



# LANGUAGE INFRASTRUCTURE MADE ACCESSIBLE

15. januar 2019

## LIA: Translitterering frå dialekt til nynorsk

Av Anneke Askeland, Kristin Hagen, Live Håberg, Janne Bondi Johannessen, Linn Iren Sjønes Rødvand og Eirik Tengesdal<sup>1</sup>

---

<b>1. RETTLEIING FOR TRANSLITTERERING AV LYDNÆR TRANSKRIPSJON TIL NYNORSK .....</b>	<b>2</b>
1.1 X-tagget .....	3
1.2 Orddanning: samansetningar og avleiingar .....	4
1.3 Funksjonsord .....	5
1.4 Pronomen .....	6
1.5 Demonstrativ .....	7
1.6 Bøyingsformer .....	7
1.7 Kjønn og kongruens .....	9
1.8 Stadnamn .....	9
1.9 Spørjeord .....	10
1.10 Sitat, avbrot .....	10
1.11 Preposisjonar .....	10
1.12 Diskurspartiklar .....	11
1.13 Infinitivmerke .....	12
1.14 Interjeksjonar .....	12
1.15 Forkortingar .....	13
1.16 Framleis spørsmål til translittereringa? .....	13
<b>2. TRANSLITTERERING MED PORTALEN .....</b>	<b>13</b>
2.1 Translitterere ny fil .....	13
2.2 Ferdigstille den translittererte fila .....	15
2.3 Opprett ny dialekt .....	16
<b>3. KORREKTURLESING .....</b>	<b>17</b>

---

<sup>1</sup> Rettleiinga bygger på dokumentet «Nordisk dialektkorpus: Oversettelse fra dialekt til bokmål» av Signe Laake i samarbeid med Ingebjørg Før Gjermundsen, Astrid Grov, Kristin Hagen, Janne Bondi Johannessen, Kari Kinn, Alexander Lykke, Eirik Olsen

# 1. Rettleiing for translitterering av lydner transkripsjon til nynorsk

## Kvifor translittererer vi?

Eit av måla for LIA er å gjere talemålsmaterialet lett å forske på. Difor får kvart lydopptak to transkripsjonar som er lenka til kvarandre: ein talemålsnær som er detaljrik og viser mykje uttalevariasjon, og ein ortografisk der uttalevariasjonen er normert og forenkla. Når alle orda i den talemålsnære transkripsjonen har ein ortografisk ekvivalent, er det enkelt å finne fram og søke i materialet. Eit søk på *eg* vil til dømes gi uttalevariantar som *e*, *æ*, *i*, *jæi*, *jei*, *æg*, *eg* og så vidare. Utan den ortografiske transkripsjonen måtte ein kjenne til all uttalevariasjon på førehand.

## Hovudretningslinjer for translitterering:

- Vi normaliserer hardt til nynorsknormalen.
- Der det er valfridom i normalen, vel vi forma som er mest lik bokmål.<sup>2</sup> Vi skriv difor *skole*, ikkje *skule*.
- Dette gjeld ikkje bøyning og kjønn – der skriv vi det som er nærast det informanten seier (t.d. *maskinen* eller *maskina*).<sup>3</sup>
- Vi normaliserer ikkje ordstilling/syntaks, men normerer konsekvent med tanke på morfologi.
- Ord som ikkje finst i Nynorskordboka markerer vi med +x\_ (sjå [1.1](#)). Unntaket er samansetningar der sisteledet står i Nynorskordboka (sjå [1.2](#)).

## Generelt om normeringsprinsippet vårt:

Normeringa er den delen av translittereringsarbeidet som kan vere vanskelegast. Ofte må vi gi slepp på morfologiske særtrekk som kan bere såpass semantisk vekt at tydinga blir endra ved translitterering. Det er viktig å hugse at poenget med den ortografiske transkripsjonen er at han skal vere ei forenkling og ein hjelpereiskap. Har forskaren bruk for fleire detaljar, er den ortografiske versjonen kopla både til den lydner transkripsjonen og til lydfila.

## Vi brukar Nynorskordboka (NNO) som mal og translittererer alltid ord for ord.

Norsk Ordbok 2014 (NO2014) kan vere til hjelp når vi skal normere eller forstå dialektord, men dersom ordet ikkje står i NNO, får det x-tagg (sjå [1.1](#))<sup>4</sup>.

---

<sup>2</sup> Merk at dette ikkje er ei språkpolitisk avgjerd. Vi ønskjer mest mogleg konsistent translitterering slik at det skal bli lettare å søke i materialet etter alle variantar av eit ord eller lemma. Det har vist seg å vere lite effektivt å operere med lister over føretrekte former. Då er det enklare å velje eit prinsipp med mest bokmålsnære former.

<sup>3</sup> Grunnen til at vi ikkje normerer bøyingsformer like hardt som ordformer, er at ein kan fange opp alle bøyingsformene av eit ord ved å søke på lemma.

<sup>4</sup> Norsk Ordbok 2014 har adressa [no2014.uio.no](http://no2014.uio.no), men inneheld førebels berre (*i-åværing*). Vi har fullstendig papirutgåve på kontoret. Ikkje bruk for mykje tid på å finne rett normering. Tagg ordet med +x\_, så blir det eventuelt retta opp i ein seinare korrektur.

Kvart ord i den lydnære transkripsjonen skal altså translittererast med eitt ord – **alltid éin til éin**. Det er ikkje mogleg å legge til ord eller ta vekk ord i berre den eine transkripsjonen. Om eitt ord er feilskriva som to ord i den lydnære transkripsjonen, kan du rette opp dette medan du translittererer i Notepad++.

**NB:** Intervjuaren er allereie transkribert på nynorsk, men translitteratoren «set om» intervjuaren sine replikkar likevel. Det fører mellom anna til ein del og/å-feil m.m. Ikkje bruk mykje tid på intervjuaren, men luk ut direkte feil.

## 1.1 X-tagg

+x\_ brukar vi på same måte som i NoTa og Nordisk dialektkorpus: Dersom ordet ikkje finst i NNO, markerer vi det med +x\_. Dette kan vere dialektord, bokmålsord eller ord frå andre språk – først og fremst leksikalske ord. Vi markerer ikkje leksikalske ord som er brukt i ei anna tyding enn den/dei som står i ordboka:

Informanten seier: *ett mistenngksåmt døttsfall*  
Vi skriv: *eit mistenksamt dødsfall*

Dette gjeld også parverb der transitive verb er brukt intransitivt (eller omvendt):

Informanten seier: *hann ha lakkt i sennga sian klåkka sju*  
Vi skriv: *han har lagt i senga sidan klokka sju*

Informanten seier: *di brannt bål på sanngtansaffi'n*  
Vi skriv: *dei brann bål på sankthansaftan*

**Merk:** Når vi translittererer i Notepad++, skriv vi understrek-teiknet mellom taggen og ordet: +x\_, slik at translitteratoren reknar taggen som ein del av ordet. Slik unngår vi problemet med ulikt tal på ord i dei to versjonane.

Vi brukar ikkje x-taggen hos intervjuaren.

### 1.1.1 Dialektord

Når vi markerer dialektord med +x\_, skal vi normere så godt det lar seg gjere når vi translittererer. Døme:

Informanten seier: *æulette*. Vi skriv +x\_ *augelette*  
Informanten seier: *skvappla*. Vi skriv: +x\_ *skvapla*

Bruk gjerne NO2014 som hjelp til å normere dialektord, men ikkje bruk for lang tid på å leite. Dersom ordet er markert med +x\_, kan vi finne tilbake og normere det seinere. Døme:

Informanten seier: *kural* (som står i NO2014, men ikkje i NNO).  
Vi skriv: +x\_ *kural*

### 1.1.2 Bokmålsord

Bokmålsord som ikkje har ein fonologisk lik ekvivalent i nynorsk, får behalde skrivemåten sin og må ha x-tagg:

+x\_*bolig*  
+x\_*foran*  
+x\_*pike*  
+x\_*rederi*  
+x\_*spise*  
+x\_*værelse*

Er du i tvil om det finst ein nynorsk ekvivalent, er det best og raskast å x-tagge.

Ord frå andre språk får også behalde skrivemåten sin:

Informanten seier: *sjitrævvel*

Vi skriv: +x\_*skittravel* (frå engelsk *travel*)

Lågtyske prefiks som *ad-*, *an-*, *be-* og *hen-* (som ikkje høyrer heime i nynorsk) beheld vi, men ordet som heilskap blir markert med +x\_. Dei andre morfema i orda som har fonologisk like ekvivalentar i nynorsk, normerer vi til den nynorske ekvivalenten. Såleis blir til dømes *begivenhet* translitterert til +x\_*begivenheit*. Døme:

+x\_*anmeldt*  
+x\_*anrop*  
+x\_*begivenheit*  
+x\_*bestyrar*  
+x\_*henstillingar*  
+x\_*utadvend*

Bortsett frå desse prefiksa (som altså utløyser x-tagg), er det *etterleddet* som bestemmer om vi skal ha x-tagg eller ikkje. Vi skriv +x\_ føre samansette og avleidde ord der etterleddet ikkje finst i NNO (sjølv om føreleddet gjer det):

+x\_*begravelse*  
+x\_*vinningsforbrytelse*<sup>5</sup>

Dersom føreleddet ikkje står i NNO, men etterleddet gjer det, markerer vi ikkje ordet, sjå kapittel 1.2.

## 1.2 Orddanning: samansetningar og avleiingar

Dei morfema i bokmålsord som har fonologisk like ekvivalentar i nynorsk, normerer vi til den nynorske ekvivalenten, sjå tabellen under. Merk at suffikset *-else* ikkje har ein slik ekvivalent, og det kan difor ikkje normerast.

---

<sup>5</sup> Merk at *-else* har eit eige oppslag i NNO, men at vi likevel x-ar ord som endar på *-else*.

Avleiingsmorfem	Bokmålsvariant:	Blir i nynorsk:	Døme:
-het		-heit	nyhet → nyheit
-lig		-leg	pluttseli → plutseleg
-som		-sam	grusåmmt → grusamt
fore-		føre-	foregikk → føregjekk
-er		-ar	bestyrer → +x_bestyrar
<b>Leksikalske ord</b>			<i>tilknytte</i> → <i>tilknytte</i>
			<i>antall</i> → +x_antal

Fleire døme: *jaseleg*, *lakreleg*, *festlegheit*, *verksamheit*, *einsamheit*, *sannheita*, *valdsam*, *førekjem*

Eit samansett eller avleidd ord skal berre ha x-tagg dersom *sisteleddet* ikkje står i ordboka. Difor skal ingen av orda i tabellen over (utan dei med prefiks som *ad-*, *an-*, *be-* og *hen-*, jf. 1.1.2) markerast med +x\_.

Andre døme utan +x\_: *holidaybrevet*, *boligkrise*, *føregjekk*, *augeblikkeleg*  
Døme med +x\_: +x\_*forberedelse*, +x\_*forbrytelse*

**NB!** Vi har valt å x-e *selvfølgelig* og *deilig* fordi dei er frekvente ord og dei nynorske ekvivalentane (*sjølvfølgeleg* og *deileg*) er ord vi ikkje finn mange døme på i bruk.

### 1.3 Funksjonsord

Når dialekten har funksjonsord som ikkje finst i NNO i den same tydinga, byter vi ut funksjonsordet med eit standardisert nynorskord utan å markere det med nokon tagg.

Informanten seier: *de va meir sny då så no.*  
Vi skriv: *det var meir snø da **enn** nå*

Informanten seier: *de va hann så kamm.*  
Vi skriv: *det var han **som** kom*

Informanten seier: *ein tå rei messt bLodige drama.*  
Vi skriv: *eit **av** dei mest blodige dramaa*

Om dialekten brukar *te* som infinitivsmärke, translittererer vi til å:

Informanten seier: *de fesstligaste va vell te gå på garen*  
Vi skriv: *Det festlegaste var vel å gå på garden*

## 1.4 Pronomen

Pronomen blir translitterert til standardformene som finst i NNO. Dativ markerer vi ikkje, i samsvar med hovudretningslinje nr. 2: *vi normerer morfologi konsekvent*.

### 1. person:

Alle variantar av 1. person eintal set vi om til **eg**.

I somme dialektar er objektsforma «oss» brukt som subjektspronomen. Dette translittererer vi også som standard nynorsk:

Informanten seier: *åss e fire mann*  
Vi skriv: *vi er fire mann*

Merk at sidan vi har valt å normere til dei formene som er mest like bokmål, translittererer vi til dømes den godkjende nynorskforma **me** til **vi**:

Informanten seier: *me likar kânn i Selljor*  
Vi skriv: *vi likar oss i Seljord*

### 2. person:

**Du** (brukt som objekt) set vi om til **deg**.

**dåkker/dikkån/dåkk** set vi om til **de** (subjektsform) eller **dykk** (objektsform).

### 3. person:

**Ho** i objektsform blir **henne** (som er den bokmålsnære varianten).

**da/ra** for «det» translittererer vi til **det**.

Når dialekten har **den** som personleg pronomen, set vi ikkje om til **han** eller **ho**, men skriv **den**.

Informanten seier: *vi hadde hytte tiliar # mænn vi fikk denn sællt*  
Vi skriv: *vi hadde hytte tidlegare # men vi fekk den selt*

Trykklett **n** og **a** blir omsett til standard nynorsk. Trykklett **n** kan bli omsett til **ein** eller **han**.<sup>6</sup> Vi prøver så godt vi kan å bruke systemet i dialekten når vi avgjer kva pronomen vi set om til, men dette kan vere vanskeleg.

Det ubundne pronomenet **man** blir translitterert til **ein**.

**Der** som formelt subjekt blir ikkje translitterert til **det**.

---

<sup>6</sup> Merk at **n** kan vere dativform av hokjønnspronomenet, og då translittererer vi **n** til **ho**. Derfor blir *æ de bra mæ n F1?* translitterert til *er det bra med ho F1?*

## 1.5 Demonstrativ

I nokre dialektar brukar ein ikkje dei proksimale demonstrativa *denne* og *dette*, men ein kan til dømes bruke *her/herre* eller *der/derre* (eller liknande former) saman med *den/det* for å få fram denne tydinga (som i «den herre»). Dei tostava formene med eit tydeleg r-element translitterer vi til *+x\_herre/derre*. Merk at NNO listar ekvivalentane *herne/derne*, men sidan desse formene er så lite i bruk, vel vi dei meir frekvente bokmålsvariantane og set på x-tag.

Informanten seier: denn **dærre** ## ee sællarn såmm kåmm  
Vi skriv: den **+x\_derre** ## e seljaren som kom

Informanten seier: hann hadde jo tjent nån tus'n på de **hærne** hær  
Vi skriv: han hadde jo tent nokon tusen på det **+x\_herre** her

Tostava former utan tydeleg -r,<sup>7</sup> til dømes *dean* i Alta, skal vi normere til anten *dette*, *denne* eller *desse*, avhengig av konteksten.

Informanten seier: *alltj dean er e rætt*  
Vi skriv: *alt dette her er rett*

Informanten seier: *ja da dennj dean dær kall vi førr N1*  
Vi skriv: *ja da den denne der kallar vi for N1*

## 1.6 Bøyingsformer

### 1.6.1 Substantiv

I det transkriberte materialet varierer bøyingsformene mykje. Vi normerer morfologi og brukar difor berre bøyingsformer som finst i NNO.

Informanten seier: *dæi ee # brukkte å fara ommkrinng i byggdo*  
Vi skriv: *dei e # brukte å fare omkring i bygda*

Dersom det er fleire bøyingsformer å velje mellom i normalen, vel vi den som er nærast det informanten seier.

Informanten seier: *de var denn massjina somm*  
Vi skriv: *det var den maskina som*

Er det vanskeleg å avgjere, vel vi den som ligg nærast bokmål.

---

<sup>7</sup> Alle dei tostava variantane kjem antakeleg opphavleg frå dei norrøne formene *hérna* og *þarna*, og ein kan sjå restar av n-en i nokre dialektar (t.d. *herrane*) (Urd Vindenes, e-post, 23.10.17).

### 1.6.2 Adjektiv

Vi normerer også kongruens:

Informanten seier: *alle var like **fattig** # og alle var like **rik** omtrent*

Vi skriv: *alle var like **fattige** # og alle var like **rike** omtrent*

### 1.6.3 Verb

Tempusformer kan vere vanskelege å normere, men vi forsøker ved å bruke konteksten:

Informanten seier: *datter hannes **ha** jo hæssta ikkje sekkert så mannge # så kannsje **ha** tatt denn ee # gådd denn veien vesst # di flesste måtte reise borrt*

Vi skriv: *dotter hans **har** jo hestar... som kanskje **hadde** tatt den ...*

I mange dialektar er det vanleg at eit verb har supinumsform (under: *sjett*) dersom det kjem etter eit verb i preteritum/preteritum perfektum (her: *hadd verre*): *det hadd verre arti å sjett dæ*. I skriftspråket skal det derimot vere ein infinitiv her, og vi translittererer derfor til infinitivsmærke med infinitiv:

Informanten seier: *det hadd verre arti å **sjett** dæ*

Vi skriv: *det hadde vore artig å **sjå** deg*

Informanten seier: *de va omm å **jorrt** å få tørrke på de*

Vi skriv: *det var om å **gjere** å få tørke på det*

Informanten seier: *æ konnj ittj ha tænngt mæ å **bodd** i Ossjlo*

Vi skriv: *eg kunne ikkje ha tenkt meg å **bu** i Oslo*

Når informanten bruker ei anna tempusform (under: *finnst*) enn det skriftspråket krev for at syntaksen skal bli riktig, normerer vi også:

Informanten seier: *dæ va de væssjte såmm **finnst***

Vi skriv: *det var det verste som **fanst***

### 1.6.4 Stivna kasusendingar

Etter *til* godtar vi stivna kasusformer som *-ar*, *-ars*, *-s* og *-e*. Merk at nokre av desse står i NNO, til dømes *til skamme*. Nokre døme:

*te bryddløps* → *til bryllaups*

*te husar* → *til husar*

*kvalt till døde* → *kvelt til døde*

Generelt kan vi behalde genitivs-s på substantiv som er normerte i NNO utan å sette på x-tag.

Informanten seier: *det er jo så hunjs langt*



Vi skriv: *det er jo så hundens langt*

Vi beheld derimot ikkje kasusendingar når det som blir sagt svarar til faste uttrykk som er leksikaliserte utan kasusending:

Informanten seier: *i lage me*

Vi skriv: *i lag med*

Når det gjeld *i mårrest*, x-taggar vi sidan substantivet her ikkje finst i NNO, men derimot i Bokmålsordboka: *i +x\_morges*.

## 1.7 Kjønn og kongruens

Vi følgjer også NNO når det gjeld kjønn.

Informanten seier: *ei stelle*

Vi skriv: *eit stelle*

For å få rettast mogleg tagging, rettar vi også kongruens:

Informanten seier: *ei fremman stelle*

Vi skriv: *eit framandt stelle*

Når vi rettar kjønn på denne måten, må vi også rette eventuelle pronomer som refererer til substantivet.

Informanten seier: *ei fremman stelle # ho lå ve æLLva*

Vi skriv: *eit framandt stelle # det låg ved elva*

Også perfektum partisipp må kongruere i tal og kjønn med substantivet det står til, både i attributiv (føre substantivet) og predikativ (etter *bli*, *vere*, *verte*) stilling:

*Den brotne armen. Han vart skoten.*

Vi har valt å gjennomføre dette for *alle* partisipp, uavhengig av om verbet er sterkt eller svakt. Dette er for å unngå problem i tagginga.

Informanten seier: *di blæi annmellt*

Vi skriv: *dei blei +x\_anmelde*

## 1.8 Stadnamn

Materialet inneheld mange stadnamn. Ofte er det snakk om små, lokale stader det ikkje er mogleg å finne på eit kart. Mange namn kan ein likevel finne i Norgeskart frå Kartverket: <http://www.norgeskart.no/#!?project=ssr&layers=1005&zoom=4&lat=7197864.00&lon=396722.00>. Ikkje bruk mykje tid på å finne ein normert skrivemåte. Kanskje finst det

ingen offisiell skrivemåte i det heile. Normaliser namnet etter vanlege rettskrivingsreglar. Vi normaliserer bøyingsformer i tråd med det offisielle namnet, slik at til dømes *Finnmarrka* blir translitterert som *Finnmark*. Hugs stor bokstav og at namn med stor bokstav ikkje skal ha x-tagg! Namn skal heller ikkje stå i hermeteikn med mindre det er eit sitat eller er brukt metaspråkleg.

## 1.9 Spørjeord

Spørjeord er ei utfordring. Dei varierer mykje og blir brukte på ulike måtar.

Informanten seier: *kø har du timmesjøkkan ?*

Vi skriv: *kor har du tømmerstokkane ?*

Viss spørjeordet er fonologisk og/eller semantisk svært avvikande frå spørjeorda i NNO, set vi dei om til spørjeord med liknande tyding i nynorsk:

Informanten seier: *kåss æi sku få takka*

Vi skriv: *korleis eg skulle få takka*

Informanten seier: *åssen ska je takke*

Vi skriv: *korleis skal eg takke*

### Nokre spørjeord:

*Å ti* → *kva tid*

*Ka ti* → *kva tid*

*Ko ti* → *kor tid*

*Hen* → *kor*

*Ko/kå* (i visse dialektar) → *kva*

## 1.10 Sitat, avbrot

Sitat og metaspråk som er markerte med hermeteikn, skal også translittererast, med mindre det dreier seg om uttalemåtar:

*ja far seier «tyve» men eg seier «tjue».*

*når dei seier «så sit mi hær» seier vi «så sidde mi hær»*

Ord i hermeteikn skal ikkje taggast med +x\_.

Avbrotne ord blir ikkje omsette.

## 1.11 Preposisjonar

Preposisjonsbruk har stor variasjonsbreidd, og kvar preposisjon har eit stort semantisk verkefelt. Preposisjonsbruk som ikkje samsvarer med ordboksdefinisjonane, lar vi stå:

*Om ferien* → *om ferien* (ikkje *i ferien*)

Samansette preposisjonar som ikkje står i NNO, translittererer vi, men vi markerer dei ikkje.

*Innigjønnå* → *innigjennom*  
*Utforbi* → *utforbi*

Preposisjonar som ikkje står i NNO, set vi om til standard nynorske preposisjonar:

*tå* → *av*

Merk at *nordan* og *sønnan* står i NNO:  
*noLa* → *nordan*  
*sunna* → *sønnan*

## 1.12 Diskurspartiklar

Norske dialektar har mange ulike diskurspartiklar. Nokre av partiklane står i NNO (*jo, vel* m.m.). Når vi har med diskurspartiklar som ikkje står i NNO å gjere, markerer vi dei som vanleg med *+x\_*:

*+x\_gitt*  
*+x\_messom*  
*+x\_jussom*  
*+x\_dama*  
*+x\_måta*  
*+x\_sjø/+x\_serdu* (sjå under)

For mange diskurspartiklar tek vi ikkje omsyn til historisk opphav. Dette gjeld den trønderske *gitt* som kjem av *gut*, og *messom* og *jussom* kjem av *mest som* og *just som*. Når det er tydeleg at orda er brukt nettopp som diskurspartiklar, og blir uttala som eitt ord, normaliserer vi til eitt ord med x-tagget.

Det er mange ulike former – både med ei og to stavingar – som har sitt opphav i *ser du*. Dei formene som tydeleg har to stavingar, normaliserer vi til *ser du* i to ord. For dei formene med berre éi staving, vel vi ei anna løysing. I filer frå Trøndelag og Nord-Møre skal vi normalisere denne diskurspartikkelen til *+x\_sjø*. Dette gjer vi fordi denne forma har blitt eit kjennemerke for denne landsdelen, og kan derfor vere interessant å søke etter. I filer frå Austlandet normaliserer vi former som *sø, su* og *se* til *+x\_serdu*. Slik får vi samla alle uttalevariantane på Austlandet under eitt.

Diskurspartiklane som blir uttala *full, fell, fill* og *fl* kjem av ordet *fulla*, som står i NNO. Bruken av desse samsvarer ofte med bruken partikkelen *vel* i standard nynorsk. Når det er slik, translittererer vi til *vel*:

*de æ fl sjlik* → *det er vel slik*

*å de e full flæire æu -> og det er vel fleire au  
de va full de mæsste-> det var vel det meste  
ja de e fell de -> ja det er vel det*

Dersom informanten seier *fulla*, skriv vi *fulla*.

### 1.13 Infinitivsmärke

Infinitivsmärke er vanlegvis *å* i dei fleste dialektane, men både variantar av *til* (t.d. *te*) og *også* (*åsså*) kan fungere som infinitivsmärke. Vi translittererer begge til *å*:

Informanten seier: *å så fårsette di dao me åsså koka # denn møso*  
Vi skriv: *og så fortsette dei da med å koke # den mysa*

Informanten seier: *de kunne kje vera nåe fælt jillt te væra leggdafållk*  
Vi skriv: *det kunne ikkje vere noko fælt gildt å vere legdefolk*

Også i tilfelle der *til* kan vere både *til* og *å*, translittererer vi til *å*:

Informanten seier: *nå o sysster ho va så svert goe te veva*  
Vi skriv: *når ho søster ho var så svært god å veve*

### 1.14 Interjeksjonar

Mange interjeksjonar eller utrop står ikkje i Nynorskordboka. I lista nedanfor har vi ramsa opp nokre interjeksjonar som likevel ikkje treng x-tagg.

<i>aha</i>	(overraska) Nynorskordboka
<i>ee</i>	(nøling - uansett lengde på ee-n)
<i>eh</i>	(avstandsindikerande)
<i>ehe</i>	(«Eg forstår» - to stavingar)
<i>em</i>	(nøling)
<i>gud a meg</i>	(overraska, utrop)
<i>heh</i>	(imponert)
<i>hm</i>	(spørjande, undrande. I Nynorskordboka i tydinga krenting)
<i>huff a meg</i>	(beklagning) Nynorskordboka
<i>hæ</i>	(spørjande) Nynorskordboka
<i>jaha</i>	(forsterkande «ja») Nynorskordboka
<i>joho</i>	(forsterkande «jo»)
<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)
<i>næ</i>	(overraska, undrande)
<i>nja</i>	(tvilande) Nynorskordboka
<i>nehei</i>	(forsterkande «nei»)
<i>ops</i>	(oi då)
<i>u</i>	(imponert)
<i>uff a meg</i>	(beklagning) Nynorskordboka
<i>ææ</i>	(konstaterande - to stavingar)
<i>å-å</i>	(«oj»)
<i>å ja</i>	(overraska)

## 1.15 Forkortingar

Etter Vinjes Skrivereglar er det valfritt med små eller store bokstavar i forkortingar av fellesnamn (TV/tv, BH/bh WC/wc og så vidare). Vi brukar likevel store bokstavar i slike forkortingar for å gjere det tydeleg at det er snakk om ei forkorting.

Informanten seier: gå *dehå* ettepå # å så bli revisor  
Vi skriv: gå *DH*<sup>8</sup> ettepå # og så bli revisor

## 1.16 Framleis spørsmål til translittereringa?

Om du er i tvil korleis eit ord bør normerast, kan du sjekke opp nettopp det ordet (eller eit liknande døme) i dette Google-dokumentet:

<https://docs.google.com/spreadsheets/d/1uafE4H05yMqkSqJvcZAiTHDgMIRyidMBhMpnWE7dd9s/edit#gid=216145278>

Her finn du i tillegg lenker til dialektressursar. (Meld gjerne ifrå om du veit om nettstader som kan vere hjelpsame.) Om du ikkje blir noko klokare av det, kan du sende ein e-post til [translit-lia@iln.uio.no](mailto:translit-lia@iln.uio.no) – då når du alle dei som jobbar med translitterering, og ein kvalifisert person vil svare deg så fort som råd. I reknearket over fører vi inn alle avgjerder som blir tatt i denne e-postlista (om det ikkje då kjem i rettleiinga).

## 2. Translitterering med portalen

### 2.1 Translitterere ny fil

1. Gå til mappa \LIA\Transkripsjon\Korrekturlest og vel deg ei .eaf-fil (lydnær transkripsjon) du vil translitterere. FLYTT fila, og alle dei tilhøyrande filene – .wav-fil (lyd) og .pfsx-fil (konfigurasjonsfil frå ELAN) – inn i di eiga mappe, slik at ingen andre startar på translitterering av same fila.
2. Gå inn på websida [www.tekstlab.uio.no/translit\\_lia](http://www.tekstlab.uio.no/translit_lia). Bruk same brukarnamn og passord som du brukar i Nordisk dialektkorpus.
3. Du kjem no inn i translitteratorportalen der du kan sjå alle filene som er lasta opp til portalen – både dine egne og andre sine filer.
4. Klikk på *Ny transkripsjon*-knappen:

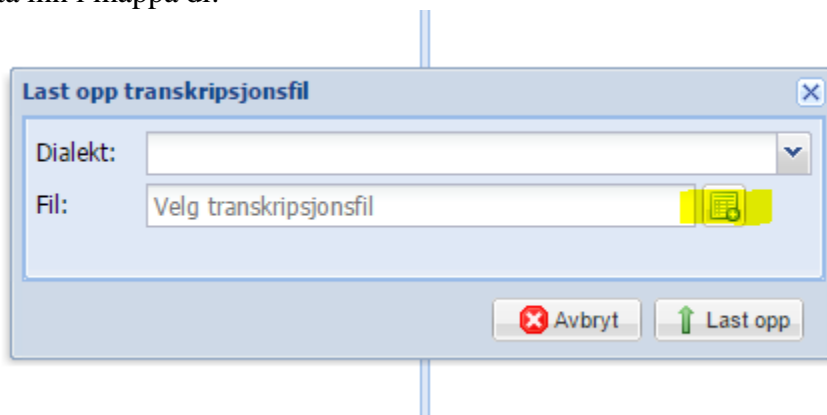
---

<sup>8</sup> Forkorting for *distriktshøgskole*

# Dialekttranslitte

Transkripsjoner	
Ny transkripsjon	Ny dialekt
Dialekt	Transkripsjonsfil
aal	aal_uio_03_et_lisp.c
flaa	flaa_uio_03_mur_lisp.c

I dialogvindauet må du velje både dialekt og transkripsjonsfil. Vel rett dialekt frå den alfabetiske lista i nedtrekksmenyen. **Dersom dialekten ikkje finst i lista, gå til punkt 2.3.** Klikk på symbolet nede til høgre og vel .eaf-fila du akkurat har flytta inn i mappa di:



Klikk på «Last opp». Fila du har valt, kjem opp i oppdelte småfiler (kvar på 200 linjer) i feltet til høgre. Klikk på A-symbolet til høgre for den første fila. Fila blir no translitterert, og når ho er ferdig, får du melding om det.

- Trykk på nedover-pila som no står ved sida av A-symbolet for å laste ned den translittererte fila:



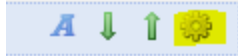
Lagre fila i mappa di. (Lag gjerne ei mappe inni mappa som heiter *Mellombelse filer* e.l. der den translittererte fila kan ligge.) Alle translittererte småfiler har .ort som etternamn.

Opne fila i Notepad++ og rett omsettinga. Alle linjene i den lydnære transkripsjonen startar med @. NB! Om du finn feil i den lydnære transkripsjonen, kan du rette feila i .ort-fila. (Tidlegare måtte ein rette den lydnære transkripsjonen i originalfila.)

- Når du har retta heile fila, skal du først lagre i Notepad++, deretter laste opp fila i portalen ved å trykke på oppover-pila:



- Klikk på det runde tannhjulsymbolet som dukkar opp ved sida av pilene for å trene translitteratoren:



8. Dersom det er ulikt tal på ord i originalen og omsettinga, vil du få ei feilmelding med linjenummer som viser kvar feilen er. Rett opp fila slik at det blir like mange ord i den lydnære og den ortografiske transkripsjonen. Last fila opp på nytt, og tren translitteratoren igjen.
9. Fila er ferdig translitterert!

## 2.2 Ferdigstille den translittererte fila

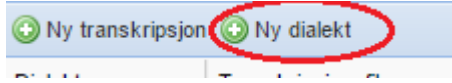
Gå gjennom punkt 2–11 i 2.1 til alle delfilene er lasta ned, translitterert, lasta opp og trent.

1. Når du har trent translitteratoren på alle filene, kjem det eit A-symbol ved sida av filnamnet i oversikta til venstre i biletet. Trykk på dette symbolet for å laste ned den translittererte fila (som .eaf-fil) til mappa di (same mappe som originalen ligg i, ikkje mellombelse filer). Når du lastar ned, blir småfilene sette saman til ei ny fil med lydnær og ortografisk transkripsjon som kvart sitt lag (*tier*). Hugs å legge til teiknet \_ og dine egne initialar før filendinga .ort.eaf (t.d. slik for Kristin Hagen: gausdal\_uio\_03\_\*\*kh.ort.eaf) når du har lasta ned fila. Om du har vore usikker på nokre av orda i omsettinga, kan du finne tilbake til desse og sjå om det hjelper å høyre kva som blir sagt. Korrekturlesinga elles skal gjerast av ein annan person.
2. Flytt den nye ELAN-fila (ort.eaf), konfigurasjonsfila (.pfsx) og wav-fila til mappa *Translitterert til korrektur*. Den originale ELAN-fila (som berre heiter .eaf, ikkje ort.eaf) skal flyttast til mappa *Translitterert til korrektur/Originalfiler*. Overflødige .pfsx-filer kan slettast.
3. Slett småfilene du har translitterert i *Mellombelse filer*.
4. Start på ein ny transkripsjon. ☺

## 2.3 Opprett ny dialekt

Dersom dialekten ikkje allereie finst i translitteratoren:

1. Gå gjennom punkta 1–3 i avsnitt 2.1.
2. Klikk på *Ny dialekt*-knappen:



3. Vel *Opprett ny dialekt* og skriv namnet på dialekten.
4. Gå tilbake (med tilbakepila i nettlesaren) og vel *Sett vektorer* på dialekten du nettopp oppretta. Dialekten som skal telle mest (som liknar mest på dialekten du skal opprette), kan ha t.d. 0.5. Andre kan ha t.d. 0.00 eller 0.1.<sup>9</sup>

---

<sup>9</sup> Forklaring frå utviklaren av translitteratoren:

«Poenget med vektene er å rangere kandidatene i de tilfellene det er mer enn en mulighet, dvs at ordet er sett i flere dialekter. 0.01 betyr egentlig bare "litt mer enn null", slik at hvis ordet kun er sett én gang totalt, så skal det telle, selv om det er fra fx. oslo-dialekt (fx. for sjeldne ord). Hvis du setter vekt 1, betyr det at det teller like mye som den dialekten man transkriberer, og det er jo litt voldsomt. Jeg gjetter på at 0.5 er fint for veldig like dialekter, men det er fint om dere merker dere feil som gjøres, og kanskje søker litt i ordlistene, slik at dere kan stille vektene nøyere etter hvert.»



### 3. Korrekturlesing

Det som er klart til korrekturlesing, ligg i *Translitterert til korrektur*-mappe. Vel ei fil til korrekturlesing. Flytt fila til di eiga mappe.

Les korrektur i ELAN. Det er ikkje meininga at ein skal bruke all verdas lang tid på denne korrekturen, men tanken er å lytte til opptaket og samanlikne med nynorskteksten, for translitteringa er i hovudsak gjort utan å lytte til lydopptaket.

Vel anten *Annotation mode* eller *Transcription mode* for dette i ELAN. Brukar du *Transcription mode*, må du velje *UtteranceType* som *column 1* og *orthography* i *column 2*. Oppsettet kan du velje sjølv. For å få oppsettet som biletet under, trykkjer du på firkantsymbolet til venstre for avspjelingsknappen og dreg *No*-feltet for *Type1: UtteranceType* til høgre.

The screenshot shows the ELAN software interface. On the left is a table with columns for 'No', 'Type2: orthography', and 'Type1: UtteranceType'. The table contains several rows of text, some with green and purple background colors. On the right is a playback control panel with a waveform, volume and rate sliders, and a settings menu with checkboxes for 'Automatic playback of media', 'Create missing annotations', 'Show tier names', 'Colors only on "No." column', 'Navigate across column', and 'Scroll current annotation to center'. A 'Configure...' button is at the bottom of the settings panel.

No	Type2: orthography	Type1: UtteranceType
	khs_orthography	khs
220	var hestane like gode til ta seg fram ?	var hestane like gode til ta seg fram ?
	aal_uio_0302_orthography	aal_uio_0302
221	nei dei var ikkje like til ta seg fram alle hestane og det var best å ha ein # hest som tok godt vegen # +u(ha føre)	næi jæ våro ikkji lik te ta sei framm addle hesstan å de va besst å ha æinn # hess såmm tok gått vegen # +u(ha fyr)
	khs_orthography	khs
222	ettersom eg skjonar # så hadde du festa deg for heile året	ettersom eg skjonar # så hadde du festa deg for heile året
	aal_uio_0302_orthography	aal_uio_0302
223	ja	ja
224	eg hadde festa meg på heile året	e hadde fessta me på hæille åre
	khs_orthography	khs
225	kor mykje hadde du da ## for eit år ?	kor mykje hadde du da ## for eit år ?
	aal_uio_0302_orthography	aal_uio_0302
226	tre hundre og fem og tju kronar	tre hundre å femm å kjue kránu
	khs_orthography	khs
227	ein som berre var borte sommarsvekene # kor mykje	ein som berre var borte sommarsvekene # kor mykje

Bruk retningslinjene frå kapittelet om «Rettleiing for translittering av lyd nær transkripsjon til nynorsk». Mange av feila i dei translittererte filene er truleg ord som ikkje har blitt translitterert, eller frekvente ord som har blitt translitterert feil. Men det kan sjølvstøtt vere andre typar feil òg.

Det er viktig å hugse på at **det skal vere nøyaktig same tal på ord i den originale transkripsjonen som i den translittererte fila**. Dersom du set inn eit nytt ord eller tek bort eit ord i den translittererte fila, må du gjere det same i den originale transkripsjonen.

**Hugs at det er ikkje den originale transkripsjonen som skal korrekturlesast, men den translittererte versjonen.**

Dersom du likevel oppdagar feil i den lydnære transkripsjonen, kan du sjølvsagt rette dei. Oppdagar du at det er mange feil i originaltranskripsjonen, bør både originaltranskripsjonen og den translittererte versjonen korrekturlesast på nytt. Noter deg det i loggen så vi kan lese ny korrektur på fila seinare. I denne omgangen prioriterer vi korrektur av dei translittererte filene.

Legg til initialane dine til sist i filnamnet når du er ferdig med korrekturen, slik som elles. Flytt alle filene til *Translitterert – ferdig korrekturlest*-mappa.