



# Fra Fuglafjørð

INFRASTRUKTUR



Erling Antoniussen



# Fjørður til dialektdatabase

Moderne færøysk er utviklet fra gammelvestnordisk og tales i et stort antall ulike dialekter. I sommer fant nordiske dialektforskere at tiden var inne for en dialektdugnad på Færøyene.

TEKST: KATRINE ZIESLER  
FOTO: TERJE HEIESTAD



I klasserom 303 på Fuglafjørður skole på Eysturoy, blant gjenglemte luer og gymposser fra i fjor, diskuterer to unge færinger fisk. "Hvis eg havi eina spann við deyðum fiskum ... so skal eg koyra alt í henda tangen... oh my dear less, tað luktar illa!"\* Marita (16) holder seg for munnen og himler med øynene mens venninnen Dina (16) fniser og skotter mot videokameraet som filmer de to. For fra klasserommet på Færøyene skal setningen om illeluktende fisk ut på en lang og omstendelig ferd. For første gang i historien skal språket fra atlantehavsøyene innlemmes i et moderne multimedialt nordisk dialektkorpus.

Utenfor rom 303 står professor i lingvistik, Janne Bondi Johannessen, med øret mot døren. Bak henne svirrer forskere og informanter og ved siden av henne står Karine Stjernholm fra Tekstlaboratoriet ved Universitetet i Oslo og poengterer at tyvlytting er vanskelig gjennom lydisolerte klasseromsdører.

– Jeg håper de snakker om kjendiser, sier Johannessen og påpeker at det er grammatiske forskjeller i det færøyske språket, avhengig av om man snakker om dyr, mennesker eller ting. – Vi gir dem en liste med forslag til temaer de kan snakke om, men de får ikke lov å snakke om kjæresten eller familie eller folk de kjenner. Derfor kan det lett ende i omfattende prat om bilmotorer når informantene våre er voksne menn. Det viktigste er likevel at praten foregår på naturlig og avslappet vis, slik at talespråket på Færøyene omsider kan samkjøres med andre nordiske dialekter, sier hun.



### «Hva heter barnet?»

Johannessen er ansvarlig for et stort infrastrukturprosjekt der målet er å samle dialekter i samtlige skandinaviskspråklige nordiske land i en stor dialektdatabase. Tidligere dialektforskning har fokusert mye på intonasjon og endelser, men de siste årene har det vokst fram en stor interesse for setningsoppbygning, altså språkets syntaks. Ved hjelp av en spesiallaget spørreskjemaundersøkelse, som går parallelt med taleinnsamlingen, er et av målene å lage kart over hele Norden påtegnet såkalte isoglosser, eller grenser for dialekttrekk. Tjukke l'er og bløde konsonanter er lokalisert fra før, men grensene for hvor de for eksempel sier «hvor bor du?» og hvor de sier «Hvor du bor?» er nye og ukjente.

– Syntaksforskning har ikke vært utbredt innen dialektologi tidligere, så vi står også overfor en terminologisk utfordring. Hva skal disse fenomenene hete, sier Johannessen.

### Klar tale

I tillegg til databasen skapes et såkalt korpus, som videotapen av jentene skal brukes i. Denne har et mye bredere bruksområde enn databasen og skal kunne benyttes av alle som ønsker å studere talespråk. Her registreres unge og gamle, menn og kvinner fra byer og bygder i de nordiske landene, som snakker fritt med hverandre om ulike temaer. I tillegg kommer mer strukturerte intervjuer med forskningsassistenter som snakker det lokale språket. Alle dataene kan brukes til å drive komparative studier og tverrspråklige undersøkelser, eller man kan se bort fra dialektene om man ønsker å studere talespråk generelt.

– Svært mange forskere vil ha bruk for dette, forteller Johannessen, som har lang erfaring med dialektkorpus. Hun har samlet inn talespråk fra den norske hovedstaden i et oslokorpus, og bidratt til at Norge nå er blant de fremste i verden på talespråkkorpus.



### Historiske endringer

Professoren er ikke redd for at materialet de bruker mye tid og krefter på å samle inn nå, skal bli utdatert om noen år.

– Språk forandrer seg. Vi har opptak i Norge som er 30 år og mer, og som fortsatt er veldig aktuelle. Å samle inn dialekter er viktig for fremtiden og gjør det mulig å spore historiske endringer. Nettopp derfor mente 55 år gamle Ebba Antoniussen at det var viktig å stille opp som informant for dialektforskerne.

– Vi må jo hjelpe de som skal forske på færøysk, sier hun og meldte like godt på mannen Erling (66) også.

– Vi hører godt på dialektene hvem som kommer fra hvilket sted, forklarer han og framhever endringene som har skjedd siden han var ung.

– Språket endres med skolegang. Jeg har bare sju års skole, men skulle gjerne

## FAKTA

Prosjektet «Scandinavian Dialect Infrastructure: Corpus, Database and Dialect Maps» har fått støtte fra NordForsk under initiativet Joint Nordic Use of Research Infrastructure.

Nordisk dialektkorpus og database utvikles i tett samarbeid med partnerne i det nordiske nettverket ScanDiaSyn og det nordiske senteret for fremragende forskning (Nordic Centres of Excellence), NORMS.

Les mer på Nordisk dialektkorpus og databases hjemmesider:  
<http://www.tekstlab.uio.no/nota/scandiasyn/>





## » Når det flunkende nye korpuset og databasen står klare om få år, vil samtlige nordiske språk ligge tilgjengelige på ett sted.

hatt mer. Det var ikke mulig da, så der borte ligger mitt «skoleskip», sier Erling og peker på et gammelt blått skrog som ligger på land ved båthavnen i Fuglafjærdur.

Siden han var barn har også den fysiske infrastrukturen på øyene endret seg. Nye undersjøiske tunneler og gode veier som binder øyene sammen, har gjort ungdomsmiljøene mye større.

– Det gjør nok noe med språket. Det blir mindre markante forskjeller mellom dialektene, mener han.

### Ordmerking krever fagkunnskap

Etter at trailertjåføren Erling og skolejentene Dina og Marita har snakket dialekt foran forskernes kamera, blir det talte språket omsatt til skrift, merket med alle interessante egenskaper og tagget med ordklasser. Hver setning (i den grad man

kan operere med setninger i talespråk – det settes ikke punktum som i skrift) inneholder informasjon om hvem som har sagt den (kjønn, alder), hva (innhold, form på ord), ortografisk form, fonetisk form og grammatisk kategorisering. Dermed blir database og korpus søkbart for grammatiske strukturer og kan avgrenses på de fleste aktuelle måter. – Dette viser hvor viktig det er at oppbyggingen av infrastruktur må være forsker drevet og forskerovervåket, sier Johannessen. – Datainnsamlingen kan ikke settes bort til folk uten forskningserfaring. Innsamlerne må forstå terminologien, se bruksområdene og kjenne feltet godt nok til å kunne lage et brukergrensesnitt som er spesiallaget for språkforskning, understreker hun.

### Dyrt og tidkrevende

Til dialektkorpuset beregner forskerne at en times opptak tilsvarer 20 timers jobb med transkripsjon og tagging. Det er med andre ord kostbart. Mens skriftspråkkorpus kun inneholder transkripsjon, skal man her matche lydfil og video med tekst. Teksten markeres med korresponderende tidspunkt i mediefilene, slik at forskerne kan se informantene uttale setningene de får opp i søk. Det er kompliserte saker.

– Alt talespråket må først transkriberes til skrift, og transkribererne må ikke bare forholde seg til det én person sier, men ofte til to som snakker i munnen på hverandre, med stønning og klikkelyder, smiler fors-

kerne på Tekstlaboratoriet ved Universitetet i Oslo. Det er her transkriberingen foregår, under tett overvåkning av overingeniør Kristin Hagen. Snart skal hun også lære opp en færøysk språkstudent i transkribering, slik at setningene fra Erling og skolejentene legges korrekt inn i databasen.

### Brukervennlig database

– Vi er veldig opptatt av brukervennlighet, forklarer Johannessen, mens hun ivrig viser fram en prøveversjon av korpuset på datamaskinen sin. – Du trenger ikke lese en omfattende manual for å bruke søkesystemet vårt eller slå opp for å finne tagger. Du søker kun ut fra synlige kategorier, på en måte som er selvfølgelig for en gjennomsnittlig databruker. Mange eldre språkdatabaser krever masse opplæring og koder, noe som gjør dem tilgjengelig kun for noen få.

Når det flunkende nye korpuset og databasen står klare om få år, vil samtlige nordiske språk ligge tilgjengelig på ett sted. Da kan forskere fra hele verden klikke seg inn på en passordbeskyttet side, søke på bestemte grammatiske strukturer eller kanskje på ordet fisk. Da vil de, blant trøff fra hele det skandinavisksktalende Norden, få opp to lattermilde jenter fra Fuglafjærdur som snakker om fisk som lukter ille. ■

Les mer på Nordisk dialektkorpus og databases hjemmesider: <http://www.tekstlab.uio.no/nota/scandiasyn/>