

# Innhold

Forord .....	3
Turid Henriksen: Betragtninger omkring lærebøker og lærebokforskning .....	5
Liv Andlem Harnæs: Vokaler hos andrespråksbrukere: en akustisk og perseptuell studie. Vokaler i det norske talespråket til tre utenlandske leger med arabisk som morsmål .....	27
Kari Tenfjord, Jon Erik Hagen, Hilde Johansen: Norsk andrespråkskorpus (ASK) – design og metodiske forutsetninger .....	52
Rita Hvistendahl: Bokomtale .....	82
Bidragstere .....	86
Abstracts and keywords in English .....	87

## NOA norsk som andrespråk

Tidsskriftet *NOA* publiserer artikler innenfor området norsk som andre- og fremmedspråk tolket i vid forstand. Målgruppe er forskere og studenter som arbeider med dette fagfeltet og med beslektede områder innenfor lingvistikk og fremmedspråksfag ved universiteter i Norge og utlandet. En annen sentral målgruppe er lærere, fra grunnskole til voksenopplæring.

Tidsskriftet har vært utgitt ved Universitetet i Oslo siden 1985, med nåværende Institutt for lingvistiske og nordiske studier som utgiver. Ansvarlige redaktører var Anne Golden, Anne Hvenekilde og Else Ryen.

Fram til 2004 var det utgitt 25 nummer av tidsskriftet. Oversikt over innholdet i disse og bestillingsadresse finnes på

<http://www.mamut.net/novus/shop/>

Fra 2005 utgis tidsskriftet av Novus forlag, for tiden med Kirsten Palm, Institutt for lærerutdanningsfag, Høgskolen i Telemark, Else Ryen, Institutt for lingvistiske og nordiske studier, Universitetet i Oslo og Hilde Sollid, Institutt for språkvitenskap, Universitetet i Tromsø, som redaktører.

Opplysninger og bestillingsadresse

<http://www.novus.no/pages/NOA.html>

## Forord

I artikkelen *Betraktninger om lærebøker og lærebokforskning* setter Turid Henriksen lærebokforskning og lærebøker i et tekst- og kommunikasjons-teoretisk perspektiv. Den kommunikasjonsmodellen hun bruker, er hentet fra den franske språkforskeren Chauradeau, og Henriksen viser hvordan både samfunnet, fagfeltet og kommunikasjonspartnerne er medbestemende for den konkrete utformingen av en læreboktekst. Hun drøfter også det hun kaller lærebøkernes doble kommunikasjonsituasjon. Den første kommunikasjonsituasjonen dreier seg om forholdet mellom lærebokforfatteren og elever/lærere som brukere av boka, mens den andre handler om kommunikasjonen mellom de ulike brukerne av boka. Henriksen viser hvordan lærebokforskning kan ta utgangspunkt i ulike sider ved denne kommunikasjonsmodellen. De eksemplene Henriksen trekker fram i artikkelen, er i hovedsak hentet fra lærebøker i fremmedspråk, men disse har en klar overføringsverdi til lærebøker i andrespråksopplæring.

Liv Andlem Harnæs' artikkel *Vokaler hos andrespråksbrukere: en akustisk og perseptuell studie. Vokaler i det norske talespråket til tre utenlandske leger med arabisk som morsmål* tar som utgangspunkt at innflytelse fra morsmålet er en viktig kilde til uttaleavvik hos andrespråksbrukere. Målsettinga er å gi en akustisk og perseptuell beskrivelse av vokalkvalitet i det norske talespråket til tre irakiske leger med arabisk som morsmål. Norsk (standard østnorsk) har ni vokalfonemer i trykksterk stilling, mens arabisk har tre par vokalfonemer. Hypotesen er at arabisktalende (og andre) vil få problemer med de norske vokalene som ikke finnes i innlærernes morsmål, og Harnæs avgrensar sin studie til de tre fremre runda vokalene /æ/, /y/ og /ø/. De tre irakiske informantene blir sammenlikna med norske morsmålsbrukere med østnorsk talemål. Studien bekrefter hypotesen, men resultatene til Harnæs viser også interessant variasjon mellom informantene.

I artikkelen *Norsk andrespråkskorpus (ASK) – design og metodiske forutset-*

#### 4 Forord

*ninger* presenterer Kari Tenfjord, Jon Erik Hagen og Hilde Johansen et nytt korpus med data fra innlærere av norsk. Dataene i ASK består av vel 1700 tekster skrevet av kandidater som har tatt *Språkprøven* og *Test i norsk – høyere nivå*. I tillegg har korpuset personopplysninger fra de samme kandidatene, samt et kontrollkorpus som består av 200 tekster skrevet av personer som har norsk som sitt morsmål. En hovedmålsetning i artikkelen er å drøfte bruk av feilkoder i korpuset og om bruk av feilkoder fører til at korpuset har havna i ”sammenlikningsfella” slik den er beskrevet av Bley-Vroman i en artikkel fra 1983. ”Sammenlikningsfella” dreier seg om at feilkoding i slike korpus gjøres på målspråkets premisser, og at mellomspråk ved en slik praksis ikke har status som selvstendige språkssystem. Tenfjord, Hagen og Johansen argumenterer for at ASK ikke kan anklages for å ha havna i sammenlikningsfella, men at feilkodinga må betraktes som en teorinøytral beskrivelse og en metode som senere teoretisk motiverte analyser kan bygge på.

Redaktørene

# Betraktninger omkring lærebøker og lærebokforskning

*Av Turid Henriksen  
Universitetet i Oslo*

## **Sammendrag**

Målet med denne artikkelen er å beskrive lærebøker og lærebokforskning i et tekst- og kommunikasjonsteoretisk perspektiv. I første del av artikkelen blir lærebøker tentativt definert og sett i forhold til pedagogiske tekster og didaktisitet. Deretter skisseres en kommunikasjonsmodell der tekstproduksjon beskrives som resultat av en dobbel prosess som tar hensyn både til samfunnet, fagfeltet og kommunikasjonspartnerne. Innenfor denne modellen vises det til hva som er særpreget for lærebøker.

I andre del av artikkelen gis det eksempler på hvordan lærebokforskningen kan knyttes til de ulike parametrene i kommunikasjonsmodellen. De konkrete eksemplene det vises til er vesentlig hentet fra lærebøker i fremmedspråk. Lærebøker i andrespråk omtales først i slutten av artikkelen, men eksemplene hentet fra fremmedspråk antas å ha relevans også for andrespråkforskningen.

*Nøkkelord: lærebøker, pedagogiske tekster, didaktisitet, fremmedspråk, kommunikasjonsmodell.*

## **Innledning**

Hva er ei lærebok? Hva er lærebokforskning? Tradisjonelt har det vært lett å besvare det første spørsmålet. Termen ”lærebok” er som oftest forbundet med skole og undervisning, og alle som en gang har vært elev og brukt slike bøker, vet hva de er. Men lærebøker er langt fra stabile størrelser, selv om

“läroböcker ärver läroböcker”.<sup>1</sup> På slutten av 1800-tallet innehold f.eks. norske lærebøker i fremmedspråk alt man trengte for en tidsmessig undervisning: tekster, grammatikkparagrafer, øvelser og ordlister. I løpet av det 20. århundre, spesielt siste halvdel, ble det vanlig å dele opp læreboka fysisk i flere deler: tekstbok, elevbok med øvelser og en spesiell bok for læreren, lærerveiledningen. Det ble også vanlig å knytte lydmaterialer til bøkene. Vi får en form for spesialiserte læremiddelpakker. Selv om vi i de aller nyeste bøkene i fremmedspråk igjen kan finne samlet i ei bok både tekster, øvelser, grammatikk og ordlister, kan dagens lærebøker ha svært kompleks utforming. Et eksempel på dette er franskboka *Ouverture* (Christensen & Wulff 2006–2008), som består av elevbok med oppgaver, elev-CD med tekster fra elevboka, lærerveiledning, lærer-CD og nettsider. Slike moderne lærebøker kan være så interaktive og åpne at elevene selv er med på å skrive lærestoffet, noe Christensen & Wulff er eksempel på. Ved bruk av den metoden som er kjent under navnet ”Storyline” (<http://www-lu.hive.no/storyline/metoden/>) skal elevene skape sine egne fiktive ”venner” og forsyne dem med identitet og liv. Spørsmålet ”Hva er ei lærebok?” er derfor ikke fullt så lett å besvare i dagens samfunn.

Hvordan man skal svare på spørsmålet om hva lærebokforskning er, er heller ikke enkelt. Innfallsvinklene er mangfoldige og avhenger både av utgangspunktet for analysen<sup>2</sup> og hensikten med den. Med de moderne, interaktive og delvis virtuelle lærebøkene blir også forskningsobjektet mer komplekst. Er det tekstboka med eventuelle andre fysiske bøker, internett-sidene eller cd-rommene som skal undersøkes, eller hele læremiddelpakken? Uansett hvordan læremidlet som objekt er utformet, fysisk og/eller virtuelt, skriver det seg inn i en historisk, geografisk og pedagogisk kontekst som er medbestemmende både for innhold og utforming, og som man ikke kan komme utenom i læremiddelforskningen.

Målet med denne artikkelen er å presentere lærebøker og lærebokforskning generelt, ikke bare i fremmedspråk, i et tekst- og kommunikasjons-teoretisk perspektiv. Jeg vil først omtale lærebøker og lærebokforskning i forhold til forskning på pedagogiske tekster og til begrepet ”didaktisitet”. Deretter vil jeg beskrive læremidler ut fra en kommunikasjonsmodell som tar hensyn både til samfunnet, lærebokforfatteren, elevene, lærerne og læreboksjangeren. Med denne modellen som utgangspunkt vil jeg vise noen innfallsvinkler til lærebokforskning, selv om jeg i denne artikkelen ikke kan omtale hele komplekset. Hensikten med denne delen av artikkelen er også å gi ideer til hva som er mulige forskningsobjekter. Siden mitt eget forsk-

ningsfelt er lærebøker i fremmedspråk, vil de eksemplene jeg viser til, vesentlig være hentet derfra, men det vil også være et par eksempler fra norsk som morsmål og religion. Først i sluttordet til artikkelen vil jeg omtale lærebøker i andrespråk. Siden dette ikke er min spesialitet, vil imidlertid det jeg har å si være begrenset. Men fordi fremmedspråklæring har mange fellestrekk med andrespråklæring, vil de eksemplene jeg har gitt uten vanskelighet kunne overføres til forskning på lærebøker i andrespråk.

## **Lærebøker, pedagogiske tekster, didaktisitet og læremiddelforskning**

### *Lærebok – en definisjon*

I en situasjon der læremidler ikke lenger bare er ei bok, kan en spørre om det er mulig å gi en definisjon av begrepet ”lærebok”? Finnes det en fellesnevner, uavhengig av den fysiske eller virtuelle utformingen, som kan hjelpe oss? Min egen tentative definisjon, som tar utgangspunkt i kommunikasjonssituasjonen, intensjonen med læreboka, og de som skal bruke den, er følgende:

”Lærebøker er multimediale tekster som skal formidle kunnskaper, ferdigheter og holdninger slik at elevene kan lære av dem og lærerne undervise etter dem.”

Denne definisjonen slår fast at mediet er komplekst, at intensjonen er kunnskapsformidling, og at brukerne utgjør to forskjellige grupper, hvorav den ene skal lære noe og den andre skal ”lære bort” noe. Ei lærebok bli på denne måten ikke bare ei lærebok, men også ei undervisningsbok, noe som man ikke alltid tar med i betraktningen når læreboksjangeren omtales. I tradisjonell klasseromundervisning er begge brukergruppene til stede, og det oppstår en samhandling mellom dem, med boka som utgangspunkt. Selv om maktforholdene mellom lærer og elever er noe annerledes i det moderne klasserommet, er det fremdeles en samhandling mellom de to gruppene. Dette utgjør, etter min mening, et av særtrekkene ved kommunikasjonssituasjonen for lærebøker.

### *Pedagogiske tekster*

I moderne undervisning brukes imidlertid ikke bare læremidler slik vi har omtalt dem ovenfor. Andre tekster, fysiske eller virtuelle, som ikke i ut-

gangspunktet er produsert for en undervisnings- eller lærings situasjon, trekkes også inn: avistekster, filmer, lydopptak, videoutdrag, f.eks. fra Youtube, nettsider etc. Slike tekster, tidligere gjerne omtalt som ”autentiske dokumenter”, har lenge vært et begrep i fremmedspråkundervisningen, der det å lære seg autentisk språkbruk, gjennom eksempler nettopp på autentisk språkbruk, var et ideal i begynnelsen av det man kaller kommunikativ språkundervisning. Slike tekster ble regnet som et viktig supplement til lærebøkene forenklete og didaktiserte språk. Lærebøkene, sammen med andre tekster, utgjør på denne måten et konglomerat som etter hvert er blitt omtalt som ”pedagogiske tekster”.

Begrepet ”pedagogiske tekster” er blitt definert av den svenske læremiddelforskeren Selander, som litt etter litt også lar det omfatte tekster som er produsert utenfor utdanningsinstitusjonene. I 1995 nevner han ”autoriteter” som skolen, arbeidsmarkedet, TV-nyhetene og helsevesenet som ”produsenter” av pedagogiske tekster, og han omtaler mottakerne som elev, klient, pasient, samfunnsborger osv. Hans brede definisjon av pedagogiske tekster er følgende: “With “pedagogic texts/educational media” we shall mean not only school textbooks but the whole range of possible text + pictorial illustrations + films + computer programmes which are produced for educational purposes, i.e. to inform and convince the reader, viewer or listener that the information and perspectives presented are true and correct.” (Selander 1995: 152). Som det framgår av denne definisjonen, ser Selander den pedagogiske teksten med retorikkens øyne, der det gjelder å overbevise mottakerne om at informasjonen er sann og korrekt. Han foreslår også å analysere pedagogiske tekster med utgangspunkt i begreper fra retorikken.

Det er lett å slutte seg til en begrepsutvidelse fra lærebøker/læremidler til pedagogiske tekster som ikke bare knyttes til utdanningsinstitusjonenes undervisning, men også til andre som er i posisjon til å overtale og belære. En grunn til dette er at læring og belæring også foregår i uformelle situasjoner. Det finnes blant annet bøker for selvstudium, instruksjonsbøker, bruksanvisninger og oppskrifter av ulike slag. Man kan derfor anta at det eksisterer spesielle forhold ved kommunikasjonssituasjonen, eventuelt også med teksten, som gjør en pedagogisk tekst til en pedagogisk tekst og et læremiddel til et læremiddel, uten at disse er laget for en utdanningsinstitusjon.

### *Didaktisitet*

Det er nettopp det spesifikke ved tekster brukt til undervisning og læring som har vært i sentrum for forskerne knyttet til det franske forskningspro-



grammet CEDISCOR.<sup>3</sup> Gruppen har interessert seg for kunnskapsoverføring<sup>4</sup> og har forsket på alle former for det de kaller ”didaktisk diskurs”. Didaktisk diskurs i ”ordinære tekster”, dvs i tekster som ikke institusjonelt er knyttet til undervisning og læring, f.eks. presse og media, har vært et privilegert område. I stedet for å ta utgangspunkt i en institusjon eller en autoritet, slik Selander gjør, tar gruppen utgangspunkt i det stipulerte begrepet didaktisitet, som defineres ut fra tre kriterier: en didaktisk situasjon, formelle trekk ved språkbruken eller teksten og den pragmatiske intensjonen med den kommunikative samhandlingen. (Moirand 1992:10). En ”didaktiske situasjon” oppstår når den ene av kommunikasjonspartnerne har større kunnskaper enn den andre og har til oppgave, eller ønsker, å dele denne kunnskapen med den andre. ”Didaktisitet” som formelle trekk kan observeres konkret i tekstene, f.eks. som definisjoner, eksemplifisering, ikoniske virkemidler etc. ”Intensjonen” knyttet til den kommunikative samhandlingen er ”å få noen til å forstå, gjøre, lære og tro”.

#### *Lærebøker i et utvidet tekstperspektiv*

Som det framgår av forskningen både til Selander og til CEDISCOR-gruppen, produseres det pedagogiske tekster eller tekster med didaktiske trekk for andre sammenhenger enn formelle utdanningsinstitusjoner. Dette faktum kan sette den tradisjonelle læreboka inn i et større tekstperspektiv, både som tekstlig fenomen og som forskningsobjekt.

Etter hvert som interessen for lærebøker og pedagogiske tekster har økt, har læremiddelkunnskap også oppstått som akademisk disiplin. I norsk sammenheng er det særlig Høgskolen i Vestfold (HIVE) som peker seg ut, både som utdanningsinstitusjon med mastergrad i pedagogiske tekster, og som pådriver for forskning på slike tekster. Forskerne ved skolen overtar Selanders definisjon av pedagogiske tekster og hevder at feltet er karakteristisk for sitt tverrfaglige og tverrvitenskapelige preg. (Knudsen, Skjelbred & Aamotsbakken 2007:7.)

### **Lærebøker og verden: kontekst og kommunikasjonssituasjon**

#### *En mulig kommunikasjonsmodell for lærebøker*

Som jeg har pekt på i innledningen, oppstår lærebøker i en historisk, geografisk og pedagogisk kontekst som er med på å bestemme innhold og utformning. For lærebokforskningen betyr det at både den ytre samfunnskonteksten og

kommunikasjonssituasjonen der læremidlet utgjør ”budskapet”, må trekkes inn. En modell som beskriver kommunikasjon vil her være klargjørende. Det eksisterer imidlertid flere kommunikasjonsmodeller, fra de enkle lineære, med en ”sender” et ”budskap” og en mer passiv ”mottaker”, til de komplekse modellene som ser kommunikasjon som prosess og samhandling, men alle er ikke like egnet for lærebokforskning.<sup>5</sup> Den franske språkforskeren Charaudeau (f.eks. Charaudeau 1995), har utviklet en modell som beskriver kommunikasjon i en dobbel prosess, en transformasjonsprosess, der en del av ”den utenomspråklige verden”, f.eks. et kunnskapsområde, blir ”språkliggjort”, og en transaksjonsprosess, der budskapets endelige utforming skjer som resultat av en samhandling mellom kommunikasjonspartnerne i en bestemt situasjon. For at en slik transaksjonsprosess kan finne sted, må det foreligge et kontraktlignende forhold mellom kommunikasjonspartnerne, de må godta hverandres roller. I transformasjonsprosessen identifiseres de fenomenene som skal kommuniseres, og i transaksjonsprosessen kan den som skal kommuniserer noe, bruke individuelle strategier, tilpasset situasjonen, for å nå fram til kommunikasjonspartneren. Av disse to prosessene er det transaksjonsprosessen som er den avgjørende for det endelige språklige produktet. De språklige produktene, tekstene, oppfattes på denne måten som resultatet av kommunikative samhandlingene som er utformet både ut fra de begrensningene kommunikasjonssituasjonen i vid forstand setter, og ut fra de individuelle strategiene den kommuniserende bruker.

Det som tiltaler meg ved modellen, er at den ivaretar både språklig og sjangermessig variasjon og individuell frihet. Kommunikasjonssituasjonen setter begrensninger, samtidig som den kommuniserende tyr til egne strategier. Disse individuelle strategiene omtales, i Charaudeaus terminologi, som ”iscenesettelser”, noe jeg vil komme tilbake til senere. Det er også en fordel ved modellen at ”språkliggjøringen av verden” ses som en dobbel prosess, der det tekstlige budskapet avhenger både av kunnskapsområdet og de begrensningene og strategiene som kommunikasjonssituasjonen gir rom for. I analysen av lærebøker, der både språk og tekster er forskjellige fra de vi finner i andre sammenhenger, kan en modell som den jeg har beskrevet, føre til en fyldigere og mer kompleks beskrivelse av studieobjektet. For begynnerbøker i fremmedspråk og andrespråk, der det er språket selv, språkformer og bruk av former, som er lærestoffet, er denne modellen spesielt godt egnet. Jf. avsnittet ”Individuelle strategier – iscenesettelse” senere i denne artikkelen.

Modellen kan belyses gjennom produksjonen av ei lærebok i engelsk i Norge etter Kunnskapsløftet 2006 (heretter K06). Vi har et fremmedspråk

som undervises i norsk skole og som skal utformes ut fra gjeldende skolelover, forskrifter og ikke minst, læreplaner. Både lærebokforfatteren<sup>6</sup> og mottakerne har her bestemte roller. Lærebokforfatterne skal lage ei lærebok som gir lærerne et grunnlag for å undervise og elevene et grunnlag for å lære. Disse kommunikasjonspartnerne må godta hverandres roller, det må eksistere en form for kontraktliknende forhold mellom dem. Lærebokforfatterne må utforme læreboka slik at den gjenkjennes som lærebok av de som skal bruke dem. Situasjonen setter begrensninger både til hva som skal formidles, og hvordan det skal gjøres. Lærebokforfatteren har likevel en individuell frihet til å utforme budskapet ut fra sine egne strategier for å fange mottakernes interesser. Det er her han kan iscenesette stoffet ut fra sine egne individuelle ideer og oppfatninger.

### *Lærebøkernes doble kommunikasjonssituasjon*

Den skissert modellen gjør rede for de to prosessene fram til ei ferdig lærebok. Men kommunikasjonen rundt lærebøkene som tekst, eller budskap, slutter ikke ved utgivelsen. Det er etter at bøkene er kommet ut på markedet at deres egentlige eksistens begynner. Lærebøker er ikke, som f.eks. romaner, laget primært kun for å leses. Man skal også lære av dem. De inngår da inn i en ny samhandling, der den opprinnelige kommuniserende, lærebokforfatteren, ikke lenger deltar. Man kan argumentere med at det også skjer med andre tekster. Romaner kan bli gjenstand for debatt, de anmeldes, og skuespill blir oppført. Lærebøker, det tenkes her på skolebøker, er likevel annerledes fordi de, rent institusjonelt, forutsetter denne nye interaksjonen. Den finner sted i undervisningssituasjonen, som tradisjonelt har foregått mellom lærer og elever, men som nå etter hvert like godt kan foregå mellom elevene. Dette er en samhandling der lærebokas innhold blir videre bearbeidet. Den komplekse samhandlingsmodellen for lærebøker kan visualiseres slik, i to etapper:

(1) lærebokforfatter > lærebok < lærere/elever

(2) lærer (elever) > lærebok < elever

Det gir etter min mening et riktigere bilde av lærebøkene dersom man ser dem som budskap i en kompleks kommunikasjonssituasjon med to etapper. I den første etappen er det viktig å peke på at lærebøker har både elever og lærere som brukere. Både i tekstbøkene, og ikke minst i lærerveiledningene, er det læreren som er den egentlige kommunikasjonspartneren for lærebokforfatteren. I et forskningsperspektiv kan man f.eks. fokusere nettopp på hvordan lærebokforfatteren ser på lærerne som skal bruke boka, og hvilke

”råd” de får. I den andre etappen er det kommunikasjonen mellom brukerne som er det viktigste. I lærebokforskningen vil man med denne siste etappen som utgangspunkt f.eks. kunne kartlegge hvordan klasseromsinteraksjonen foregår, og i hvor høy grad læreboka legger opp til det. Jf. avsnittet ”Kommunikasjonspartnerne – lærebokforfatteren og brukerne” senere i denne artikkelen.

I tillegg til lærebokas innhold og utforming, kan alle faktorene ved kommunikasjonsprosessen bli tatt som utgangspunkt: den situasjonelle konteksten læreboka oppstår i, lærebokforfatteren i sin rolle som lærebokforfatter, brukerne og kommunikasjonen dem imellom.

## **Lærebøker og lærebokforskning i lys av kommunikasjonsmodellen**

### *Lærebøker og samfunnet*

La oss begynne med den ytre konteksten.<sup>7</sup> Lærebøker er produkter både av de samfunnsforholdene de oppstår i på et bestemt tidspunkt, og av det kunnskapsområdet de forholder seg til. Slik sett vil selve kunnskapsområdet også være en del av den ytre konteksten og gjenstand for variasjon. La oss ta fremmedspråk som eksempel nok en gang. Vitenskapen om språk har endret seg gjennom tidene, og ei lærebok i fremmedspråk fokuserer derfor på andre språklige fenomener nå enn for femti år siden. I lærebokforskningen kan dette gjøres rede for ved å sammenlikne lærebøker i et historisk perspektiv.

Det finnes lærebøker som har sterkt tilknytning til kunnskapsområdet og løser tilknytning til det samfunnet de oppstår i. Med det mener jeg å si at disse bøkene ikke er spesielt avhengige av ytre forhold og reguleringer, som læreplaner og pedagogiske direktiver. Dette gjelder f.eks. generelle lærebøker i akademiske fag for universitetsstudier. Her vil sannsynligvis faglig nivå normalt, men ikke alltid, spille en større rolle enn type institusjon og lokale fagplaner. Det gjelder f.eks. også lærebøker i sjakk, esperanto eller andre kunnskapsfelt eller ferdigheter som ikke er vanlige skolefag. Ved nærmere ettersyn kan det nok likevel vise seg at også slike mindre kontekst-avhengige lærebøker bærer preg av å være skrevet, om ikke for undervisning, så i alle fall for læring. De vil på en eller annen måte være tilrettelagt for at eleven skal kunne forstå hva han skal tilegne seg. Det vil si at det er selve intensjonen med kommunikasjonshandlingen ”å få noen til å forstå” som gir det ferdige produktet deres særtrekk.

Et eksempel på dette er ei bok med tittelen *Fransk handelskorrespondanse. Til bruk for handelsgymnasier og forretningsmenn* (Chauvin 1941), som inneholder modelltekster for handelskorrespondanse. Tittelen tyder på at boka både var laget for undervisning i en utdanningsinstitusjon og for forretningsmenn i praktiske situasjoner. Uansett brukere har boka didaktiske trekk som også knytter den til det norske samfunnet. Den inneholder en ordliste fra norsk til fransk, og mønsterbrevene omtaler norske forhold. Norske institusjoner, f.eks. Den norske Creditbank (Chauvin 1941: 14), forekommer i brevene. Avsenderne for noen av dem er norske firmaer, og brevene er sendt fra norske byer. Det handler om eksport og import av produkter, f.eks. tran og vin, mellom Norge og Frankrike. Lærebokforfatteren Chauvin<sup>8</sup> har iscenesatt sitt lærestoff slik at det tilpasses de potensielle brukerne og det samfunnet læreboka er produsert i.

De lærebøkene som brukes i en institusjonell undervisningssituasjon, er på en helt annen måte innskrevet i en kontekst med læreplan, fagets status, skoletype, timeantall etc., som varierer i tid og rom. K06 kan nok en gang tjene som eksempel. I forarbeidene til K06 ble det bestemt at alle fag skulle inneholde fem grunnleggende ferdigheter. Denne bestemmelsen kan vi gjenfinne både i læreplanene og i lærebøkene, f. eks. i fremmedspråk, der fokus er satt på tall og beregninger på en helt annen måte enn i tidligere bøker. Det ble videre bestemt at læremålene skulle beskrives som kompetansemål, noe som i flere skolefag førte til at det i læreplanene ikke er noen beskrivelser av det konkrete innholdet i undervisningen. Det som er interessant for lærebokforskeren når det gjelder fremmedspråk, er da å undersøke hvordan kompetansemålene konkret får sin utforming i lærebøkene.

Læreplaner oppstår heller ikke bare i et nasjonalt rom, men følger gjerne internasjonale tendenser når det gjelder syn på fagfelt, læring og undervisning. Her er det mange eksempler å vise til. Da den direkte metoden ble innført for faget engelsk i Norge i 1925, var det som resultat av internasjonale påvirkninger. Et av de viktigste trekkene ved den direkte metoden var at man ikke skulle bruke oversettelse til morsmålet som tilgang til meningen med tekstene, noe som igjen fikk konsekvenser for lærebøkernes utforming. Et mer aktuelt eksempel er Europarådets rammeverk for språkundervisning (*The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*), som har satt et sterkt preg på norske læreplaner i engelsk og fremmedspråk etter K06. Engelskplanen er likevel forskjellig fra læreplanen i "fremmedspråk", fordi engelsk har en annen status i norsk

skole, og i det norske samfunnet forøvrig, enn de andre fremmedspråkene. Også andre tekster, f.eks. offisielle rundskriv o.l. tilhører lærebøkens samfunnskontekst. Lærebøkene inngår på denne måten både i et nasjonalt og internasjonalt tekstunivers som i noen tilfeller må trekkes inn dersom man fullt ut skal forstå læreboka.

En vanlig innfallsvinkel til lærebokforskning er derfor nettopp å se innholdet i lærebøkene i lys av læreplanen, som ”den oppfattede læreplan” i forhold til ”den formelle læreplan”. Lærebokforskning kan på denne måten bli en del av læreplanforskning, som igjen kan være en del av forskning på skolens innhold. (Gundem 1989).

I lærebokforskning som ser læreboka opp mot læreplanene, kan det være fristende å kritisk gå på jakt etter mangler. Man finner da sannsynligvis raskt ut at læreboka ikke oppfyller læreplanen, kanskje fordi den ikke inneholder de og de momentene som er nevnt i planen. Men det spørres om denne tilnæringsmåten alltid er den optimale. En alternativ måte å gå fram på er å undersøke *hvordan* lærebokforfatteren har tolket sin oppgave, på hvilken måte han har oppfattet læreplanen, og hvilke valg han har gjort. De læreplanene vi har i Norge etter K06 er åpne planer, der det er sluttkompetansen som beskrives, ikke det konkrete innholdet. Med slike planer kan det være vanskeligere å påvise at læreplanene ikke er oppfylt.

Et relevant eksempel på denne problematikken er de tolkningene som er gjort av læreplan for fremmedspråk etter K06. I denne planen, som er delt i hovedområdene *Språklæring, Kommunikasjon og Språk, kultur og samfunn*, er et av kompetansemålene under *Språk, kultur og samfunn*, nivå II, at eleven skal kunne ”drøfte sider ved livsvilkår og samfunnsforhold i språkområdet”. Flere av bøkene som er skrevet eller revidert etter K06, inneholder faktaopplysninger om mållandet på norsk. For den som analyserer læreboka, vil det være relevant å spørre seg hvorfor lærebokforfatteren har valgt norsk, og om dette betyr at det anses for legitimt å lære seg ”kulturkunnskap” på norsk, noe som igjen har betydning for forholdet mellom språk og kultur i fremmedspråkundervisningen. Man kan også spørre seg om faktaopplysninger om landet, på norsk eller på fremmedspråket, gir elevene godt nok grunnlag for å kunne ”drøfte sider ved livsvilkår”, uansett om drøftingene foregår på norsk eller på fremmedspråket. Og da kan det være grunn til å se på andre sider ved læreboka, f.eks. oppgavene, for å undersøke hvorvidt de gir elevene øvelse i å drøfte samfunnsforhold. Dette er gjort f.eks. i masteroppgavene til Nygård (2006) og Mandt (2008).

Men konstituerende tekster som læreplaner og direktiver er bare én av innfallsvinklene til lærebokforskning ut fra samfunnskonteksten. Det er ikke uvanlig at lærebøkene ses opp mot samfunnet generelt, uavhengig av læreplaner. Spesielt gjelder dette lærebøker i historie- og samfunnsfag, der kunnskap om samfunnet selv utgjør fagstoffet. I mye av det man kan kaller ideologisk lærebokforskning, er lærebøkene sett opp mot den historiske eller samfunnsmessige virkeligheten. (Johnsen 1993).

Hvis lærebokforskningen ikke knyttes til den umiddelbare samfunnskonteksten med skole, læreplaner etc., men til nasjonen eller verden generelt, er det bare fantasien som setter grenser for hva som kan bli gjenstand for forskning. Oxfeldt (2008) ser f.eks. på hvordan kunnskap knyttet til "det hjemlige" framstilles i forhold til kunnskap om andre samfunn, blant annet med utgangspunkt i hjemlige og fremmede dyr i P.A. Jensens *Læsebog for Folkeskolen og Folkehjemmet*.

I mange slike tilfeller vil lærebokforskningen knyttes til overordnede moralske, etiske eller samfunnsmessige spørsmål. Breilid & Nicolaisen (2002) er et eksempel på dette. Forfatterne har sett på hvordan lærebøker og lærerveiledninger i religion framstiller kvinnen *Sita* i hinduismen, og hvordan henholdsvis jøder og egyptere framstilles i fortellingen om exodus.

Vi kan også nevne noen eksempler fra forskning på lærebøker i fransk som fremmedspråk. I likhet med lærebøker i historie og samfunnsfag som omhandler andre land, formidler også lærebøker i fremmedspråk et bilde av målkulturen. Det hender imidlertid at det ikke bare er målkulturen som innskrives i lærebøkene, men også kulturen i det landet der bøkene brukes. Slik har f.eks. Schopp (1985) undersøkt sovjetisk ideologi i lærebøker i fransk, og konkluderer med at den sovjetiske kulturen har en mye større plass enn den franske i materialet, og at Frankrike framstår som et blekt bilde av Sovjetsamveldet. Hun forklarer sine resultater ut fra hensikten med språkstudiet i det tidligere sovjetiske samfunnet, som ikke var å åpne seg overfor fremmede kulturer, men tvert imot en måte å spre sovjetisk propaganda på.

Zarate (1995) er inne på noe av det samme. Hun hevder at beskrivelsen av målkulturen ofte utviskes for at læremidlet skal tilpasses verdinormene i det hun kaller *den lokale kulturen*, og at disse normene anses som universelle. F.eks. kan man tenke seg at svinekjøtt ikke blir omtalt i lærebøker laget for muslimske land. Et av hennes eksempler er ei lærebok for araberlandene der et bilde av en strand ble fjernet i den endelige versjonen fordi det ikke passer å vise mennesker i badedrakt. (Zarate 1995: 17). Hun nevner

også eksempler på amerikanske og kanadiske lærebøker, der alkohol, nakenhet og kroppskontakt mellom personer av forskjellig kjønn er tabuemner. Disse studiene er interessante eksempler på at lærebøker også formidler holdninger, moral og oppdragelse knyttet til det landet de produseres i, som blant annet Johnsen & al.(1999) og Selander (1988) viser til.

La oss vende tilbake til kunnskapsområdet og se det i forhold til lærebøker. Lærebøker er laget for å formidle et fag, og dette faget er igjen en del av et kunnskaps- eller vitenskapsområde. Jeg vil igjen bruke et eksempel fra fremmedspråk. Malmberg (1981) var inne på forhold vitenskap – læring da han for noen tiår siden betraktet fremmedspråklæring som anvendt språkvitenskap:

We are in fact in a position to follow, step by step, this general evolution of the theory of language as it is reflected within the fields of applications. No manual of a foreign language starts today, as was so often the case half a century ago, with a series of isolated sentences which only served to give examples of unknown phonemes, unknown forms, and elementary rules of word order. A text, spoken or written, with a reasonable meaning and in a reasonable extra-linguistic context, now introduces the beginner to the world of the new language. (Malmberg 1981: 12)

Når det er vitenskapen eller fagfeltet som står i sentrum og formidles til undervisning og lærebøker, slik Malmberg påpeker i dette sitatet, får vi det Bjørndal (1982:40) kaller ”vitenskapssentrert læreplantenkning”. Ut fra en slik tenkning blir forskning på lærebøker i fremmedspråk forskning på hvordan språkvitenskapelige forskningsresultater anvendes i undervisningen.

Ideelt forandrer lærebøker seg med utviklingen av vitenskapen eller kunnskapsområdet, men ikke alltid i takt, gjerne senere. Dette var situasjonen for lærebøker i fremmedspråk på det tidspunktet da språkvitenskapelig forskning ble utvidet fra formell lingvistisk forskning i snever forstand til språkforskning i vid forstand, med vekt på språkbruk og kommunikasjon. Som en konsekvens av utviklingen ble lærebokforskernes fokus utvidet til varierte aspekter ved språk og språklig atferd. Dette førte til at språkbruken i lærebøkene, spesielt dialogene, ble et yndet forskningsobjekt. Nye parametre som kommunikasjonssituasjon, deltakere, emne, og konversasjonsprinsipper ble trukket inn i analysen. Resultatet var nedslående, for lærebokdialogene nådde selvsagt ikke opp mot de kravene som



stilles til ”normale” samtaler eller ”normale” kommunikasjonssituasjoner. Se f.eks. Gschwind-Holtzer (1981) og Fløttum (1984).

Å se fagstoffet i lærebøker kun i forhold til den vitenskapsgrenen de hører inn under, kan imidlertid fort bli for snevert. En vitenskapsgren kan ikke overføres direkte til en undervisningssituasjon, og “En lærebok blir et produkt med både faglig, pedagogisk og ideologisk innhold.” (Johnsen & al. 1999: 10). Det er med andre ord mange andre fagfelt som også kommer inn i overføringen fra vitenskap til skolefag. I Charaudeaus terminologi har man, i en streng vitenskapssentrert tilnærming, ikke tatt hensyn til transaksjonsprosessen mellom partnerne i kommunikasjonen. I lærebøker er stoffet didaktisert, valgt ut fordi læreren skal kunne undervise det og elevene skal kunne lære av det. Dermed skal det mye til at f.eks. dialoger i lærebøker i fremmedspråk er nøyaktig lik de man finner i verden utenom læreboka. Av den grunn ble den kritiske lærebokforskningen til Gschwind-Holtzer (se ovenfor) imøtegått av Besse (1985), som fant det ulogisk at lærebokdialogene ble kritisert nettopp for det som var hensikten med den, nemlig å presentere språkformer i en lærings- og undervisningssituasjon.

### **Kommunikasjonspartnerne - lærebokforfatteren og brukerne**

I den kommunikasjonsmodellen jeg har skissert, er de primære kommunikasjonspartnerne lærebokforfatteren som henvender seg til en brukergruppe bestående av elever og lærere. Men kommunikasjonspartnerne for læreboka er ikke konstante. Dette gjelder først og fremst på det konkrete plan, dvs. lærebokforfatternes identitet og posisjon i samfunnet. Mens man tidligere hadde markerte lærebokforfattere, f.eks. Gunnar Høst<sup>9</sup>, er det nå mer og mer forlagene som er den reelle kommuniserende, og den erfarne læreren som skriver sin lærebok etter mange års lærergjerning, er mindre til stede. Ei av de nye lærebøkene i fransk etter K06 er blitt til fordi forfatterne så en annonse fra forlaget i lærertidsskriftet *Utdanning* (Lidenlaub 2008: 35). En mulig innfallsvinklet til lærebokforskningen kan nettopp være å undersøke forholdet mellom lærebokforfatter og lærebok for å finne ut hvordan forfatteren har forholdt seg til både læreplan, andre relevante tekster og den generelle konteksten læreboka inngår i. Man kan, gjennom intervjuer, få nærmere innblikk i hvordan lærebokforfatterne har arbeidet, hva han har lagt vekt på og hva som har vært viktig. Lidenlaubs masteroppgave tar blant annet for seg dette forholdet gjennom intervjuer med to lærebokforfattere etter K 2006.

En annen måte å fokusere på lærebokforfatteren er å ta for seg en av de tradisjonelle læreboktekstene som ikke er lærestoff, nemlig forordene. Tradisjonelt hadde ikke forordene noen uttalt mottakergruppe, selv om de sannsynligvis ble skrevet for lærerne. I den senere tid har det også vært vanlig å skrive forordet direkte for elevene. Enkelte lærebøker, f.eks. Bødtker & Høst (1946), har både et forord for lærerne og en annen forordliknende tekst for elevene. I tillegg til forord har lærebøker etter hvert fått lærerveiledning som standardtillegg. Også disse kan være gjenstand for lærebokforskningen, se f.eks. Henriksen (2008) og Skjelbred (2007). Å undersøke slike tekster, som ikke er direkte lærestoff, men som kommenterer lærestoffet, kan være fruktbare bidrag både til forståelsen av enkeltbøker og til faghistorien, hvis en ser utviklingen over tid.

Søkelyset kan også rettes mot transaksjonen mellom lærebokforfatteren og kommunikasjonspartnerne og det sjangermessige resultatet av kontrakten mellom dem. Som jeg allerede har pekt på, kan læreboka være løst knyttet til en bestemt sosial kontekst, men likevel oppfattes eller brukes som lærebok. For at det tekstlige produktet forfatteren gir fra seg skal oppfattes som ei lærebok, må det, som nevnt, gjenkjennes som sådan, både ut fra intensjonen med kommunikasjonen og ut fra formelle trekk som kommunikasjonssituasjonen legger opp til. Det vil i realiteten si at læreboka tilhører en sjanger. Den har visse trekk, som selvsagt er underlagt forandringer som alle andre trekk ved sjangrer, men som likevel er gjenkjennelig.

Når lærebokforfatteren tilrettelegger sitt produkt for undervisning og læring, bruker han konvensjoner som er spesielt frekvente i nettopp lærebøker. Det kan dreie seg om tilrettelegging i forhold til alder og utviklingsstadier, f.eks. ved bruk av et enklere språk og om progresjon i lærestoffet. Det kan også bety eksempler, som igjen kan være av bestemte typer, og det kan være øvelser og kontrollspørsmål som viser at stoffet er forstått. Det kan være spesielle grafiske utforminger. Et annet trekk ved lærebøker er resymerende framstillinger eller andre former for framheving av lærestoffet. Kort sagt, de trekkene som CEDISCOR har identifisert som didaktisitet.

Å finne ut over tid hvordan læreboksjangeren forandrer seg ettersom den ytre situasjonen forandrer seg, kan være en tilnæringsmåte til disse aspektene ved lærebøker. I lærebøkene i fremmedspråk laget etter K06, der "språklæring" er et av hovedområdene, og der ansvaret for læringen hviler på elevene selv, finner man f. eks. selvevalueringsskjemaer som elevene skal bruke for å vurdere sin egen framgang. Tidligere hadde man kanskje

andre kontrollspørsmål som elevene kunne bruke for selvtesting, og som kan danne grunnlag for en komparativ studie.

Vi har hittil i dette avsnittet tatt for oss lærebokforfatterne, kontrakten mellom lærebokforfatteren og brukerne, og sjangermessige aspekter ved lærebøker. Men fokus kan også settes på brukerne, lærere og elever, og spesielt på den sekundære kommunikasjonen som oppstår mellom dem. Her kan forskningsspørsmålene være hvordan læreboka oppfattes, hvordan den brukes, og ikke minst *om* den brukes. Klasseromobservasjon, intervju med lærere og elever er velegnete instrumenter for dette. Heyerdahl-Larsen (2008) er et eksempel på slike studier. Hun har undersøkt bruken av lærebøker etter Reform 97, og finner f. eks. at engelsk er det faget som i størst grad bruker lærebøker i undervisningen. Et annet eksempel, av større omfang, er Skjelbred, Solstad & Aamotsbakken (2005). I dagens situasjon, der både lærebøker og metodefrihet overlates til det lokale nivået, dvs. til kommuner, skoler og enkeltlærere, kan undersøkelser om bruken av lærebøker i undervisning og læring være nyttige. Ikke minst når man tar i betraktning de moderne, åpne og interaktive læremidlene.

Som jeg har påpekt, inngikk læreboka tradisjonelt i en samhandling mellom læreren og klassen. Slik dagens syn på læring legger opp til, kan interaksjonen like gjerne foregå mellom elevene. Gjør de det, og hvor står i så fall læreboka i denne prosessen? Finnes det f. eks. øvelser, instruksjoner o.l. i læreboka selv som legger opp til elevsamarbeid? Tidligere, markante eksempler på klasseromsaktiviteter innskrevet i lærebøker finner vi i norske fremmedspråkbøker fra spesielt den første fasen av den dirkede metoden. Eksempler her kan være Bødtker & Høst (1929) og ikke minst Skotvedt & Pande (1931). Begynnertekstene i disse lærebøkene imiterte aktiviteten i klasserommet, der standardåpningen var ”Goddag, jeg er lærer, dere er elever, vi er i et klasserom”. I Henriksen (2005) kaller jeg disse tekstene ”speiltekster”. De synliggjør overfor lærerne, kanskje mer enn overfor elevene, hvordan klasseromsinteraksjonen skal foregå ut fra ideen om språkundervisning uten bruk av oversettelse til morsmålet.

### **Individuelle strategier – iscenesettelse**

Utvelgelsen av lærestoffet gjøres både ut fra kunnskapsområdet, den øvrige samfunnskonteksten og brukerne av læremidlet. Når utvalget er foretatt, kan man betrakte tilretteleggelsen av det som en iscenesettelse av stoffet,

slik jeg har nevnt tidligere. Forfatteren kan velge å framstille stoffet akkurat som han vil, bare han respekterer visse rammer som sjangeren setter. En slik iscenesettelse kan innebære at man blander faktastoff med fiksjon, og at lærestoffet formidles gjennom oppdiktete personer. Lærebøker i fremmedspråk er velegnet for denne type iscenesettelser fordi språklæring de fleste tilfeller også innebærer en tilnærming til en eller flere målkulturer som kan gjøres gjennom møte med mennesker.<sup>10</sup>

Som nevnt tidligere utgjør språket selv, dvs. grammatikk, ordformer, setninger og sammenbinding av setninger til tekster, lærestoffet i fremmed- og andrespråkbøkene. Men det er ikke mulig å vise eksempler på språk uten at setningene eller teksten også har et ikke-språklig, la oss kalle det ”kulturelt” innhold. Dette kulturelle innholdet kan imidlertid være mer en iscenesettelse av det språklige lærestoffet enn et selvstendig kulturelt lærestoff. Det innebærer at når fiktive personer forekommer i lærebøkene, kan de ha som hovedfunksjon nettopp å presentere språkformer, grammatikk og ikke minst vokabular, i praksis. For noen tiår siden var det f.eks. populært å iscenesette språket gjennom en gjennomgående kjernefamilie, vennene deres, matvanene, deltakelse i skole, arbeid og fritidsaktiviteter, slik at det oppsto fiktive lærebokunivers som til en viss grad skulle representere målsamfunnet, men som selvsagt ikke var fullgode representanter for dette samfunnet. Men gjennom kjernefamiliene fikk lærebokforfatterne presentert et utvalg av nyttige ord og vendinger fra dagliglivet.

Både iscenesettelsene og det lærestoffet de skal formidle, utgjør velegnede studieobjekter for lærebokforskeren. Ved f.eks. å studere dialogene til de iscenesatte personenes, kan man kartlegge hva som er det egentlige lærestoffet, dvs. de språkformene som ligger til grunn for replikkene, og se både hva som er valgt ut som lærestoff og hvordan det er gjort. Et eksempel fra begynnerbøker i fransk kan belyse dette. I fransk finnes det en spesiell artikkel for ubestemt mengde og ubestemt antall, den såkalte ”delingsartikkelen”: *du, de la, des*, som til alt overmål erstattes av preposisjonen *de* etter presisering av mengde eller antall: eks. *du beurre (smør) > un kilo de beurre (et kilo smør)*. Det er typisk for lærebøker i fransk å koble delingsartikkelen til mat og drikke, enten som innkjøpsscener eller som måltider og restaurantbesøk. Det er altså i slike tilfeller ikke måltidet eller innkjøpene i seg selv som er det primære læremålet, men grammatikk og til dels vokabular. Selvsagt formidles det også noe kulturkunnskap gjennom slike iscenesettelser, men det er likevel språket som er hovedsaken. Jf. Henriksen (2001).

Alle disse didaktiske faktorene som gjør læreboka til ei lærebok, må etter min mening være en av grunnforutsetningene for forskningen. Med det mener jeg å si at hensikten med lærebokforskningen, i alle fall når det gjelder fremmedspråk, ikke kun kan være. f.eks. å avdekke at lærebokspråket er enklere enn såkalt vanlig språk, eller å finne ut at det samfunnsbildet læreboka gir, ikke er i full overensstemmelse med de sosiale forholdene i landet. Et enkelt språk er et sjangertrekk ved læreboka, og et utvalg av sosiale forhold likeså. I lærebokforskningen må læreboka vurderes ut fra sine egne premisser, ikke i opposisjon til dem. Men det betyr selvsagt ikke at språk eller samfunnsforhold i læreboka ikke kan analyseres kritisk, så lenge man tar hensyn til at transformasjonen av kunnskapsfeltet til ferdig språklig produkt også er avhengig av transaksjonen mellom kommunikasjonspartnerne.

## Sluttord

Det finnes mange innfallsvinkler til lærebokforskning, og det finnes mange måter å omtale lærebokforskning på. I denne artikkelen har jeg valgt å se lærebokforskning som en del av forskning på pedagogiske tekster, samtidig som jeg har villet peke på slektskapet til forskningen på didaktisk diskurs og didaktisitet. Jeg har videre valgt å plassere lærebøker innenfor en kompleks kommunikasjonsmodell der både samfunnet, fagfeltet og kommunikasjonspartnerne er medbestemmende for den konkrete utformingen av teksten. Jeg har også vist til at lærebøker inngår i en dobbel kommunikasjonsituasjon som jeg mener er karakteristisk for læreboksjangeren. Jeg har videre ønsket å vise hvordan lærebokforskning kan ta utgangspunkt i samfunnskonteksten, deltakerne i kommunikasjonsprosessen, lærestoffet og iscenesettelsen av det. På denne måten blir lærebokforskningen, uansett fag, primært plassert innenfor en kommunikasjonsteorisk ramme.

Selv om denne artikkelen er ment å omhandle lærebokforskning generelt, er de fleste av eksemplene mine hentet fra fremmedspråk. Dette har jeg ikke gjort bare fordi det er mitt eget felt, men også fordi lærebøker i fremmedspråk, det sier seg selv, har fellestrekk med lærebøker i andrespråk. I begge tilfeller dreier det seg om språkformer som skal læres: uttale, grammatikk og ikke minst vokabular, men også bruk av språket, skriftlig og muntlig, med alt det innebærer av stil- og registervariasjon. Det er imidlertid visse forskjeller mellom undervisning i fremmedspråk og andrespråk

som sannsynligvis vil gi seg utslag i lærebøkene, og som man vil måtte ta hensyn til i lærebokforskningen. Undervisningen i andrespråk har et helt annen forhold både til målpråket og målkulturen enn fremmedspråkundervisningen har. Undervisningen foregår i målpråklandet og språkelevne er omgitt av språket både auditivt og visuelt i det fysiske landskapet. Dette har som konsekvens at den delen av språklæringen som foregår utenfor klasserommet og lærebøkene, nødvendigvis må være mye større for andrespråkelevne enn for elever som lærer et fremmedspråk kun gjennom læreboka og klasserommet i sitt eget land. Andrespråkundervisning er dermed mye mer virkelighetsnær enn fremmedspråkundervisning. Der fremmedspråkelevne har behov for grunnbetydningen av et ord for å forstå en tekst, har andrespråkelevne behov for bruksvalørene for å kunne forstå kulturen de er omgitt av. Det vil igjen si at både språket og målkulturen i prinsippet også må trekkes inn i lærebøkene på en mer autentisk måte. Likevel vil lærestoffet, både det språklige og det kulturelle, være iscenesatt og didaktisert, i likhet med det man finner i fremmedspråkbøkene, om enn kanskje ikke på samme måte.

## Noter

1. Selander 1994: 46
2. I det følgende vil jeg bruke termene ”forskning” og ”analyse” om hverandre, selv om ”lærebokanalyse” ikke alltid er identisk med ”lærebokforskning”. ”Lærebokforskning” er et mer omfattende begrep, og det finnes lærebokforskning som ikke innebærer dybdeanalyse av enkeltverk. Termen ”analyse” brukes imidlertid ikke så sjelden i overskrifter, jf. litteraturlista for denne artikkelen. På den andre siden brukes termen ”lærebokanalyse” også hyppig i forbindelse med studier, f.eks. i studieplaner, gjerne i didaktikk, og kan også innebære studentoppgaver, som da ikke nødvendigvis er forskningsbaserte.
3. *Centre de recherches sur la didacticité des discours ordinaires*, senere omdøpt til *Centre d'études sur les discours ordinaires et spécialisés*.
4. transmission des connaissances
5. For en oversikt, se Askeland 2008
6. I moderne læreboksproduksjon dreier det seg nok både om lærebokforfatter og forlag.
7. Se Tønnesson (2008) for situering av lærebøker i forhold til ulike kontekster.
8. Jeg har sett på sjuende opplag ved Jens Vindenes, og kan derfor ikke bekrefte hvorvidt det er Chauvins eller Vindenes sine iscenesettelser det dreier seg om.
9. Gunnar Høst redigerte og forfattet flere lærebøker i fransk og hadde et markert syn på fremmedspråklæring som også gikk på tvers av allmenne oppfatninger av fremmedspråkpedagogikk i tida. Hans lærebøker har preget mange tiår av franskundervisningen i det 20. århundre.

10. Se Risager (2003) for en inngående drøfting av forholdet mellom språk og kultur i fremmedspråkundervisningen.

## Litteratur

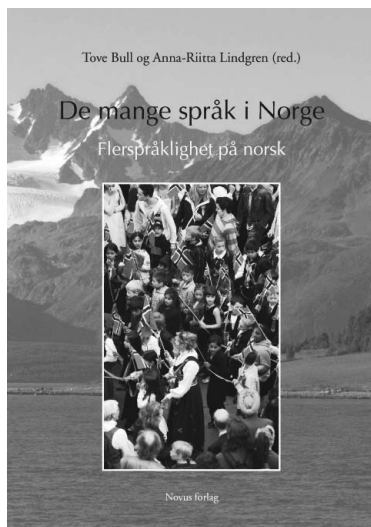
- Askeland, N 2008: *Lærebøker og forståing av kommunikasjon*. Oslo: Universitetet i Oslo. Digitale utgivelser ved UIO. <http://www.duo.uio.no> [Avhandling for ph. d. graden, Det humanistiske fakultetet]
- Besse, H. 1985: *Méthodes et pratiques des manuels de langue*. Paris: Didier.
- Bjørndal, B. 1982: *Et studium i lærebøkernes didaktikk. Lærebøkene og skolens innhold*. Rapport nr. 1. Oslo: Pedagogisk forskningsinstitutt. Universitetet i Oslo.
- Breilid, H. & Nicolaisen, T. 2002: Values in Religious Narratives as Presented in Norwegian RE Textbooks: Focus on Gender and Ethnicity. I Mikk, J., Meisalo, V., Kukemelk, H. & Horsley, M. *Learning and Educational Media. The Third IARTEM Volume*. Tartu: Tartu University Press, 139–147.
- Bødtker, A.T. & S. Høst 1929: *Lærebok i fransk for begynnere*. Utgave B. Oslo: Aschehoug.
- Bødtker, A.T. & S. Høst 1946. *Fransk for begynnere*. Utgave C ved dr.philos. Gunnar Høst. Oslo: Aschehoug.
- Charaudeau P. 1995: Une analyse sémiolinguistique du discours. *Langages* 117, 96–111.
- Chauvin, T.1941: *Fransk handelskorrespondanse til bruk for handelsgymnasier og foretningmenn*. Sjuende opplag ved Jens Vindenes. Oslo: Gyldendal norsk forlag.
- Christensen, H. & Wulff, T. 2006 - 2008: *Ouverture. Fransk for ungdomstrinnet. 8, 9, 10*. Oslo: Forlaget Fag og Kultur.
- Council of Europe 2001: *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University press.
- Fløttum, K. 1984: Spørsmål/svar-sekvens eller samtale? Analyse av en type læreboktekst innen fransk som fremmedspråk. *Moderna språk*. Vol. 78, nr. 3, 229–237 og 4, 319–224.
- Gundem, B.B. 1989: Innholdet i utdanningen: et viktig felt for forskningen. *Temaer for utdanningsforskningen. Innlegg og arbeidsnotater fra*

- planleggingsdag 27. april 1989*. Oslo: NAVF, 29–36.
- Gschwind-Holtzer, G. 1981: *Analyse sociolinguistique de la communication et didactique. Application à un cours de langue*. De Vive Voix. Paris: Hatier.
- Henriksen, T. 2001: "La mise en scène de l'article partitif," i *Romansk Forum*, 13. Oslo: Universitetet i Oslo, 3-18.
- Henriksen, T. 2005: *Begynnerbøker i fransk fra 1896 til 1974. A. Trampe Bødtker og Sigurd Høst, Lærebog i fransk for begyndere og dens arvtakere. Innhold, tekststrategier og iscenesettelser*. Oslo: Unipub forlag.
- Henriksen, T. 2008: Fremmedspråklæring som skapelsesmetafor. Grunnleggende ideer bak språklig utvelgelse og kulturelt innhold i Bødtker, A.T. & S. Høst 1946. Fransk for begynnere. Utgave C ved dr.philos. Gunnar Høst. I Skjelbred, D. & Aamotsbakken, B. (red.) *Norsk lærebokhistorie - en kultur- og danningshistorie*. Oslo: Novus forlag, 67–82.
- Heyerdahl-Larsen, C. 2000: *Læreboken – tvangstrøye eller helsetrøye? En teoretisk og empirisk framstilling av lærebokens rolle i undervisningen*. Oslo. [Hovedoppgave. Universitetet i Oslo.]
- Johnsen, E.B. 1993: *Textbooks in the Kaleidoscope. A Critical Survey of Literature and Research on Educational Texts*. Oslo: Scandinavian University Press.
- Johnsen, E.B. & al. 1999: *Lærebokkunnskap. Innføring i sjanger og bruk*. Oslo: Tano-Aschehoug.
- Knudsen, S. V, Skjelbred, D. & Aamotsbakken, B. (red.) 2007: *Tekst i vekst. Teoretiske, historiske og analytiske perspektiver på pedagogiske tekster*. Oslo: Novus forlag.
- Lidenlaub, J. Grimestad 2008: *Intention, Création, Action*. Oslo: Universitetet i Oslo. Digitale utgivelser ved UIO. <http://www.duo.uio.no> [Masteroppgave, Institutt for litteratur, områdestudier og europeiske språk]
- Malmberg, B. 1981: Applied linguistics: past, present and future. *AILA 81, Proceedings II, Studia linguistica* 35, no. 1–2, 7–14.
- Mandt, L. 2008: *Culture dans la Promotion de la Connaissance (K06). Du programme scolaire au matériel didactique*. Universitetet i Oslo [Masteroppgave, Institutt for litteratur, områdestudier og europeiske språk]
- Moirand, S. et al. 1992: Un lieu d'inscription de la didacticité. Les catastrophes naturelles dans la presse quotidienne. *Les carnets du Cediscor* 1. Paris: Presses de la Sorbonne Nouvelle.



- Nygård, G. 2007: *Kultur i Kunnskapsløftet - fra læreplan til lærebok. På hvilken måte legger Mundos nuevos 1 opp til at elevene skal kunne nå kompetansemålene for språk, kultur og samfunn i K06?* Oslo: Universitetet i Oslo. Digitale utgivelser ved UIO. <http://www.duo.uio.no> [Masteroppgave, Institutt for litteratur, områdestudier og europeiske språk]
- Oxfeldt, E. 2008: Pædagogiske heterotopier: Vidensepistemer fra 1800-tallets læsebog til dagens læremidler. I Skjelbred, D. & Aamotsbakken, B. (red.) *Norsk lærebokhistorie - en kultur- og dannelseshistorie*. Oslo: Novus forlag.
- Risager, K. 2003: Det nationale dilemma i sprog- og kulturpædagogikken: et studie i forholdet mellem sprog og kultur. Kbh.: Akademisk Forlag.
- Schopp, M. 1985: Le discours d'une méthode. La place de l'idéologie dans les manuels de français utilisés dans les écoles spéciales soviétiques. *Études de linguistique appliquée* 60, 70–88.
- Selander, S. 1988. *Lärobokskunskap*. Lund: Studentlitteratur.
- Selander, S. 1994: Pedagogiska texter och retorik. Selander, S. & B. Englund (red.). *Konsten att informera och övertyga. En antologi om pedagogik, text och retorik*. Stockholm: HLS Förlag, 41–53.
- Selander, S. 1995: Research on pedagogic texts: An approach to the institutionally and individually constructed landscapes of meaning. Skyum-Nielsen, P. (red.). *Text and Quality. Studies of Educational Texts*. Oslo: Scandinavian University Press, 152–170.
- Skjelbred, Dagrun 2007: Lærerveiledningen og dens forløpere - et teksthistorisk perspektiv. I Knudsen, S. V, Skjelbred, D. & Aamotsbakken, B. (red.) *Tekst i vekst. Teoretiske, historiske og analytiske perspektiver på pedagogiske tekster*. Oslo: Novus forlag, 119–135.
- Skjelbred, D, Solstad, T. & Aamotsbakken, B. 2005: *Kartlegging av læremidler og læremiddelpraksis*. Rapport nr. 1. Høgskolen i Vestfolds skriftserie. [<http://www-bib.hive.no/tekster/hveskrift/2005.html>]
- Skotvedt, C. & P. Pande 1931: *Le premier livre de français*. Oslo: Gyldendal.
- Tønnesson, J. L. 2008: Lærebokhistorie: en topografisk analysemodell. I Skjelbred, D. & Aamotsbakken, B. (red.) *Norsk lærebokhistorie - en kultur- og dannelseshistorie*. Oslo: Novus forlag, 165 – 178.
- Zarate, G. 1995: *Représentations de l'étranger et didactique des langues*. Paris: Didier.

# Nytt fra Novus forlag



Denne boka er et resultat av et følt savn fra forfatternes side. Det har lenge eksistert litteratur om minoritetsspråk i Norge, både om språkene til de gamle og til de nye minoritetene.

Generelle framstillinger og framstillinger om enkeltspråk finnes, og mye er selvsagt skrevet om norsk. Men noen samlet framstilling av norsk språkmangfold har ikke eksistert før nå.

Boka tar mål av seg til å gi et generelt overblikk over det språklige mangfold og de språklige variasjoner i Norge. Det betyr ikke at alle språk som tales i Norge, blir omtalt. Det ville vært umulig! Til det er mangfoldet for stort. Et slikt mangfold gir store muligheter, både for enkeltindividet og samfunnet når den språklige rikdommen blir positivt tatt imot. Språkpolitiske implikasjoner av mangfoldet berøres derfor også.

[www.novus.no](http://www.novus.no) / [novus@novus.no](mailto:novus@novus.no)  
tlf 2271 7450

# Vokaler hos andrespråksbrukere: en akustisk og perseptuell studie

## Vokaler i det norske talespråket til tre utenlandske leger med arabisk som morsmål

*Av Liv Andlem Harnæs  
Universitetet i Oslo*

The best way of describing vowels is not in terms of the articulations involved, but in terms of their acoustic properties

(Ladefoged 2003: 105)

### **Sammendrag**

I artikkelen presenteres funn fra en akustisk og perseptuell studie av vokalproduksjon i det norske talespråket til tre utenlandske leger med arabisk (irakisk) som morsmål.

Studien tar utgangspunkt i en norsk uttalenorm (målspråksnorm) som etableres gjennom tre informanter med norsk som morsmål.

I diskusjonene rundt uttaletilegnelse av andrespråk hos voksne har kontrastive hypoteser stått sterkt, og i denne studien vil jeg trekke inn kontrastive forklaringsmåter, men den individuelle variasjonen kan være stor, og flere forklaringer kan være aktuelle, som for eksempel påvirkning fra skrift. En viktig erfaring fra arbeidet med vokalanalyse er at akustisk analyse av vokalkvalitet er fordelaktig fordi vi kan tallfeste hva vi hører, noe som er viktig i vitenskapelig sammenheng. Et funn som gir retning for videre utforskning av materialet, er at det ikke er tilstrekkelig å søke forklaringer på vokalavvik på fonemplanet. Uttalen av vokalen /ʌ/ hos de tre irakiske informantene var spesielt målspråksnær i noen ord (eller stavelser) som

*plutselig* og *slutt*, og man må lure på hva det er ved disse ordene (eller stavelsene) som genererer den målspråkskorrekte uttalen av /t/ her.

*Nøkkelord: uttale, vokaler, akustisk analyse, arabisk.*

## Innledning

Innflytelse fra morsmålet er uten tvil en viktig kilde til uttaleavvik hos andrespråksbrukere (Odlin 1990, Bohn 1995, Bohn og Flege 1997, Husby og Kløve 1998). Morsmåletets fonetikk og fonologi påvirker andrespråksuttalen. Ulike utenlandske aksenter er et vitnemål om det: Vi sier for eksempel at hun snakker norsk med fransk aksent, eller med kinesisk aksent osv, og ofte gjetter vi riktig med hensyn til personens førstespråk, men ikke alltid (se Harnæs 2001 for hvordan en gruppe norske morsmålsbrukere gjenkjente aksentene til ti utenlandske studenter). Usikkerhet når det gjelder gjenkjennelse av morsmålssubstrat, kan skyldes at ikke alle andrespråksbrukere har de typiske uttaleavvikene som er karakteristiske for morsmålsgruppen. Det kan også skyldes at vi som morsmålsbrukere har ulike erfaringer med og kunnskaper om andre språk, og derfor har ulike forutsetninger for å gjenkjenne ulike substrater. Likevel har vi en mer eller mindre klar oppfatning av at personer med samme morsmål har noen felles språklige trekk som er avvikende i forhold til et målspråk. Hvilke språklige trekk det dreier seg om, og om og i hvilken grad personer med samme morsmål har de samme avvikene, kan vi ikke svare fullgodt på. Av lærererfaring vet vi at prediksjoner om at personer med samme morsmål har de samme uttaleavvikene, ofte er riktige; for eksempel forkorter polske studenter lange vokaler i norsk, og tyske studenter uttaler de stemte plosivene /b/, /d/ og /g/ ustemt i utlyd og sier /ba:t/ istedenfor det norske /ba:d/ med stemt konsonantlyd til slutt.

Intonasjonen i norsk er nær knyttet til tonemene. Tonemer eller tonelag finnes i norsk og svensk, men er ellers sjeldne i europeiske språk. Tonemene er bundet til trykkstavelene og til den tydelige forandringen i tonehøyde som skjer ved overgangen mellom trykksterk og trykksvak stavelse. Forløpet av denne forandringen er både avhengig av tonemet i stavelen og av hvor i landet den som snakker, kommer fra. Norske intonasjonsmønstre er altså bundet til dialektområder, og vi kan gjenkjenne ulike dialekter på intonasjonen. Når nordmenn snakker for eksempel engelsk, kan vi ofte høre

norske intonasjonsmønstre. Også en utenlandsk intonasjon blir en del av en fremmed aksent i et andrespråk, men aksent skapes på ulike nivåer, og også segmentavvik skaper aksent (for eksempel en engelsk eller amerikansk /t/ i norsk). For andrespråksbrukere kan en sterk utenlandsk aksent være et handikap selv om den ikke er til direkte hinder for forståelsen. Mange undersøkelser har konkludert med at jo sterkere aksenten er, desto mer negativt blir taleren vurdert på alle områder (Cunningham-Andersson 1990, 1993, 1997; Husby og Kløve 1998; Kulbrandstad 2006).

Andrespråksbrukere kan ha vansker med å lære både enkeltlyder (vokaler og konsonanter) og prosodi (trykk, tone, intonasjon), og kanskje er det i noen grad morsmålsavhengig om det er segmenter eller prosodi som byr på de største utfordringene for den som lærer seg et andrespråk eller fremmedspråk (Odlin 1990).

Det er også viktig å huske på at morsmålet påvirker vår evne til å oppfatte språklyd. For eksempel kan det være vanskelig for studenter med arabisk som morsmål å skille mellom /i/ og /e/ i norsk. I arabisk tilhører [i] og [e] samme fonem, og det er derfor rimelig å anta at det var morsmålet vokalkategorisering som lå til grunn for et spørsmål fra en student med arabisk som morsmål: Han hadde skrevet ned de to ordene *litt* og *lett* og spurte læreren hvorfor man skriver samme ord på to forskjellige måter. I en artikkel om persepsjon og produksjon av en ny vokalkategori hos voksne skriver Bohn & Flege (1997) at de finner støtte i sine undersøkelser for at erfaring med et andrespråk ser ut til å ha dypere innvirkning på produksjon enn persepsjon av en ny lyd. Dette er kontroversielt, men mange hevder at det er raskere eller lettere å lære å uttale en ny lyd enn det er å oppfatte den samme lyden.

Men vi kan ikke regne med at alle uttaleavvik skyldes påvirkning fra morsmålet. Andre språklige kilder til uttaleavvik kan være: et annet andrespråk eller fremmedspråk, sjeldne eller uvanlige trekk i målspråket (*markerthet*), såkalte utviklingsfaktorer (faktorer som er virksomme en begrenset tid i læringsforløpet) og påvirkning fra skriftspråksformer. I tillegg kommer alle påvirkningsfaktorer som ikke er internt språklige (Husby og Kløve 1998, Arnfast 2007).

Denne artikkelen omtaler en pilotundersøkelse. Jeg studerer vokaler i det norske talespråket til voksne andrespråksbrukere med arabisk som morsmål. Målet her er å gi en akustisk og perseptuell beskrivelse av vokalkvalitet i det norske talespråket til tre irakiske leger med arabisk som morsmål. Den akustiske beskrivelsen gis i form av tallverdier (i hertz, Hz) for de tre første

formantene (se nedenfor) i vokalene. Gjennom den perseptuelle beskrivelsen (det vil her si bare min individuelle persepsjon) knyttes formantverdiene til mitt hørselsinntrykk av en bestemt vokal.

Jeg vil sammenlikne vokalene til de tre legene fra Irak med de tilsvarende vokalene hos tre norske morsmålsbrukere som snakker østnorsk.

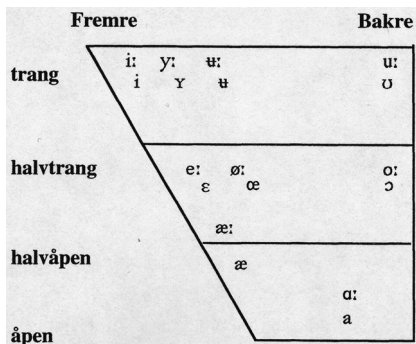
### Vokaler i norsk og arabisk

Det er vanlig å regne med at norsk (standard østnorsk) har ni vokalfonemer i trykksterk stavelse (Kristoffersen 2000). De ni vokalfonemene har en kort og en lang variant, som i skrift markeres indirekte med dobbeltkonsonant eller to konsonanter etter kort vokal, for eksempel (å) *takke*, /a/, og enkelt konsonant etter lang vokal, for eksempel (et) *tak*, /a:/. Regelen gjelder bare i trykksterke stavelser, trykklette stavelser har alltid kort vokal. Inkonsekvenser når det gjelder skriftlig markering av kort og lang vokal, for eksempel at vokalen er lang i ord som *rart*, *pent*, *rikt*, *gult*, *søkt(e)*, *lånt(e)*, gjør at skriftspråket ikke er en umiddelbart god rettesnor for uttalen. Det gjelder også i særlig grad de mange ordene i norsk som ikke har trykk på første stavelse, og det gjelder sammensatte ord. Ord som *vokal*, *apotek*, *kvittering*, *homogent*, *privatliv*, *fåmælt* skaper usikkerhet hos andrespråksbrukere både med hensyn til trykkplassering og kort og lang vokal.

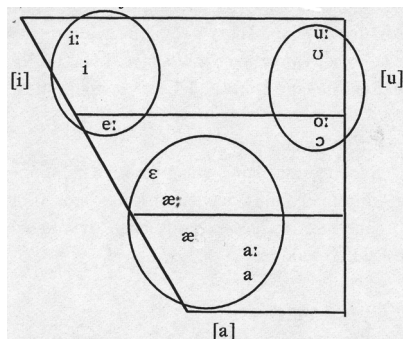
Arabisk har tre par vokalfonemer. De har som norsk korte og lange varianter med betydningsskillende funksjon. De tre vokalparene er kort og lang /a/, kort og lang /u/ og kort og lang /i/. Korte vokaler skrives vanligvis ikke i arabisk, men markeres med diakritiske tegn i opplæringstekster. Lange vokaler har sine egne skrifttegn. Awde & Samano (1986: 27) hevder at arabiske vokaler er “rene” (“pure”), at de ikke diftongeres slik som for eksempel engelske vokaler ofte gjør. Vi kommer tilbake til dette punktet i spørsmål om morsmålspåvirkning når det gjelder kvalitetsendring i noen lange vokaler.

Vokalene inni ellipsene i figur 2 (neste side) har fonemisk funksjon i norsk, se figur 1, men fonetisk i arabisk. Det vil si at vokalene i en ellipse er allofoner av samme fonem i arabisk; de tre fonemene i arabisk er [i], [a] og [u]. (Illustrasjonen er hentet fra Husby og Kløve (1998)).

Figur 1: Norske vokaler



Figur 2: Arabiske vokaler, fonetisk og fonemisk



### Prediksjoner og hypoteser som gjelder arabisktalende

I likhet med alle de andre språkene som ikke har fremre rundede vokaler, må det forventes at andrespråkstilegnere med arabisk som morsmål vil få problemer med de norske fremre vokalene. Tilsynelatende skulle ikke de tre vokalparene som både norsk og arabisk har, være noe problem for arabisktalende under tilegnelse av norsk. Vi kan få positiv transfer fra lang og kort /i/, /a/ og /u/ til de "tilsvarende" vokalene i norsk, mens alle de andre vokalene vil by på problemer.

Det må imidlertid også forventes problemer med de tre vokalparene som vi har felles, fordi i arabisk er alle vokalkvaliteter i området for de norske /i:/, /i/ og /e:/, /e/ allofoner av samme vokal. På samme måte er alle vokalkvaliteter i området /æ/ og /a/ allofoner av samme vokal i arabisk. Det samme gjelder alle avskygninger av /u/ og /o/. Kvaliteten av disse vokalene i arabisk avhenger av tilliggende konsonanter, mens det i norsk er fonemisk distinksjon mellom disse vokalene. Med arabisk som morsmål vil det være vanskelig å tilegne seg et språk som norsk der det er svært små forskjeller mellom ulike fonem (Husby og Kløve 1998: 49).

Uttalevansker når det gjelder /i/, /a/ og /u/ vil ikke bli drøftet her; her vil jeg undersøke nærmere i hvilken grad prediksjonen er gyldig for de norske vokalene /ɥ/, /y/ og /ø/, de fremre rundede vokalene i norsk, som ikke bare kan være vanskelige å oppfatte og uttale for dem som har arabisk som mors-

mål, men også for alle andre som ikke har vokalene i sitt morsmål. Jeg vil prøve å finne ut i hvilken grad hypotesen er riktig for mine irakiske informanter, og om den eventuelt kan nyanseres: Gjelder vanskene alle ord hvor vokalene forekommer, eller er det noen ord eller stavelser hvor vokalene oftere eller alltid uttales riktig eller galt? Er det forskjell på uttalen om vokalen står i et enkelt ord uttalt alene, eller om vokalen forekommer i et ord i en setning eller tekst? Hvilken forklaring på vokalavviket er mest sannsynlig i det enkelte tilfellet?

At ord som inneholder disse tre vokalene, kan være vanskelige å uttale for denne gruppen og andre morsmålsgrupper, er riktig. Som lærer har jeg ofte opplevd at disse tre vokalene uttales på en måte som skaper uklarhet eller misforståelser, for eksempel har jeg notert på kurs for utenlandske leger at de som har arabisk som morsmål, sier /ø:rer/ slik at det høres ut som /o(:)rer/, (og mer komplisert når /blu:o:rer/ nærmest blir /blo(:)ə(:)rer/), og at /øyetryk/ blir /øyetrik/. De tre ordene *undersøkelse*, *urin* og *utvikle* (fremkvente ord på dette kurset) uttales for eksempel svært ofte med en /u/-lyd som første lyd, det gjelder også vanlige ord som *under*, *du*, *skulle* og *kunne*. Dersom tonemene ikke er på plass, kan uklarheter forsterkes ytterligere i ord som *årene* (av en *åre*), *årene* (av et *år*) og *ørene* (av et *øre*) og *håret* (av et *hår*) og (å) *høre*.

## Informanter

Den foreliggende studien er tenkt som den første av flere undersøkelser av talespråket til i alt 17 voksne mannlige leger fra Irak med arabisk som morsmål. De har alle gått på fagspråk- og kommunikasjonskursene for utenlandske leger ved tidligere Institutt for lingvistiske fag (ILF) og nåværende Institutt for lingvistiske og nordiske studier (ILN) ved Universitetet i Oslo (UiO). Det har blitt gjort lydopptak av de 17, og opptakene har blitt overført til lydfiler på min PC (se nedenfor). Det er også gjort opptak av sju norske menn oppvokst på Østlandet og ansatt ved ILF/ILN.

Til den pilotstudien som presenteres her, har jeg valgt ut seks av informantene (tre med arabisk og tre med østnorsk som morsmål). De tre fra Irak ble valgt ut på grunnlag av sin generelle norskuttale slik at én syntes jeg generelt hadde dårlig uttale (irakA), én ganske god uttale (irakB), og én var litt midt imellom (irakC). De tre norske informantene (norsk1, norsk2 og norsk3) valgte jeg tilfeldig.



De irakiske informantene har lært norsk som voksne. De tre som er med i denne studien, var 40 (irakA), 36 (irakB) og 28 (irakC) år gamle da de begynte å lære norsk, og de oppgir å ha lært norsk (gått på kurs og selvstudium) i henholdsvis to, fem og ett år da lydopptakene ble gjort.

Informantene ble informert om - og aksepterte velvillig - at lydopptakene skulle brukes til utforskning av uttalen deres, uten spesifisering av hva ved uttalen som skulle fokuseres.

## **Materiale**

Lydopptakene av de seks informantene er gjort på et vanlig kontor på Blindern (UiO) på dat-kassetter og så overført til lydfiler (wav-filer) på min PC med god hjelp fra HF's Av-tjeneste og programmet Cool Edit 96. Senere opptak, men ingen av dem som er med her, er overført direkte til PCen uten veien om datkassetene. Alle informantene har lest inn de samme enstavelsesord, flerstavelsesord, sammensatte ord, løsrevne setninger, en tekst (*Nordavinden og sola*) og fortalt hvordan de har lært norsk (et fremmedspråk for de norske deltakerne). De arabiske deltakerne har også lest inn en tekst på arabisk, *Nordavinden og sola* i arabisk skrift, (IPAs *Handbook* 1999) med tanke på at vokalene i morsmålet kan testes. I denne undersøkelsen har jeg bare vurdert vokaler i (fonologisk) trykksterk stavelse.

## **Akustisk og perseptuell metode**

Når det gjelder den akustiske studien av vokalkvalitet, har jeg benyttet programmet Praat og målt verdiene for de tre første formantene, F1, F2 og F3 (se nedenfor). Når det gjelder den perseptuelle studien, har artikkelforfatteren vært den eneste bedømmeren, og jeg har utelukkende registrert lyd-kvaliteten jeg har hørt, som tilhørende et fonem. I uttalen av for eksempel ordet *vunnet* har jeg plassert vokallydene jeg har hørt i trykksterk stavelse som tilhørende ett av to vokalfonemer: enten /ʌ/ eller /u/, andre alternativer har ikke vært aktuelle. Siden dette er en pilotundersøkelse, og jeg har ønsket å knytte formantverdiene til lyd-kategorier, har jeg funnet dette meningsfullt og tilstrekkelig i denne omgang. Hvis man studerer den grafiske fremstillingen av de to forekomstene av *vunnet* (se nedenfor), er verdien av andreformanten, F2, omtrent 500 hertz lavere hos informant irakA enn hos de

fem andre. Hos irakA hører jeg en /u/-lyd, hos de andre en /ʊ/-lyd. Mer omfattende analyser av vokalkvalitetene er planlagt i et senere prosjekt for blant annet å undersøke nærmere hvordan irakernes uttale av /ʊ/ plasserer seg i forhold til morsmålsdeltakernes uttale av den samme vokalen og andre trange, rundede vokaler som /u/ og /y/. Se for øvrig tabeller over noen formantverdier for /u:/, /ʊ:/ og /ʊ/ i vedlegget helt til slutt i denne artikkelen.

### Formanter og tolkinger av formantverdier

Vokalanalysen har jeg gjort i taleanalyseprogrammet Praat. Akustisk er en vokal en bestemt fordeling av formanter i et klangspektrum. En formant er en gruppe forsterkede deltoner i klangspekteret. Hvor formantene ligger i klangspekteret, bestemmes av formen og størrelsen på resonansrommet, som er hulrommet mellom stemmerissen (glottis) og munnleppene.

I den følgende figuren, Figur 3 (s. 35), ser vi bølgeform (øverst) og spektrogram (nederst) av ordet *plutselig* (mellom de vertikale stiplede linjene), uttalt av en av de norske morsmålsbrukerne. Spektrogrammet viser formanter som mørke bånd markert med prikkede linjer.

I figur 4 (s. 36) er /ʊ/ skilt ut mellom de stiplede, vertikale linjene (man kan også regne at /ʊ/ begynner lenger til venstre der førsteformanten begynner), og de tre formantene F1, F2 og F3 ligger som tre mørke (og prikkete), horisontale bånd.

### Formanter og tolkinger av formantfrekvenser

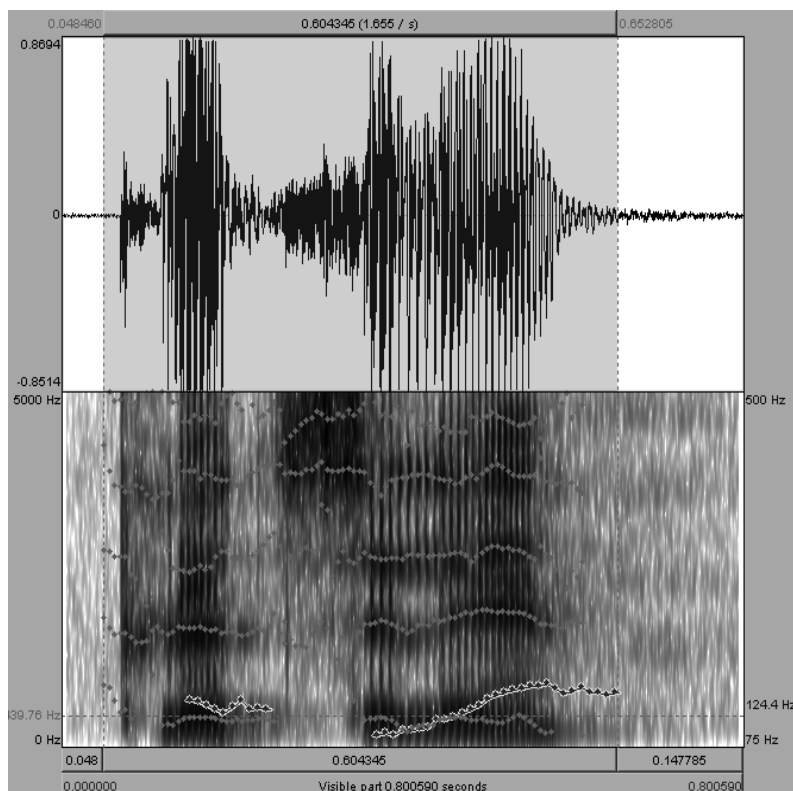
Formanter måles i hertz, Hz. Når F1-frekvensen minker, er det vanlig å regne med at tunga har beveget seg til en høyere posisjon (en trangere lyd). Når F2-frekvensen øker, er det vanlig å regne med at tunga har beveget seg til en mer fremre posisjon, lavere F2-verdier betyr altså en mer bakre lyd.

Effekten av lepperunding er en senkning av alle formantfrekvenser. Rundede vokaler har vanligvis lavere formantfrekvenser enn urundede fordi lepperunding oftest forlenger munnhulen.

Den prinsipielle forskjellen mellom /i/ og /y/ er at tredjeformanten er lavere for /y/. Andre (og tredjeformanten) er lavere for /ø/ enn for /e/, og /u/ har lavere verdier enn /ʊ/ når det gjelder andreformanten. I ordet *plutselig*, se nedenfor, er tallverdiene for /ʊ/-lydene godt i underkant av

Figur 3:

p l u t s e l i



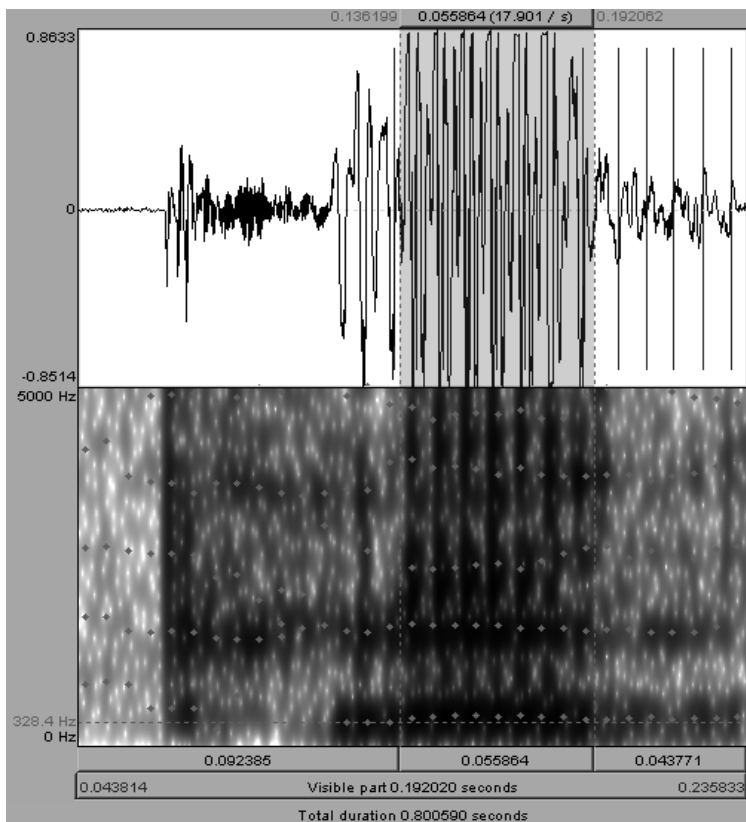
500Hz for F1, i overkant av 1500Hz for F2 og rundt 2500Hz for F3 hos de norske informantene. Det finnes en individuell variasjon hos de norske morsmålsbrukerne (se tabell over eksakte verdier i vedlegget helt til slutt i artikkelen), men variasjonen har ikke innvirkning på min oppfatning av vokalkvaliteten.

### Måling av formanter

Alle vokalene jeg har målt, har jeg avgrenset manuelt i bølgeform og spektrogram (jf Figur 3 og Figur 4). Det er vanlig å måle formantene midt

Figur 4:

p l u

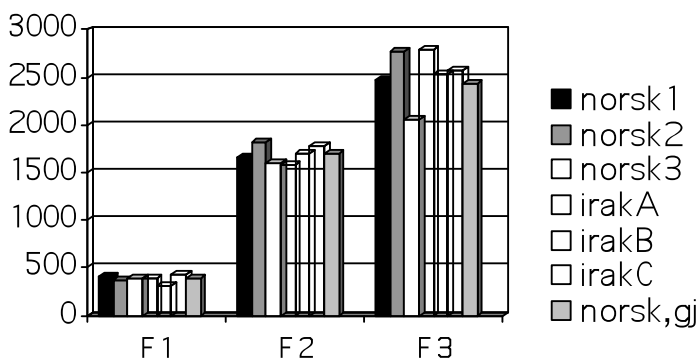


i vokalen, i et område hvor de er jevne og stabile. I stedetfor å måle ett sted midt i vokalen (det er ofte vanskelig å finne et sted eller område hvor alle tre formanter er stabile samtidig) har jeg valgt å måle formantene to steder i vokalen: først én tredjedel inn i vokalen (når den vanligvis har stabilisert seg) og så etter to tredjedeler av vokalforløpet. Den første målingen brukes som hovedmåling og er den verdien som brukes i målingene i denne studien (den andre målingen angis av og til i parentes hvis de to er svært ulike). Ved å sammenlikne de to målingene kan man se eventuell bevegelse eller retning innen vokalen (se for eksempel målingen av /y/ i *dyd – s(etning)* hos informanten irakA hvor det er en klar kvalitetsendring innenfor vokalen).

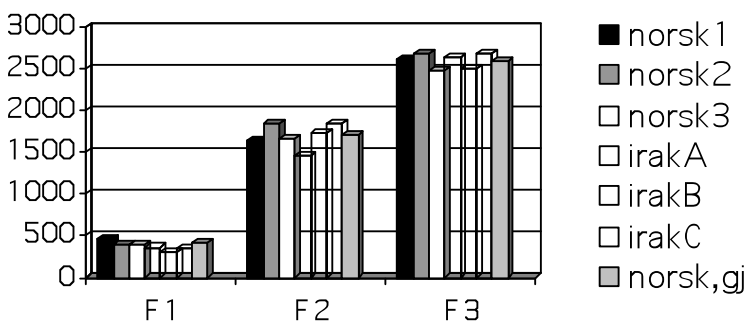
I det følgende vises resultater av målinger av de tre vokalene /ʉ/, /y/ og /ø/; ”e” etter ordet vokalen er hentet fra, betyr at ordet ble lest som enkeltord, ”s” at ordet forekommer i en løsrevet setning, ”t” at ordet står i teksten *Nordavinden og sola* (t1: første forekomst, t2: annen forekomst osv), og ”f” at ordet er hentet fra opptak av fri tale. Tallverdiene i den vertikale akse angir antall *hertz*, Hz, som er måleenheten for frekvens. Den sjuende søylen i hver gruppe representerer gjennomsnittsverdiene for de tre norske informantene.

### Målinger av /ʉ /-lyder

Figur 5: Kort /ʉ / (plutselig – e):

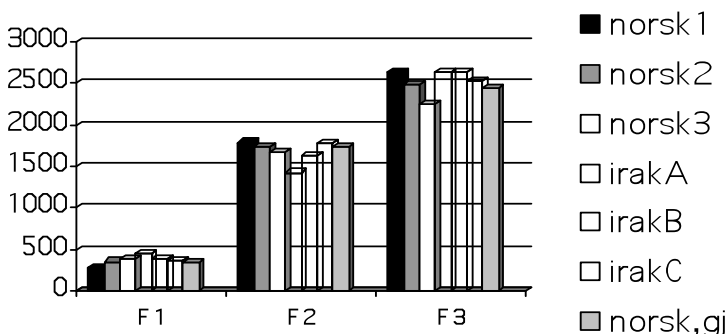


Figur 6: Kort /ʉ / (plutselig – t1):



Tabeller med nøyaktige formantverdier for vokalmålingene finnes i vedlegg 1-6 til slutt i artikkelen, men ved å studere og sammenlikne høyden på de grafiske søylene får man en oversiktlig informasjon om formantverdiene.

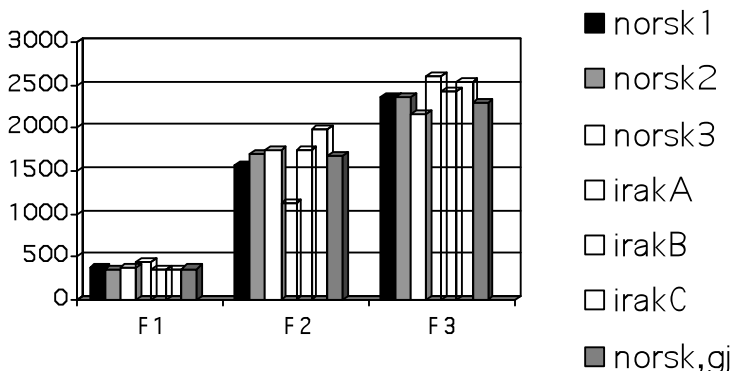
Figur 7: Kort /ʌ/ (plutselig – t2):



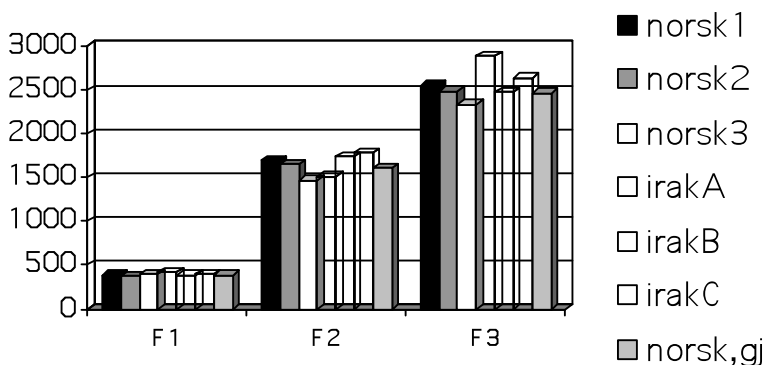
Når det gjelder de tre forekomstene av ordet *plutselig* i materialet, hører jeg /ʌ/-lyder både hos de irakiske og de norske deltakere. Formantmålingene viser ikke stor variasjon selv om andreformanten (F2) til irakA gjennomsnittlig er litt lavere enn hos de andre. Det tyder på at hans /ʌ/-lyd er en litt mer bakre lyd. Norsk3 har en lavere tredjeformant (F3). Trekket går igjen i alle målingene av /ʌ/-lyder hos ham (se søylediagrammene nedenfor).

I de følgende fire søylediagrammene vises også målinger av kort /ʌ/-lyd:

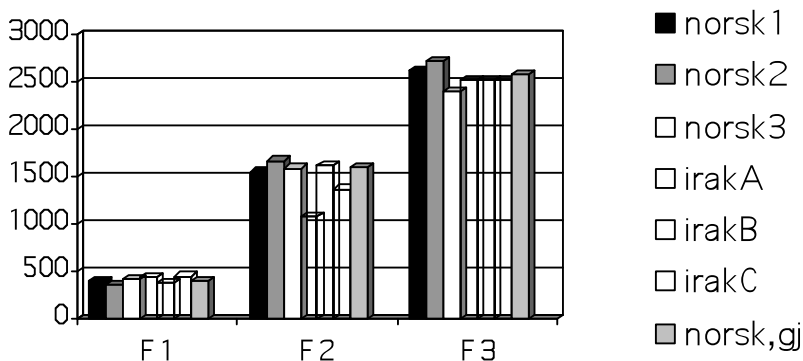
Figur 8: Kort /ʌ/ (skutt – s):



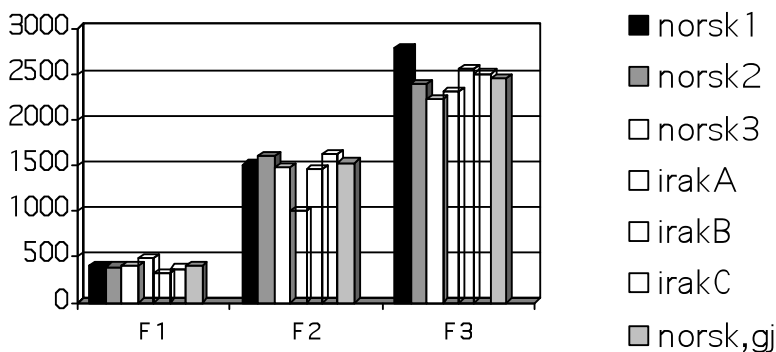
Figur 9: Kort /ʊ/ (slutt – t):



Figur 10: Kort /ʊ/ (vunnet – e):



Figur 11: Kort /ʊ/ (vunnet– t):







Når det gjelder de to ordene med lang /ʉ/-lyd, *uttalen-e*(nkeltord) og *mulig-t*(ekst), avviker andreformanten til både irakA og irakC i *uttalen*, men det er bare irakA som har markert lavere andreformant i *mulig*. I disse tre forekomstene hører jeg en /u/-lyd. I ordet *uttalen* uttaler både irakB og irakC /ʉ/ som en kort lyd. Det kan virke som den lange /ʉ/-lyden i prefikset *ut-* er spesielt vanskelig å mestre enda det er et frekvent prefiks i norsk.

### Hva viser målingene av /ʉ/-lydene?

Formantverdiene indikerer at de tre irakerne uttaler /ʉ/ i de tre forekomstene av ordet *plutselig* med en /ʉ/ som ikke skiller seg vesentlig fra de norske informantene (for nøyaktige formantverdier se vedlegg 1). Det samme gjelder ordet *slutt* fra teksten *Nordavinden og sola*. Målingene viser ”norske” verdier for /ʉ/, og lydene i disse ordene hos de tre irakerne høres også ut som norske /ʉ/-lyder (selv om irakA har noe lavere verdier for F2 i de to forekomstene av *plutselig* i teksten, 1454Hz og 1426Hz).

Når det gjelder de øvrige målingene av kort /ʉ/, så viser de at både irakB og irakC har formantverdier som er på nivå med de norske selv om variasjonen i andreformanten til irakC er ganske stor: 1370Hz i *vunnet-e* og 1988Hz i *skutt-s* som ytterpunkter. De tilsvarende gjennomsnittsverdiene for de norske deltakerne er 1604Hz (*vunnet-e*) og 1671Hz (*skutt-s*). Likevel oppfatter jeg vokalen i de tre ordene *skutt-s*(etning), *vunnet-e*(nkeltord) og *vunnet-t*(ekst) hos både irakB og irakC som akseptable /ʉ/-lyder. IrakA derimot har mye lavere F2-verdier i disse tre ordene: 1079Hz (*vunnet-e*), 997Hz (*vunnet-t*) og 1122Hz (*skutt-s*). Lavere F2-verdier betyr som sagt en mer bakere lyd, og /ʉ/-ene til irakA i disse tre ordene høres også mer ut som /u/-lyder. I motsetning til hos irakC kan vi legge merke til at den avvikende uttalen av /ʉ/ hos irakA, er mer stabil: Målingene viser nærliggende verdier enten ordet uttales som enkeltord eller forekommer i en tekst. IrakB er den av de tre irakerne som har andreformantverdier som ligger nærmest de norske: 1752Hz (*skutt-s*), 1614Hz (*vunnet-e*) og 1451Hz (*vunnet-t*).

For irakA ligger andreformantene i underkant av 1000Hz i de to ordene med lang /ʉ/: *uttalen-e*(nkeltord) har 920Hz og *mulig-t*(ekst) har 956Hz. Målingene indikerer en bakre lyd, og de to lange /ʉ/-ene høres også ut som /u/-lyder. Når det gjelder uttalen av de to lange /ʉ/-ene hos irakB og irakC, så er /ʉ/-ene i *mulig-t*(ekst) nær de norske selv om både F2 og F3 ligger litt over det norske gjennomsnittet. Når det gjelder /ʉ/-ene i *uttalen-e*(nkeltord),

så er det bare irakB som har tilnærmet morsmålnorske verdier (F2 er 1654Hz hos irakB, det norske gjennomsnittet her er på hele 1887Hz, svært høyt i forhold til andre målinger av /u/). Andreformanten hos irakC er 831Hz, og uttalen hans høres også ut som en /u/-lyd. Både irakB og irakC uttaler /u/ -en i *uttalen*-e(nkeltord) som en kort lyd. Kanskje er det de to -ene som følger vokalen i skrift, som gjør vokalen kort? Partikkelen *ut-* er ganske frekvent i norsk (ordet *uttale* brukes jo ofte i undervisning) og forekommer hyppig på fagspråkskurset for leger, som i ord som for eksempel *utskrive* en pasient, *utbrudd* av sykdom, *utskrift* av journal, *utbrent*, *utbetaling* av trygd osv, så det kan synes påfallende at uttalen ikke er mer fokusert her.

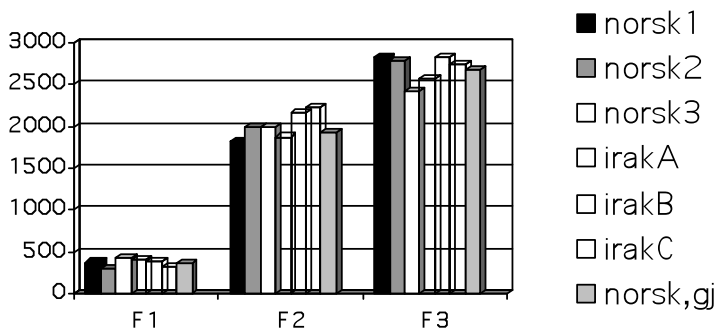
Sammenfatningsvis så kan alle de tre irakerne uttale /u/ slik at den høres ut som en morsmålnorsk /u/ i stavelsene *plu-* og *slu-*. Bare irakB, som er den av informantene som har bodd lengst i Norge (fem år), har akseptabel kvalitet på alle forekomster av /u/.

En kompliserende faktor når det gjelder uttalen av vokalen /u/ i norsk, er at bokstaven *u* uttales både som en /u/- og en /u/-lyd, og vi kan ikke gi gode og enkle regler for fordelingen. Hvorfor heter det for eksempel *skuff*, *truffet*, *bukke*, *ung*, *ungdom*, *pung*, *bukser*, *sukker* osv med /u/-lyder, mens det heter *skuffelse*, *hun*, *du*, *luft*, *skummel*, *uvanlig*, *vindu*, *minutt*, *ubevisst*, *forundring*, *kunne*, *putte*, *skulle* med /u/-lyder? Alle ordene i den siste gruppen har jeg hørt uttalt med /u/-lyd hos legene med arabisk som morsmål.

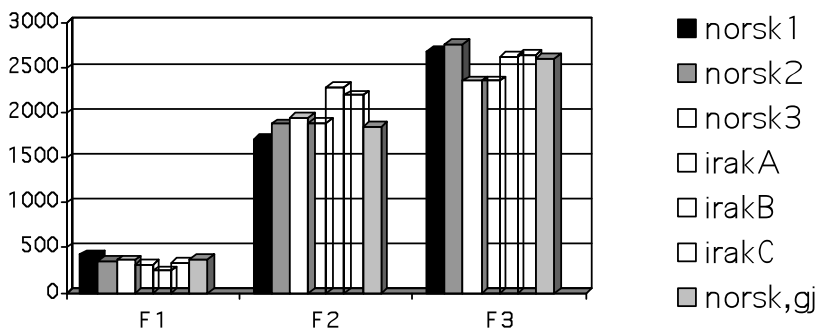
Jeg har ikke målt alle forekomster av /u/-lyder i materialet, men når jeg hører gjennom alle opptakene av de tre irakiske deltakerne, hører jeg ikke andre erstatningslyder enn /u/-lyder for /u/. Kanskje plasserer /u/ seg nærmere området til /y/ enkelte ganger, de høye verdiene til andreformanten hos irakC kan tyde på det (hos ham er F2 1988Hz i ordet *skutt*). En akustisk analyse med eksakte målinger av flere ord med /u/ hos disse og andre av de irakiske legene, kunne fortalt oss mer nøyaktig hva slags /u/-lyder erstatningslydene for norsk /u/ er, og hvordan de forholder seg til deres uttale av /u/. Dette er et tema for videre utforskning.

### Målinger av /y/-lyder

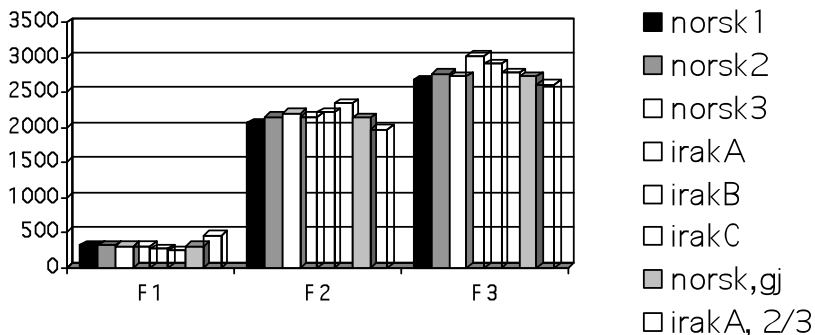
Figur 14: Kort /y/ (knytte – e):



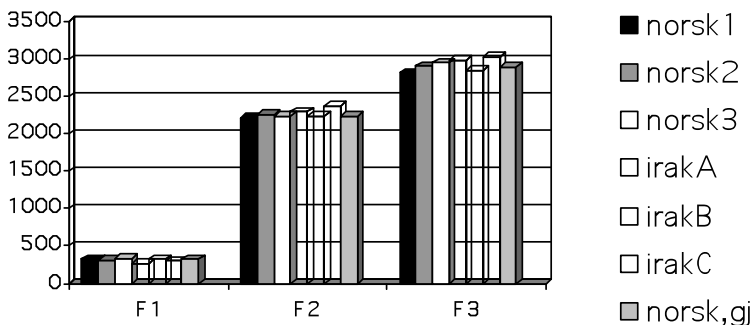
Figur 15: Kort /y/ (knyttet – t):



Figur 16: Lang /y/ (dyd – s):



Figur 17: Lang /i/ (tid – s):



### Hva viser målingene av /y/-lydene?

Det er ikke så mange forekomster av /y/ i materialet, men hvis vi sammenlikner formantverdiene for de korte /y/-lydene i *knytte-e*(nkeltord) og *knyttet-t*(ekst), er forskjellene så pass små at de ikke indikerer noen klar forskjell i uttalen. Likevel kan vi notere at andreformanten (F2) til både irakB og irakC er høyere enn både det norske morsmålsgjennomsnittet og hver enkelt av de tre norske deltakerne. Det kan tyde på mindre lepperunding og en mer fremre lyd, det vil si en mer i-aktig lyd. Lytter man på de to ordene hos de tre irakerne, så hører jeg likevel seks ord med hva jeg oppfatter som greie, korte /y/-lyder.

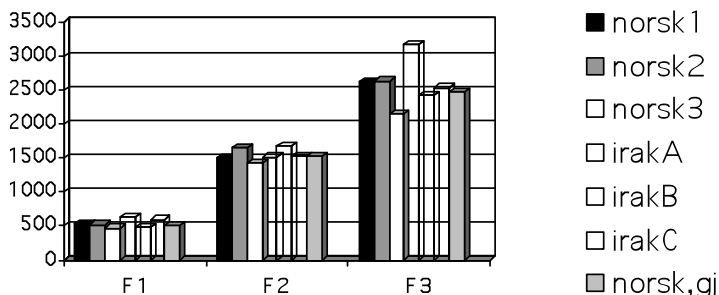
Når det gjelder lang /y/-lyd, så har jeg bare én måling som er fra ordet *dyd*-s(etning), som forekommer som siste ord i en setning i materialet. Studerer vi formantverdiene til den lange /y/-en i ordet, ser vi at de tre norske morsmålsbrukerne har noe lavere verdier for både andre- og tredjeformanten i *dyd*, men forskjellene er igjen små og mindre enn hva de var for andreformanten ved kort /y/-lyd. Lavere verdier her kan igjen peke mot sterkere lepperunding og en noe mer bakre lyd hos de norske informantene; irakB demonstrer nærmest at han kan uttale /y/ riktig ved å gjøre vokalen ekstra lang. IrakA har igjen sin særegne uttale, og hos ham kan jeg høre et tydelig kvalitetsskifte i uttalen: siste del av lyden høres ut som en /e/-aktig lyd. Formantverdiene målt totredjedeler inn i vokalen er 450Hz for F1, 1965Hz for F2 og 2605Hz for F3, verdier som kan karakterisere en mer åpen fremre lyd. Dersom det er riktig at vokalene i arabisk er ”rene” (s. 30), skyldes det tydelige kvalitetsskiftet innenfor denne lange vokalen ikke

morsmålpåvirkning. Kanskje er uttalen hans her heller et eksempel på at det er vanskeligere å uttale en lang ny og fremmed lyd enn en kort fremmed lyd?

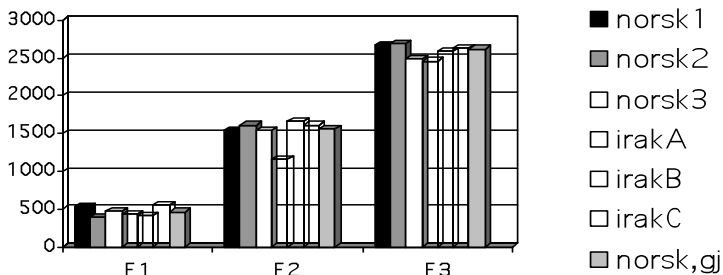
Sammenlikner vi lang /y/-lyd og lang /i/-lyd hos de seks deltakerne, ser vi at målingene for /i/-lydene er svært like og svært jevne for alle seks.

### Målinger av /ø/-lyder

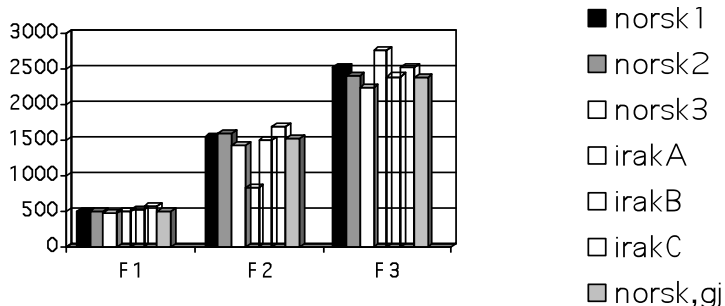
Figur 18: Kort /ø/ (først – t1):



Figur 19: Kort /ø/ (døtrene – e):



Figur 20: Lang /ø/ (prøve – t1):



## Hva viser målingene av /ø/-lydene?

Målingene av både de to korte /ø/-lydene og de lange /ø/-lydene (se tallverdier både for korte og lange /ø/-lyder til slutt i artikkelen) viser igjen at informant irakA har sin egen profil, mens formantverdiene hos både irakB og irakC avviker lite fra de morsmålsnorske, og uttalen deres av /ø/-lydene høres også ut som /ø/ med riktig lengde. IrakA er mindre stabil i sin uttale av /ø/. I ordene *først*, *prøve*, *øre* og *gjøre* høres /ø/-en ut som en /o/-lyd, og formantavvikene er tydelige for både første og tredjeformanten. I *døtrene* høres uttalen hans nærmest ut som en lang /u/-lyd, annenformanten er her betydelig lavere (1156Hz) enn hos de fem andre.

Den vanligste erstatningslyden for /ø/ er /o/, men irakA bruker en /u/-lyd i *døtrene*. Når irakA forteller hvordan han har lært norsk, bruker han også en tydelig /u/-lyd for /ø/ i ordene *bøkene*- f(ri tale) og *bøker*- f(ri tale), men i disse ordene kan /u/-lyden kanskje være en overføring fra entallsformen *bok*? Når man leser, kan vel også skriftbildet *ø* forbindes med skriftegnet *o* og dermed gi uttalen /u/.

Artikulatorisk karakteriseres /ø/ og /o/ ved trekkene halvåpne og rundede. De skiller ved at /ø/ er fremre mens /o/ er bakre. Arabisk har ikke kombinasjonen fremre og rundet. En alternativ erstatningslyd for /ø/ kunne tenkes å være /e/ som er fremre, halvåpen og urundet. Men så langt jeg har kunnet se, finnes ikke /e/ som erstatningslyd for /ø/. Hovederstatningslyden for /ø/ er /o/, med /u/ som en mulighet. Det er grunn til å framheve at to av de tre irakerne har en fullt ut akseptabel uttale av /ø/ i disse ordene.

## Oppsummering og avslutning

Avvikende uttale når det gjelder kvaliteten i de tre fremre, rundede vokalene /*u* /, /*y*/ og /*ø*/, finnes hos alle de tre irakiske deltakerne, irakA, irakB og irakC, i denne studien. Prediksjonen om at andrespråkstilegnere med arabisk som morsmål ”vil få problemer med de norske fremre [rundede] vokalene” stemmer. Men avvikene er svært ulikt fordelt de tre imellom. Oppholdstid i Norge kan være én forklaring. Den av dem som har vært lengst her i landet, har også færrest vokalavvik. Det er særlig hos irakA (som er den eldste av informantene, 42 år da opptaket ble gjort) at vi kan registrere mange avvik. Men det er faktisk slik at to av de tre irakerne, irakB (som har vært fem år i Norge) og irakC (som er yngst, 29 år, og bare har lært norsk i ett år) mestrer

alle de tre vokalene i alle omgivelser, det vil si enten ordet vokalen forekommer i, leses som et enkelt ord, om ordet står i en setning, i en tekst eller forekommer i fri tale. Det er flere kvalitetsavvik i (fonologisk) lang vokal enn i kort vokal.

Ett funn i denne studien synes jeg det ville være spesielt interessant å undersøke nærmere: Ordet *plutselig* forekommer 18 ganger i materialet. Alle de seks deltakerne leser ordet én gang som enkeltord, og ordet forekommer to ganger i teksten *Nordavinden og sola* som informantene leser høyt. I de 18 forekomstene av ordet *plutselig* viser formantmålingene omtrent de samme verdiene hos de tre irakiske deltakerne som hos de tre norske. Og i akkurat dette ordet (eller i stavelsen *plu-*) uttaler alle de tre irakerne /ʊ/ slik som morsmålsinformantene gjør. Det samme gjelder den ene forekomsten av ordet *slutt* (eller stavelsen *slu-*). Formantverdiene for /ʊ/ i *slutt* er også ganske like for alle de seks informantene, og uttalen av ordet høres morsmålsnorsk ut hos de tre irakerne. Dersom de øvrige fjorten irakerne jeg har opptak av, også viser spesielt målspåksnær uttale av /ʊ/ i nettopp disse ordene (eller stavelsene), kan man begynne å spekulere på hva det er ved disse ordene (eller stavelsene) som genererer den målspåkskorrekte uttalen av /ʊ/ her. Det forteller oss i hvert fall at det sannsynligvis ikke er tilstrekkelig å søke forklaringer på vokalavvik på fonemplanet.

Helt til slutt vil jeg framheve at en viktig erfaring for meg i denne studien er den åpenbare fordelene som ligger i *akustisk* analyse av vokal-kvalitet: Vi kan tallfeste hva vi hører, noe som er svært viktig i vitenskapelig sammenheng. Det vil likevel alltid være nødvendig å sammenholde den akustiske analysen med en persepsjonsanalyse, for det er kvaliteten på hørselsinntrykket som er det endelige målet for disse studiene.

## Litteratur

- Arnfast, Juni Söderberg 2007. Hvem tenker på accent? Om motivasjon og udtaletilegnelse hos voksne inlærere. *Nordand. Nordisk tidsskrift for andrespråksforskning*, årg. 2 nr. 2.
- Awde, Nicholas and Putros Samano 1986. *The Arabic Alphabet. How to read and write it*. London: Al Saqi Books.
- Bohn, Ocke Schwen 1995: Cross-Language Speech Perception in Adults. First Language Transfer Doesn't Tell it All. I W. Strange (ed). *Speech Perception and Linguistic Experience Issues in Cross Language Re-*

- search*. Baltimore: York Press.
- Bohn, Ocke Schwen and James Emil Flege 1997. Perception and production of a new vowel category by adult second language learners. I A. James and J. Leather (eds.): *Second-Language Speech. Structure and Process*. Berlin, New York: Mouton de Gruyter.
- Cunningham-Andersson, Una 1990. Native reactions to non-native speech. *Phonum* 1.
- Cunningham-Andersson, Una 1993. Stigmatized pronunciations in non-native Swedish. *Phonetic experimental research*, Institute of Linguistics, Universitetet i Stockholm
- Cunningham-Andersson Una 1997. Native speaker reactions to non-native speech. I A. James and J. Leather (eds.): *Second-Language Speech. Structure and Process*. Berlin, New York: Mouton de Gruyter.
- Handbook of the International Phonetic Association. A guide to use of the International Phonetic Alphabet*. University Press. Cambridge 1999.
- Harnæs, Liv Andlem 2001. Reaksjoner på utenlandsk aksent i norsk. *NOA norsk som andrespråk*, 2001, nr. 23.
- Husby, Olaf og Marit Helene Kløve 1998. *Andrespråksfoniologi. Teori og metodikk*. Oslo: Ad Notam Gyldendal.
- Kristoffersen, Gjøert 2000. *The Phonology of Norwegian*. Oxford: Oxford University Press.
- Kulbrandstad, Lars Anders 2006. Holdninger til “utenlandsk” aksent. I H. Sandøy og K. Tenfjord (red.). *Den nye norsken? Nokre peilepunkt under globaliseringa*. Oslo: Novus Forlag.
- Ladefoged, Peter 2003. *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques*. Malden, Mass.: Blackwell Publishing.
- Odlin, Terence 1990. *Language Transfer. Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.



**Vedlegg****Vedlegg 1:**

Formantverdier for **kort /ʌ/-lyd** i *plutselig* – e(nkeltord) og t(ekst) 1 og 2, målt en tredjedel inn i vokalen (verdier for totredjedeler noen steder i parentes):

Informant	F1	F2	F3
Norsk1	410,462,272(384)	1658,1641,1788	2466,2606,2636
Norsk2	368,384,351	1818,1838,1743	2772,2674,2483
Norsk3	391,396,380	1593,1652,1666	2062,2478,2244
IrakA3	90,359,446	1573,1454,1426	2785,2633,2633
IrakB	306,305,383	1694,1729,1626	2537,2495,2645
IrakC	418,350,370	1771,1831,1784	2570,2682,2523
Norsk,gj.*	390,414,334	1690,1710,1732	2433,2586,2444

\*(gjennomsnitt)

**Vedlegg 2:**

Formantverier for **lang /ʌ/-lyd** i *uttalen* – e(nkeltord) og *mulig* – e(nkeltord). Første tall er for *uttalen*, det andre for *mulig*:

Informant	F1	F2	F3
Norsk1	309, 345	1886, 1461	2651, 2281
Norsk2	302, 326	1891, 1700	2679, 2357
Norsk3	326, 374	1883, 1425	2269, 2183
IrakA	387, 429	920, 956	2661, 2399
IrakB	344, 285	1654, 1758	2545, 2609
IrakC	366, 288	831, 1701	2669, 2345
Norsk,gj	313, 348	1887, 1529	2533, 2274

**Vedlegg 3:**

Formantverdier for **kort /y/-lyd i *knytte*** – e(nkeltord) og ***knyttet*** – t(ekst). Første tall er for *knytte* – e, det andre for *knyttet* – t:

Informant	F1	F2	F3
Norsk1	375, 442	1828, 1709	2829, 2681
Norsk2	299, 348	1994, 1882	2782, 2765
Norsk3	431, 365	1987, 1951	2418, 2360
IrakA	401, 313	1875, 1887	2580, 2359
IrakB	379, 244	2170, 2291	2834, 2631
IrakC	325, 338	2237, 2195	2739, 2649
Norsk,gj	368, 378	1936, 1847	2676, 2602

**Vedlegg 4:**

Formantverdier for **lang /y/-lyd i *dyd*** – s(etning) og **lang /i/-lyd i *tid*** – s(etning). Første tall er for *dyd*, det andre for *tid*:

Informant	F1	F2	F3
Norsk1	308, 321	2054, 2211	2675, 2815
Norsk2	309, 309	2150, 2265	2762, 2909
Norsk3	302, 333	2194, 2240	2726, 2955
IrakA	304, 271	2154, 2289	3013, 2984
IrakB	260, 327	2211, 2241	2899, 2853
IrakC	239, 304	2335, 2373	2776, 3027
Norsk,gj	306, 321	2132, 2239	2721, 2893
IrakA, 2/3	450	1965	2605

### Vedlegg 5:

Formantverdier for **kort /ø/-lyd i først** – t1 og **døtrene** – e(nkeltord). Første tall er for *først*, det andre for *døtrene*:

Informant	F1	F2	F3
Norsk1	537, 527	1499, 1537	2615, 2676
Norsk2	518, 390	1653, 1612	2632, 2691
Norsk3	469, 474	1432, 1546	2153, 2485
IrakA	625, 434	1511, 1156	3170, 2463
IrakB	482, 411	1678, 1667	2418, 2593
IrakC	594, 552	1530, 1607	2535, 2631
Norsk,gj	508, 464	1528, 1565	2467, 2617

### Vedlegg 6:

Formantverdier for **lang /ø/-lyd i prøve** – t1, **øre** – e(nkeltord) og **gjøre** – s(etning). Første tall er for *prøve*, det andre for *øre* og det tredje for *gjøre*:

Informant	F1	F2	F3
Norsk1	496, 474, 512	1552, 1597, 1597	2526, 2545, 2429
Norsk2	504, 403, 469	1594, 1603, 1514	2410, 2710, 2340
Norsk3	473, 451, 480	1430, 1585, 1424	2230, 2349, 2244
IrakA	501, 578, 660	837, 1228, 1041	2770, 2659, 2260
IrakB	533, 499, 446	1494, 1667, 1597	2393, 2427, 2440
IrakC	583, 586, 477	1691, 1484, 1599	2513, 2535, 2374
Norsk,gj	491, 443, 487	1525, 1595, 1512	2389, 2535, 2338

# Norsk andrespråkskorpus (ASK) – design og metodiske forutsetninger

*Av Kari Tenfjord, Jon Erik Hagen og Hilde Johansen  
Universitetet i Bergen*

## **Sammendrag**

Artikkelen gir en presentasjon av Norsk andrespråkskorpus (ASK) – et nytt elektronisk innlærerkorpus som er utviklet ved Universitetet i Bergen. Korpuset inneholder personopplysninger om og tekster skrevet av kandidater som har gått opp til to ulike tester i norsk: *Språkprøven i norsk for voksne innvandrere* og *Test i norsk – høyere nivå*. Artikkelens hovedhensikt er å diskutere bruk av feilkoder i korpuset, og hvorvidt bruk av feilkoder i seg selv betyr at en havner i ”sammenlikningsfella”, som Bley-Vroman (1983) advarer mot i artikkelen ”The Comparative Fallacy”, altså at vår behandling av språkdataene skjer på målspråkets premisser og dermed undergraver mellomspråk som egne, selvstendige språkssystem. I artikkelen argumenteres det for at feilkoding slik det er gjort i ASK, må betraktes som en teorinøytral beskrivelse, en metodisk forutsetning for en teoretisk basert analyse av mellomspråk. ’Anklagen’ om å havne i sammenlikningsfella, et angrep som gjerne blir rettet mot klassisk feilanalyse eller enhver annen befatning med *feil*, har sitt utspring i at en ikke skiller klart nok mellom metode og teori i andrespråksforskningen.

*Nøkkelord: innlærerkorpus, sammenlikningsfella, feil, feilkoder.*

## **Introduksjon**

I denne artikkelen vil vi presentere Norsk andrespråkskorpus (ASK), et elektronisk korpus av innlærertekster som er utviklet ved Universitetet i Bergen.<sup>1</sup>

Hovedmålet med utviklingen av ASK har vært å skape et korpus av innlærertekster som kan styrke mulighetene for empiriske studier av norsk som andrespråk spesielt, og andrespråklæring generelt. En viktig forutsetning for å starte dette utviklingsarbeidet var muligheten for tverrfaglig samarbeid. Dataene er hentet fra Norsk språktests arkiver, språkteknologikompetansen er hentet fra Aksis (Avdeling for kultur, språk og informasjonsteknologi) og andrespråksforskningskompetansen er hentet fra tidligere Nordisk institutt, nå en del av Institutt for lingvistiske, litterære og estetiske studier. Utviklingen av ASK inngikk i NFR-prosjektet *Parallele korpus* som pågikk ved Aksis i perioden 2003-2006. Faglig prosjektleder for ASK var Kari Tenfjord.

Korpuset består ikke bare av innlærernes skriftlige tekstproduksjon, men inneholder også annotering av grammatiske kategorier, avvik fra målspråksnormen (feilkoder) og persondata knyttet til informantene.

Med denne presentasjonen av ASK vil vi gjøre rede for hvordan vi har gått frem ved utviklingen av korpuset, og ta opp noen grunnleggende spørsmål av mer epistemologisk art som er relevante for alle andrespråksstudier, nemlig forholdet mellom teori og metode innen vårt fagfelt, som til tider kan være problematisk. Vårt mål er å argumentere for den vitenskapsteoretiske holdbarheten i å bruke et korpus som dette i studiet av andrespråk, spesielt med tanke på at korpuset inneholder feilkoder.

Artikkelen er delvis ment som et motsvar til innsigelser vi har møtt i form av spørsmål om hvordan ASK håndterer kjente mellomspråksfenomener som variasjon, transfer, strategisk språkbruk, overgeneraliseringer osv., og antydninger om at vi ved å feilkode innlærerspråkene gjør oss skyldig i ”The Comparative Fallacy” (Bley-Vroman 1983), altså at vår behandling av språkdataene skjer på målspråkets premisser og dermed undergraver mellomspråk som egne, selvstendig språkssystem. Dette er spørsmål og innvendinger som i og for seg er velmotiverte og fornuftige, men med denne artikkelen ønsker vi å vise at det å kode en tekst for feil *i seg selv* ikke er å gå i en slik sammenlikningsfelle som Bley-Vroman med rette advarer mot.

I tillegg til spesiell vektlegging av feilkodingen i korpuset, ønsker vi også å gjøre rede for korpusets bruksmuligheter. Til tross for at feilkodingen har tatt mye oppmerksomhet og ressurser i utviklingsarbeidet, er søk etter ulike feiltyper bare en del av korpusets muligheter.

### Andrespråkskorpus – hva og hvorfor

Et korpus er ikke bare en samling data, slike som vi ofte har brukt i norsk andrespråksforskning, for eksempel 20 eller 50 tekster hentet fra testbesvarelser eller fra egne elever/studenter evt. med noen personopplysninger om informantene. Den vanlige prosedyren har vært at forskeren har transkribert håndskrevne tekster til maskinskrevne, og behandlet dem videre manuelt, både søk etter spesielle strukturer eller ord, og kvantifisering. Eventuelt har man lagt tekstene inn i regneark for å effektivisere behandlingen. Dette er tidkrevende og i forhold til elektroniske korpora mindre effektivt, og mengden av tekster begrenser seg selv i forhold til den tiden forskeren har til disposisjon. Det at vi nå har muligheten til å behandle store tekstmengder elektronisk, sparer mye tid og gir et sikrere grunnlag for statistiske analyser. Men for at et korpus skal kunne brukes som forskningsverktøy, må det tilfredsstillende visse kriterier. Grangers definisjon av et andrespråkskorpus som er basert på Sinclairs (1996) definisjon av (ordinære) korpora, er mye sitert:

Computer learner corpora are electronic collections of authentic FL/SL [Foreign Language/Second Language] textual data assembled according to explicit design criteria for a particular SLA/FLT [Second Language Acquisition/Foreign Language Teaching] purpose. They are encoded in a standardised and homogeneous way and documented as to their original and provenance (Granger 2002:7).

Korpuslingvistiske undersøkelser er ikke et eget forskningsfelt, heller må det kalles en metode.

Corpus linguistics can best be defined as a linguistic methodology which is founded on the use of electronic collections of naturally occurring texts, viz. corpora. It is neither a new branch of linguistics nor a new theory of language, but the very nature of the evidence it uses makes it a particularly powerful methodology, one which has the potential to change perspectives on language. (Granger 2002:4)

Granger påpeker videre at bruk av korpus heller er en komplementær datakilde enn en som erstatter andre, slik som for eksempel introspeksjon og elisitering. Men hun sier også at det er generell enighet om at korpus er "[...] the only reliable source of evidence for such features as frequency."

(McEnery & Wilson 1996:12). Når man bygger et *andrespråkskorpus* (*learner corpus*), må man også ta hensyn til at det er stor variasjon i andrespråksbruk (som delvis skiller seg fra variasjon man finner i vanlige monolingvale korpora) og at det er mange ulike faktorer som påvirker både språklæring og språkbruk.

### Datamaterialet i ASK

ASK består foreløpig av noe over 1700 tekster på 200-400 ord som er skrevet av kandidater som har vært oppe til *Språkprøven i norsk for voksne innvandrere*<sup>2</sup> og *Test i norsk – høyere nivå*. I henhold til definisjonene i *The Common European Framework of Reference for Languages* (CEFR) måler disse to testene et nivå i nærheten av henholdsvis B1 (*Threshold level*) og B2 (*Vantage level*). Dessuten består ASK av et kontrollkorpus på 200 tekster av samme type skrevet inn av personer med norsk som morsmål. Kontrollkorpuset skiller seg uunngåelig fra hovedkorpuset fordi tekstene ikke er skrevet i en testsituasjon.

En opplagt fordel ved å bruke testbesvarelser som data, er at slike tekster er rimelig homogene når det gjelder både tekstsjanger og produksjonskontekst, i og med at alle kandidatene har tatt testen under de samme fysiske rammebetingelser, med den samme tidsbegrensningen, og at deres testprestasjoner er blitt evaluert av sensorer med den samme type instruksjon og opplæring. Dessuten tilfredsstillende de også krav til homogenitet i den forstand at det kun er tekster fra beståtte besvarelser som er lagt inn i korpuset. Det vil si at tekstene i prinsippet er *på* eller *over* et visst funksjonelt nivå.<sup>3</sup> I tillegg gir persondataene som er lagt inn, opplysninger om flere variabler som kan ha påvirkning på språklæringen.

Utvalget av besvarelser er foretatt etter bestemte kriterier, motivert av følgende betraktninger:

- 1 Informantene skulle representere en rimelig typologisk variasjon av førstespråk (*morsmålskriteriet*).
- 2 Hver informant skulle ha en definert kommunikativ minimumskompetanse i norsk, noe som er sikret ved at de har bestått testen de har avlagt (*bestått-kriteriet*).
- 3 For hver informant skulle det foreligge personopplysninger som

kunne tenkes å være relevante for språkutviklingen, og som skulle kunne lenkes til kandidatens besvarelse (*info-kriteriet*)

Vi vil nå gå nærmere inn på de tre utvalgsriteriene.

### *Morsmålskriteriet*

Det viktigste kriteriet for utvalg av tekster til korpuset har vært innlærernes morsmål. Dette er et teoretisk motivert kriterium siden virkningen av nett-opp denne variabelen kanskje er den som har vært mest diskutert innenfor vårt fagfelt, samtidig som det ikke på noen måte er dristig å påstå at det i dag forventes av en hvilken som helst moderne teori om andrespråklæring at den skal kunne si noe om morsmålets virkning på språklæringsprosessen, språkbruk og språkutvikling.

Det er imidlertid metodiske problemer forbundet med det å teste hypoteser hvor morsmålet inngår som en forklarende faktor. Vi mener at en potensielt fruktbar måte å imøtekomme disse problemene på, er å utføre statistiske analyser av språk produsert av innlærere med ulik morsmålsbakgrunn, samtidig som man holder andre faktorer konstante. Dette er også i tråd med moderne metodiske innsikter i transferforskningen (jf. f.eks. Jarvis 2007 og Jarvis & Pavlenko 2008). ASK-korpuset muliggjør transferforskning etter slike metodiske retningslinjer, både fordi mengden av data, bredde i morsmålsgrupper og fordi produksjonsforholdene har vært like for alle informantene.

Av statistiske grunner ønsket vi å ha så mange som 100 tekster fra hver morsmålsbakgrunn på hvert av testnivåene, men på grunn av et ønske om typologisk variasjon blant førstespråkene, mangler vi fremdeles tekster fra høyere nivå for språkene albansk, vietnamesisk og somali.

Til sist var det ønskelig at de ti språkene som ble valgt, skulle representere rimelig store minoritetsgrupper i Norge. I utvalget ble til slutt følgende morsmålsgrupper representert: tysk, nederlandsk, engelsk, spansk, russisk, polsk, bosnisk-serbisk-kroatisk, albansk, vietnamesisk og somali.

### *Bestått-kriteriet*

Informantenes skriftlige prestasjoner er alle bedømt til å ligge *på* eller *over* visse definerte språkferdighetsnivåer. På begge de to testnivåene er *kommunikativ funksjonalitet* lagt til grunn for bedømmelsen. For å bestå disse testene måtte altså kandidatene være i stand til å kommunisere sitt intenderte innhold etter kravene på hvert nivå. Dette kravet er svært viktig, det er nem-



lig ikke bare et spørsmål om hva som er praktisk, men det er faktisk helt essensielt for den metodologiske validiteten til ASK. Dette poenget vil bli drøftet nærmere når vi gjør rede for feilkodingsprosessen, og i den bredere diskusjonen av sammenligningsfella lenger ned.

### *Info-kriteriet*

Selv om testomstendighetene for alle kandidatene prinsipielt er lik, gjelder ikke dette for innlæringskonteksten, og i så måte er det forskjeller på kandidatene. Derfor har vi i tillegg til selve testbesvarelsene opplysninger om hver av kandidatene som kan være relevant for forskning på andrespråklæring eller tospråklighet, slik som alder, oppholdstid i Norge, antall timer med språkundervisning osv. Norsk språktest begynte tidlig med å innhente persondata fra testtakerne med tanke på muligheten for videre forskning på dataene. Det var frivillig for kandidatene å gi slike opplysninger til forskningsformål. Ved søknad til Datatilsynet om å legge inn tekster og personopplysninger i korpuset ble det stilt spesifikke krav til hvordan dataene skulle behandles.

Siden vårt korpus omfatter denne typen informasjon om alle kandidatene, er det mulig å kontrollere for en rekke ulikheter både i innlæringskontekst og i innlærerbakgrunn, noe som kan være av stor betydning. Vi er derfor rimelig trygge på at korpuset består det kriteriet som Granger (2002:9) foreslår: ”The usefulness of a learner corpus is directly proportional to the care that has been exerted in controlling and encoding the variables.”

## **Sammenligningfella i andrespråksforskning**

Bley-Vroman (1983) påpeker innsiktsfullt at et *målspråksfiksert* språklæringssyn som går ut på at språktilegnelse handler om å tilegne seg flere og flere av målspråkets strukturelle og leksikalske entiteter helt til en er i mål, ikke åpner opp for mellomspråkets autonomi, og dermed hindrer en i å identifisere de mønstre og krefter som former dets utvikling og de faktorer som påvirker dette. Hvis man i forskningen begrenser seg til en systematisk analytisk sammenlikning mellom morsmålet og målspråket ved hjelp av *målspråkets* strukturelle kategorier, vil man ikke være i stand til å se andre, mer grunnleggende mønstre og strukturer i mellomspråket. Man vil simpelthen ikke kunne identifisere dets strukturelle og funksjonelle egenart.

Bley-Vroman understreker dermed hvor viktig det er å bevare mellom-språkets prinsipielle uavhengighet av *både* morsmålet og målspråket i struktur, virkemåte og utvikling.

Som påpekt av f.eks. Year (2004), er det en grunnleggende feil å tro at man mer eller mindre kan "lese av" informantens mellom-språkskompetanse med denne typen målspråksfiksert metodikk. Lakshmanan & Selinker (2001) utvikler også dette poenget, i tillegg til White (2003), Pienemann (1998), Larsen-Freeman & Long (1991). Som denne lista over henvisninger viser, er synet på *sammenlikningsfella* som en metodisk og teoretisk feil, noe det er bred enighet om innen andrespråksfeltet, til tross for høyst ulike teoretiske tilnærminger til feltet ellers. Som antydnet ovenfor, deler også vi denne enigheten fullt ut.

Den form for feilkoding en kan finne i ASK, kan nok tilsynelatende virke som en øvelse i nettopp den type andrespråksforskning som det så sterkt advares mot. Når vi hevder at dette bare er tilsynelatende, er det fordi vi mener man må skille mellom den analytiske og *preteoretiske* informasjonen som korpuset gir, og de *teoretiske modeller* vi konstruerer for å *forklare* denne informasjonen. Det første er et spørsmål om *observasjon*, dvs. å identifisere fenomener som krever forklaring, mens selve forklaringen er et spørsmål om *teori*. Vi mener at disse to nivåene ofte er blitt forvekslet i andrespråksfeltets nyere historie, noe vi kommer nærmere inn på i slutten av artikkelen.

Vi håper at vår presentasjon av ASK og de observasjoner og refleksjoner vi gjør oss i den forbindelse, vil demonstrere at termen *språkfeil* i andrespråksforskning, til tross for det er en upopulær term blant mange av dagens forskere, er en fullstendig legitim metodisk begrepsanvendelse innen for rammen av en reflektert forskningsprosess. Det samme gjelder feilkoding som praksis. Det er *ikke* slik at det å registrere og kategorisere feil nødvendigvis forutsetter en teori som søker innsikt i andrespråkskompetanse ved hjelp av å systematisk beskrive andrespråksperformansen gjennom målspråkets strukturelle kategorier alene, eller som kategoriserer et andrespråktrekk som noe som *samsvarer* eller *avviker* fra målspråket. Vi vil komme tilbake til dette etter å ha beskrevet prinsippene som ligger til grunn for feilkodingen i ASK.

## Feilkodingsprosessen

*Helhetlig tolkning av teksten forut for rekonstruksjon og koding.*

Hver enkelt tekst i korpuset har blitt lest og innholdsmessig tolket så nøy-

aktig som mulig. For å kunne bruke denne typen tekst som forskningsdata, må vi kunne forutsette at informantene uttrykker et rimelig klart, identifiserbart og koherent innhold, og at dette innholdet i tillegg er forståelig med innfødt kompetanse i normert norsk som eneste språklige ressurs. Siden kandidatens språkbruk er vurdert av sensorer og sertifisert som kommunikativt adekvat, kan vi i utgangspunktet konkludere med rimelig sikkerhet at kandidatens kommunikative intensjon er begripelig for en hvilken som helst innfødt norsktalende, selv om den logisk nok er formulert i hennes individuelle mer eller mindre målspråklige mellomspråk. Et viktig poeng her har vært at systematisk mellomspråksanalyse *ikke* skulle være nødvendig for en vellykket tolkning, og en slik analyse har til og med vært eksplisitt forbudt i retningslinjene for feilkodingen.

Generelt vil en hvilken som helst tekst, enten den er produsert av en innfødt språkbruker eller ikke, ikke bli fortolket bare på grunnlag av sitt bokstavelige innhold, men i høy grad også på grunnlag av dens kontekst. Både den umiddelbare språklige konteksten, den litt mer omfattende situasjonskonteksten og den i siste instans brede kulturelle konteksten, er alle relevante nivåer for presis tolkning av forfatterens kommunikative intensjon. Vi mener derfor at denne typen betraktninger er relevante forut for rekonstruksjon og feilkoding, mens systematisk mellomspråksanalyse derimot *ikke* er relevant.

Grunnen til at bestått-kriteriet har vært så viktig, er at det å identifisere språkbruksfeil er meningsløst og logisk umulig hvis man ikke i utgangspunktet har en rimelig klar oppfatning av hva som er ytringens intenderte innhold. Denne fundamentale innsikt er en grunnpillars for Corders klassiske algoritme for feilanalyse (Corder 1973). Vår dokumentasjon av kandidatens generelle kommunikative kompetanse kombinert med en rimelig kontekstuell tolkning imøtekommer etter vår mening, i alle fall til en viss grad, den generelle innvendingen mot feilanalyse som Lakshmanan & Selinker's (2001) har, nemlig at teksten kanskje ikke gir en adekvat gjengivelse av forfatterens kommunikative intensjon.

### *Prinsipper for valg av rekonstruksjon*

Vi forutsetter altså at vi utelukkende på grunnlag av morsmålskompetanse i norsk i hovedsak har klart å identifisere en koherent intensjon fra forfatterens side. Når denne delen av jobben er gjort, har neste steg vært å identifisere de stedene hvor teksten skiller seg fra målspråkssnormen og velge en rekonstruksjon som kan danne grunnlaget for selve feilkodingen. Tekst-

utdrag med mer enn én mulig rekonstruksjon, har ofte forekommet, og ideelt sett burde man naturligvis registrere alle logisk tenkelige rekonstruksjoner, og kode hvert enkelt alternativ. Dette ville imidlertid ikke være en gangbar løsning fra et praktisk og økonomisk synspunkt. I stedet har vi ved utvalget fulgt to grunnleggende prinsipper som skal sikre en mer eller mindre homogen praksis blant ulike kodere:

*Det pragmatiske probabilitetsprinsippet (PP-prinsippet)*

Velg den tolkningen av innholdet som er *mest sannsynlig* fra et pragmatisk synspunkt når man tar i betraktning teksten som helhet, i tillegg til den situasjonelle og kulturelle konteksten.

*Det minimale modifikasjonsprinsippet (MM-prinsippet)*

Velg den rekonstruksjon av teksten som utgjør *minst mulig endring* av originalen. Med andre ord: Velg det alternativet som medfører minst korreksjon.

Disse to prinsippene er ikke alltid kompatible, noe som til tider har framstått som et dilemma, men dette er etter vårt syn uunngåelig uansett hvilken strategi som velges. Følgende eksempler med bruk av 'rød vin' kan illustrere hvordan PP-prinsippet trekker i en retning og MM-prinsippet trekker i en annen, men en skjønsmessig vurdering tilsier at vi gir PP-prinsippet forrang:

\*Det er vanlig å ha litt rød vin med måltiden.

Her vil to innfødte rekonstruksjoner være mulige:

Rekonstruksjon 1: Det er vanlig å ha litt *rød vin* til måltidet.

Rekonstruksjon 2: Det er vanlig å ha litt *rødvin* til måltidet.

Alternativ 1 forutsetter at det er snakk om *all rødfarget vin*, uansett *vinkategori*, mens alternativ 2 forutsetter at det er snakk om *rødvin*, uansett *farge*. Som innfødte språkbrukere vet vi at *rød vin* og *rødvin* betyr to forskjellige ting, og denne betydningsforskjellen er reflektert i skriftspråket ved særskrivning versus samskrivning. Velger vi alternativ 1, er jo frasen egentlig feilfri som den står, og derfor vil MM-prinsippet si at det er dette alternativet vi må velge - ingen korreksjon kan vel være mer minimal enn

ingen modifikasjon i det hele tatt! Men det er langt mindre sannsynlig at kandidaten her har ønsket å vektlegge farge alene, snarere enn vinkategori, så på dette pragmatiske grunnlag vil være legitimt å velge alternativ 2, i tråd med PP-prinsippet, til tross for at det er i strid med prinsippet om minimal modifikasjon.

Mens eksempelet ovenfor gjør det naturlig å gi PP-prinsippet forrang, er det MM-prinsippet som foretrekkes i kodingen av følgende eksempel:

\*Denne retter bruker vi flest i vårt daglige livet.

Denne setningen kan rekonstrueres på minst tre forskjellige måter:

Alternativ 1: Disse rettene bruker de fleste av oss *i vårt dagligliv*.

Alternativ 2: Disse rettene bruker de fleste av oss *i vårt daglige liv*.

Alternativ 3: Disse rettene bruker de fleste av oss *i dagliglivet*.

De tre alternativene har ulike adverbialledd i slutten av setningen, og er altså strukturelt sett ganske forskjellige. Semantisk og pragmatisk er de derimot nokså like, og derfor kommer ikke PP-prinsippet til anvendelse her, men i stedet fremhever MM-prinsippet alternativ 2 som det naturlige valg, da det er den rekonstruksjonen som medfører minst endring fra originalen. Beskrivelsen av den feilen som er begått, kan da begrenses til galt valg av bøyingsform på substantivet ("livet" i stedet for det korrekte "liv"), mens de to andre medfører mer omfattende korreksjoner som går over flere ord.

Her kan det innvendes at selv om alle tre alternativer i prinsippet er mulige, må man likevel kunne si at alternativ 3 er den mest idiomatiske varianten, men denne blir altså vraket av våre prinsipper. En kan likevel ikke se bort fra at enkelte kodere ville tillegge denne idiomfaktoren så stor vekt at den likevel ville bli foretrukket ut fra selve hovedprinsippet om at en innfødt rekonstruksjon skal være *innfødt*. Så her har det vært et visst rom for skjønn, og våre rekonstruksjonsprinsipper eliminerer derfor ikke arbitrær bedømmelse fullt ut. Dette rommet for skjønn er naturligvis en ulempe, men er ikke katastrofalt. For uansett kriterier vil forskeren ha mulighet til å vurdere og eventuelt forkaste vår valgte rekonstruksjon. Dette vil alltid være tilfellet uansett teknologisk hjelpemiddel - forskeren må selv inspisere sine data.

*Klassifisering av avvik ved hjelp av feilkoder*

Ved at alle tekstene i korpuset ble systematisk fortolket og rekonstruert i overensstemmelse med prinsippene skissert ovenfor, ble det mulig å identifisere feilene, dvs. de delene av teksten der mellomspråket ikke følger målspråksnormen, og kategorisere dem ved hjelp av *feilkoder*. Vi minner om at en klar tolkning av hver enkelt setning er en logisk forutsetning for å kunne rekonstruere og feilkode den. Tekstpassasjer som er konstruert slik at det intenderte innhold ikke lar seg tolke overhodet, kan derfor ikke gjøres til gjenstand for standard feilkoding. I slike tilfeller er hele passasjen kodet med "X" (*Uidentifiserbar feil*).

Der innlærertekstene skiller seg fra målspråket, er det blitt tildelt en feilkode og et forslag til korreksjon (se vedlegg 1). Feilkoden kategoriserer feilen ved grammatiske og leksikalske begreper som er relevante for beskrivelsen av målspråket, dvs. normert bokmål, men målspråkets kategorier er *ikke* brukt fordi vi tror at dette reflekterer den "rette" strukturen i mellomspråket, men av metodiske årsaker: For i det hele tatt å kunne snakke om språk trenger man et *metaspråk*, og som vi skal komme tilbake til, har vi valgt målspråket til denne funksjonen fordi vi ønsket et *teorinøytralt* metaspråk.

På samme metodiske grunnlag hadde vi heller ingen som helst motforestillinger mot å kategorisere ordforrådet i tekstene etter målspråkets ordklassekategorier. Ved hjelp av Oslo-Bergen-taggeren<sup>4</sup> har hver ordforekomst blitt automatisk tagget for ordklasse, morfologiske og syntaktiske trekk. Dette har gjort tekstene søkbare på samme måte som monolingvale norske korpora. Det er imidlertid problemer knyttet til å bruke en automatisk tagger som er konstruert for målspråket, på innlærerspråk, ettersom ord med ortografiske feil og andre former for avvik fra målspråksnormen, naturligvis er vanskelig å kjenne igjen for taggeren. I ASK er dette problemet løst ved at taggeren arbeider på *korreksjonen* av ortografiske feil (feiltyperen ORT), og ikke på innlærerens ortografi. Men både syntaktiske og morfologiske forskjeller mellom norsk og innlærerspråkene er naturligvis med på å 'forvirre' taggeren, noe som kan føre til at taggeren gjør feil også når det gjelder ordklassetagging. Det har derfor blitt utviklet muligheter for å redigere den automatiske taggingen manuelt, og dette har løst problemet, iallfall til en viss grad. Redigeringen har først og fremst vært av ordklassetaggingen, mens vi har unnlatt å redigere syntaktiske og morfologiske kategorier. Sistnevnte er verken mulig eller ønskelig, siden det faktisk forutsetter analyse av mellomspråkene. Likevel har både den morfologiske taggingen en viss

verdi, særlig i mer eksplorative søk. Men her er det viktig å ikke falle i sammenligningsfella.

For illustrasjon, gjengir vi en liste av våre feilkategorier her (se for øvrig Kodeboken på [www.ask.uib.no](http://www.ask.uib.no))

<p><b><u>LEKSEMFEIL</u></b></p> <p>W    galt ord</p> <p>ORT  ortografisk feil</p> <p>PART  sammensetningsfeil</p> <p>SPL  særskrivingsfeil</p> <p>DER  avledningsfeil</p> <p>CAP  galt valg av stor/liten bokstav</p> <p>FL   ord fra andre språk enn norsk</p>	<p><b><u>SYNTAKSFEIL</u></b></p> <p>M    ord eller frase mangler</p> <p>R    ord eller frase er overflødig</p> <p>O    ordstillingsfeil</p> <p>Feilkategorien O har følgende underkategorier:</p> <p>  INV    inversjon mangler</p> <p>  OINV   overinversjon</p> <p>  SCA    setningsadverbial i leddsetninger           plassert etter det finite verbet</p> <p>  MCA    setningsadverbial i hovedsetninger           plassert før det finitt verb</p>
<p><b><u>MORFEMFEIL</u></b></p> <p>F    gal morfosyntaktisk kategori</p> <p>INFL  gal form, men riktig morfosyntaktisk kategori</p>	<p><b><u>TEGNSETTINGSFEIL</u></b></p> <p>PUNC  gal tegnsetting</p> <p>PUNC M  mangler tegnsetting</p> <p>PUNC R  overflødig tegnsetting</p>
<p>X            uidentifiserbare feil, dvs. vanskelig/umulig å tolke tekstutdraget</p>	

### Illustrasjon 1: Feilkategorier i ASK

I tillegg til feilkodene som klassifiserer feilen, har koderne angitt et forslag til korreksjon. På grunnlag av denne korreksjonen, genereres det en rekonstruert utgave av hver enkelt tekst, som i teorien skal være grammatisk korrekt ifølge bokmålsnormen. På denne måten er ASK også et *parallelkorpus*, bestående av et korpus av originale innlærertekster og et korpus

med rekonstruksjoner av disse. De to parallelle korpusene er representert med hver sitt brukergrensesnitt, henholdsvis ASK og ASK-korrekt. Begge versjonene kan, sammen eller hver for seg, gjøres til gjenstand for empiriske studier gjennom de søkemulighetene som brukergrensesnittene gir anledning til. I visningen av søk har man også muligheten til å se originalsetningen parallellstilt med den rekonstruerte utgaven (se vedlegg 7).

### Feilkoding som teorinøytral beskrivelse

Feilkoding har lenge vært etablert som en standardprosedyre i innlærerkorpora<sup>5</sup>, en konsekvens av at korpora med denne typen tekster trenger egne teknikker i forhold til andre typer korpora, slik Granger påpeker:

[...] computer learner corpora quite naturally call for their own techniques of analysis [...] such as error tagging, which are specially designed to cater for the anomalous nature of learner language. (2002:18)

En slik mer eller mindre mekanisk analyseprosedyre gir ganske mye informasjon om språket i teksten, i den forstand at den skaffer til veie systematisert data om kandidatenes *språkbruk*. Men det er veldig viktig å være klar over at slike teknikker må være i overensstemmelse med den relativt moderne innsikt at innlærerspråk har sin autonomi som egne språkssystem, og at deres strukturelle egenart ikke uten videre kan tolkes som et derivat av morsmål eller målspråk. Det er derfor vi så sterkt vi bare kan, understreker at våre termer *feil* og *feilkoding* ikke benevner strukturelle trekk ved mellomspråket som sådant, men kun er tekniske termer for rent analytiske begreper som er empirisk og preteoretisk definert uten at noen spesiell teori om språklæring, språkbruk eller språkutvikling forutsettes. Vårt mål er å kunne bruke tekstkildene til å skaffe til veie og tilrettelegge data som det så er opp til andrespråksforskere, ikke oss som korpusutviklere, å foreslå gode teorier til å redegjøre for.

Det er også viktig å merke seg at ved feilkodingen er det *feilene*, ikke strukturtrekk ved mellomspråket, som har blitt kategorisert og definert, og derfor må ASK ikke forstås som en slags maskin som gir anledning til å avlese mellomspråksstrukturen hos informantene. I stedet kan en si at kodingen, både den automatiske ordklassetaggingen og feilkodingen, påfører tekstene en strukturell beskrivelse som er rimelig veldefinert og systematisk,



men i prinsippet *teorinøytral*. Det er en ekstern analyse som er påført tekstene og som gir en systematisk, sammenliknbar beskrivelse av hvordan mellomspråk kan avvike fra måten innfødte språkbrukere bruker målspråket på. Feilkodingen i korpuset er i så måte kun et redskap for identifikasjon og klassifikasjon av avvik fra den skriftlige målspråksnormen, en måte å påvise forskjeller mellom innlærerspråk og målspråket.

Vi kan altså si at feilidentifikasjon er en praktisk snarere enn en teoretisk oppgave, og at feilanalyse er et *metodisk hjelpemiddel*, ikke en *teori* om andrespråkslæring. Skillet mellom teori og metode har dessverre ofte blitt sammenblandet i vårt fagfelts historie, og når en leser faglitteraturen, kan en av og til få inntrykk av at det å befatte seg med feil overhodet, er å gå rett i Bley-Vromans sammenlikningsfelle. Men vi mener dette bare er en holdbar anklage så lenge man ikke skiller mellom feilanalyse som teori og feilanalyse som metode. Det er av stor viktighet at man ikke blander disse to nivåene sammen, noe Corder også poengterte så tidlig som i 1973, jf. følgende sitat, hvor temaet er utelatelse av målspråkets obligatoriske ubestemte artikkel i den engelske mellomspråkssetningen *\*there is bus stop*:

The omission of the article (in this case) is only the surface evidence for an erroneous or idiosyncratic linguistic system. A full description of the error involves ‘explaining’ it in terms of the linguistic processes or rules which are being followed by the speaker [...]. Of course, superficial description is a necessary condition for linguistic explanation but it is not a sufficient one [...]. (Corder 1973:277)

Vårt korpus er, med sin feilkoding, nettopp denne typen overflatebeskrivelse som Corder påpeker som en *nødvendig*, men ikke *tilstrekkelig* betingelse for forståelse av teksten og språket i den. Vår koding er en strukturell beskrivelse av mellomspråkstekstene ut fra målspråksgrammatikkens kategorier, *ikke* av mellomspråkets indre grammatiske struktur. Feilkodingen er derfor ikke et *resultat* av en mellomspråksanalyse, men et metodisk ledd som logisk nok må gå *forut* for en slik analyse i forskningsprosessen.

## Håndtering av kjente mellomspråksfenomener

Vi har ved noen anledninger fått spørsmål om hvordan korpuset behandler kjente mellomspråksfenomener som transfer, feilkategoriseringer, over-

generaliseringer, kommunikasjonsstrategier osv. Vårt enkle svar er at vi *ikke* håndterer dem. Etter vår mening er termer som disse ikke empiriske teorinøytrale termer, altså *explanandum-termer*, men i stedet termer som forutsetter en spesifikk teori om hvordan spesifikt identifiserbare språktrekk kan forklares, altså *explanans-termer*. Oppgaven med å gi en teoretisk forklaring på de data ASK skaffer til veie, er forskerens, og ikke vår. Vi presenterer kun data som en hvilken som helst gangbar teori om andre-språklæring har som oppgave å forklare.

Dette springende punkt kan illustreres med følgende konkrete eksempel: Innlærere med engelskspråklig morsmålsbakgrunn produserer gjerne substantivfraser av typen: *\*det hus*, *\*den dag* og *\*den kirke*. Dette er strukturer som korresponderer med engelskspråklige uttrykk som *the house*, *the day* og *the church*, og det ville være nærliggende for noen å klassifisere disse innlærerformene som opplagte tilfeller av *transfer*, siden de så tydelig avspeiler de tilsvarende engelske formene. Andre igjen vil alternativt forklare denne mellomspråksforekomsten som et eksempel på *overgeneralisering*, siden man i norsk faktisk bruker 'den/det/de' som foranstilt bestemt artikkel på stort sett samme måte som den engelske 'the', men med et mer innskrenket bruksområde da de ikke kan brukes umiddelbart foran kjernessubstantivet. Noen kunne til og med påpeke at fenomenet kan være et eksempel på det såkalte *multiple effects principle* (jf. Laksmanan & Selinker 2001, Gass & Selinker 1992), siden to angivelige andrespråklæringsprosesser, både *transfer* og *overgeneralisering* trekker i samme retning og gir samme resultat.

Disse formene blir i korpuset likevel ikke kodet som transferformer, overgeneraliseringer eller *multippel effect*-former, siden dette er termer som ikke er analytiske, men teoretiske. De er begreper som hører hjemme i teorier som er nyttige som redskap for å *forklare* våre data, men ikke for å *beskrive* dem. Vår jobb som korpusutviklere er å beskrive *hvordan*, men ikke *hvorfor*, de skiller seg fra målspråksgrammatikken, og det gjør vi ved hjelp av det målspråksgrammatiske metaspråket. I ASK ville eksempelvis ordet *det* i *\*det hus* bli kodet som "R" (redundant), mens ordet *hus* ville bli merket "F" (feil morfosyntaktisk form), og korrigert til *huset*. Videre kan det påpekes at både *det* og *hus* er gangbare ord i målspråksnorsk, hvor de har en ordklassetilhørighet som de blir tagget for i overensstemmelse med målspråkets leksikon. Følgelig vil en forsker som bruker korpuset og ønsker å teste en hypotese om hvordan engelskspråklige innlærere uttrykker bestemthet, kunne søke etter relevante avvik, enten det er strykninger,

tilføyelser eller omrokkeringer, og kan på den måten få bekreftet eller avkrefte hypotesen sin.

Vårt mål er som nevnt å være strengt teorinøytral i den rene beskrivelsen, og derfor tar vi overhodet ikke stilling til om den bruken av bestemt artikkel som er beskrevet ovenfor, skyldes transfer, overgeneralisering, den kombinerte effekt av begge eller ingen av delene. Vår måte å kode selve avviket på, er og blir den samme uansett hva feilen kan forklares med.

Ett annet eksempel kan illustrere dette: Setningen *\*I går jeg dro til byen*, inneholder et klassisk tilfelle av inversjonsfeil av typen underinversjon. Den ukontroversielt rette rekonstruksjon må her være: *I går dro jeg til byen*. Som vi vet, finnes det høyst ulike, til dels inkompatible, tilnærminger til dette fenomenet, og enten en vil forklare det ut fra en teori om grammatisk kompetanse (jf. Hammarberg & Viberg 1979), en teori om informasjonsprosesser (jf. Håkansson 2001) eller en teori om kommunikativ funksjonalitet (jf. Lund 1997), så er det vårt håp at ASK skal være nyttig for alle andrespråksforskere *uansett teoretisk orientering*. Korpuset har som mål å skaffe til veie eksempler på selve fenomenet, uten å si noe om hva det skyldes, og bruker derfor ikke termer som er spesielle for noen av disse tre teoretiske tilnærmingene i sin beskrivelse av avviket.

### Det epistemologiske grunnlaget for vår tilnærming

Prinsippet om teorinøytral beskrivelse er et metodisk grunnprinsipp for ASK, og de begrepene vi bruker er empirisk snarere enn teoretisk forankret, noe som gjør at det blir meningsløst å diskutere om korpuset er ”riktig” eller ”galt”. Det som det derimot er meningsfullt å spørre om, er om den gir data som er relevante eller ikke for en gitt problemstilling. Siden teorinøytral beskrivelse av dataene er et så viktig grunnlag for oss, ønsker vi å gi vår argumentasjon en mer generell vitenskapsteoretisk forankring.

Vår grunnleggende vitenskapsteoretiske premiss er følgende: For at et hvilket som helst observert fenomen kan forklares, må det identifiseres og beskrives så presist som mulig uavhengig av teoretiske fordommer, og selve det begrepsmessige rammeverket for observasjonen (*observasjonsteorien* i Poppers terminologi) må være uten teoretisk problematiske begreper. Dette popperianske prinsippet er etter vårt syn en grunnleggende forutsetning for å teste ut en hvilken som helst teori i et hvilken som helst vitenskapsgren (jf. Popper 2002), og når vi bruker målspråksbeskrivelsen og en generalisert

form for feilklassifisering for å beskrive tekstene, blir det mulig å identifisere fenomener som enhver gangbar teori om andrespråklæring bør kunne forklare. Vårt deskriptive apparat utgjør i så måte ikke en teori, men en *modell* i Poppers forstand, dvs. et begrepsapparat generert av observasjonsteorien.

Av disse grunnene har det vært veldig viktig for oss at vi så langt som mulig kunne beskrive språket i tekstene med et begrepsapparat som er empirisk definerbart og teorinøytralt. Vårt eneste mål har vært å tilby en objektiv og preanalytisk beskrivelse av informantenes språkbruk, og for dette har vi funnet målspråksgrammatikken som velegnet, både for kategorisering av språklige kategorier og for beskrivelse av avvik fra målspråksnormen. I korpuset er dette den *eneste* rollen målspråksgrammatikken innehar.

Vi deler Poppers oppfatning rent prinsipielt, at med mindre en eller annen form for empirisk identifikasjon og rimelig presist definert preteoretisk beskrivelse av fenomener som trenger forklaringer er mulig, så blir empirisk validering av en hvilken som helst teori om dem ikke bare umulig, men kanskje til og med meningsløs. Så hvis det ikke skulle være mulig å identifisere språkbruk som avviker fra målspråksnormen, enten dette skjer i form av ”feil” eller på annen måte, hva slags grunnlag skulle en da ha for å evaluere teorier om andrespråklæring? Det må være mulig, iallfall i prinsippet, å observere, beskrive, klassifisere og referere til det en hvilken som helst teori om andrespråklæring/-utvikling/-bruk bør kunne forklare, uten dermed å binde seg til noen aprioriske forestillinger om innholdet i en slik teori.

### **Målspråksgrammatikken som modell for empirisk observasjon.**

Det å kunne beskrive fakta i en teorinøytral ramme er således ikke bare en mulighet, det er en forutsetning at det som skal forklares, lar seg identifisere uavhengig av den teori som måtte ha som mål å forklare det. Vi mener at målspråksgrammatikken er det opplagte valg for rollen som deskriptiv modell. Som nevnt ovenfor, forutsetter vi at datakilden, altså mellomspråkstekstene stort sett lar seg tolke semantisk og pragmatisk med innfødt kompetanse i norsk som eneste språklige ressurs. Vi mener at dette ikke bare er et tilgjengelig, men også et nødvendig og vitenskapsteoretisk velmotivert valg.

En kan derfor si at våre feilkodinger er en systematisk målspråksbasert beskrivelse av observerbare overflatefenomener, inklusive identifikasjon

og klassifikasjon av entiteter og strukturer som avviker fra målspråksnormen, uten at vi tar noen som helst stilling til hva som måtte være forklaringen på de avvik vi registrerer. Vårt mål er derfor ene og alene å tilrettelegge data, dvs. potensielle informasjonskilder, som gjør utprøving av teorier innenfor vårt fagfelt enklere og mer effektivt enn det som kanskje ellers ville vært mulig.

Bruk av målspråksgrammatikken som begrepsmodell er imidlertid ikke den eneste tenkelige løsningen. En alternativ tilnærming kunne være å utføre en uttømmende mellomspråksanalyse av alle tekstene i henhold til en eller annen teoretisk modell, for så å kode tekstene i samsvar med de enhetene og strukturelle begrepene som en slik teori motiverer og forutsetter. Vi nekter ikke for at et slikt prosjekt kunne være teoretisk mulig, men det ville for det første av mange grunner kreve større ressurser enn det som det overhodet ville være mulig å skaffe finansiering til i all overskuelig framtid. Men en mer interessant innvending er at anvendeligheten av en slik modell etter all sannsynlighet ville ha begrenset nytteverdi: Ulike teorier innenfor andrespråksforskningen er så ulike, så foranderlige, så usammenliknbare og så heterogene at et korpus bygd på grunnlag av en enkelt kontroversiell teori ville bli foreldet som rammeverk, og ville være mindre fleksibel enn vår av en rekke forskjellige grunner.

Vi mener at den store fordelen med vår tilnærming er at vår beskrivelsesmodell, målspråksgrammatikkens begreper, er velkjent, ikke bare som en skisse, men som en rik, detaljert og rimelig entydig norm. Dette gjør modellen velkjent for alle som har jobbet med feilkodingen, siden alle har utdanning i norsk som andrespråk, noe som ikke bare er praktisk, men som også sikrer en enhetlig referanseramme for alle koderne og for alle tekstene, noe som er av stor verdi, både for feilkodingens validitet og reliabilitet.

### **Hvorfor ASK-databasen ikke er en eneste stor sammenlikningsfelle**

Siden vår selvforståelse altså går på at utvikling av korpuset ikke er et teoretisk men et preteoretisk anliggende, erklærer vi oss *ikke skyldige* i en anklage om at vi begår den metodefeilen det er å sammenlikne innlærerspråket med målspråket. Vi mener at vi ikke går i sammenlikningsfella! Vi mener at Bley-Vromans og andres metodologiske problematisering av alle andrespråksteorier som ser mellomspråket med målspråksøyne, er en viktig kritisk påpekning. Men man har fullstendig misforstått hvis man inspirert av denne advarselen om gangbar andrespråksteori bannlyser enhver

sammenlikning med målspråket som *forskningsmetode*. Det ville ikke bare være å kaste ut barnet med badevannet, det ville være å begå en fundamental kategorifeil på begrepsnivå! Siden ASK med sine ulike kodinger er en datakilde, og ikke en teoretisk avhandling om andrespråklæring, kan den ikke avvises med Bley-Vroman argumentasjon.

Dette manglende skillet mellom teori på den ene side og forskningsmetode på den andre, mener vi er en begrepsfeil som har ridd andrespråksfeltet som en mare i mange år. *Kontrastiv analyse*, for eksempel, framstilles ofte både som en teori og som en forskningsmetode. Som metodisk redskap har jo språksammenlikning slett ikke utspilt sin rolle, selv om den teoretiske forestillingsverden som gjorde det særlig nærliggende å bruke denne metoden ikke lenger lever. Det som har utspilt sin rolle er naturligvis kontrastiv analyse som *teori*. På samme måte kan det til og med godt hende at de såkalte *morfemstudiene* rapportert i en serie publikasjoner av Dulay & Burt på 70-tallet (jf. f.eks. Dulay & Burt 1974) i prinsippet fremdeles kan være en relevant metode for andrespråksforskning, fortutsatt at de resultatene denne framskaffer, kan tolkes produktivt av en moderne teori om andrespråksforskning, noe som den naive universalismen som i sin tid motiverte denne metoden, *ikke* var.

Poenget vårt er at den meget rettkomme kritikken mot de teoretiske forestillingene som metodisk kontrastiv analyse, morfemstuder og feilanalyse var motivert og inspirert av, ikke må forveksles med den potensielle nytten disse metodene har som forskningshjelpemidler i seg selv. Da Schachter (1986) påviste "*an error in error analysis*", beviste hun ikke at feilanalyse var "galt", men at den som metode hadde visse begrensninger i utforskning av hennes teori om unnvikelsesstrategisk språkadferd, noe som i sin tid var en meget viktig innsikt. Poenget er at denne påpekningen av feilanalysens uanvendelighet for dette spesifikke formålet ikke på noen som helst måte diskvalifiserer feilanalyse som metode på et mer generelt plan.

Vi synes at denne tendensen til å forveksle teori og metode i SLA har forårsaket mye unødig kontrovers. Derfor er det viktig å understreke at ASK essensielt er et metodeverktøy som ikke forutsetter noen spesielle teoretiske føringer. Derfor er det å utnytte korpuset som *kilde* til potensiell viten om andrespråksfenomener ikke noen *comparative fallacy*. Metodisk sammenlikning mellom innlærerspråk og målspråk er ikke i seg selv å gå i sammenlikningsfella - med mindre man skulle mene at *enhver* referanse til målspråket i andrespråksforskningen er en metodefeil, en posisjon vi mener er fullstendig uholdbar.

Som det framgår av vår redegjørelse for strukturen i korpuset, har vi både i dens arkitektur og måten vi har arbeidet på søkt å unngå dette helt fra grunnen av, i det fromme håp at korpuset skal bli nyttig for forskere som arbeider med norsk eller andre nordiske språk som andrespråk i Norge og verden for øvrig.

### Søkemuligheter i ASK

Siden artikkelens hovedhensikt har vært å diskutere bruk av feilkoder i andrespråkskorpus har dette fått mye oppmerksomhet i artikkelen. Søk på feil er imidlertid ikke noe hovedpoeng for mange forskere, og vi vil derfor også rette fokus mot andre søkemuligheter som korpuset gir. Det er nemlig fullt mulig å bruke ASK, uten å ta utgangspunkt i feil og feilkoder. Leech har utviklet sju maksimer for annotering av korpus (Leech's Maxims of Annotation), og den første maksimen lyder: "It should be possible to remove the annotation from an annotated corpus in order to revert to the raw corpus" (Leech 1993 parafrasert i McEnery & Wilson 1996:25). ASK oppfyller dette kriteriet, da korpusets datamateriale lett kan brukes uten å utføre søk på feilkoder, uten at feilkodene kommer til syne eller på noen måte forvansker tilgangen til rådataene, dvs. tekstene i sin opprinnelige form. For eksempel kan man søke etter forekomster av bestemt ordklasser, enkeltord eller lemma, eller strenger av ord, for eksempel *verb + preposisjon, substantiv + possessiv* eller omvendt: *possesiv + substantiv, ordet det + lemma av verbet være*. Vi kan også søke på bestemte uttrykk som *i hvert fall, i dag* osv. I tillegg er det mulig å bruke regulære søkeuttrykk, dvs. søkeuttrykk som ikke er lagt inn i korpuset søkemeny. Alle søk kan vises i konkordansformat, eller man kan få frem bruksfrekvensen. Frekvensen kan sorteres etter ulike personopplysninger som morsmål, alder, oppholdstid i Norge osv. Men kan også begrense søkene til å gjelde bestemte morsmål, bare en av testtypene, eller etter andre personvariabler, slik at man ikke søker i alle 1700 tekstene. Se vedlegg 2-7 for illustrasjoner av ulike søk og visningsformat.

Koblingen mellom språkdata, persondata og et effektivt brukergrensesnitt gjør det mulig å bruke ASK til å belyse et bredt spekter av problemstillinger innen andrespråksbruk, andrespråksutvikling, andrespråkslæring og testing av språkkompetanse, ikke minst slike problemstillinger som er generert gjennom tidligere studier i norsk og nordisk andrespråksforskning.

Vi håper imidlertid også at korpuset skal åpne for testing av nyere teorier innen feltet, og ikke minst håper vi at det skal gi rom for spennende eksplorative studier, både av grammatiske, leksikalske, pragmatiske og tekstlige fenomener, i tillegg til undersøkelser av språkesterne faktorerers betydning for språklæringsprosessen; som informantenes alder, kjønn, antall år i Norge, opprinnelsesland, morsmål, utdanningsbakgrunn osv.

### Noen avsluttende merknader

Selv om korpuset har rike muligheter til søk uten å bruke feilkodene i det hele tatt, vil vi likevel avslutte med at feilkodingen også er en viktig ressurs i ASK. Vi vil dessuten minne om at referanse til målspråket i andrespråkforskningen er omfattende. Flere av de veletablerte termene i forskningsfeltet vårt forutsetter faktisk en form for målspråkssammenlikning som i høy grad kan diskuteres: både begrepene *tilegnelse*, *fossilisering*, *unnvikelse* og *overgeneralisering* tar utgangspunkt i en form for metaforikk som forutsetter at målspråket er en gjenstand for tilegnelse, selv om mellomspråk i dag betraktes som noe som *skapes* snarere enn tilegnes. Det er altså ikke bare begrepet *feil* som er problematisk. Når vi tenker oss om, finner vi kanskje at hele metaforstrukturen selv i moderne andrespråksteori i større eller mindre grad forutsetter sammenlikning med målspråket, ikke minst selve termen *mellomspråk* gjør det, siden den antyder en reise fra et utgangspunkt til et mål.

### Noter

1. Denne artikkelen baserer seg i stor grad på en tidligere artikkel (Tenfjord, Hagen og Johansen 2006) publisert i *Rivista di Psicolinguistica Applicata (RiPLA)*. Korpuset er også presentert i Tenfjord (2004, 2007) og Tenfjord, Meurer og Hofland (2006).
2. Tilsvaret Norskprøve 3 etter den nye læreplanen fra 2005.
3. ASKeladden-prosjektet studerer transfer fra morsmålet i innlærerspråk. Studier er korpusbasert og ASK er den viktigste datakilden. Prosjektleder er Kari Tenfjord, og Norges forskningsråd finansierer prosjektet for perioden 2008–2012. Videreutvikling av ASK er et viktig delmål i ASKeladden, bl.a. blir 1200 av tekstene vurdert på nytt for å få en sikrere og mer detaljert plassering på grunnlag av beskrivelsene i CEFR. Dette delprosjektet ledes av post.doc Cecilie Carlsen ved ASKeladden.
4. Oslo-Bergen-taggeren er en automatisk tagger for norsk bokmål og nynorsk, og en beskrivelse av den finnes på følgende adresse: <http://omilia.uio.no/obt/les.html>
5. Feilkoding er brukt i de største innlærerkorpusene vi kjenner til, for eksempel *The*



*Internationale Corpus of Learner English (ICLE) og The Cambridge Learners' Corpus (CLC).*

## Litteratur

- Bley-Vroman, Robert 1983. The Comparative Fallacy in Interlanguage Studies: The Case of Systematicity. *Language Learning. A Journal of Applied Linguistics*, 2006, volume 33,1, s. 1 – 17.
- Corder, Stephen Pit 1973. *Introducing Applied Linguistics*. Harmondsworth: Penguin Education.
- Council of Europe 2002. *The Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Dulay, Heidi & Burt, Marina 1974. Natural sequences in Second Language Acquisition. *Language Learning* 24, s. 37-53.
- Gass, Susan & Selinker, Larry 1992. *Second Language Acquisition: An Introductory Course*. Hillsdale NJ: Erlbaum.
- Granger, Sylviane 2002. A bird's-eye view of learner corpus research. I: S. Granger, J. Hung og S. Petch-Tyson (red.). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam and Philadelphia: John Benjamins.
- Hammarberg, Bjørn & Åke Viberg 1979. The place holder constraint, language typology, and the teaching of Swedish to immigrants. *Studia Linguistica* 31 s. 106 – 133.
- Håkansson, Gisela 2001. Tense Morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition* 4,1 s. 85 – 99.
- Jarvis, Scott 2007. Theoretical and Methodological Issues in the Investigation of Conceptual Transfer. *Vigo International Journal of Applied Linguistics* 4, s. 43 –71.
- Jarvis, Scott & Aneta Pavlenko 2008. *Crosslinguistic Influence in Language and Cognition*. New York and London: Routledge.
- Lakshmanan, Usha & Larry Selinker 2001. Analysing Interlanguage: How do we know what learners know? *Second Language Research* 17, 4, s. 393 – 420.
- Larsen-Freeman, Diane & Michael H. Long 1991. *An introduction to second language acquisition research*. London and New York, Longman.
- Leech, Geoffrey 1993. Corpus Annotation Schemes. *Literary and Linguistic*

- Computing* 8, 4, s. 275 – 281.
- Lund, Karen 1997. *Lærer alle dansk på samme måde? En længdeundersøgelse af voksnes tilegnelse af dansk som andetsprog*. København: Special-pædagogisk forlag.
- McEnery, Tony & Andrew Wilson 1996. *Corpus Linguistics*. Edinburgh University Press.
- Pienemann, Manfred 1998. *Language Processing and Second Language development: Processability Theory*. Studies in Bilingualism 15, Amsterdam/Philadelphia, Benjamins.
- Popper, Karl R. 2002. *The Logic of Scientific Discovery*. Revised edition: Original 1959, London: Routledge.
- Schachter, Jacquelyn 1974. An error in error analysis *Language Learning* 27, s. 205 – 214.
- Tenfjord, Kari 2004. ASK – A Computer Learner Corpus. *CALL for the Nordic Languages*, s. 147 – 158. Copenhagen Studies in Languages 30, Samfundslitteratur.
- Tenfjord, Kari 2007. ”ASK and you will find what you seek”. I Cecilie Carlsen & Eli Moe (red.). *A Human Touch to Language Testing. A collection of essays in honour of Reidun Oanes Andersen on the occasion of her retirement*. Oslo: Novus Press.
- Tenfjord, Kari, Jon Erik Hagen & Hilde Johansen 2006. The hows and whys of coding categories in a learner corpus (or How and why an error-tagged learner corpus is not ipso facto one big comparative fallacy). *Rivista di Psicolinguistica Applicata (RiPLA)* VI.3, s. 198 – 208.
- Tenfjord, Kari, Paul Meurer & Knut Hofland 2006. The ASK corpus: A language learner corpus of Norwegian as a second language. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*, s. 198 – 208.
- White, Lydia 2003. On the Nature of Interlanguage Representation. I Catherine J. Doughty & Michael H. Long (eds.): *The Handbook of Second Language Acquisition*, s. 19 – 42. Blackwell.
- Year, JungEun 2004. Instances of the Comparative Fallacy. I *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* 4,1.

## Vedlegg

Vedlegg 1: Eksempel på feilkoding og korreksjon i programmet Oxygen.

Vedlegg 2: ASKs søkegrensesnitt med eksempel på søk med feilkode.

Vedlegg 3: Visning av treff i Kwic-konkordans (Key Word In Context-konkordans)

Vedlegg 4: Eksempel på søk etter sekvens av ord.

Vedlegg 5: Visning av treff i Kwic-konkordans.

Vedlegg 6: De mest frekvente verbene i ASK (kolligasjonsøk).

Vedlegg 7: Visning av treff i setningsparallellestilt konkordans.

**Vedlegg 1:** Innlærertekstene er skrevet inn i programmet Oxygen, og ved hjelp av HTML-koding er språk som avviker fra norsk norm kodet med en feilkode og et forslag til korreksjon. Utdraget her viser en feil av typen SPL – særskrivingsfeil.

```
<p>Jeg bor i fjerde etasje. Vi er glade i <sic type="SPL" corr="panoramautsikt">panorama utsikt</sic>. Vi ser hele &sted1; gjennom vinduet. Jeg har to barn og jeg har alltid ønsket meg å bo på et sånt sted. To skoler og mange barnehager befinner seg i området. Jeg synes at det er veldig viktig for folk som har barn. Hvis jeg ville kjøpe en leilighet, så skulle jeg legge vekt på alle disse punktene.</p>
```

**Vedlegg 2.** Søk på substantiv med feilkoden F (feil morfosyntaktisk kategori) i Språkprøvetekstene.

Søkeuttrykk så langt: **[type='.\* F .\*' & features='.\* subst .\*' & testtype='Språkprøven']**

Oppdater    Tilbakestill skjemaet

+ 1. ord
target
- +

ord

ignorer stor/liten bokstav

**attributter:** (skjul)

feiltype

undertype

korleksjon

lemma

grammatiske trekk: (skjul)

ordklasse

morfologi

syntaks

**repetisjon**

fra  til

oppgavetittel

dokument

korpusvalg

**persondata:** (skjul)

testtype

hjemland

språk

alder

kjønn

[flere persondata ...](#)

Vedlegg 3: Utdrag fra Kwic-konkordansen for søket i vedlegg 2.

Vis kun ett treff per side | forrige 35 treff | neste 35 treff | treff: | bredde: 300px | Last ned

3 2 1 KWIC 1 2 3		korleksjon
rdi uten frihet har ikke liv. <S> Jeg mener at alle	<SIC> Landet	</SIC> i denne verd må gir til dene landet som
mna. <S> Lærerne anbefaler ham for å studere	<SIC> litteraturen	</SIC> for å fortsette på universitet. </S> </P> Etti
ar politik, historie, litteratur og den <SIC> daglige	liv	</SIC> av dette landet. </S> </P> Barneoppdragels
for vi mye av tid hadde vi på krig, for <SIC> dette	liv	var </SIC> ikke så bra for dem. </S> Men nå er jeg
problemer med helse. <S> De får problemer med	<SIC> lunger	</SIC> og en del får kreft, det betyr at alle som ik
gså vanskelig, de skal bytter skolen, venner og	<SIC> lærerne	</SIC> </S> De er trist, fordi skal misste vennene
ham. <S> De siste dagene for han dør, ber han	<SIC> mamma	</SIC> si til å gå hos Bernardo for å få tak i falke
orsk land, norsk historie. <S> Jeg lærte at første	<SIC> man	</SIC> som var for første gang i Amerika var fra N
ir første gang i Amerika var fra Norge, og første	<SIC> man	</SIC> som gikk først på dene sydpolen var fra N
uligheter å ta utdanninge som man vil, fordi var	<SIC> mangler	</SIC> i penger på skolene, dårlig økonomi, mang
e med menneskene de ikke kjenner osv., <S> Men	<SIC> mann	</SIC> har (overstr.) rett til å gjøre alt, som å gi t
ne som var på jobb. <S> I hele fabriken var bare	<SIC> mann	</SIC> så når de jentene kommer der på jobb all
r stort bosjet og har mulighet for å hjelpe disse	<SIC> mennesker	</SIC> som har sykdom eller ulykke, som det er ik
har godt framtid. <P> <S> Framtida? det betyr fri	<SIC> mening	</SIC> og ønsker for alle uten religion og farges t
dene vakre trærne, elver, jeg skal fortelle til alle	<SIC> menneske	</SIC> som har ikke opplevet naturen her. </S> <D
. <S> Det er min forslag og anbefaller jeg til alle	<SIC> menneske	</SIC> som bor i Norge for å bruke fritida sine fo
ige steder i verden hvor det er krig. <S> Det er?	<SIC> menneskes	</SIC> hjerne og hjerte? som gjør at det skal bli i
d krig nå? <P> <S> Nå har probleme i Kosovo, har	<SIC> mennesker	</SIC> der mat, klær? </S> </P> Og jeg er glad for
are med selskap. <S> Det var fint å være med de	<SIC> mennesker	</SIC> fordi hadde de morselige tinger å si og jeg
le mulighetene for å bli rik land eller livet til alle	<SIC> mennesker	</SIC> som bor der. </S> </P> Religion? Det er ikke
in skal jobbe for en god framtid, jeg vil at noen	<SIC> mennesk	</SIC> i verden skal blir fatig. </S> Til slut vil jeg si
im dagen. <S> Tenk å bo i nord Norge når det er	<SIC> midnattsola	</SIC> hva skal de gjøre. </S> Men jeg vill ikke sie
nske foreldre. <S> Når han blir voksen, går han i	<SIC> militær	</SIC> </S> Der han blir veldig film. Da han oppdi
: fodselen. <S> Det var den DATO klokka 700 om	<SIC> morgen	</SIC> når hun skulle til sentrallykshuset på STI
som de må passe på. <S> Når muslimene gå på	<SIC> moske	</SIC> kinnfolk kan ikke være på sammen sted m
este problem er kriminaliteten som stopper alle	<SIC> mulighetene	</SIC> for å bli rik land eller livet til alle menneske
bli store og studere her i Norge og hun har alle	<SIC> mulighetene	</SIC> til det. </S> Det som vi ikke hadde vi når v
ette jeg slutter og jeg vil bare takke folkene og	<SIC> myndigheten	</SIC> som passer på naturen. takk. </S> </P> Opt
gikk fortore enn før så da NAVNM ble to og halv	<SIC> måneder	</SIC> fikk vi oppholstatelse og den var det bes

**Vedlegg 4:** viser søk etter forekomster av 'i dag' i tekstene hentet fra Språkprøven i norsk for voksne innvandrere

Søkeuttrykk så langt: [word='i' %c & document1='no.*' & testtype='Språkprøven'] [word='dag' %c & document1='no.*' & testtype='Språkprøven'] Oppdater: Tilbakestill skjemaet	
1. ord <input type="text" value="i"/> <input type="button" value="+"/> <input type="button" value="-"/> ord <b>attributter:</b> (sliu) <input checked="" type="checkbox"/> ignorer stor/liten bokstav feiltype <input type="text" value="CAP"/> <input type="text" value="DER"/> <input type="text" value="F"/> undertype <input type="text" value="AGR"/> <input type="text" value="INV"/> <input type="text" value="MCA"/> korreksjon <input type="text"/> lemma <input type="text"/> grammatiske trekk ... <b>repetisjon</b> <input type="text" value="1"/> <input type="button" value="fra"/> <input type="text" value="1"/> <input type="button" value="til"/> <input type="text" value="1"/>	2. ord <input type="text" value="dag"/> <input type="button" value="+"/> <input type="button" value="-"/> ord <b>attributter:</b> (sliu) <input checked="" type="checkbox"/> ignorer stor/liten bokstav feiltype <input type="text" value="CAP"/> <input type="text" value="DER"/> <input type="text" value="F"/> undertype <input type="text" value="AGR"/> <input type="text" value="INV"/> <input type="text" value="MCA"/> korreksjon <input type="text"/> lemma <input type="text"/> grammatiske trekk ... <b>repetisjon</b> <input type="text" value="1"/> <input type="button" value="fra"/> <input type="text" value="1"/> <input type="button" value="til"/> <input type="text" value="1"/>
oppgavetittel <input type="text"/> dokument <input type="text"/> korpusvalg <input type="text" value="andrespråk"/> <input type="button" value="▼"/> <b>persondata:</b> (sliu) <input type="text" value="Språkprøven"/> <input type="button" value="▼"/> hjemland <input type="text" value=""/> <input type="button" value="▼"/> språk <input type="text" value=""/> <input type="button" value="▼"/> alder <input type="text" value=""/> kjønn <input type="text" value=""/> <input type="button" value="▼"/> flere persondata ...	
<input type="button" value="Søk innenfor: 100 ord"/> <input type="button" value="▼"/>	

## Vedlegg 5: Viser KWIC-konkordans for søket i vedlegg 4

; **Søk:** [word='i' %c & document!='no.\*' & testtype='Språkprøven'] [word='dag' %c & document!='no.\*'  
 6.  Vis kun ett treff per side |  | treff:  | KWIC  | bredde:  |  |

3 2 1 KWIC 1 2 3

å forklare. <s> Det gjelder rett og slett samfunnet som lever nå, i dag, dette året, i begynnelsen i av den nye alderen. </s> Men hvis v  
 fisjon og språk. Vi diskuterte og snakket sammen om alt. <p> <s> I dag jeg er glad at jeg har så mange venner fra mange land og jeg f  
 rmt leilighet, mat, venner som er nødvendige for samfunnet. <s> I dag vi har veldig modern medisin som er veldig viktig for samfunnet.  
 in man klæs seg, hvordan verdens samfunn utviklet seg. <p> <s> I dag er det ikke så mange som er flinke til å male, selv om har vi mye  
 l barna? For var foreldrene mest opptatt med oppdragelsen. <s> I dag begge foreldrene er opptatt hele dagen med jobb. </s> Derfor v  
 r. Norsk kultur | <p> For fire år siden, traff jeg en norskmann. <s> I dag er jeg gift med ham og bor i Norge. </s> Jeg måtte integrere meg  
 iktig hvis man kan pleier bli hjelpsom med andre også. <p> <s> I dag er det ikke så vanskelig for man å finne jobb i fleste land. </s> M  
 anter er også veldig viktig. <p> Først er god jobb muligheter. <s> I dag i vår verden er det ikke sikkert at vi vil ha våre jobber i morgen.  
 i gammel uttalelse redaktør seir " dårlig nyheter selges ". <p> <s> I dag er det mulig å få nyheter i mange forskjellige former. </s> Vi kan  
 993. Det koster kr 10 000. Faktura var kom fra Tyskland. <p> <s> I dag, mobiltelefoner er billige å kjøpe og billige å bruke. </s> Også di  
 m det. | <p> Mobiltelefoner ble innført cirka 15 til 20 år siden. <s> I dag har mange mennesker mobilen, du kan alltid hører dem i forsjell  
 er i Norge. Det finnes mobiltelefoner overalt, i hvert hjørnet. <s> I dag er det rett og slett vanskelig å finne et sted hvor det er inge  
 in mus s ikk og kolleger. Det er godt til sitt egen inspirasjon. <s> I dag er det ikke bare viktige mennesker som reiser men nesten hele  
 er veldig forskjellige, det tar tid for man kan bestemme seg. <s> I dag har jeg både marka og sentrum i nærheten og det betyr mye fo  
 nders. Som jeg kan huske meg, var det ett paradisi for barn. <s> I dag er det ikke mulig å spille fotball på gata men at jeg var 10 år ga  
 jeg huske meg at jeg var veldig fornøyd med rockmusikk. <p> <s> I dag har ungdom mye mer penger. </s> Jeg hadde bare nok å kjøpe i  
 iere dette problemet kan politi å kontrollere oftere trafikk. <s> I dag både i byen og utenfor byen er biltrafikk så stor at mennesker ε  
 hagen, gris og andre dyr og de hadde ikke så mye fritid. <p> <s> I dag er vi også veldig opptatt med familie, men vi har ikke så mye tid t  
 iaskin-. [særlig en mulighet til å sen g de e-poster] og andre. <s> I dag har samfunnet svært lett adgang til alt dette som skjer i verden  
 ham til å følge. For og nå | <p> Livet står ikke på stedet hvil. <s> I dag kommer nye teknologier. </s> Alt rundt oss utvikler seg. Samfunn  
 om fikk riktig oppdragelse er god eksempel for sine barn. <p> <s> I dag er det vanskelig å gi viktig oppdragelse til barn. </s> De Barna ei  
 ker har mange muligheter, men det er ikke så lett å leve nå. <s> I dag er det veldig viktig hva du har og hvor mye du har. </s> Mange r  
 r, og vi trenger ikke å go ut. Det muligheten kaller vi telefon. <s> I dag spiller telefonen veldig viktig rolle i vår livet. </s> <p> Hvis vi tre  
 lde sett sammen, bytte t s med forskjellige synspunkter. <p> <s> I dag kjenner jeg ikke så mange norske fjernsynsprogrammer for bar

**Vedlegg 6:** Liste over de 21 mest frekvente verbene i tekstene hentet fra Språkprøven i norsk for voksne innvandrere.

Søket ga 55754 treff. Det ble funnet 1042 forskjellige kolligasjoner.

<b>match lemma</b>	<b>absolutt frekvens</b>	<b>relativ frekvens</b>
være	11683	0.20955
ha	4956	0.08889
kunne	3014	0.05406
måtte	1543	0.02768
bli	1462	0.02622
skulle	1247	0.02237
gå	1089	0.01953
bo	1082	0.01941
få	1029	0.01846
gjøre	973	0.01745
komme	928	0.01664
ville	877	0.01573
se	867	0.01555
lære	830	0.01489
synes	737	0.01322
like	674	0.01209
si	601	0.01078
ta	566	0.01015
snakke	552	0.00990
tro	525	0.00942
tenke	507	0.00909



**Vedlegg 7:** En setningsparallellestilt visning av søk på ordstrengen 'jeg er', hvor verbet 'er' skal erstattes med et annet verb, gjør det mulig å se både originalsetningen og den rekonstruerte utgaven av den.

Treff 1 - 11 av 11. |  Vis kun ett treff per sic | setningsparallellestilt | Last ned | Nytt søk

**ord**

Jeg er på idrettslinja, så vi må reise litt for mye, fordi vi har masse aktiviteter.  
 Jeg går på idrettslinja, og vi må reise mye, fordi vi har masse aktiviteter.

Når jeg er på kollektivt transport føler jeg meg tryggere hvis jeg har min mobiltelefon.  
 Når jeg reiser kollektivt, føler jeg meg tryggere hvis jeg har min mobiltelefon.

Så hver gang jeg er fri kjører jeg til STED.

Så hver gang jeg har fri, kjører jeg til STED.

På jobb har vi nå skolekjøring nå, og jeg er veldig berørt når jeg ser barna i toget og på plen.  
 På jobb har vi nå skolekjøring nå, og jeg blir veldig rørt når jeg ser barna på toget og på plenen.

Når jeg tenker på det jeg er grusomt sint og tenker " hvorfor har det måtte skjedde til oss ".  
 Når jeg tenker på det, blir jeg grusomt sint og tenker: " Hvorfor har dette måttet skje med oss? "

Jeg håper at jeg kan få, når jeg er bestått alle eksamen.

Jeg håper at jeg kan bli det, når jeg har bestått alle eksamenene.

Så når jeg er sluttet å studere, så kan jeg tenke om framtiden.

Så, når jeg har sluttet å studere, kan jeg tenke på framtiden.

Men jeg tror verre skal dette bli med de andre vennene jeg er ikke så nær til.

Men jeg tror det kan bli verre med de andre vennene jeg ikke har så nært forhold til.

Jeg er veldig lyst å tjene Gud.

Jeg har veldig lyst til å tjene Gud.

Jeg mener at jeg er rett for å velge dette yrket.

Jeg mener at jeg gjorde rett i å velge dette yrket.

Det tror jeg er lettere å ta ut penger i framtida.

Jeg tror det blir lettere å ta ut penger i framtida.

## Bokomtale

*Av Rita Hvistendahl  
Universitetet i Oslo*

Anne Marit Vesteraas Danbolt og Lise Iversen Kulbrandstad: *Klasseromskulturer for språklæring. Didaktisk fornying i den flerkulturelle skolen*. Oplandske Forlag 2008. 182 sider.

I 2005 ble norskopplæringen for elever fra språklige minoriteter i Oslo endret ut fra en bekymring for elevenes gjennomgående svake prestasjoner på en rekke lokale og nasjonale kartleggingsprøver i lesing. Alle elever skulle undervises etter en felles læreplan i norsk, og ”forsterket tilpasset opplæring i norsk” erstattet faget norsk som andrespråk. Samtidig satset Utdanningsetaten i Oslo på etterutdanning av flere hundre lærere, og evalueringen av endringen kom i form av et aksjonsforskningsprosjekt som Høgskolen i Hedmark gjennomførte. Begge forfatterne deltok som aksjonsforskere i prosjektet.

*Klasseromskulturer for språklæring. Didaktisk fornying i den flerkulturelle skolen* er en bearbejdet versjon av rapporten fra prosjektet. Boka har ti kapitler. Innledningsvis gjøres det rede for bakgrunnen for prosjektet og forskningstilnærmingen, før de deltakende skolene presenteres, og de gjennomførte delprosjektene beskrives og analyseres. Til slutt oppsummeres lærernes refleksjoner. Det siste kapitlet ”Å bygge klasseromskulturer for læring” utdyper på et forskningsbasert grunnlag utvikling av elevenes lesekompetanse, organisering av norskopplæringen og arbeidsmåter i norskfaget, før kapitlet munner ut i fem råd for utvikling av norskopplæringen, spesielt arbeidet med lesekompetanse.

Aksjonsforskningsprosjektet hadde et dobbelt formål, det skulle både utvikle ny kunnskap og videreutvikle klasseromspraksis. Lærerne som deltok,

var med i et nettverk av skoler og fulgte en kursrekke ved høgs skolen i tillegg til å samarbeide med forskere om norskopplæringen. I boka beskrives og analyseres fire prosjekter på to skoler. På "Hvitveisen" brukes lekespråk for å utvikle språklig bevissthet, og språkverkstedet organiseres på nytt med vekt på innholdet i lesing og skriving, begge prosjekt på 2. trinn. På "Blåveisen" står systematisk arbeid med ordforråd og avissjangrer på 4.-5. trinn og arbeid med H. C. Andersens eventyr på 6.-7. trinn i fokus. Alle prosjektene illustrerer ulike måter å styrke elevenes språk- og lesekompetanse på.

På Hvitveisen skole observerte forskeren at flere av elevene fra språklige minoriteter strevde med å lage rimord på norsk. For å stimulere elevenes språklige bevissthet satte hun sammen med lærerne i gang prosjektet "Lek med isbjørnspråket", et forsøk med lekespråk med fokus på lydstrukturer og ordlagning knyttet til arbeidet med dyrelivet i polområdene. Sammen med lærerne fant elevene opp et eget språk for isbjørnene. Utdragene fra samtaler i klassen viser at elevene fra språklige minoriteter kom med relativt få forslag til "isbjørnspråket", og at lærerne i liten grad brukte mulighetene til språksammenlikning. I stedet styrte de elevenes oppmerksomhet mot den grammatiske strukturen i lekespråket de lagde, og norsk grammatikk dannet malen. Lærerne fikk se utskrifter av samtaler med elevene og reflekterte over egen undervisning i etterkant sammen med forskeren. Å gjennomføre undervisning med språk som tema viste seg å være mer krevende enn både lærerne og forskeren hadde forutsett, for verken elever eller lærere var vant til å snakke om språk. I diskusjonen av prosjektet påpeker forfatterne at undervisning som dette nettopp stiller store krav til lærernes språkkompetanse.

Observasjoner i språkverkstedet viste at elever med et svakt norskspråklig utgangspunkt hadde lite utbytte av arbeidet på lese- og datastasjonene. Når elevene ble trent i å samtale om det de hadde lest etter at lærerne hadde modellert for dem hvordan det skulle gjøres, økte både motivasjonen for lesingen og det språklige utbyttet, og på datastasjonen skrev elevene om bøkene og lærte seg å bruke det digitale verktøyet. Gjennom dette eksempelet viser forfatterne hvordan lærere kan legge til rette for språklig samhandling mellom elevene, og hvordan lesing, samtale og skriving kan knyttes sammen. Forfatterne peker også på hvor viktig samtalen med medelever innenfor en strukturert ramme er, særlig for elever som ellers er tilbakeholdne med å ta ordet på andrespråket.

I arbeidet med aviser på 4.-5. trinn på Blåveisen skole var det lagt mindre vekt på elevenes muntlige samhandling. Her dominerte den lærer-

styrte gruppesamtalen og det systematiske arbeidet med ordlæring, og gjennom samtalen med læreren ble elevene trent i å snakke om selve læringsprosessen. Forfatterne beskriver grundig og detaljert et sekstimers undervisningsopplegg. De gjengir og kommenterer flere utdrag fra gruppesamtalene slik at lærerens strategier trer tydelig fram. Selv om det ble arbeidet strukturert og flittig med å skrive avisartikler og lære ord som overskrift, journalist, redaksjon og pressefrihet, satt ikke alt like godt hos alle elevene da arbeidet ble avsluttet. Flere av dem hadde et stort behov for fortsatt systematisk muntlig og skriftlig arbeid med språket, understreker forfatterne.

Eventyrprosjektet på 6.-7. trinn blir noe mer summarisk beskrevet. Her er det elevenes framføringer av H. C. Andersens eventyr som viser deres utbytte av arbeidet, og flere får demonstrert talenter som verken lærere eller medelever visste at de hadde. På Blåveisen skole skjer den tilpassete norskopplæringen i en baseorganisering, og leseren får innsyn i lærernes nyanserte og reflekterte vurdering av de pedagogiske fordelene og ulempene med denne måten å organisere opplæringen på. Det mest positive er kanskje den ene lærerens refleksjon: Når alle elevene deles inn i forskjellige grupper, opplever ingen at de blir "tatt ut" av undervisningsfellesskapet.

Forfatterne gir fylldige beskrivelser av arbeidet i klasserommene. Som forskere kan de bygge framstillingen på et rikt materiale som består av både løpende notater og logger, elevarbeider, lyd- og videoopptak og fotografier fra klasserommene. Dette materialet kommer også til nytte i selve framstillingen, som fortsatt bærer noe preg av å være en bearbeidet rapport. Utdrag av klasseromssamtaler, intervjuer med lærere, elevenes tekster og tegninger og bilder fra klasserommene bryter opp en ellers tett tekst, lar flere stemmer komme til orde og legger til rette for at leseren kan trekke sine egne konklusjoner og etterprøve forfatternes analyser.

Læreren som har vært med i aksjonsforskningsprosjektet, uttaler seg gjennomgående positivt om å ha forskere til stede i undervisningen, ikke bare som "fluer på veggen", men som "klegger på kroppen". De har deltatt i faglig kvalifiserte samtaler om egen undervisning med nye perspektiv og nye ideer som utbytte, og de har delt erfaringene sine med andre lærere på nettverksmøter. De har blitt stimulert til faglig refleksjon, og de har i større grad enn før rettet oppmerksomheten sin mot de særskilte opplæringsbehovene som elever fra språklige minoriteter har.

Framstillingen i boka er faglig velfundert. Denne leserens erfaring er at utbyttet av de enkelte kapitlene øker ved flere gangers gjennomlesing. Boka

er innholdsrik, tett på faglig og pedagogisk informasjon og gir innsikt i læringsprosesser som sjelden gjøres synlige for andre enn de som deltar i dem. Observasjonene fra undervisningen kommenteres og diskuteres i lys av annen relevant nasjonal og internasjonal forskning. Slik blir de noe mer enn interessante glimt fra praksis. Det er den forskningsbaserte framstillingen som gjør dette til en bok studenter, lærere og andre forskere kan lære noe av. Fagpolitiske beslutningstakere kan ta forfatterens råd til etterretning. Det trengs utvikling av kvaliteten på leseprøvene, gjennomtenkning av bruken av nivåbaserte grupper, utvidet samhandling mellom elevene i språklæringen og inkludering av andrespråkstemaer og flerkulturelle perspektiver i arbeidet med felles læreplan i norsk for at elever fra språklige minoriteter skal få bedre utbytte av opplæringen.

## Bidragstere

**Jon Erik Hagen**, førsteamanuensis, Institutt for lingvistiske, litterære og estetiske studier, Universitetet i Bergen.

Kontaktadresse: [jon.hagen@lle.uib.no](mailto:jon.hagen@lle.uib.no)

**Liv Andlem Harnæs**, universitetslektor, Institutt for lingvistiske og nordiske studier, Universitetet i Oslo.

Kontaktadresse: [l.a.harnas@iln.uio.no](mailto:l.a.harnas@iln.uio.no)

**Turid Henriksen**, førsteamanuensis emerita, Institutt for litteratur, områdestudier og europeiske språk, Universitetet i Oslo.

Kontaktadresse: [t.l.henriksen@ilos.uio.no](mailto:t.l.henriksen@ilos.uio.no)

**Rita Hvistendahl**, førsteamanuensis, Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.

Kontaktadresse: [r.e.hvistendahl@ils.uio.no](mailto:r.e.hvistendahl@ils.uio.no)

**Hilde Johansen**, stipendiat, Institutt for lingvistiske, litterære og estetiske studier, Universitetet i Bergen.

Kontaktadresse: [Hilde.Johansen@lle.uib.no](mailto:Hilde.Johansen@lle.uib.no)

**Kari Tenfjord**, professor, Institutt for lingvistiske, litterære og estetiske studier, Universitetet i Bergen.

Kontaktadresse: [kari.tenfjord@lle.uib.no](mailto:kari.tenfjord@lle.uib.no)

## Abstracts and keywords in English

Turid Henriksen  
University of Oslo

### **Reflections on textbooks and textbook research**

#### *Abstract*

The purpose of this article is to describe textbooks and textbook research within a framework of text and communication theory. In the first part of the article, textbooks are given a tentative definition and seen in relation to pedagogical texts and “didacticity”. Then follows a model of communication in which text production is seen as a result of a double process, taking into consideration both society, subject field and the communication partners. What is considered specific for textbooks is described within this model.

In the second part of the article, textbook research related to the different parameters of the communication model is described. The examples shown are almost exclusively from research on textbooks in foreign language learning. Textbooks in second language learning are treated briefly only at the end, but the examples from foreign language textbook research are believed to be relevant also for research on second language textbooks.

*Keywords: textbooks, educational texts, didacticity, foreign language, research design.*

[t.l.henriksen@ilos.uio.no](mailto:t.l.henriksen@ilos.uio.no)

Liv Andlem Harnæs  
University of Oslo

**Vowels in the language of L2 users: an acoustical and perceptual study  
Vowels in the spoken Norwegian of three foreign physicians with Arabic  
L1**

*Abstract*

The article presents findings in an acoustical and perceptual study of vowel production in the spoken Norwegian of three foreign physicians with Arabic (Iraqi) L1. The basis of the study is a Norwegian norm for pronunciation (target language norm) which is established through three informants with Norwegian L1.

Contrastive hypotheses have been dominant in discussions about acquisition of pronunciation of L2s in adults, and in the present study, I will include contrastive interpretational models, but individual variation may be large and several interpretations may be relevant, e.g., influence from spelling. An important experience from the work with vowel analysis is that acoustical analysis of vowel quality is beneficial because we can quantify what we hear, which is important scientifically. A finding that indicates a direction for further investigation of the material is the fact that seeking explanations of vowel deviation on the phonemic level is not sufficient. The pronunciation of the vowel /ʉ/ in the language of the three Iraqi informants was particularly close to the target language norm in certain words (or syllables) such as *plutselig* and *slutt*, and one asks oneself what it is about these words (or syllables) that generates a pronunciation that accords with the target language norm of /ʉ/ here.

*Keywords: pronunciation, vowels, acoustical analysis, Arabic.*

[l.a.harnas@iln.uio.no](mailto:l.a.harnas@iln.uio.no)



Kari Tenfjord, Jon Erik Hagen, Hilde Johansen  
University of Bergen

### **Norwegian Learner Corpus (ASK) design and methodological prerequisites**

#### *Abstract*

The article presents Norwegian Learner Corpus (ASK) - a new corpus of second language data developed at the University of Bergen. The corpus consists of systematically annotated written language test essays originally produced by candidates sitting for official tests of Norwegian as a second language at two distinct proficiency levels. In addition to purely linguistic annotation of their recorded written performance, the corpus also provides personal data pertaining to the candidates themselves. The main focus of the article is to discuss whether the practice of error recording and error coding is in itself theoretically misguided by virtue of the so-called "comparative fallacy" argument adduced by Bley-Vroman in his article "The Comparative Fallacy" (Bley-Vroman 1983), that description of learner language as systematic deviations from target language structure and usage fails to appreciate, and even notice, the independence and autonomy of the interlanguage itself.

The error tagging of the ASK corpus is a theoretically neutral description of significant empirical features that any viable theory should be able to explain, and is defended in the context of a broader discussion of relations between theory and data in science generally and in SLA specifically, and our overall conclusion is that error recording and error coding is *not* methodologically misguided since error analysis is not in itself a *theory* of SLA but rather a *research method*, that can, in principle, service any theory at all. We contend that the so-called "comparative fallacy" charge often levelled against classical error analysis, or any preoccupation with errors at all, stems from this failure to distinguish properly between these two notions within SLA.

*Keywords: learner corpus, comparative fallacy, error, error tagging.*

[kari.tenfjord@lle.uib.no](mailto:kari.tenfjord@lle.uib.no)

[jon.hagen@lle.uib.no](mailto:jon.hagen@lle.uib.no)

[hilde.johansen@lle.uib.no](mailto:hilde.johansen@lle.uib.no)

## Til bidragsyterne

Tidsskriftet *NOA - norsk som andrespråk* publiserer originalartikler innenfor fagfeltet norsk som andre- og fremmedspråk tolket i vid forstand.

Tidsskriftet har referee-ordning. Artikler blir anonymisert før de blir vurdert av minst to fagfeller. Innsendte artikler må ikke være publisert andre steder eller være sendt til andre vitenskapelige tidsskrifter/forlag mens de er til vurdering i NOA. Det publiseres artikler på svensk og dansk i tillegg til norsk. Bare unntaksvis vil det trykkes artikler på engelsk.

Manuskript sendes redaksjonen som vedlegg på e-post.

Krav til manuskriptet:

Manuskriptet sendes inn i Word, skrives i Times New Roman med halvannen linjeavstand og 12 punkt skrift.

Lengden bør ikke være mer enn 25 sider medregnet tabeller, illustrasjoner, figurer, noter og referanseliste. 25 sider tilsvarer ca. 56 000 tegn inkludert mellomrom.

Norsk sammendrag på maksimum 15 linjer og 4-6 nøkkelord plasseres mellom tittelen og innledningen.

Engelsk sammendrag på maksimum 15 linjer, inkludert tittelen på engelsk, pluss 4-6 engelske nøkkelord sendes inn som separat vedlegg.

Tittelen skal være kort og beskrivende for artikkelen. Undertitler (avsnittsoverskrifter) skal ikke ha mer enn to nivåer. Underavsnitt skal ikke nummereres. Nivåene skilles fra hverandre ved at det høyeste nivået har 12

punkt halvfet. Det neste nivået skrives 12 punkt kursivert.

Tittelarket skal omfatte artikkelens tittel, eventuell undertittel, forfatterens navn, institusjonstilknytning og e-postadresse.

Sitat i teksten som overstiger to linjer, markeres med innrykk, uten bruk av anførselstegn. Det skal være én blank linje foran og én blank linje etter sitatet. Referansen plasseres etter sitatet. Kortere sitat skrives inn i den løpende teksten og settes i anførselstegn.

Noter begrenses og plasseres samlet etter artikkelen. Notene markeres med vanlige tall.

Referansehenvisninger i teksten skal være med forfatternavn, publiseringsår og eventuelt med sidetall, slik: ... også der kommentarene til det som blir sagt er kritiske (Foucault 1999:17). Eller: I Lindblad og Salströms bok (2001:20) omtales denne observasjonen som ...

Litteratur skal stå samlet og i alfabetisk rekkefølge til slutt i artikkelen og settes opp slik:

For bøker:

Baker, Colin 2001. *Foundations of Bilingual Education and Bilingualism*. Clevedon: Multilingual Matters.

For kapittel i bok:

Hvenekilde, Anne og Elizabeth Lanza 2001. Applying social network analysis to the Filipino community in Oslo. I A. Hvenekilde og J. Nortier (red.). *Meetings at the Crossroads. Studies of Multilingualism and Multiculturalism*. Oslo: Novus forlag, 296 – 313.

For tidsskriftsartikler:

Laursen, Pia Helle 2006. Andetsprosperspektiver på tematiske mønstre i naturfagsundervisningen. *NORDAND Nordisk tidsskrift for andrespråksforskning*, 2006, årg. 1 nr. 1, 51 – 72.

Manuskript til bokanmeldelser:

Lengden på anmeldelsen bør normalt ikke overstige 4 sider. Tittelen på bok-

anmeldelsen skal inneholde: Forfatternavn, årstall for utgivelse. Boktittel med undertittel (i kursiv). Forlag og utgivelsessted. Sidetall. Anmelderens navn.

Illustrasjoner: Grafiske framstillinger, tabeller o. l. skal sendes sammen med manuskriptet som regnearkfil eller som bilde (min. 240 dpi). De kan eventuelt sendes som reproduksjonsklare originaler (klare til gjengivelse i tidsskriftet).