Dette vil i fyrste

omgang bli gjort til bokmål.

Når vi vel å skrive bokmål og ikkje nynorsk i desse to tilfella, er det hovudsakleg av økonomiske omsyn: Alt det språkteknologiske verktøyet som skal brukast på materialet etter transkripsjonen, til dømes ordklassetaggning og parsing, finst berre for bokmål.

Vi planlegg å omsette transkripsjonane til nynorsk ortografi på eit seinare tidspunkt.

• Omsetjinga til bokmål skjer halvautomatisk. Vi er avhengige av at det er like mange ord i dialekttranskripsjonen som det vil vere i den omsette versjonen. Ofte vil ein difor måtte særskrive ord som informantten uttalar som eitt ord, men som ikkje er oppført som eitt ord i *Bokmålsordboka*. Vanlege døme er *ja vell* (*jævell*), *næi da* (*næida*) og *va fall* (*vaffal*). Sjekk med ordboka om du er i tvil! Samansetningar som ikkje står i ordboka, skal likevel skrivast i eitt ord. Med samansetningar meiner vi her samansette ord i dei opne ordklassane, som substantiv og adjektiv, men det same gjeld også for samansette preposisjonar. Døme: *guLgennsærn*, *kjemmpenydli*, *borrtåt*, *oppijøno*.

3.4.2 Tal

Tal: skriv vi med bokstavar. Slik får vi med om t.d. *1600* blir uttala *ett tus'n sekks hundre* eller *sekkst'n hundre*. Tal under hundre skriv vi i eitt ord. Unntaket er tal etter den gamle tellemåten, som vi skriv som tre ord. Alle tal over hundre skriv vi i fleire ord.

```
kjuefire
fir å kjue
hundre å kjuefire
hundre å fir å kjue
tre tus'n to hundre å kjuefire
```

Unntaket er *årstal*, som vi skriv i eitt ord uansett:

```
nitt'nhundreåttrettiåtte
nitt'nåttåtræddve
åttåtræddve
```

Ordenstal: skriv vi også i eitt ord med bokstavar. Ordenstal over hundre skriv vi i fleire ord.

```
kjuefjære
hundre å kjuefjære
```

Brøkar: skriv vi også med bokstavar:

```
to treddjedels
```

3.4.3 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). Namna får fortløpande nummer etter kvart som dei dukkar opp i teksten:

Informanten seier:

jæi så att Kari Hanns'n ga Nina dokumennt'ne till Nills'n

]

Vi skriv:

jæi så att F1 ga F2 dokumennt'ne till E1

Namn på offentlege personar treng vi som regel ikkje anonymisere.

Fotballag, stadnamn, namn på kjæledyr og organisasjonsnamn m.m. skriv ein normalt inn dersom ein ikkje vurderer at det er sensitive personopplysningar involvert, sjå [3.2](#). Bør eit slikt namn anonymiserast, skriv *N1*, *N2* osv.

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 *sisst'n*, skriv vi med liten bokstav. Dersom slike namn består av meir enn eitt ord, set vi også hermeteikn rundt for å synleggjere at orda høyrer saman: "*båkks'n går*", "*sjlå på rinng*", "*jeppe pinne*".

Dersom eit namn skal skrivast som namn og ikkje som kode, skal det skrivast slik det blir uttala, men vi må passe på at det er like mange ord i dialekttranskripsjonen og bokmålsomsetjinga (sjå [3.4.1.8](#)). Merk at vi bruker stor forbokstav ved særnamn: *Teveto (TV2)*, *Firehå (4H)*, *Priss'n Breik (Prison Break)*. Dersom det er vanskeleg å oppfatte namnet, skriv vi berre *N1*, *N2* osv. NB! Ikkje bruk mykje tid på namn!

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_04gk* osv.) i staden for dei vanlege *F* og *M*.

3.4.4 Ikkje-språklege lydar

Ikkje-språklege lydar som latter, hosting og liknande som er del av ei ytring, skal vere med i transkripsjonen når dei skapar mening i samtalen (sjå [3.3](#)). Dette kan til dømes vere eit host som markerer ueinigheit. Latter skapar nesten alltid mening, men eit ukontrollert host skal ikkje vere med.

Lydar skal spesifiserast som eventen **noise**. Skriv **Ctrl-d** og vel korrekt event. Det same gjeld ikkje-språklege lydar som går saman med ei ytring (t.d. *leende*, *gjespende*, osv.), her bruker vi **Start of event** og **End of event** som skildra i [3.3](#):

```
[gjespende-] jeg er så trøtt [-gjespende]
```

Latter er også ein event, men finn stad såpass ofte at vi har ein eigen hurtigkommando for denne eventen: **<Alt-o>**:

```
[latter] det var morsomt
```

Latter saman med ei ytring transkriberast med *leende*-eventen, som tilfellet *gjespende* ovanfor.

3.4.5 Sitat og ord som er stava eller lesne

3.4.5.1 Sitat

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

da kjørte jæi denn "jæi? hæ?" da jüge jæi

Dersom ein informant snakkar om uttalen av eit ord, regnar vi også dette som sitat dersom ordet ikkje er stava.

æ sier ikke "snø" æ sier "sne"

3.4.5.2 Ord som er stava eller lesne

Staving av ord: I ord som er stava, skriv vi bokstavane ein for ein, og rammar inn bokstavane med eventen **pronounce**:

kutt skrives *[pron=stavet-]* c u t *[-pron=stavet]* på engelsk

Skriv <Ctrl-r> og vel **stavet** og **Start of event** føre med <Ctrl-r>, **stavet** og **End of event** bak. Merk at vi ikkje markerer noko med **stavet** dersom det dreier seg om t.d. vanlege forkortingar, produktnamn eller firmanamn med forkortingar (DVD blir *devede A/S* blir *aess*, NHH blir *Ennhåå*, *AG3* blir *agetre*).

Høgtlesing: Dersom noko blir lese høgt, skal dette markerast med **pronounce** <Ctrl-r> og **høytlesing**.

3.4.6 Avbrot, pausar og uklare passasjar

3.4.6.1 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 gjenteke 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- -søttiätte*.

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 eventen **stønning**, andre gonger kan ein velje å markere nølinga med taggen *{uforståelig}*.

3.4.6.2 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. Vi kan også skrive tre punktum etter bindestrek for avbrot:

```
høres ut såmm sånn hær ...  
høres ut såmm sånn hæ- ...
```

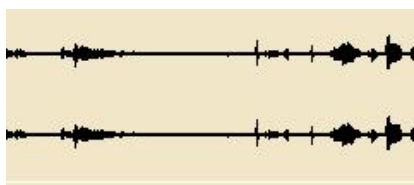
Dersom ein talar avbryt seg sjølv og fortset med ei anna ytring, skal vi ikkje skrive tre punktum. Skriv den påbegynte ytringa og fortset fortløpande med den neste:

```
du har du har du har ikke jorrt de
```

Dersom det er pausar i ytringa, sjå avsnitt [3.4.6.3](#).

3.4.6.3 Pausar

Korte pausar markerer vi med #, mellomlange pausar med ## og svært lange pausar med ###. 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 eller 1 cm i spektrogrammet, sjå under). Ein kort pause er kortare. Det kan vere vanskeleg å avgjere når ein skal markere korte pausar. Ein hugseregulering kan vere at det skal vere eit markert opphald utan lyd. Ofte vil ein pause kunne sjåast som ein rett strek på spektrogrammet. Det er likevel verdt å merke seg at ein rett strek ikkje alltid vil vere ein pause. Enkelte konsonantlydar vil ikkje gi utslag på spektrogrammet. Nedanfor er først ein kort og så ein mellomlang pause:



```
å # jæi tror ## det æ sjlik att
```

Ein ekstra lang pause ### finn vi sjeldan i våre intervju og samtalar. Vi kan definere ein ekstra lang pause som ein pause som er så lang at samtalepartnarane ville synest det begynte å bli pinlig stille. Den mellomlange pausen er den med størst variasjon, og kan vare ganske kort eller ganske lenge.

Dersom informanten pustar, vil dette kunne sjåast som lyd på spektrogrammet. Dersom det ikkje dreier seg om eventar som **latter** eller **sukking**, vil vi markere pusting som pausar (vi bruker ikkje eventen **pusting**). Ulike typar nøling som *ee* osv. markerast med ord frå lista i [5.3](#). Hosting, snufsing og klikkelydar markerast med eventen **noise** (sjå [3.3](#)) 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 [3.4.6.2](#)).

Pausar finn vi som regel inne i segment. Mellom segment og turar (sjå [3.5](#)) vil det som regel også vere pausar, men desse pausane markerer vi ikkje.

3.4.6.4 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 *uforståelig*-kommentar, <Alt-u> i staden:

```
men takk for at du {uforståelig}
```

Uklart: Skulle det vere passasjar som du trur du forstår, men som du ikkje er sikker på, set du inn ein *uklart*-event. Dersom det dreier seg om fleire ord, set du inn eventen føre og bak (Vel **Ctrl-r** saman med **uklart** og **Start of event** og **End of event** eller marker dei uklare orda og vel hurtigkommando <Alt-k>):

```
han sa at [pron=uklart-] han visste [-pron=uklart] at de skulle komme
```

Eitt uklart ord markerer vi med <Alt-k> slik:

```
1: e dere så jo litt på det det skjemaet i stad da # nei e lista der  
2: * mm +[pron=uklart]
```

3.4.7 Teiknsetjing og stor bokstav

Stor forbokstav bruker vi ikkje i starten av ei setning. Namn skriv vi derimot med stor bokstav. Semikolon, kolon, komma, utropsteikn og punktum bruker vi ikkje. Spørjeteikn bruker vi som vanleg ved spørsmål, også dersom spørsmålet blir avbrote. Ved spørsmål ein ikkje ventar svar på (*de ær sjlik ikke sannt de bare blir sånn*), kan spørjeteikn droppast.

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.

```
ji mæi en sæigmann  
næi menn du kann få en Håbbi viss du jir mæi Nokian minn
```

3.5 Tur, segment og meiningseining

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 *segment* og *meiningseining*. Segment og meiningseiningar deler ein tur opp i mindre einingar. Dei neste avsnitta skal forsøke å klargjere korleis vi skal bruke desse tre omgrepa, og vi har brukt Linell og Gustavsson (1987) som referanse.

3.5.1 Tur

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.

Det er ikkje alltid innlysande korleis ein skal dele ein samtale inn i turar. Kanskje kan det hjelpe å 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 vil tilsvare to turar, sjølv om det er den same informanten som snakkar.

Ein tur (*turn* i *Transcriber*) set vi inn i dokumentet med hurtigkommandoen <Ctrl-t>.

3.5.1.1 Replikk og ikkje-replikk

Ofte vil stafettpinnen gå uproblematisk fram og tilbake mellom samtalepartnarane, men mange gonger kan det vere vanskeleg å avgjere kven som har turen. Her strekk ikkje stafettpinnemetaforen til, og vi innfører omgrepa *replikk* og *ikkje-replikk* (frå Linell og Gustavsson, 1987) for å klargjere turomgrepet:

Ein *replikk* er ei ytring som blir sagt av den som har turen. Ein *ikkje-replikk* blir sagt av den som ikkje har turen. Under vil vi gå nærare inn på kva som kan seiast å vere ein *ikkje-replikk*. Dersom vi har eit klart bilete av kva ein *ikkje-replikk* er, kan vi definere *replikk* som det motsette. Den samtalepartnaren som har turen, er dermed den personen som ytrar replikkar. Når den andre samtalepartnaren ytrar noko som kan definerast som ein replikk, skiftar turen over på vedkommande.

Ein *ikkje-replikk* er ikkje eit konversasjonelt bidrag, og 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 avbrotforsøk*: 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*:

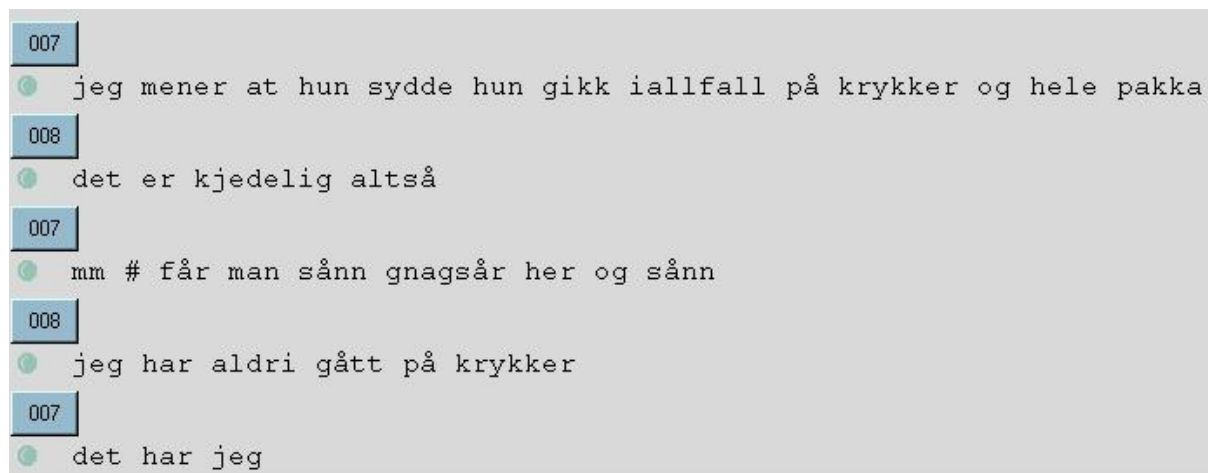
- ikkje svar på spørsmål (svar på spørsmål er *replikk*)
- kort latenstid (pause) eller overlapp mellom A og B sine ytringar
- svak lydstyrke eller lite omfang

- ofte omformuleringar eller korte spørsmål

Altså: Dersom noko ikkje er feedback eller mislukka avbrotforsøk, er det ein replikk.

3.5.1.2 Tur og turskifte

For å oppsummere: Når vi transkriberer, skal vi dele samtalen opp i *turar* (<Ctrl-t>). Ein *tur* inneheld ein eller fleire *replikkar* frå same talar:



Turskifte oppstår 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. Det er ikkje turskifte om talar B kjem med ulike former for feedback, sjå [3.5.1.1](#).

Dersom talar A responderer på B sine ytringar, er det snakk om ein replikk hos B. Dersom A ikkje responderer, er det derimot snakk om ein ikkje-replikk. Småord som *mm*, *ja* o.l., som ofte opptre som ikkje-replikkar, kan altså i visse samanhengar vere replikkar. Det er verdt å merke seg at det oftast er respons, eller mangel på respons, som avgjer om ei ytring får status som replikk eller ikkje-replikk. Alle ytringar som får respons, må vere replikkar.

At samtalepartnarane snakkar i munnen på kvarandre, er svært vanleg i ein 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. Sjå [3.5.3](#) for å lese meir om overlappende tale.

Latter er som regel ikkje nok til å halde på turen. Dersom 007 i illustrasjonen er talar 1 i ein tur, seier noko og deretter ler, kan 008 ta over turen ved å ytre ein replikk sjølv om 007 framleis ler. Då vil 008 bli talar 1 i ein ny tur, medan latteren til talar 007 vil bli markert som overlappende ikkje-replikk:

007

- mm # man får sånn gnagsår her og sånn [latter]

008+007

- 1: jeg har aldri gått på krykker
- 2: [latter]

007

- det har jeg

3.5.2 Meiningsseining og segment

Vi bruker omgrepet *meningsseining* for den minste eininga ein samtale skal delast opp i. Ifølge Linell og Gustavsson (1987) er ei *meningsseining* ei formelt og/eller semantisk avgrensa eining som utgjer ei ytring eller ein del av ei ytring. Vi kan bruke følgjande kriterium for å dele ytringar inn i *meningsseiningar*:

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

I *Transcriber* tilsvarar ei *meningsseining* nesten alltid eit *segment*. Eit segment startar med ein grøn runding. Eit nytt segment startast ved å taste *Enter*.

Monica

- ja fordi at jeg går rundt og prøver være venner med alle
- og det er ikke noe jeg har bevisst prøvd å gjøre
- jeg har prøvd å være hyggelig mot folk
- for det er ingen her inne jeg direkte misliker så hvorfor skulle jeg gå rundt da å være
- [snufsing] ja, selvfølgelig er det noen jeg liker mer enn andre men det er # jeg kjenner ikke alle

Føre kvart segment, altså ved kvar grøne runding, blir det sett inn ein tidskode i *Transcriber* som bind lyden saman med det som er transkribert. I sluttproduktet til ScanDiaSyn vil segmentet vere den minste eininga brukaren kan få som svar dersom han eller ho søker etter spesielle konstruksjonar i korpuset. Er du i tvil om kvar du skal setje ei segmentgrense etter å ha brukt kriteria over, kan det kanskje hjelpe å sjå for seg ein søkesituasjon. Er segmentet svært kort, vil brukaren få lite informasjon om konteksten til det han søker etter. Er segmentet svært langt, vil det vere tungt å jobbe med. I eksemplet under er *ja* sett saman med det påfølgande segmentet fordi eit *ja* aleine ville blitt eit svært kort segment. Pausemarkøren # viser at det er ein kort pause mellom *ja* og resten av segmentet:

030 + 029

- 1: ja # eller jeg var veldig keen +[lang=English] på å ta n før
- 2:
- 1:
- 2: mm

Kor langt eit segment bør vere, vil variere noko. Ein kan tenke seg eit ”idealsegment”, som varer i om lag 10 sekund og har tekst tilsvarande ei linje i Transcriber. Dette vil sjølvstøtt ofte vere umogleg å få til, ettersom informantane snakkar med ulik fart. Samstundes skal ein forsøke å dele ytringane inn i segment og samstundes ta vare på meiningseiningane. Likevel kan det vere nyttig å ha ”idealsegmentet” i bakhovudet som ein peikepinn.

Som hovudregel deler vi aldri opp meiningseiningar i fleire segment. Eitt unntak er når talar 1 og talar 2 snakkar i munnen på kvarandre utan at meiningseiningane deira er synkroniserte. I slike tilfelle let vi talar 1 sine meiningseiningar styre oppdelinga av samtalen i segment, og deler opp talar 2 sine meiningseiningar. For å ta vare på talar 2 sine meiningseiningar, set vi inn taggen **me** føre og bak meiningseininga. Meir om dette i [3.5.3.4](#).

007+008

- 1: ja for de regner med at det går sin gang
- 2: [pron=me-] ja det det går
- 1: jeg hadde skikkelig traumer for at jeg ikke kunne klokken og sånn
- 2: sikkert [-pron=me]

3.5.3 Innskot av ikkje-replikkar, overlapping og delte meiningseiningar

Som framstillinga over kjem inn på, går ikkje segment- og turskifte alltid føre seg så smertefritt. Under ser vi ei oversikt over korleis vi løyser ulike kompliserte tilfelle.

3.5.3.1 Innskot av ikkje-replikkar

Det er vanleg at talar 2 kjem med ein eller fleire ikkje-replikkar som innskot i løpet av talar 1 sin tur. Dette er som regel feedback av typen *mm*, *oi* og liknande. Slike innskot behandlar vi slik:

Sett inn ein tidskode ved hjelp av **Enter**. Trykk <Ctrl-t> og vel første talar i skjemaet. Kryss så av for **Overlapping speech** og vel deretter andre talar. (Nokre gonger vil programmet sjølv gjette seg til kven som er første og andre talar. Ofte vil det også bli feil, så her må ein passe på.) Etter å ha trykka **OK**, ser du at begge talarane kjem opp i namnefeltet. Under er talarane nummerert med første og andre talar. Talar 1 har turen, og talar 2 er anten taus eller kjem med ikkje-replikkar medan talar 1 er taus.

008+007

- 1: husker vi m dere fikk sånn mat på døren
- 2:
- 1:
- 2: ja
- 1: og så var jeg med og levere det en gang så var jeg så misunnelig
- 2:
- 1:
- 2: mm

Av og til vil talar 2 sine ikkje-replikkar overlappe med talar 1 sine replikkar, sjå [3.5.3.3](#).

Legg merke til at ein tur dermed kan markerast på to måtar: Anten som ein tur der den som har turen, har ordet aleine slik som her:

```
Monica
● ja fordi at jeg går rundt og prøver være venner med alle
● og det er ikke noe jeg har bevisst prøvd å gjøre
```

Eller ein tur der talar 2 har ikkje-replikkar som innskot eller overlapping under heile eller delar av talar 1 sin tur slik som det førre eksemplet over.

I nokre av samtalanene er det eksempel på at intervjuaren også kjem med ikkje-replikkar. Desse transkriberer vi vanlegvis ikkje, på same måte som vi ikkje transkriberer ikkje-replikkane til intervjuaren i intervjuet (sjå [3.5.3.3](#)). Som bileta over viser, opererer *Transcriber* berre med to talarar under ein tur, difor er det enklast å kutte den tredje talaren sine ikkje-replikkar. Dersom intervjuaren eller andre utanforståande kjem med replikkar, må ein prøve å få dei med i transkripsjonen.

3.5.3.2 Ikkje-replikk eller del av replikk?

Ved turskifte kan det vere vanskeleg å avgjere om det som blir sagt i overgangen mellom turane, tilhøyrer den første eller den andre turen:

```
● 1: skada av å sitte på fly ikke sant|
  2:
029 + 030
● 1: ja # men det er sånn du får i Norge òg ved å fly små fly
  2:
```

I dette eksemplet kan *ja*-et i den andre turen kanskje også seiast å vere ein ikkje-replikk i den første turen. Vi har likevel valt å sjå det slik at talar 2 aldri disponerer det siste segmentet i talar 1 sin tur aleine. Difor blir det siste talar 2 seier aleine, alltid ein del av replikken i neste tur.

3.5.3.3 Overlapping

Vi markerer overlapping på same måte som innskot. Ein overlappingstur lagar vi ved å trykke **<Ctrl-t>** og velje første talar i skjemaet. Kryss så av for **Overlapping speech** og vel deretter andre talar.

I ein slik overlappingstur er talar 1 den som har turen og ytrar replikkar, talar 2 kjem med ikkje-replikkar, sjå klassifiseringa av *replikkar* og *ikkje-replikkar* i [3.5.1.1](#) over.

```
● 1: jeg husker jeg hadde mistet e jeg mistet så mye når jeg var liten
  2: vet det
```

Det er ikkje så nøye akkurat kvar talar 2 kjem inn med ikkje-replikkar i talar 1 sin tur, men vi markerer at ein ikkje-replikk kjem litt ut i eller på slutten av talar 1 sitt segment ved å setje inn ei stjerne. Dersom det kjem fleire ikkje-replikkar med pause mellom frå talar 2 i løpet av talar 1 sitt segment, kan vi skilje ikkje-replikkane med stjerne:

- 1: fordi før så var det sånn man var ikke helt sikker på hvem som egentlig var vennene sine og
2: mm * m
- 1: hvert fall sånn på ungdomsskolen så tar det i hvert fall et par år før du # merker hvem du liksom nære da
2: * m

Av og til bruker vi også ei stjerne i talar 1 sitt segment. Dette er i så fall i starten av segmentet dersom talar 2 sin ikkje-replikk startar før talar 1 sin replikk. Andre pausar i talar 1 sitt segment markerer vi som vanleg med #.

All overlapping skal med i samtalanene – sjølv om det er tidkrevjande å gjennomføre. I intervjuet har vi bestemt oss for å ikkje transkribere intervjuaren sine oppmuntrande *hm* og andre små ikkje-replikkar i løpet av informantens sin tur. Vi tek sjølvstund med reelle replikkar frå intervjuaren.

3.5.3.4. Delte meningseiningar

Når talar 1 og talar 2 snakkar i munnen på kvarandre utan at meningseiningane deira er synkroniserte, let vi talar 1 sine meningseiningar styre oppdelinga av samtalen i segment, det vil seie at talar 1 alltid får ”snakke ferdig”. Talar 2 sine meningseiningar kan dermed bli delte opp over fleire segment. For å ta vare på talar 2 sine meningseiningar, rammar vi dei inn med eventen **pronounce** og **me** (meningseining). Det er to hurtigkommandoar for dette: **<Alt-m>** for start på meningseining, og **<Alt-n>** for slutt på meningseining.

007 + 008

- 1: ja for de regner med at det går sin gang
2: [*pron=me-*] ja det det går
- 1: jeg hadde skikkelig traumer for at jeg ikke kunne klokken og sånn
2: sikkert [*-pron=me*]

Talar 2 si meningseining kan også vere delt over to turar:

008 + 007

- 1: sånn i julekalender og sånn fikk sånn brukte ting og sånn
2: [*pron=me-*] jeg husker jeg fikk en sjampo

007 + 008

- 1: for e farget hår jeg hadde ikke farget hår en gang [*latter*]
2: * [*latter*]

Dette skjer i tilfelle der talar 2 snakkar samtidig med talar 1, og der talar 2 i løpet av meningseininga overtar turen.

NB! [*pron=me-*], altså eventen som markerer starten på ei meningseining, kan du plassere ved

starten av meiningseininga. [-pron=me], eventen som markerer slutten på meiningseininga, må derimot, av tekniske grunnar, alltid stå i slutten av eit segment. Av og til vil ein difor inkludere ytringar som strengt ikkje høyrer til i meiningseininga, i denne eventen. Slik unngår ein til dømes for korte segment.

Som vi har sett tidlegare (3.5.3.3), bruker vi stjerne som pauseteikn i talar 2 sine ikkje-replikkar. Når talar 2 si meiningseining er delt opp over to turar slik at talar 2 overtar turen, bruker vi vanleg pauseteikn i talar 2 si meiningseining. Dette er fordi det som er inne i meiningseininga, er ein replikk:

```
004+003
1: ja nei men hvis du er fornøyd så kan vi godt gjøre det ja da
2: * ja * mm * [pron=me-] ja ikke sant # det

003+004
1: synes jeg også [-pron=me]
2:
```

3.5.3.5 Skjulte turar

Som vi har sett over, vil ein *tur* bestå av talar 1 sine replikkar og talar 2 sine ikkje-replikkar. Når talarane snakkar i munnen på kvarandre, kan det likevel skje at talar 2 seier noko som talar 1 reagerer eller svarar på. Talar 2 har då ifølge definisjonen i 3.5.1.1 ytra ein replikk som eigentleg burde hatt sin eigen tur. Problemet er at replikken blir ytra samtidig med at talar 1 også har ein replikk, og at talar 1 svarar talar 2 umiddelbart. Vi har valt å markere dette ved å lage ein ny tur – sjølv om talar 1 og talar 2 er den same i begge turane:

```
004+003
1: nei e det er ikke så veldig søtt
2: er det ikke det?

004+003
1: nei det er sånn man har i brød og særlig utlandet Italia Frankrike og sånn så liker de gjerne å
brødet med karvefrø # og det e jeg synes det er helt forferdelig
2: trodde det var ... * ja * ja
```

Når vi har to turar etter kvarandre der talar 1 og talar 2 er dei same, tyder det altså at talar 2 har hatt ein replikk på slutten av den første turen slik som over, der *er det ikke det?* blir ytra av talar 2 samtidig med *nei e det er ikke så veldig søtt* av talar 1. Talar 1 svarar umiddelbart på talar 2 sitt spørsmål, og dette er altså skilt ut som eigen tur.

4. Korrektur

4.1 Organisering av filene

Alle filene skal korrekturlesast av ein annan transkribør enn den som opphavleg transkriberte fila. Når du skal velje ei fil for korrektur, går du inn i *Scandiasyn – Til_korrektur*. Legg ein kopi av fila du skal korrekturlese inn i *Originalfiler* før du flytter fila inn i di eiga transkripsjonsmappe. Slett fila frå *Til_korrektur*. Opne fila som vanleg. Lydfila finn du i lydfil-mappa til den originale transkribøren (sjå på initialane i filnamnet). Når du lagrar fila, legg du til dine egne initialar på slutten av filnamnet, etter transkribøren sine:

```
stavanger_03gm-04gk_ahl_lh.tr
```

Av og til må ei fil korrekturlesast fleire gonger. Kvar korrekturlesar skal legge til sine egne initialar i slutten av filnamnet, slik at ein alltid kan sjå kor mange og kven som har vore involverte i transkribering og korrekturlesing av fila. Dersom du korrekturles ei fil, og vil at nokon andre skal sjå på henne når du er ferdig, legg du fila i mappa *Til_andrekorrektur* i *Scandiasyn*.

Sjå meir om filorganisering på ScanDiaSyn-prosjektet i [2.2](#). I [7.3](#) finn du ei hugseliste du bør bruke når det er du som utfører den siste korrekturen på ei fil.

4.2 Lese korrektur

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 [2.2](#)).

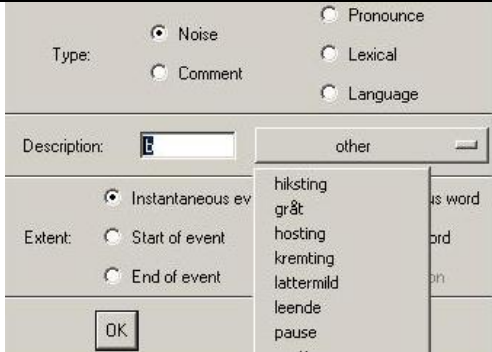
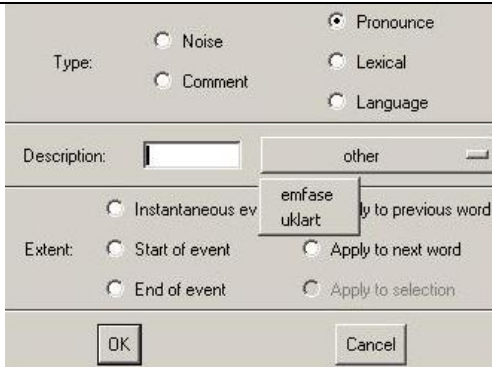
Når ei fil er ferdig korrekturlese, flyttar du fila til *NoTa – Transkripsjon – Ferdig korrekturlest* og formsjekkar (<http://omilia.uio.no/cgi-bin/nota/wf.pl>). Når fila er ferdig formsjekka, eksporterer du fila til HTML-format (**File – Export – Export to HTML format...**). Lagre i *Scandiasyn – html-filer*. Deretter flyttar du fila til *Korrekturlest*-mappe i *Scandiasyn*.


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

5.1 Hurtigkommandoar

avbrot	<Alt-a>	{avbrudd}
copy	<Ctrl-c>	
create turn	<Ctrl-t>	<
cut	<Ctrl-x>	
delete breakpoint	<Ctrl-Backspace>	
engelsk	<Alt-e>	+{lang=English}
find/replace	<Ctrl-f>	
fremre klikkelyd	<Alt-f>	
innledende bemerkningar	<Alt-i>	{innledende bemerkningar}
insert breakpoint	<Return>	
language	<Ctrl-l>	
latter	<Alt-o>	[latter]
meningsenhet, start	<Alt-m>	[pron=me-]
meningsenhet, slutt	<Alt-n>	[-pron=me]
new trans	<Ctrl-n>	
noise	<Ctrl-d>	
nyord/slang/dialekt ...	<Alt-x>	+{lang=X}
open trans	<Ctrl-o>	
paste	<Ctrl-v>	
play around cursor	<Alt-Space>	
play/pause	<Tab>	
play segment	<Shift-Tab>	
pronounce	<Ctrl-r>	
quit	<Ctrl-q>	
save	<Ctrl-s>	
sensitive pers	<Alt-s>	{sekvens med sensitive personopplysningar}
uforståelig	<Alt-u>	{uforståelig}
uklart	<Alt-k>	+{pron=uklart}
undo	<Ctrl-z>	

5.2 Skjema over eventar

<p>noise</p>	<p><Ctrl-d></p>	<p>Fremre klykkelyd bakre klykkelyd sugelyd labial frikativ labial vibrant sibilant latter lattermild * leende gjespende gjesping gråt hosting hviskende kremting leende lydmalende ord nysing pause pusting * snufsing stønning syngende sukking trekker pusten</p> <p>* ikkje i bruk</p>	
<p>pronounce</p>	<p><Ctrl-r></p>	<p>høytlesing me stavet uklart</p>	

language <i>(berre for intervjuar)</i>	<Ctrl-l>	arabisk ... engelsk ... svensk ... X	
comment			Bere brukt med hurtigkommandoane <Alt-p>, <Alt-s>, <Alt-u>, <Alt-i> og <Alt-a> for høvesvis <i>problem</i> , <i>sekvens med sensitive personopplysninger</i> , <i>uforståelig</i> , <i>innledende bemerkninger</i> og <i>avbrudd</i>

5.3 Liste over interjeksjonar

I talemål er det som regel fullt opp av lyd målende ord som *eh, aha, em, ææææ, oiiii* osv. Nokre av desse orda er interjeksjonar som vi finn i Bokmålsordboka (*BMO* i oppstillinga under), andre er ikkje standardiserte. Under har vi laga vår eiga standardiserte liste over interjeksjonar vi skal bruke. Legg merke til at vi ikkje markerer lengde på vokalane. To like bokstavar står for to stavingar, med unntak av nølelyden *ee*. Alle interjeksjonane er sett opp med tyding. Vel den interjeksjonen som ligg nærast det informanten seier, både når det gjeld uttale og tyding. Enkelte av interjeksjonane, som nølelydane, skriv vi på same måte uansett korleis informanten uttalar dei (t.d. *ee* for nøling av ulike vokalkvalitetar). Andre tilpassar vi uttalen til informanten (t.d. *off a mæi* for *huff a meg*). Pass på at uttrykka er delt inn etter ordgrensene i lista. Dette er viktig av omsyn til omsetjaren (sjå [3.4.1.8](#)). Dersom det er intervjuaren som kjem med interjeksjonane, skriv vi dei som i skjemaet under.

Interjeksjonar vi ikkje endrar stavemåten på:

<i>ee</i>	(nøling - uansett lengde på een)
<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>BMO</i> 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) <i>BMO</i>
<i>gud a meg</i>	(overraska, utrop)
<i>huff a meg</i>	(beklagning) <i>BMO</i>
<i>hæ</i>	(spørjande) <i>BMO</i>
<i>jaha</i>	(forsterkande "ja") <i>BMO</i>
<i>næ</i>	(overraska, undrande)
<i>nja</i>	(tvilande) <i>BMO</i>
<i>næhei</i>	(forsterkande "nei")
<i>ops</i>	(oi då)
<i>u</i>	(imponert)
<i>uff a meg</i>	(beklagning) <i>BMO</i>
<i>ææ</i>	(konstaterande - to stavingar)
<i>å-å</i>	("oj")
<i>å ja</i>	(overraska)

Nokre ord er berre lyd målende utan spesiell tyding. Dette er ord som rett og slett skildrar lyd, som til dømes *uuuuuuuuu* for ulinga til ein ambulanse eller *ææææææææææ* som skrikings til eit barn. Slike ord skriv vi ikkje inn, men karakteriserar dei med eventen **noise** <Ctrl-d> og **lydmalende ord**.

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

1. Transcriber-ikon på skrivebordet

Transcriber skal vere lagt inn på alle maskinene, men det finst gjerne ikkje noko hurtigikon på skrivebordet. Slik gjer du:

- Gå inn på Start/Programs/Transcriber. (Frå Start-knappen nedst til venstre på dataskjermen.)
- Høgreklikk på Transcriber-teiknet og vel 'Create shortcut'.
- På 'do you want the shortcut to be placed on the desktop instead' svarar du yes/ja :)

2. Hente inn scandiasyn

- Gå inn i My Computer.
- Klikk på Tools og Map Network Drive.
- Skriv inn [\\hf-lager\scandiasyn](#) (der streken står og blinkar, etter 'Folder:') og trykk Finish.
- Gjer det same for [\\foni\nota](#) også.

3. Finn du ikkje 'speaker' på heimeområdet?

- Gå (t.d.) inn i My Computer, trykk på Tools og Folder Options.
- Trykk på fana View og kryss av for 'Show hidden files and folders' eit stykke ned på lista. ('speaker' er nemleg ei 'hidden file' :)

4. Hurtigkommandoar i Transcriber

- Gå inn på fyrstesida på scandiasyn og *kopier* den fila som heiter [kristiha.cfg](#)
- Gå ut av scandiasyn og inn i Local Disc (C:) og vidare til Documents and Settings.
- Lim inn fila [kristiha.cfg](#) i den mappa som har ditt brukarnamn (t.d. ashildso eller signal)
- Endre 'kristiha' til ditt brukarnamn (t.d. til ashildso.cfg eller signal.cfg)
- Opne Transcriber og trykk på Options og Load configuration file.
- Etter dette må du opne ei transkripsjonsfil i Transcriber og trykke på Save configuration og deretter lagre som vanleg (altså berre vanleg 'save'). For å få lagra må du ofte gjere ei eller anna endring i transkripsjonen i tillegg.
- Dersom du ikkje får brukt hurtigkommandoane likevel, etter dette, kan du gå inn i Options og sjå på Bindings. Dersom det er ei lang liste der med hurtigkommandoar du kjenner igjen, pleier det å hjelpe (merkeleg nok) å trykke på ein av dei, kven som helst, og så trykke på ok. Dersom lista over hurtigkommandoar ikkje står der, må du prøve på nytt, (altså gjere om igjen dei to siste punkta over).

7.2 Hugseliste – når ein startar på ein ny transkripsjon:

1. Finn ei ledig lydfil. Flytt den over til mappa med ditt brukarnamn i *Lydfiler*-mappa.
2. Opne lydfila i Transcriber.
3. Under *Fil* vel du *Edit episode attributes*. Fyll inn ditt eige namn (for- og etternamn, ikkje brukarnamn) under *Transcriber's name*.
4. Klikk på den raude firkanten der det står ”report”. Vel *New topic* og skriv inn ein av desse:

Intervju

Samtale – familie

Samtale – venner

Samtale – bekjente

Samtale – ukjente

5. Hugs at alle informantkodane skal vere like i transkripsjonen som i filnamnet. Heiter fila gausdal_01um-02uk, skal informantkodane i transkripsjonen vere gausdal_01um og gausdal_02uk.
6. Lagre fila i di eiga mappe under Scandiasyn. Legg til dine eigne initialar på slutten av filnamnet (berre transkripsjonsfila, ikkje lydfila).

Gratulerer, du er no klar til å transkribere! :-)

7.3 Hugseliste – når ei fil er ferdig korrekturlese:

1. Legg til dine egne initialar i filnamnet.
2. Sjekk at alle talarane har fått rett informantkode
3. Sjekk at transkribøren har fylt ut rett namn under *Fil* → *Edit episode attributes*.
4. Sjekk at *report*-feltet er rett utfylt.
5. Formsjekk
6. Eksporter til HTML-format (*Scandiasyn* → *html-filer*)
7. Flytt fila til ”Korrekturlest”-mappa.

