

Ja takk, begge deler

*En undersøkelse av variasjonsmønstre
i anvendelse av invertert ordstilling
i norske mellomspråkstekster*

Hilde Johansen

Mastergradsoppgave
Nordisk institutt
Universitetet i Bergen
2007

Innholdsliste

Kapittel 1: INNLEDNING.....	5
1.1 Bakgrunn for oppgaven	5
1.2 Problemstilling	5
1.3 Metodiske opplysninger	6
1.3.1 <i>Datamaterialet</i>	6
1.4 Oversikt over avhandlingen	7
Kapittel 2: TEORI	8
2.1 Mellomspråk og andrespråksforskning	8
2.1.1 <i>Studiet av mellomspråk – en historisk gjennomgang</i>	9
2.1.2 <i>Studiet av mellomspråk i dag</i>	10
2.2 Variasjon i mellomspråk	11
2.2.1 <i>System og variasjon</i>	11
2.2.2 <i>Variasjonsfenomenet</i>	12
2.2.3 <i>Systematisk variasjon og årsaker</i>	12
2.2.4 <i>Systematisk variasjon - et paradoks?</i>	15
2.2.5 <i>Hvorfor studere variasjon?</i>	16
2.3 Inversjon i mellomspråk	18
2.3.1 <i>Inversjon og typiske innlæreravvik</i>	18
2.3.2 <i>Inversjon og utvikling</i>	19
2.3.3 <i>Inversjon og morsmål</i>	21
2.3.4 <i>Inversjon og variasjon</i>	22
2.3.5 <i>Å lære seg inversjon</i>	24
2.3.6 <i>Inversjon og invertering - prosess eller struktur?</i>	24
Kapittel 3: VARIASJON OG INVERSJON.....	26
3.1 Fins det variasjon og hvor vanlig er det?	26
3.1.1 <i>Definisjon av variasjon</i>	26
3.1.2 <i>Hva regnes som inversjonskontekster?</i>	27
3.1.3 <i>Inversjonsmønster i tekstene</i>	28
3.1.4 <i>Utvalg av tekster til undersøkelsen</i>	29
3.2 Inversjonsutfordringen	29
3.2.1 <i>Tre 'sannheter' om inversjon i mellomspråk</i>	30
3.2.2 <i>Hva det kan være, og hva det nok ikke er</i>	32
3.3 Variasjon og språklig prosessering	34
3.3.1 <i>Prosessering og lokalisering av variasjon</i>	34
3.3.2 <i>Kunnskap, kontroll og variasjon</i>	36
3.3.3 <i>To hypoteser om variasjon i anvendelse av invertert ordstilling</i>	37
3.4 Metode	39
3.1.2 <i>Datamateriale for variasjonsundersøkelsen</i>	39
3.2.3 <i>Individuelle variasjonsmønster</i>	40

Kapittel 4: IDENTIFISERINGSHYPOTESEN 41

4.1 Opparbeiding av kunnskap om inversjonens distribusjon	41
4.2 Foranstilt ledd som signal om invertert ordstilling	43
4.2.1 <i>Prosessering av invertert innputt</i>	44
4.2.2 <i>Identifiseringsgrunnlag</i>	46
4.2.3 <i>Syntaktisk uklarhet</i>	48
4.2.4 <i>Prediksjoner for materialet</i>	49
4.3 Operasjonalisering og gjennomføring	49
4.3.1 <i>Gjennomføring</i>	49
4.3.2 <i>Prediksjoner og profiler</i>	50
4.3.3 <i>Frekvenskategorier</i>	52
4.3.4 <i>Andre sjekkpunkt</i>	54
4.4 Resultater	56
4.4.1 <i>Resultat etter frekvensfordeling</i>	56
4.4.2 <i>Eksempler</i>	57
4.4.3 <i>Profiler etter å ha tatt tvetydighet i betraktning</i>	59
4.4.4 <i>Andre sjekkpunkter</i>	60
4.5 Konklusjon	62
4.6 Metodiske forbehold	63
4.6.1 <i>Identifiseringsfaktoren</i>	63
4.6.2 <i>Operasjonaliseringene</i>	64
4.6.3 <i>Den praktiske gjennomføringen</i>	64

Kapittel 5: KAPASITETSHYPOTESEN 66

5.1 Kontroll og kapasitetsbelastning	66
5.2 Operasjonalisering: Lingvistisk kontekst og kapasitetsbelastning	67
5.2.1 <i>Avgrensning av den lingvistiske kontekst</i>	68
5.2.2 <i>Overinverterte kontekster</i>	69
5.2.3 <i>Operasjonalisering av kapasitetsbelastning</i>	69
5.2.4 <i>Mål for kapasitetsbelastning</i>	70
5.2.5 <i>Operasjonalisering av hvert enkelt ledds relative kapasitetsbelastning</i>	70
5.3 Gjennomføring av undersøkelsen	75
5.3.1 <i>Tildeling av kategorier og ekstrarfaktorer</i>	75
5.3.2 <i>Sammenligning av kontekster</i>	76
5.3.3 <i>Kompleksitetsprofiler</i>	77
5.4 Resultater	77
5.4.1 <i>Oversikt over tekstenes profiler når kun kompleksitet tas i betraktning</i>	77
5.4.2 <i>Eksempler</i>	78
5.4.3 <i>Oversikt over tekstenes profil med både kompleksitet og ekstra faktorer</i>	83
5.4.4 <i>Konklusjon til undersøkelsen av kapasitetshypotesen</i>	84
5.6 Vurdering av metoder og operasjonalisering (og positive)	84
5.6.1 <i>Avgrensningen av den lingvistiske konteksten:</i>	84
5.6.2 <i>Operasjonalisering av kapasitetsbelastning</i>	85
5.6.3 <i>Undersøkelsesmetoden</i>	85

Kapittel 6: DRØFTING 87

6.1 Identifisering eller kapasitet	87
6.1.1 Oppsummering av hypotesene	87
6.1.2 Oppsummering av resultatene	88
6.1.3 De ulike tekstenes stilling til de ulike hypotesene	89
6.1.4 Identifisering eller kapasitet? En diskusjon om problemer med hypotesene	89
6.1.5 En oppsummering	92
6.3 Synkroni og diakroni	93
6.3.1 Metodiske problemer med dette synet på variasjon	94
6.4 Metodiske forbehold	96
6.4.1 Feilkilde: bruk av skriftlige tekster	96
6.4.2 Hypotesenes begrensninger	97
6.4.3 Comparative fallacy	98
6.5 Noen sluttbemerkninger om variasjon og koherens i tekst	99

Kapittel 7: UTLEDNING 101

LITTERATURKISTE..... 103

Kapittel 1: INNLEDNING

1.1 Bakgrunn for oppgaven

Den såkalte subjekt/verb-inversjonen¹ i norsk og enkelte andre nordeuropeiske språk, har rykte på seg for å være en av de vanskeligste strukturer å tilegne seg for andrespråksinnlærere. Det har også vært fremstilt som en gåte at man ikke kan forklare hvorfor dette er så vanskelig å få på plass. Karen Lund har påpekt at dette kan skyldes den manglende kommunikative funksjonen denne strukturen har i deklorative setninger, til forskjell fra i spørsmål, hvor en slik invertert ordstilling fungerer som signal om at det er et spørsmål (Lund 1997:161).

Min interesse for inversjon ble vekket i en grunnfagsoppgave i Norsk som andrespråk, hvor jeg utførte en feilanalyse av en mellomspråkstekst. Her fikk jeg øynene opp for fenomenet *variasjon*, og spørsmålet om hvorfor innlærere tilsynelatende mestrer og ikke mestrer en regel i én og samme tekst. Man kan for eksempel støte på setninger som *'i dag jeg kjøpte epler på butikken'*, og *'på butikken var det mange folk'* i én og samme tekst. Jeg syntes dette var et interessant fenomen, og en vinkling til læring av språk som kunne vise mye. Jon Erik Hagen har gjort en pilotundersøkelse av denne typen variabel mestring av inversjon (Hagen 1992), og denne har vært et viktig springbrett for min egen undersøkelse.

1.2 Problemstilling

Jeg ønsker i denne oppgaven å undersøke denne typen variabel anvendelsen av inversjon, med det hovedmål å se om jeg kan finne frem til de egenskapene som *styrer* om vekslingen mellom det vi i problematiske ordelag kan kalle 'korrekt' og 'avvikende' form av inversjonsstrukturen. Hvorfor bruker innlæreren invertert ordstilling i noen av de deklorative setningene som krever det, men ikke i andre?

Undersøkelsen min er deduktiv, eller hypotesetestende, og den overordna hypotesen min er at det er forhold ved den umiddelbare språklige konteksten, og prosesseringen av den, som er hovedpåvirkningskilden til variasjonen. Rent metodisk skiller jeg mellom to hypoteser, *identifiseringshypotesen*, som hevder at variasjonen er et resultat av en ufullstendig kunnskap

¹ Når en indikativsetning innledes av et ikke-subjekt, skal subjektet stå *bak* den finitte verbet, i midtfeltet. (Hagen 1998: 272)

om *når* det skal være invertert ordstilling; og *kapasitetshypotesen*, som har som utgangspunkt at variasjonen er styrt av en varierende belastning av innlærerens oppmerksomhetsressurser. De ulike setningene belaster kapasiteten i ulik grad i ytringsøyeblikket, og kan ha positiv eller negativ innvirkning på hvorvidt innlærerne klarer å benytte kunnskapen de har.

1.3 Metodiske opplysninger

Undersøkelsen tar utgangspunkt i *skriftlige* tekster skrevet av voksne innlærere av norsk, og er en *flerkasusstudie*, hvor jeg har vektlagt å se på variasjonsmønstre på individnivå, ikke kollektivt, slik for eksempel Anne Brautaset har gjort (Brautaset 1996).

1.3.1 Datamaterialet

Datamaterialet mitt er hentet fra Norsk Andrespråskorpus (ASK)², som er et elektronisk korpus av tekster hentet fra *Språkprøven i norsk for voksne innvandrere* og *Test i norsk – Høyere nivå*. I tillegg er det et kontrollkorpus av 200 tekster skrevet av innfødte nordmenn. Selve ASK sto ikke klart før ved slutten av 2006, men siden desember 2005 har en testutgave av korpuset vært tilgjengelig, *ASK-test*, og det er her jeg har hentet mine tekster. ASK-test består av hundre tekster, skrevet av innlærere med 10 forskjellige morsmål³.

I ASK har alle originaltekstene blitt transkribert, annotert med personlige variabler og en *feilkoding*, hvor alle strukturer som avviker fra innfødt norsk er kategorisert med en feilkode, som sier noe om hva slags type avvik det er snakk om, samt forslag til *korreksjon*. I tillegg er all tekst behandlet av en automatisk tagger, *Oslo-Bergen-taggeren*, som har identifisert alle ord med ordklasse og morfologiske og syntaktiske funksjoner.

Denne tilretteleggingen av tekstene har gjort det mulig å utføre ulike typer søk i materialet, både på ord, språklige kategorier og avvik. Jeg har dessverre ikke hatt mulighet til å bruke alle funksjonene som ASK tilbyr, siden undersøkelsene mine startet før selve søkegrensesnittet til ASK var ferdigutviklet og fordi undersøkelsen min på noen områder har krevd en mer manuell tilnærming til materialet. Jeg har hovedsaklig brukt ASK til å finne tekster med den typen variasjon jeg ønsker å undersøke, mens resten av analysen er gjort manuelt. Jeg er i ettertid, nå som ASK er ferdig og jeg har lært mer om mulighetene for søk, blitt klar over at det hadde vært et ypperlig verktøy som kunne ha lettet meg for mye av det jeg har gjort

² <http://spraktek.aksis.uib.no/projects/ask>

³ Albansk, engelsk, nederlandsk, polsk, russisk, bosnisk-serbisk-kroatisk, spansk, somali, tysk og vietnamesisk

manuelt. Jeg har i alt hentet ut 19 tekster fra dette korpuset, de tekstene som har *variasjon* slik jeg definerer dette i neste kapittel.

Tekstene i ASK er langsvarsoppgaver hentet fra *Språkprøven i norsk for voksne innvandrere*.⁴ Språkprøven måler ferdigheter på det såkalte *mellomnivået*, B1, og blir som regel tatt av innlærere som har avsluttet grunnopplæringen i norsk, en svært heterogen gruppe med variert yrkes- og utdanningsbakgrunn, og med ulike formål med å ta testen (Saarik 1004: 13) Langsvarsoppgavene skal være på rundt 200 ord, og innlærerne kan velge blant ulike tema. I mitt materiale har emnene vært blant annet bomiljø, konkurranse, barneoppdragelse, alkohol og alkoholvaner, yrker, framtida m.m. Denne typen testbesvarelser har vært brukt i en rekke tidligere hovedfagsoppgaver om norsk som andrespråk.

1.4 Oversikt over avhandlingen

Denne avhandlingen har sju kapitler inkludert innledningen. I kapittel 2 ser jeg på den teoretiske bakgrunnen for oppgaven, om mellomspråk, variasjon og andre studier av tilegnelsen av inversjon. I kapittel 3 ser jeg nærmere på hvor utbredt variasjon er i tekstene i ASK-test. På grunnlag av generelle teorier om mellomspråksvariasjon og det vi vet fra før om tilegnelse av inversjon, underbygger jeg så de to hypotesene som er utgangspunkt for undersøkelsene mine, *identifiseringshypotesen* og *kapasitetshypotesen*. Undersøkelsene av disse to presenteres i henholdsvis kapittel 4 og 5. I kapittel 6 drøfter jeg resultatene og sammenligner de to hypotesene, i tillegg til å drøfte problemstillinger rundt hva variasjon kan si oss om mellomspråksutvikling generelt. I kapittel 7 kommer en kort oppsummering, og konklusjon.

⁴ Denne testen ble i juni 2005 erstattet av *Norskprøve 3 for voksne innvandrere* (nivå B1) (<http://www.alte.org/members/norwegian/norwegian/intermediate.php>)

Kapittel 2: TEORI

I dette kapitlet vil jeg gjøre rede for fremveksten av *innlærerspråk* som studieobjekt, og ulike aspekter ved studiet av andrespråklæring. Jeg vil også gå inn på andre mer spesifikke begreper og fenomener som er viktige i oppgaven.

2.1 Mellomspråk og andrespråksforskning

I de lingvistiske tilnærmingene til andrespråklæring står innlærernes variant av andrespråket sentralt. Når man skal lære seg et nytt språk, står man overfor en prosess hvor man skal forstå, bygge opp og anvende et fremmed språkssystem. I begynnelsen er gjerne de språklige ressursene få, og systemet er relativt begrenset og enkelt, men etter hvert som større deler av andrespråket blir tilgjengelig for innlæreren, og en lærer mer, vokser det frem et komplekst system som i større grad kan brukes til kommunikasjon. Det er disse innlærervariantene av et språk, fra det spede anslag til et mer utbygd system, vi gjerne kaller for *mellomspråk*. Dette er en oversettelse av Larry Selinkers term *interlanguage*, som for første gang så dagens lys i en artikkel med samme navn i 1972. Mellomspråk representerte et mellomstadium mellom morsmålet og det språket som skulle læres, *målspråket*, og for Selinker ble mellomspråkene formet av et sett med kognitive strategier;

"... a single system composed of rules that have been developed via different cognitive strategies- for example, transfer, overgeneralization, simplification, and the correct understanding of the target language. At any given time, the interlanguage grammar is some combination of these types of rules." (MacLaughlin 1987: 63)

I ettertid har andre teorier om hva som former mellomspråket det kommet til, og det er en utbredt oppfatning at mellomspråket er et naturlig språk, med et eget system uavhengig av målspråket. Når man i dag bruker termen mellomspråk, er det ofte uavhengig av hva man måtte mene former mellomspråket, man bruker det teorinøytralt, som benevnelse på det språket innlærer har på et bestemt tidspunkt i innlæringa, eller som overlappende systemer over tid (Bergreen & Tenfjord 1999: 18). Enkelte ønsker å unngå denne termen, siden det har tilhørighet i en spesiell teori, Selinkers mellomspråksteori, og foretrekker derfor heller å bruke ordet '*innlærerspråk*'. Teorinøytral bruk av termen mellomspråk er allikevel så utbredt at man ikke lenger nødvendigvis legger noen teoretiske forutsetninger i det, noe heller ikke jeg gjør i denne avhandlingen, hvor jeg kommer til å bruke både '*innlærerspråk*' og '*mellomspråk*'.

Når man snakker om mellomspråk er det gjerne med to forskjellige betydninger; både om den faktiske, observerbare *språkbruken* (mellomspråk som produkt), og om prosessene og *mellomspråkssystemet* som ligger bak språkbruken (mellomspråk som prosess) (Sharwood Smith 1994: 35).

Det er en utbredt antakelse i dag at innlærerspråk er *systematiske* og *dynamiske*. De utvikler seg over tid, fra enkle til mer komplekse systemer, og de utvikler seg på en mer eller mindre *systematisk* måte. En stor del av den lingvistiske tilnærminga til andrespråksforskning søker å avdekke denne utviklingen og systematikken i mellomspråk. Både for å skjønne hvordan språklæring foregår, for å skjønne mer om språk generelt, og for å kunne legge opp til bedre språkundervisning.

2.1.1 Studiet av mellomspråk – en historisk gjennomgang

I studiet av mellomspråk har man gjennom årene hatt ulike vinklinger. Ulike paradigmer og tilnæringsmetoder har regjert til ulike tider, og vokst frem som kritikker og videreføringer av hverandre, til vi i dag har kommet det stadiet der vi faktisk kan se nytten i å skulle studere mellomspråket fra ulike, men komplementære vinkler, slik at vi kan få et komplett bilde av den prosessen det er å lære et annet språk enn morsmålet.

Den første teoretiske aktiviteten rundt andrespråklæring vokste frem i 1950- og 60-årene med utgangspunkt i pedagogiske behov, og var altså som Mitchell og Myles formulerer det ”...*still very much adjunct to the practical business of language teaching*”. (Mitchell and Myles 1998: 23). Språklæringspedagogikken på denne tiden var sterkt påvirket av datidens ledende lingvistiske og læringsteoretiske paradigmer. Med et strukturalistisk syn på språk som et avgrenset sett med strukturer som fungerte som modeller for setninger – og et behavioristisk syn på læring som innarbeiding av vaner ved hjelp av repetisjon - var språklæringa i praksis et spørsmål om imitasjon og repetisjon (Bergreen og Tenfjord 1999: 48). I tillegg fokuserte man undervisninga på de strukturene som var regna som vanskelige å lære. Hva som var vanskelig, kom man frem til ved hjelp av *kontrastiv analyse*, som da var den første teoretiske tilnærminga til andrespråklæring. Ideen var at de strukturene som morsmålet og målspråket hadde til felles, var lette å lære, mens de som var nye i målspråket, var vanskelige. Andrespråksstudier i disse dager hadde derfor fokus på likheter og ulikheter mellom morsmål og målspråk, med det mål å kunne forutsi hva som ville være lett og vanskelig, og kunne

legge opp undervisningen deretter. Noe begrep om innlærernes språkproduksjon som et selvstendig språkssystem fantes ikke i denne perioden. Men mot slutten av 1960-tallet beveget man seg bort fra strukturalismen og behaviorismen, blant annet med fremveksten av den generative lingvistikken. Språk gikk fra å være vaner til å være et sett med regler. I tillegg opplevde man at kontrastiv analyse ikke fungerte tilfredsstillende som metode for å forutsi vansker, noe som førte til at man i enda større grad begynte å fokusere på det språket innlærerne faktisk produserte. I 1967 lanserte Corder sin *Error Analysis*, feilanalysen, som var et metodisk verktøy for å klassifisere de feilene innlærerne gjorde. Corder påpekte at de færreste feilene kunne forklares med morsmålspåvirkning, og det store spørsmålet ble hva som egentlig var årsaken til feilene. Corder var også den første til å benevne innlærerspråket med termen *approximate systems*. I kjølvannet av feilanalysen og fokuset på innlærerspråket, kom så Selinker med sitt *interlanguage*, og ideen om mellomspråket som et eget system, uavhengig av både morsmålet og målspråket.

2.1.2 Studiet av mellomspråk i dag

Man har altså i løpet av disse årene beveget seg fra å fokusere på likheter og ulikheter mellom morsmål og målspråk - og ideen om at man kunne bruke disse til å forutsi vansker - til en tilnærming til andrespråklæring som fokuserer på det språket innlærerne faktisk produserer, fra en deduktiv til en mer induktiv, utforskende tilnærming. Samtidig har man beveget seg fra å være opptatt av mellomspråket som *produkt*, til å være mer opptatt av det som *prosess*, og å ville si noe om det systemet og de prosessene som ligger bak språkbruksproduktene, og påvirker utviklingen over tid. En tredje viktig utvikling i studiet er at man har beveget seg fra et ønske om å forstå feilene og deres årsaker, og til å ha tilnærminger som ikke nødvendigvis fokuserer på feil, men heller på måten innlærerne utnytter det språklige materialet de har til rådighet for å kommunisere effektivt. Språklæring er ikke lenger nødvendigvis tilegnelse av strukturer, men uttrykking av funksjoner.

Mitchell & Myles setter midten av 1980-årene som tidspunktet da andrespråksfeltet oppnådde status som et selvstendig forskningsfelt, løsrevet fra språkpedagogikken (Mitchell & Myles 1998: 40). Andrespråksforskningen har vokst til å bli et multifasettert fagområde, et felt hvor ulike fag, ulike teoretiske retninger har smeltet sammen for å belyse ulike aspekter av språklæring. Studiet av selve mellomspråket er i dag bare ett av flere tilnærminger innen feltet andrespråklæring, hvor lingvistiske og læringsteoretiske teorier er de mest sentrale.

Andrespråksforskningen fokuserer både på det språklige aspektet av læringen, studiet av

mellomspråket, og på andre, mindre språkrelaterte, forhold som man mener påvirker innlæringa.

I dag har man spesielt tre hovedområder i andrespråksforskningen; 1) studiet av system og variasjon i mellomspråk, 2) studiet av tilegnelse og utvikling, og 3) morsmålets rolle i innlæringa (McLaughlin 1987). Denne undersøkelsen vil først og fremst fokusere på system/variasjonsaspektet, og ikke like mye på morsmål og utvikling over tid, så jeg vil gå nærmere inn på denne tilnærminga til mellomspråk i neste del.

2.2 Variasjon i mellomspråk

2.2.1 System og variasjon

Helt siden mellomspråksteorien ble lansert med ideen om at språket innlærerne produserte hadde sitt eget system uavhengig av morsmål og målspråk, har en av de mest grunnleggende antakelser om mellomspråk vært at de er *systematiske*. Berggreen og Tenfjord skriver følgende: ”... *mellomspråk kan ses og beskrives som egne språkssystem med en intern logikk der elementer står i bestemt forhold til hverandre, forutsetter hverandre og danner mønstre: mellomspråk er med andre ord regelbaserte*” (Berggreen & Tenfjord 1999: 19). Elaine Tarone stiller imidlertid et interessant spørsmål i forordet til boka ’Variation in Interlanguage’: ”*But what does it mean to say that learner language is systematic when there is massive evidence that the organization of learners’ language performance varies from one situation to another.*” (Tarone 1998: forordet). Mens systematikken ofte har vært en slags pre-empirisk antakelse, viser empirien at språket til innlærerne faktisk er varierende og ustabilt, ikke bare over tid og i ulike situasjoner, men også og i én og samme kommunikative situasjon.

Dette fenomenet kalles gjerne *variasjon*, og er egentlig et sosiolingvistisk begrep, brukt om måten morsmålsbrukere veksler mellom to ulike former som betyr det samme. Innen andrespråksforskningen brukes begrepet litt annerledes, og i dette kapitlet vil jeg gå nærmere inn på hva variasjon er, hvordan det kommer til syne i mellomspråk, hvilke årsaker det har, og hvorfor det er et relevant og interessant fenomen å studere.

2.2.2 Variasjonsfenomenet

Om Tarone spør seg hva det egentlig vil si at mellomspråk er systematisk, følger jeg opp med å spørre; hva vil det egentlig si at det varierer? Roger W. Anderson definerer variasjon kort og greit som ”*The same person using different linguistic forms or constructions for what appears to be the same intended meaning.*” (Anderson 1989: 46). Det er vanlig å skille mellom to typer variasjon i mellomspråk; variasjon mellom ulike innlærere (*inter-learner variation*), og variasjon hos én og samme innlærer (*intra-learner variation*) (Ellis 1989). Innen andrespråksforskningen er det som regel denne siste typen variasjon som har vært interessant i studiet av variasjon, at språket til en og samme innlærer varierer. Man deler videre inn i *horisontal* variasjon, på ett bestemt tidspunkt⁵, og *vertikal* variasjon, som er variasjon over tid, i pakt med at mellomspråket utvikler seg og blir mer komplekst⁶ (Ellis 1989:22). I tillegg kan man skille mellom horisontal variasjon i *ulike situasjoner*, og i *en og samme situasjon*, og det er denne typen variasjon jeg har planlagt å undersøke. At innlærerne i én og samme tekst eller samtale både anvender og ikke anvender invertert ordstilling der det i målspråket er obligatorisk.

2.2.3 Systematisk variasjon og årsaker

En utbredt antagelse om variasjonen er at den er *systematisk*, vekslingene mellom former er ikke tilfeldig, den er motivert av andre faktorer, og derfor forutsigbar. Anderson skriver: ”*All such surface variation can be reduced to invariance and a set of principles that govern distribution of forms and change (i.e. restructuring) over time.*” (Anderson 1989:46).

Variasjonen er derfor bare et overflatefenomen, og ved nærmere undersøkelser kan man finne en eller annen form for systematikk i vekslingen.

Det er vanlig å operere med tre hovedtyper av slike styrende prinsipper. Variasjonen kan være parallell, med variasjon i enten den *lingvistiske*, den *psykolingvistiske* og/eller den *sosiale/situasjonelle* konteksten (Bergreen & Tenfjord 1999:233). Dette innebærer at når disse faktorene forandres, vil det oppstå en tilsvarende variasjon i bruk av et gitt trekk. Variasjon etter *lingvistisk kontekst* innebærer at vekslingen mellom ulike former er påvirket av den nærmeste språklige omgivelsen. Variasjon i forhold til den *psykolingvistiske konteksten*, er

⁵ Synkron variasjon

⁶ Diakron variasjon

måten graden av press i kommunikasjonssituasjonen påvirker innlærerens veksling mellom former, som regel med tanke på oppmerksomhetsfokus og planleggingstid. Enkelte språklige uttrykk krever mer av innlærerens kapasitet enn andre, og noen typer situasjoner legger mer press på innlæreren, noe som kan føre til variasjon. Variasjon i forhold til den *sosiale/situasjonelle konteksten* er en mer ekstern påvirkning på variasjonen. Variasjonen kan da avspeile sosiale forhold eller stilistiske valg, eventuelt innlærerens særegne tildeling av form til ulike funksjoner. *Labovianske modeller*, hevder at innlæreren sitter inne med ulike registre av målpråket, som brukes i ulike situasjoner. Disse ulike registrene skiller seg fra hverandre i graden av oppmerksomhet mot form, og noen er for eksempel mer utsatt for morsmålpåvirkning enn andre (Tarone 1988: 39).

Denne inndelingen i årsaker er overført fra sosiolingvistikken, det finnes også andre taksonomier. Suzanne Romaine foretrekker å dele inn i *eksterne* og *interne* påvirkningsfaktorer, hvor de interne er lingvistiske, knyttet til markerthet og transfer, mens de eksterne er sosiale (Romaine 2003:411).

2.2.4 Årsakenes relevans for ulike typer variasjon

Det er klart at de forskjellige årsakene har ulik relevans som forklaring til de ulike typene variasjon. De sosiale teoriene egner seg best til å forklare - og er vel først og fremst ment til å forklare - variasjon fra situasjon til situasjon og over tid. De er mindre egnet som forklaring på variasjon i én og samme tekst/samtale, ettersom slike forhold da holdes mer eller mindre konstant⁷. Det samme gjelder til en viss grad de psykolingvistiske faktorene, som kanskje først og fremst vil forklare variasjon i ulike oppgavetyper, såkalt *task variation* (Tarone 1988:12), hvor ulike typer forhold rundt ytringsøyeblikket legger ulikt press på innlærerens oppmerksomhet og kapasitet, og dermed fører til varierende språklig utfall. *Den språklige konteksten*, derimot, virker mer egnet til å forklare variasjon i en og samme kommunikasjonssituasjon, da dette er en faktor som sjelden er konstant i løpet av en tekst/samtale. Tvert imot vil de språklige omgivelsene hele tiden endre seg, og det kan godt tenkes at dette fører til variasjon i bruk av et gitt trekk. Berggreen og Tenfjord skriver følgende: ”*Variasjon etter lingvistisk kontekst later til å være et grunnleggende trekk ved mellomspråk. Denne typen variasjon lar seg f.eks. ikke lett påvirke av oppgavetype eller datainnsamlingsmetode. Variasjonen fins dessuten på alle språklige nivåer...*” (Berggreen og

⁷ Med mindre det for eksempel er en samtale med flere samtalepartnere e.l.

Tenfjord 1999: 235). Ellis fremhever også den lingvistiske konteksten som mer sentral enn de andre: ”*Irrespective of whether there is a social or communicative motivation for language choice, users will perform variably according to the amount of effort required to process one variant of a structure as opposes to another*” (Ellis 1989: 33)

a) Mer om språklig kontekst

Berggreen og Tenfjord viser til to ulike opphav til den språklige kontekstens påvirkning. For det første *egenskaper ved innputt*, som frekvens og saliens, knyttet til i hva slags kontekster man lettest legger merke til og først lærer et trekk. Det andre opphavet er relatert til *produksjonsforholdene*, og nært knyttet til psykolingvistiske forhold. Noen forskere har slått sammen disse to til en *kapasitetsbetinget variasjon* (Berggreen og Tenfjord 1999: 235), som forklares slik:

”*Typisk for innlærere er at de har skort på oppmerksomhetsressurser og prosesseringskapasitet. Når informasjonen skal hentes fram og aktiviseres, belastes bearbeidingsmekanismene. Dersom informasjonsinnhentingen ikke kan foregå nokså automatisert, blir det avgjørende for prosesseringen at den språklige konteksten ellers er enkel og lite krevende.*” (Berggreen og Tenfjord 1999: 235)

Her møtes altså det språklig uttrykkets struktur og innlærerens språkbearbeidingskapasitet som årsak til variasjon. Det er vanlig å snakke om den språklige kontekstens *tyngde* eller *vekt*, og at én form brukes i såkalt lette kontekster, mens en annen brukes i de tunge.

b) Fri variasjon

I tillegg til systematisk variasjon, mener noen at det også finnes en form for ikke-systematisk, *fri variasjon*. Ellis mener dette forekommer på et tidlig stadium, idet innlærerne begynner å få teft om en målspråksregel (Ellis 1989: 23). Den frie variasjonen oppstår som resultat av at mellomspråkets mer naturlige, universelle regler kommer i konflikt med målspråkets regler, eller fordi reglene i målspråket selv er variable. Fri variasjon er imidlertid en omdiskutert term, og mange nøyer seg med å konkludere at det bare er resultat av for dårlig leting (Berggreen & Tenfjord: 257). Det virker allikevel på meg som om noen teoretikere har bitt seg for hardt fast i antakelsen om mellomspråkets selvstendige systematikk, at de kanskje ikke er villig til å innse at noe variasjon faktisk er tilfeldig. Det er metodisk nødvendig å avgrense

noen faktorer, slik at man ikke leter seg i hjel etter en enkelt styrende kraft. Så vidt jeg kan se definerer Ellis selv fri variasjon som den variasjonen man ikke har funnet andre årsaker til (Ellis 1989:26), det er en slags negativ definisjon, ikke en påstand om at noe ikke styrer variasjonen, annet enn innlærernes utprøving av hypoteser om målspråksreglene.

2.2.4 Systematisk variasjon - et paradoks?

De ulike typene variasjon kan altså føres tilbake til en eller flere styrende faktorer, og det på en systematisk måte. Dette kan i seg selv virke som et paradoks, for om det finnes et system, hvorfor kaller man det da *variasjon*?

Den mest åpenbare forklaringen er man ikke vet at det er et system før man har undersøkt nærmere og faktisk funnet et, noe som ikke alltid lar seg gjøre. Variasjon blir derfor et utgangspunkt, et *overflatefenomen*, som skal undersøkes og forklares, og spørsmålet blir da hvordan man identifiserer språkbruk som variabel framfor systematisk.

Tarone (1988:11) skriver at det har vært vanlig med to ulike typer definisjoner av variasjon; *eksterne* og *interne*⁸. Disse to er ulike med hensyn til hvordan man identifiserer variasjonen, altså hva som kvalifiserer til å bli kalt variabelt. Variasjon kan være *ekstern*, det vil si i forhold til målspråket, eller *intern*, i forhold til seg selv i en annen situasjon eller på et annet tidspunkt i utviklinga. Den eksterne tilnærminga kjennetegnes ved at man ser på om et trekk er anvendt eller ikke i såkalte *obligatoriske kontekster*, det er altså fokus på mestring eller ikke-mestring av et trekk i forhold til målspråksstandarden. Ved en intern tilnærming ser man bort fra målspråkets norm, og ser heller på hvordan innlæreren distribuerer et trekk på sin egen systematiske måte, i tråd med sitt eget idiosynkratiske system. Den eksterne tilnærmingen har blitt kritisert blant annet av Huebner og Bickerton (Huebner 1983), som mener at man kommer til kort i å finne systematikken i mellomspråk om man bare undersøker obligatoriske kontekster, blant annet fordi man da ikke fanger opp eventuelle *overgeneraliseringer*, tilfeller hvor trekket anvendes i andre kontekster enn de obligatoriske. Bickerton uttrykker det slik: "... (the criterion) pays no heed to occasions on which the feature in question is used in environments where its absence, rather than its presence, would be obligatory in the target language..." (Huebner 1983: xxiii). En ekstern tilnærming vil også komme feil ut i forhold til problemstillingene Bley-Vroman (1983) ønsker å belyse med sin

⁸ Må ikke forveksles med Suzanne Romaines skille mellom eksterne og interne årsaker til variasjon, se 2.2.3, s. 12

artikkel om ”*The comparative fallacy*”. Han advarer mot å skulle se mellomspråk med ”målspråksøyne”, fordi mellomspråk ikke nødvendigvis er avvikende utgaver av målspråket, men språk styrt av egne prinsipper.

Begge tilnærmingene tar allikevel utgangspunkt i målspråket, på den måten at de velger en bestemt målspråksstruktur, og ser på hvordan den er brukt. Forskjellen er at den ene ser på om distribusjonen av strukturen stemmer overens med målspråket, mens den andre fokuserer på hvordan den er distribuert i innlærerens eget system. Kort sagt, den ene ser på variasjon som variabel mestring, mens den andre ser på det som uttrykk for at mellomspråket har sitt eget system, og at strukturen faktisk *er* mestret, i forhold til dette egne systemet.

Systematisk variasjon er altså ikke et paradoks, fordi variasjonen og systematikken egentlig befinner seg på ulike plan. Variasjon er overflatefenomenet som oppstår fordi en har en forventning om at det skal være konsekvent i forhold til en norm, enten mellomspråket selv, eller målspråket. Om en nærmere undersøkelse viser at det er en *underliggende* systematikk, kan man si at variasjonen er systematisk, selv om den på *overflaten* er varierende. Prinsippene bak systematikken vil kunne si noe om *hvorfor* det er variabelt på overflaten.

Konklusjonen må være at variasjon er et overflatefenomen, og at hvilken norm man baserer sine forventninger på er avgjørende for hva det egentlig vil si at noe varierer. Er det variasjon som variabel mestring av målspråksstrukturer, eller er variasjonen egentlig et idiosynkratisk system hvor formene har sin spesielle distribusjon? Disse to er på ingen måte utelukkende, og mellomspråket vil være både systematisk i forhold til seg selv, og variabelt i forhold til målspråket. Det er bare to sider av samme mynt, to ulike måter å beskrive på, og ikke egentlige egenskaper ved språket i seg selv. Variasjonen vil alltid være der, fordi den vil være definert av uavhengige faktorer, og det er stor forskjell på å si at den er *forventet* å være systematisk, og at den faktisk er *påvist* å være det.

2.2.5 *Hvorfor studere variasjon?*

Hovedmålet med å se på mellomspråksvariasjon vil være å finne ut hva som skiller konteksten til den ene formen, fra konteksten til den andre formen, for så å knytte dette til en forklaring av hvilke faktorer som ligger bak og styrer vekslingen. Ofte kan den ene formen beskrives som *korrekt* i forhold til målspråket, og den andre som *avvikende*, og en kan si noe om hvorvidt formen er tilegnet eller ikke på grunnlag av dette. En undersøkelse av variasjon

vil derimot ikke *først og fremst* være opptatt av om trekket er tilegnet eller ikke, men heller forsøke å beskrive en dypere dimensjon ved oppbygging av et språkssystem og læring av målspråkets strukturer. Ellis belyser dette når han trekker et skille mellom to ulike typer utvikling i språklæring; *development as sequence* og *development as growth* (Ellis 1989: 36). Førstnevnte går på om et trekk er lært eller ikke lært, etter "*some predetermined criterion of performance*" (Ellis 1989:36). Et slikt forutbestemt kriterium kalles gjerne et *aquisition point*, og sier noe om hvor stor del av kontekstene som skal være korrekte for at man skal kunne si at innlæreren har tilegnet seg trekket. Ellis mener dette er et forenklet syn på tilegnelsen av språk. Han mener det er nødvendig med en mer dynamisk tilnærming, og uttrykker det slik: "*Investigating development-as-sequence, however, is only half the story, we also need to consider development-as-growth by finding out what happened to linguistic forms once they have entered interlanguage.*" (Ellis 1989: 43).

Dette vekstaspektet av utviklingen ivaretar den viktige forskjellen mellom å *ha kunnskap* om en regel, og å *kunne bruke* den i alle kontekster og under alle forhold. Det tar med i betraktningen det faktum at et trekk sjelden faller 100% på plass med en gang, at det som regel er snakk om en gradvis prosess, hvor innlæreren opparbeider større og større kontroll med regelen (*automatisering*), slik at den etter hvert kan anvendes effektivt og smertefritt både i ulike språklige kontekster og under ulike psykolingvistiske forhold. En variasjonstilnærming som har som mål å finne ut hva som skiller en kontekst fra en annen - enten det er produksjonsforholdene, det språkstrukturelle, eller kapasitetsbelastning - vil ivareta en slik nyanse, og kan derfor gi en god, dypere innsikt i den delen av mellomspråkutviklinga som skjer fra et trekk ikke brukes i det hele tatt, og til det blir anvendt 100% målspråklikt. I følge Ellis vil også synkron variasjon hos en innlærer være et fingeravtrykk av den diakrone, og studier av variasjon kan derfor gi oss indirekte innsikt i utviklingsprosessen (Ellis 1989:40), noe jeg vil komme tilbake til i kapittel 6.

Vurderingen av hvor stor innsikt studier av variasjon kan gi oss i andrespråklæring er imidlertid ulik i de ulike faglige retningene. Variasjon regnes gjerne som alt fra et uvesentlig bruksfenomen, til å være så viktig del av mellomspråket og språklæring at enhver seriøs teori om andrespråklæring bør kunne gjøre rede for det. Ellis skriver at "*... variability is an integral characteristic of interlanguage and that in order to understand how SLA proceeds it is necessary to recognise this and to build explanatory models that account of it*" (Ellis 1989: 23). Gregg (1990) har kritisert Ellis og Tarones oppfatninger av variasjon, og hevder at

variasjon er noe som hører språkbruksfæren til, og som ikke har noe med språkkompetanse å gjøre; og ”at redegjørelser for kompetanser innebærer at en ser vekk fra tilfeldigheter ved språkbruk, ..” (Berggreen & Tenfjord 1999: 261). Denne uenigheten er altså én side av en annen grunnleggende uenighet innen variasjonsforskningen: om variasjon skal regnes som et *kompetansefenomen* eller et *performansfenomen*. Er variasjonen en avspeiling av en tilsvarende variabel kompetanse, eller er den resultat av forhold som har med produksjonssituasjonen å gjøre, som ikke sier noe om hva innlæreren egentlig kan og ikke kan? Konflikten ligger altså i om hvorvidt man kan trekke konklusjoner om innlærerens kompetanse på grunnlag av variasjon. Dette er spørsmål jeg vil se nærmere på i kapittel 3.

2. 3 Inversjon i mellomspråk

2.3.1 Inversjon og typiske innlæreravvik

Det finnes en rekke andrespråksstudier av inversjonsstrukturen/ V2-regelen i norsk og andre V2-språk. Fenomenet har blitt studert fra ulike teoretiske utgangspunkt og med ulike metoder, men de fleste er stort sett enige i at dette trekket er en utfordring for innlærerne, fordi det faller sent plass, om i det hele tatt. Hovedproblemstillingen er innlærerkonstruksjoner som **Forrige torsdag det var fint vær*, som altså ikke er grammatisk korrekt i norsk. Denne typen avvik kalles ofte for *underinversjon* (Hagen 1992: 29). Hagen viser også til to andre typer inversjonsfeil; *feilinversjon* og *overinversjon*. Feilinversjon er når to andre ledd enn subjekt og verb har byttet plass, og er ifølge Hagen lite vanlig. Overinversjon, som er mer vanlig, er en overgeneralisering av inversjonsregelen, hvor invertert ordstilling brukes der det ikke er vanlig i målpråket, ofte etter konjunksjoner og subjunksjoner. Inversjonsstrukturen finnes ikke bare i tematiserte deklarativsetninger⁹, men også i avhengige og uavhengige interrogativsetninger (spørsmål), og derfor har mye av forskningen sett på disse to inversjonstypene - spørreinversjon og fremflyttingsinversjon - i sammenheng med hverandre. Jeg vil nå oppsummere noen studier som har forsøkt å belyse tilegnelsen av inversjon fra ulike perspektiver. Jeg har valgt å følge MacLaughlins (1987) tredeling for hovedområder innen andrespråksforskningen i fremstillingen av studiene; først studier som hovedsaklig vil si noe om utviklingsprosessen, dernest de som ser på hvordan morsmålet påvirker læringen, og til

⁹ Dvs. deklarativsetninger som har et annet ledd enn subjektet i forfeltet

slutt studier som - i likhet med min undersøkelse - har som mål å si noe om system og variasjon.

2.3.2 Inversjon og utvikling

Det er spesielt to studier som er sentrale og grunnleggende på dette området, *ZISAs*¹⁰ *multidimensjonale modell* (Meisel, Clahsen & Pieneman 1983) og *Karen Lunds funksjonelle tilegnelseshypotese* (Lund 1997). Disse to har ulike forklaringer på hvorfor inversjon kan være et vanskelig trekk å lære, og hvorfor det læres sent.

a) ZISA-prosjektets multidimensjonale modell

ZISA-prosjektets multidimensjonale modell er et forsøk på vise når inversjonsstrukturen tilegnes i forhold til andre syntaktiske konstruksjoner, med vekt på innlærerens interne kognitive begrensninger, og eksterne individuelle forhold. I følge denne modellen læres inversjon i deklarativsetninger sent pga innlærerens begrensede prosesseringskapasitet. Når innlæreren skal formulere seg utover den kanoniske leddstillingen, krever de alternative ordstillingene ulike grader av prosesseringskapasitet siden leddene i setningen må *flyttes* på. Innlærerne følger ifølge ZISA-prosjektet en generell læringsrekkefølge i pakt med at innlæreren gradvis får mer kapasitet til å prosessere og utføre mer kompliserte flyttinger. Subjekt/verb-inversjonen blir definert som en omrokering av setningsinterne konstituenten, og kommer derfor sent i løypa, lenge etter spørsmålsinversjon, som regnes som mindre kompleks, ettersom det bare er snakk om en omrokering av 'kantkonstituenten'.

En annen viktig faktor hos ZISA er inndelingen i ulike innlærertyper; de *korrekthetssøkende* og de *kommunikasjonsorienterte*. Den første typen innlærere forventes å unngå foranstilling, for slik å oppnå høy grad av korrekthet, mens den andre typen vil anvende foranstilling på bekostning av korrekthet.

Manfred Pienemanns *Prosessabilitetsteori* (Pienemann 1998) er en videreføring av de utviklingsmønstrene ZISA-prosjektet avdekket, og belyser innlæring av syntaks i forhold til utviklingen av leksikalske og morfologiske regler. Modellen består av fem nivåer for utvikling av et mellomspråk, og det er avgjørende å beherske de laveste nivåene før en

¹⁰ Zweitspracherwerb Italienischer und Spanischer Arbeiter

beveger seg videre til neste nivå. Inversjon kommer først på nivå 4, og dette tolkes som grunnen til at det faller på plass så sent, bare etterfulgt av leddsetningsordstilling.

Sentralt i begge disse modellene er at inversjon regnes som en komplisert *flyttingsprosess*. Denne prosessen krever stor mengde kapasitet fra innlærerens side, og det er nødvendig å ha prosessert seg gjennom flyttinger på lavere nivåer og bygd opp kapasitet, før de kan mestre inversjonsstrukturen.

b) Karen Lunds funksjonelle hypoteser

Karen Lund (1997) argumenterer for en *funksjonell* tilnærming til språktilegnelse, med fokus på språkets innholdsside fremfor den tradisjonelle orienteringen mot form og struktur. Hun mener innlæringen av inversjon må ses i forhold til den funksjonen strukturen har, og ikke i forhold til strukturens form. Hun argumenterer mot ZISA-prosjektets påstand om at foranstillingsinversjonen er en krevende *flyttingsprosess*, og mener grunnen til at den læres sent er mangelen på en klar semantisk eller pragmatisk funksjon: "*Den grammatiske struktur, inversion, er ikke per se vanskelig at tilegne; det er den kun i den kontekst hvor semantikken og pragmatikken stort sett er fraværende: i den fremsættede sætning.*" (Lund 1997: 23)

I deklarativsetninger har ikke inversjonen noen egen funksjon, slik den har i ja/nei-spørsmål, hvor den markerer et pragmatisk skille mellom spørsmål og utsagn, og det er derfor mindre kommunikativt press på innlæreren til å lære seg foranstillingsinversjon. Denne teorien støttes av at inversjon i spørsmål læres tidligere enn i deklarativer, både hos Lund og ZISA-prosjektets informanter. Inversjonen i seg selv, det å plassere subjektet bak verbet, er altså ikke vanskelig, men mangelen på en klar funksjon fører til at strukturen ikke blir lagt merke til i starten av tilegnelsen. Lund påpeker også at innlærere ofte overgeneraliserer inversjon til andre kontekster, og bruker dette mot ZISA; hvorfor overgeneralisere en så vanskelig struktur til en kontekst som er enda mer komplisert, altså leddsetninger?

Lunds holdning til utviklingen videre i inversjonstilegnelsen er at innlærerne danner seg hypoteser om *når* det skal inverteres. Hun ser ut til å mene at det skjer en gradvis prosess, det hun kaller *stagingprosess*, hvor inversjonen først faller på plass i noen kontekster for så å spre seg: "*Generelt i tilegnelsen af et nyt sprog er det således at nye strukturer først tilegnes i nogle faste kontekster, fx efter specielle adverbialer og/eller i bestemte verbale kontekster,*

hvorefter repertoiret gradvis udvides og stabiliseres” (Lund 1997: 155) Dette er noen jeg vil komme tilbake til i kapittel 3, fordi det er vesentlig i forhold til mine egne hypoteser.

c) Brautaset 1996

En større undersøkelse av inversjon i norsk mellomspråk ble utført av Anne Brautaset i 1998. Hovedmålet var å sjekke holdbarheten til den gjengse vurderinga om underinversjon som typisk for mellomspråk og spesielt vanskelig å tilegne seg. Undersøkelsen ser på feilkvoter og utviklingsmønster over tid, og Brautasetts materiale viser at det faktisk er dominans for korrekte strukturer, og at samtlige informanter har inverterte setninger allerede fra første innsamling, etter to måneder. I tillegg er det en klar progresjon i innlæringa hos alle informantene hennes. Brautaset konkluderer derfor med at inversjonsfeil ikke ser ut til å være verken dominerende i norsk mellomspråk, eller vanskelige å overvinne. Dette er interessant fordi det strider imot en rekke andre studier.

2.3.3 Inversjon og morsmål

Flere studier har sett på morsmålets rolle i tilegnelsen av inversjon. Karen Lund har i samme undersøkelse som er nevnt over, sett på forskjellene mellom tre ulike morsmålsgrupper; nederlandsk, engelsk og spansk. Det viktigste funnet er at selv om alle følger den samme utviklingstendensen, så er det kun de nederlandske informantene som ved siste innsamling har oppnådd en viss stabilitet for inversjonsstrukturen (Lund 1997:163). Lund mener dette har sammenheng med at disse informantene har en lignende inversjonsstruktur i sitt eget morsmål, noe som gjør innlæringa lettere for dem. Dette kommer av at bruk av inversjonsstrukturen allerede er automatisert hos innlæreren gjennom morsmålet, og det eneste som kreves er å oppdage at denne strukturen også finnes i målspråket.

”Når lørneren først har opdaget sammenhængen mellem udtryk og inhold, og hvis der er strukturelle overenskomster mellem L1 og L2, kræver det få kognitive ressourcer at anvende denne strukturen i L2, i og med den allerede er operationaliseret og automatiseret i L1”
(Lund 1997: 164)

Katrin Saarik (2002) har sett på forskjellen mellom estiske og finske innlæreres tilegnelse av inversjon. Estisk har en lignende V2-regel som norsk, men den er ikke helt lik den norske, i

og med at den er pragmatisk motivert, og ikke grammatisk som i norsk, noe som gjør at verbet kan stå på tredjeplass i enkelte setninger. Hun konkluderer med at finnene, som ikke har en V2-regel i morsmålet, har større problemer med inversjon enn esterne, som kan overføre regelen fra morsmålet, og dermed få mange korrekte strukturer. Allikevel er denne formen for transfer negativ, mener Saarik, ettersom den fører til at esterne, i motsetning til finnene, ikke lærer seg den norske regelen. Esterne følger nemlig den estiske regelen, ikke den norske, noe som fører til at bruken av pragmatiske motiverte V3-konstruksjoner fra estisk også overføres til norsk, og fører til avvik. Så mens nederlenderne ser ut til å nyte godt av å ha en liknede V2-regel som norsk, kan dette på sett og vis være negativt for esterne.

2.3.4 Inversjon og variasjon

Flere av studiene som er omtalt over, har indirekte sagt noe om variasjon i mellomspråk. Både Lund og ZISAs forklaringer på hvorfor spørreinversjon læres før foranstillingsinversjon, sier noe om hvorfor inversjonsstrukturen er anvendt i *spørsmål*, men ikke i *deklarative setninger*. Mens ZISA mener variasjonen er styrt av samspillet mellom strukturens kompleksitet og innlærernes interne prosesseringsforhold, mener Lund at det er skillet mellom den funksjonelle spørreinversjonen og den kommunikativt *tomme* foranstillingsinversjonen som styrer denne variasjonen.

Det man ikke har gått inn på i disse undersøkelsene, er den vekslingen mellom invertert og ikke-invertert ordstilling som ofte bare er å finne i deklarative setninger i mellomspråk. Inversjon mangler pragmatisk og semantisk motivasjon i *alle* deklarative setninger, så man kan ikke lenger ty til en funksjonell forklaring når man ønsker å si noe om hvorfor innlæreren veksler mellom inverterte og ikke-inverterte deklarative setninger.

Bolander (1988) og Hagen (1992) har gjort forsøk på å kartlegge denne typen variasjon i deklarative setninger. De har begge undersøkt hva som skiller den lingvistiske konteksten til inverterte og ikke-inverterte setninger, ved å se på hvilke foranstilte ledd, subjekter og finitte verbaler som kan se ut til å være positive og negative for inversjonsmestring. Bolander kommer blant annet fram til at objekter i forfeltet er positivt, mens leddsetninger er hemmende. For subjekter er det positivt om disse er substantiver eller førstepersonspronomen. Ellers fremhever hun verb som uttrykker mening og antakelse som positivt. Bolander har dessuten et viktig poeng om at *formularlæring* kan være årsak til noen av resultatene hennes. Det at foranstilte objekter ofte forekommer sammen med invertert ordstilling i hennes

materiale, kan ha vært resultat av formularlæring av det-setninger, og setningsvedheng som 'tycker jag.'

Jon Erik Hagen har undersøkt skriftlige arbeider fra tre ulike tidspunkt, og kom til følgende konklusjon: *"Det underinverterte setninger tenderer mot å ha felles, synes å være at de inneholder trekk som belaster språkbrukerens informasjonsbehandlings-kapasitet i sterkere grad enn de korrekte inverterte setningene han får til."* (Hagen 1992: 34). Hagen mener altså at variasjonen er kapasitetsbetinget, styrt av kapasitetsbelastningen til hver enkelt kontekst, og han legger frem en rekke egenskaper ved de tre leddene som kan sies å fremme og hemme inversjonsmestringen:

a) Pragmatiske trekk som er underinverteringsutløsende

- tunge (antall ord) foranstilte ledd
- sterk tekstbinding mellom foranstilt ledd og foregående kontekst
- koordinering av setninger øker sjansene for underinvertering i siste helsetning

b) Pragmatiske trekk som fremmer korrekt inversjon

- pronomensubjekter
- lange tunge subjekter
- rematiske subjekter (subjekter som inneholder "ny informasjon")
- frekvent verb som finitt verb
- kort finitt verb

(Berggreen og Tenfjord 1999: 268)

Hagen oppfordrer til en mer systematisk utforskning av disse trekkene, og en rekke andre undersøkelser har tatt oppfordringen og utført del-undersøkelser av variasjon. Vikøy (2004) ser på det foranstilte leddet, og Brautaset (1996) ser på foranstilt ledd og subjekt. men ingen av disse ser ut til å finne noen bekreftende tendenser for forfeltet, bortsett fra at Brautaset sier at pronomensubjekt ofte forekommer med korrekt inversjon. Jeg vil komme tilbake til disse studiene senere i kapittel 4 og 5.

Et problem med Hagen er at han ikke sier noe direkte om hva som gjør disse trekkene mer eller mindre belastende for informasjonsbehandlingskapasiteten. Dette kan være en svakhet. En styrke ved Hagens undersøkelse er at han ikke ser på de ulike leddene hver for seg, men sammenstiller dem, og ser på samspillet mellom positive og negative faktorer i hver setning.

Hagens studie har også vært utgangspunkt og inspirasjon for min undersøkelse, men jeg ønsker ikke å anvende hans funn direkte, slik Vikøy og Brautaset har gjort. Jeg har for det første ønsket å se om det stemmer at det faktisk er en systematikk i variasjonen styrt av den lingvistiske konteksten, og for det andre undersøke hvilke egenskaper ved konteksten som står bak det Hagen omtaler som ”belastning av informasjonsbehandlingskapasiteten”, noe som ikke er operasjonalisert i hans artikkel. I kapittel 3, 4 og 5 vil jeg bygge opp mine egne hypoteser og operasjonaliseringer i tilknytning til variasjon.

2.3.5 Å lære seg inversjon

Det er altså gjort en rekke undersøkelser av den utfordringen inversjon er for innlærere av V2-språk. Noen har gjort forsøk på å forklare hvorfor det faller på plass så sent, og noen sier noe om morsmålets påvirkning på innlæringa, og gjennom tilnærminger som ser på variasjon, har vi fått presentert en rekke lingvistiske konteksters positive eller negative påvirkning på anvendelse av inversjonsstrukturen. Som alle andre trekk som skal læres er det klart at innlæringsprosessen er et resultat av både kognitive faktorer, morsmålsbakgrunn og rent lingvistiske forhold, og at alle disse på belyses for å få et helhetlig bilde av prosessen.

2.3.6 Inversjon og invertering - prosess eller struktur?

Innenfor noen teoretiske retninger er det vanlig å tenke på inversjon som en form for flytteprosess hvor subjektet og verbet bytter plass, slik som ZISA-modellen fremstiller det. Dette har også vært en vanlig måte å beskrive det for innlærere på, enten at subjekt og verb *byter plass*, eller at subjektet *flytter bak* verbet. Det er allikevel også mulig å bruke denne termen uavhengig av hva som er opphavet til en slik struktur, både i innfødt norsk og i norsk mellomspråk, og det er det jeg vil gjøre. Jeg har valgt å først og fremst bruke formuleringen *invertert ordstilling*, for å fremheve at det er en beskrivelse av en overflatestruktur (finit verb etterfulgt av subjekt), og ingen flytteprosess. Ordstillingen er invertert på den måten at den avviker fra den kanoniske¹¹, i og med at subjektet står bak verbet. *Inversjon* og *invertering* er allikevel mye brukte termer, så jeg kommer til å bruke disse på samme teoriuavhengige måte, uten å legge til grunn teorier som forutsetter flytteprosesser.

¹¹ Som i norsk som kjent er subjekt, verbal, objekt (SVO)

Jeg vil dessuten benytte termene *overinversjon* og *underinversjon* som beskrivelser av ikke-målspråkslike strukturer i mellomspråket, men ikke nødvendigvis for å fokusere på feilaspektet ved dem. Hovedinndelingen min er uansett i *invertert* og *ikke-invertert*, siden noen av de ikke-inverterte ikke nødvendigvis er underinverterte eller avvikende i norsk grammatikk, og noen av de inverterte ikke nødvendigvis er korrekte fordi de er overinverterte, noe jeg kommer nærmere inn på i de neste kapitlene.

Kapittel 3: VARIASJON OG INVERSJON

I dette kapittelet vil jeg gå nærmere inn på forutsetningene for undersøkelsen min. Etter å ha sagt litt mer om hva jeg legger i begrepet *variasjon*, ser jeg litt på hvor utbredt slik variasjon er i materialet mitt fra ASK-test. I 3.3 prøver jeg å trenge inn hva som kan være inversjonsutfordringens natur, for så å knytte dette til teorier om variasjon og språklig prosessering, og avslutter med å presentere di to hypoteser som er grunnlaget for mine undersøkelsene av variasjon.

3.1 Fins det variasjon og hvor vanlig er det?

I første omgang, som grunnlag for hovedundersøkelsen, har jeg vært interessert i å finne ut hvor utbredt variasjon egentlig er i materialet mitt. Målet har vært å få oversikt over inversjonens status i de 100 tekstene i ASK-test, og arbeidet har først og fremst vært en oppregning av *obligatoriske kontekster*. Korpuset består av hundre tester, fra ti ulike morsmål. Undersøkelsen har blitt gjort ved at jeg har lest gjennom tekstene og telt kontekster, etter kriterier for inversjonskontekster, som jeg gjør rede for i kapittel 3.2.

3.1.1 Definisjon av variasjon

Variasjon vil hos meg være definert som en *veksling i én og samme tekst mellom invertert og ikke-invertert ordstilling der en i målspråket har et grammatisk krav om invertert*. Det er altså på sett og vis en ekstern definisjon, ettersom jeg tar utgangspunkt i obligatoriske kontekster og definerer variasjonen i forhold til målspråkets norm, men jeg vil i tillegg ta med overgeneraliseringer og steder der det ikke er overgeneralisert, for å kunne ta forbehold om at invertert ordstilling brukes også utenom obligatoriske kontekster, som i en intern tilnærming. Dette fordi jeg skiller mellom *fenomenet* jeg ønsker å forklare; variabel anvendelse av inversjonsregelen i tematiserte deklarativer setninger, i én og samme tekst; og *metoden* jeg bruker for å undersøke dette, som selvfølgelig må ta forbehold om at mellomspråket har sitt eget system og sin egen distribusjon, fordi innlæreren kan tenkes å overføre strukturer til andre kontekster enn de obligatoriske. Fokuset vil som sagt ligge på vekslingen mellom invertert og ikke-invertert, og ikke mellom korrekt og avvikende, som i en tradisjonell ekstern tilnærming.

3.1.2 Hva regnes som inversjonskontekster?

Jeg har konsentrert undersøkelsen min rundt det jeg har kalt *inversjonskontekster*, som da er setninger i innlærertekstene hvor det enten *er* invertert ordstilling eller *skulle vært* det ifølge norsk grammatikk. Jeg har regnet med to typer av inversjonskontekster; *obligatoriske* og *overgeneraliserte*. Siden jeg skal se på foranstillingsinversjon i hovedsetninger, vil ikke avhengige og uavhengige interrogativsetninger, altså spørresetninger, regnes som obligatoriske kontekster, kun deklarativer setninger. Jeg vil skille mellom obligatoriske kontekster og overinverterte kontekster, ettersom de obligatoriske kontekstene er det som skal undersøkes, og de overgeneraliserte bare skal belyse bruken generelt.

Obligatoriske kontekster er altså deklarativer setninger som innledes av andre ledd enn subjektet. I tillegg har jeg regnet med eventuelle leddsetninger med forfelt. Leddsetninger har vanligvis ikke forfelt, men noen typer setninger kan allikevel ha det, og disse vil ha samme krav om invertert ordstilling som hovedsetninger. Det kan dessuten tenkes at det finnes eksempler på leddsetninger med forfelt som ikke kan ha det i norsk, men også disse regnes som obligatoriske kontekster.

Noen typer setninger er ikke obligatoriske kontekster i strengeste forstand, da de kan stå både med og uten invertert ordstilling i norsk. Dette gjelder setninger med *ekstraposisjon*¹², som ikke egentlig er en foranstilling av et ledd, men plassering av ledd i et løst forfelt utenfor setningen (Faarlund, Lie og Vannebo 1997: 904). Dette er vanlig etter uttrykk som *heldigvis*, *for eksempel*, *forresten*, *dessverre* m.m., som i *Heldigvis, jeg fikk lov*. Ved slike setninger kan det være vanskelig å vurdere hva innlæreren egentlig har tenkt; er det et bevisst valg, eller er det en underinversjon? Jeg vil allikevel regne med denne typen setninger, både de som er inverterte og de som ikke er det, fordi jeg ikke fokuserer på korrekt og avvikende, men på invertert og ikke-invertert.

Når det gjelder *overinversjonskontekster* har jeg tatt med disse der de finnes, som regel er dette i leddsetninger, etter konjunksjoner, og etter venstredislokerte subjekter.

¹² Ledd som ligger utenfor setningsskjemaet. Slike ledd står i et løsere forhold til setninga, og skilt ut med komma/pause. Når det kommer først i setningen, kalles det gjerne *løst forfelt*. (Faarlund, Lie og Vannebo 1997: 904)

3.1.3 Inversjonsmønster i tekstene

Med hensyn til invertert ordstilling kan man forvente fire typer innlærertekster, hvor type 4 er det som blir regnet som variasjonstekster:

1. Tekster *uten* obligatoriske inversjonskontekster.
2. Tekster med *bare inverterte* obligatoriske inversjonskontekster
3. Tekster med *bare underinverterte* obligatoriske inversjonskontekster
4. Tekster med *både inverterte og underinverterte* obligatoriske inversjonskontekster

Det er viktig å tilføye at jeg her ikke opererer med noe som helst mål for *tilegnelse*, og bevisst ikke vil si noe om hvorvidt jeg regner inversjon for tilegnet eller ikke. De fire kategoriene er altså bare en *beskrivelse*, med utgangspunkt i om en tekst har variasjon eller ikke. Type 3 - bare inverterte kontekster – betyr altså ikke nødvendigvis at innlæreren som har skrevet teksten har tilegnet seg inversjon. For å kunne trekke en slik konklusjon, må en rekke andre forhold være avklart, for eksempel hvilke typer av inversjonskontekster som er representert, og om det er noen overinversjoner.

Følgende tabell viser fordelingen av de fire typene tekster, både totalt og etter morsmål.

Tekster med	Type 1	Type 2	Type 3	Type 4
Hele ASK-test	0	55	1	44
Albansk	0	2	0	8
Engelsk	0	7	0	3
Nederlandsk	0	9	0	1
Polsk	0	4	1	5
Russisk	0	7	0	3
Serbokroatisk	0	3	0	7
Somali	0	3	0	7
Spansk	0	4	0	6
Tysk	0	10	0	0
Vietnamesisk	0	6	0	4

De mest tydelige resultatene her, er at det ikke finnes noen tekster som *mangler* inversjonskontekster, selv om 0 av tekstene har så lite som bare én kontekst. Videre har over halvparten av tekstene, 55 stykker, alle obligatoriske kontekster invertert, mens 44 tekster viser variasjon. Størstedelen av variasjonstekstene, 25 stykker, har bare én eneste ikke-

invertert kontekst. 17 av de 55 av tekstene som har alle obligatoriske kontekster invertert, har også overinversjon, noe som viser at deres inversjonsregel enda ikke er helt målspråkslik.

De tyske og nederlandske tekstene er de eneste som ikke har variasjon, og disse er de eneste språkene i korpuset som har en V2-regel som tilsvarer den norske. Alle disse tekstene har 100% inverterte setninger, bortsett fra en av de nederlandske, som ikke har invertert ordstilling i én av 15 kontekster. Denne setningen innledes av *for eksempel*, og kan altså regnes som grammatisk uten inversjon. Dette skulle være i tråd med Karen Lunds resultater, hvor de nederlandske informantene raskere oppnådde stabilitet i inversjonsmestringen (Lund 1997:163).

3.1.4 Utvalg av tekster til undersøkelsen

Denne undersøkelsen av korpuset har gitt grunnlag for en tredeling av tekstene,

1. Tekster uten variasjon
2. Tekster med variasjon og én ikke-invertert kontekst
3. Tekster med variasjon og minst to ikke-inverterte kontekster

44 av tekstene har variasjon slik jeg har definert det i kapittel 3.1.1., og av disse har størstedelen bare én kontekst uten invertert ordstilling, noe som kan være resultat av en glipp. På grunn av sannsynligheten for at en feil er en ren glipp, og ikke knyttet til innlærerens egentlige kunnskap om inversjon, vil jeg i den videre variasjonsundersøkelsen utelate tekster med bare én ikke-invertert kontekst, og heller se på den siste gruppen tekster, som har minst to kontekster uten invertert ordstilling.

3.2 Inversjonsutfordringen

Selv om man helst skal holde seg til en beskrivelse av mellomspråk som ikke er styrt av målspråkets kategorier, så vil alltid statusen som korrekt eller avvikende i forhold til målspråket være en mulig beskrivelse. Det som er avgjørende er hvorvidt man fokuserer på det som avviker som ”feil”, eller om man bare anser det som uttrykk for et underliggende idiosynkratisk system (jf. Corder 1982: 37). På grunn av dette kan kunnskap om hva som gjør inversjon til en utfordring være sentralt i forhold til å finne ut hva som styrer variasjonen. De kontekstene som kan beskrives som avvikende representerer tross alt en stor del av kontekstene som utgjør variasjonen slik jeg har definert den her.

Jeg vil nå se på de tre mest sentrale observasjonene, eller 'sannhetene', om inversjon i mellomspråk, og bruke dette for å finne ut mer om hva som kan være inversjonsutfordringens natur.

3.2.1 Tre 'sannheter' om inversjon i mellomspråk

Av de studiene som jeg har omtalt i kapittel 2, er det spesielt tre observasjoner/konklusjoner som er sentrale:

1. Inversjon er vanskelig, og læres sent (Hagen 1992, Lund 1997)
2. Inversjon er et trekk hvor innlærerne viser variasjon (Bolander 1989, Hagen 1992)
3. Inversjon overgeneraliseres (Hagen 1992, Lund 1997)

Hagen skriver i sin artikkel at "*Dette at norsk krever finitt verb på andre plass i fortellende setninger er, som vi lærere vet, notorisk vanskelig å tilegne seg for fremmedspråklige innlærere*" (Hagen 1992: 28). At inversjon er vanskelig, er altså ikke bare noe som forskning viser, det er noe som alle vet. Allikevel, hva vil det si at det læres sent og at det er vanskelig? De fleste studier som har sagt noe om dette, har sett på om inversjonen er falt på plass i obligatoriske kontekster i målspråket. De ser på det som Ellis kaller *development as sequence* (Ellis 1989:36, jf. kap. 2.2.6). I slike tilnærminger er det vanlig å sette en grense for hvor mange kontekster som skal være korrekte for at trekket skal regnes som lært, et såkalt *acquisition point*. Dette kan være alt fra krav om bare én korrekt kontekst, til krav om 90% eller 100% korrekthet. Det er viktig å merke seg at slike tilegningspunkt bare er metodiske virkemidler, ikke egentlige egenskaper ved innlærerens kompetanse, og konsekvensen av at ulike forskere opererer med ulike krav, er at man kan ende opp med ulike konklusjoner om hvorvidt et trekk er lært eller ikke, og om det er vanskelig eller ikke.

De ulike tilegningspunktene som velges representerer altså utvalgte punkt på veien fra aller første anvendelse av et trekk, til avvik ikke lenger forekommer. Men de fleste trekk læres gradvis, og det vil alltid gå en stund fra tidspunktet en innlærer begynner å invertere, til han/hun ikke lenger gjør feil. På grunn av dette mener Ellis at det er meningsløst å operere med en oppdeling i lært og ikke lært. Dette er punkter man trekker for metodens del; "*Here interlanguage rules cannot be dichotomised as acquired or not acquired according to some pre-selected criterion level of performace, but must be viewed dynamically*" (Ellis 1989: 36)

Det han peker på her, er det faktum at man ofte ikke finner 100% mestring før etter lang tid, og det som er imellom, er nettopp det som utgjør variasjon. Inversjonsstrukturer kan dukke opp tidlig, men det blir gjort 'feil' lenge, og noen ganger fossiliseres strukturen. I mellomtiden vil man ha variasjonsmønster, hvor innlæreren veksler mellom inverterte og underinverterte setninger i sterkere eller svakere grad. Denne variable mestringen er altså et tegn på at det er et trekk som læres gradvis, noe Hagen også understreker; ”*Det synes å være hevet over enhver tvil at korrekt anvendelse av inversjonsregelen er en ferdighet som kun gradvis tilegnes i kommunikativ språkbruk.*” (Hagen 1992:36).

De to første sannhetene om inversjon, at det læres sent, og at det er vanlig å finne variasjon, er altså til en viss grad to sider av samme sak. Bedre enn å si at inversjon læres sent, er å si at *inversjonsfeil overvinnes sent* (Lund 1997, Hagen 1992), og inntil de er overvunnet, vil en ofte oppleve å finne inverterte og underinverterte setninger i en og samme tekst.

Den siste observasjonen er av en litt annen natur. *Overinversjon*, altså invertert ordstilling i kontekster der det ikke skal være det, er funnet i de fleste studiene som er nevnt over, om enn ikke særlig frekvent. Lund påpeker at hun synes det er merkelig at inversjon, som jo regnes for å være en svært vanskelig prosess av ZISA, allikevel overgeneraliseres (Lund 1997: 167). Dette mener hun er et tegn på at inversjonsstrukturen i seg selv ikke er vanskelig.

Flere har utelatt overinversjoner fra undersøkelsene, og både Hyltenstam og Hagen mener ifølge Bergreen og Tenfjord (1999: 269) at overinversjon ikke er noe avgjørende steg som innlæreren må gjennom for å lære inversjon, og det ser ut til at det heller ikke er noe som alle gjør, og i tillegg overvinnes det ganske raskt (Hagen 1992: 29). Både Brautaset og Saarik har utelatt overinversjoner, fordi de har funnet så få eksempler i sitt materiale. Jeg har valgt å ta med de få forekomstene i mitt materiale, av de metodiske grunnene jeg har vært inne på.

Det at det overgeneraliseres kan være et tegn på at innlærerne opererer med en hypotese om distribusjon av inversjonsstrukturen, noe som vitner om at de har fått teften om en regel. Som sagt over, overgeneraliseringer kan være viktige kilder til mellomspråkets eget system og gi kunnskap om innlærerens hypoteser om målspråket (Huebner 1983:xiii).

Det gjenstår allikevel å vise at overinversjon faktisk er en *overgeneralisering*, i betydningen ”overbruk” av en regel. Bruk av invertert ordstilling i leddsetninger trenger faktisk ikke å være en ren overgeneralisering av en regel. Om det resulterer strukturer som allerede finnes i målspråket, kan det like gjerne være en form for helfraselæring, hvor setningskonstruksjoner fra målspråket brukes både i hovedsetninger og leddetninger.. Lund og Hagen har vektlagt at dette oftest skjer etter subjunksjoner som er homonyme med enkelte foranstilte adverbialer. Jeg skal se nærmere på dette i kapittel 4.

3.2.2 Hva det kan være, og hva det nok ikke er

Basert på disse tre punktene, og forhold som er relatert til dem, har jeg skapt meg en forestilling av hva som kan være inversjonsmestringens problematiske natur, og som danner grunnlaget for de hypotesene jeg stiller i denne avhandlingen.

a) Hva det nok ikke er:

Både Hagen og Lund m.fl. konstaterer altså, i motsetning til ZISA-prosjektet og Pienemann, at det ikke ser ut til å være inversjonsstrukturen/den inverterte ordstillingen i seg selv som er problematisk. For det første finner man den tidlig i spørsmål, for det andre finner man den ofte i enkelte kontekster, og for det tredje overgeneraliseres den. Det som gjør inversjon i deklarativer problematisk, må altså være relatert til *andre* faktorer enn selve inverteringen.

Lund viser at det er den manglende pragmatiske funksjonen som gjør foranstillingsinversjon vanskelig i forhold til spørsmålsinversjon, men hennes funksjonelle tilegnelseshypotese kan som sagt bare si noe om hvorfor inversjon i deklarativer er vanskeligere enn i spørsmål, men ikke hvorfor det tilsynelatende er vanskeligere i noen typer deklarativer enn i andre.

b) Hva det kan være - to mulige forklaringer

Jeg mener det finnes to mulige grunner til at noen inversjonskontekster er vanskeligere å gjennomføre enn andre. Den ene er knyttet til ufullstendig avkoding av distribusjonen til inversjonsstrukturen, og den andre er knyttet til innlæreres begrensede prosesseringskapasitet.

Overinversjon tatt i betraktning, kan det virke som at innlærerne har problemer med å skjønne når det skal være invertert ordstilling, de inverterer jo når de ikke skal. I tråd med dette kan det være nærliggende å tro at også underinversjon kan være utslag av en slik usikkerhet. Lund

antyder at overinversjon kan være resultat av feilaktige hypoteser hos innlærerne, uten egentlig å direkte si noe om at dette kan være svaret på underinversjon også. Jeg mener at en mulig forklaring på variasjonen i inversjonsbruk, både for underinversjon og overinversjon, er at det er resultat av en *ufullstendig avkoding av målspråket*. Den første forklaringen på variabel anvendelse av invertert ordstilling er altså at det er resultat av et **tolkningsproblem**, hvor innlærerne har problemer med å identifisere *når* det skal inverteres og når det ikke skal det.

Hagen derimot har en annen forklaring på variasjonen og underinverteringen. Han mener at det kan være snakk om *et kapasitetsproblem*, at underinversjon skiller seg fra det han kaller overinversjon og feilinversjon, som han anser som tolkningsbetinga. Han skriver følgende:

”Hittil har vi betraktet inverteringsfeil som et strukturelt interpretasjonsproblem, og som sådant et lingvistisk kompetanseproblem av språkssystematisk art. Men det er klart at strukturell interpretasjon er bare en side av saken når det gjelder språklæring. Like viktig er i hvilken grad våre studenter faktisk makter å anvende sin kompetanse i performanse.” (Hagen 1992: 32)

Hagen ser ikke ut til å mene at det er tolkning av regelen som skaper variasjon, men at det er kapasitetsbegrensinger. Innlærerne vet at det skal inverteres, men mangel på oppmerksomhetsressurser fører til avvik, slike avvik som Corder kaller *postsystematiske* feil, i motsetning til *presystematiske* feil som oppstår fordi innlæreren ikke kjenner regelen (Corder 1982: 18). Den andre forklaringen på variabel anvendelse av invertert ordstilling er altså at det er resultat av en varierende kapasitet, som gjør at innlærerne ikke alltid får anvendt den kunnskapen de har.

Her er det altså to mulige forklaringer på variasjonen:

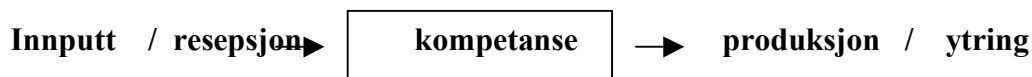
Forklaring 1:Tolkning Forklaring 2: Kapasitet
--

Implisitt i disse forklaringene ligger det påstander om den relative vanskegraden til ulike inversjonskontekster; noen er lettere å avkode eller anvende enn andre, og det er disse som kan forventes å stå *med* invertert ordstilling i tekstene.

3.3 Variasjon og språklig prosessering

Disse to mulige forklaringene impliserer også ulike ting når det kommer til hvor i den språklige bearbeidingsprosessen variasjonen er lokalisert. I hovedsak representerer de de to hovedutfordringene en innlærer står overfor i møtet med det nye språket: *avkodning* av målspråket og *anvendelse* av kompetanse, noe som gjerne blir omtalt som *resepsjon* og *produksjon* (Bergreen og Tenfjord 1999: 86).

Innen psykolingvistikken snakker man gjerne om *språklig prosessering* som en bearbeiding av språk over flere stadier. Innlæreren møter språklig *innputt*, som avkodes og analyseres og tas opp som kunnskap om målspråket, eller *kompetanse*. Kompetansen blir så utgangspunktet for innlærerens egne ytringer på andrespråket.



For å skille ut den delen av innputt som potensielt kan tas opp i kompetansen og den som faktisk gjør det, har C. Chaudron foreslått å skille mellom innputt og faktisk *inntak* (Berggreen og Tenfjord 1999: 134). Det er en grunnleggende todeling av denne prosessen i *resepsjon* og *produksjon*, som langt på vei tilsvarende skillet mellom *kompetanse* og *performanse*, mellom kunnskap om språket og bruk av språket.

3.3.1 Prosessering og lokalisering av variasjon

Når det gjelder lokalisering av hvor i denne prosessen variasjonen har sitt opphav, står det tidligere nevnte skillet mellom variasjon som *kompetansefenomen* og *performansefenomen* (språkbruksfenomen) sentralt. Her er det vanlig å skille mellom to paradigmer (Tarone 1988); på den ene siden *the homogenous competence paradigm* som har som utgangspunkt at kompetansen er stabil og homogen, og på den andre siden *the heterogenous competence paradigm* som mener at den er heterogen og *variabel*. Om det heterogene paradigmes syn på innlærernes kompetanse skriver Sharwood Smith:

”... , *they have a competence which contains usage informations such that rules are stated in terms of the various manifestations they have according to the situation and context and the speaker’s communicative requirements.*” (Sharwood Smith 1994: 107)

Det heterogene paradigmet vil derfor lokalisere variasjonen i den *reseptive* delen av prosessen, knyttet til oppbyggingen av kompetansen gjennom avkoding av innputt, og som en del av en tilsvarende *variabel kompetanse*, mens det homogene paradigmet vil lokalisere den i *produksjonsfasen*, styrt av performanseforhold, og variasjon er derfor ikke knyttet til hva innlæreren kan, men i hvor stor grad han greier å benytte kunnskapen.

Det er viktig å poengtere at dette ikke har like mye å gjøre med hvor i prosessen variasjonen egentlig er lokalisert, som hvilket syn de ulike teoretiske retningene har på språklæringsprosessen i seg selv. Det er altså ikke først og fremst egenskaper ved språket i seg selv som er avgjørende, det avgjørende her er hva man definerer *kompetanse* som; er kompetanse bare *kunnskap* om språkstrukturene, eller er det også evnen til å bruke kunnskapen under ulike forhold? Dette er den velkjente forskjellen mellom kunnskap om språket, og det som noen kaller *prosesseringskontroll* (Sharwood Smith 1994), automatisering (MacLaughlin 1987), eller *development as growth* (Ellis 1989).

Ellis og MacLaughlin ser er blant dem som anser oppbyggingen av kontroll over bruken av strukturer som *en del* av kompetansen, og som derfor ikke opererer med noe skille mellom performanse og kompetanse. Andre, for eksempel Gregg (1990), mener at dette ligger utenfor kompetansen, og altså er performanse, og derfor ikke er relevant i en studie av kompetanse. (Bergreen og Tenfjord 1999: 25). Men som Sharwood Smith poengterer, dette er en metodisk avgrensning av studieobjektet, det er ikke en påstand om at kompetanser ER homogene (Sharwood Smith 1994:107). Språkbruk er en del av det innlæreren skal lære seg, men det er relevant for teorier om språkbruk/performance, ikke for teorier om språklig kompetanse (Bergreen og Tenfjord 1999: 261). Så om variasjon er et kompetansefenomen eller et performansefenomen, kommer egentlig an på hvordan man har valgt å avgrense sitt studieområde:

”... *one can distinguish models which present one system within language as unaffected by other systems or levels and following different principles and, on the other hand, models which see different systems as interacting, as following the same basic design and therefore not really separate or autonomous.*”(Sharwood Smith 1994:107)

Hovedkjernen i skillet er altså om bare kunnskapen, eller både kunnskap og kontroll over/automatisering av anvendelsen skal regnes inn i begrepet kompetanse. Selv ønsker jeg ikke å henge meg opp i dette skillet mellom kompetanse og performanse. I min sammenheng finner jeg det mer relevant å operere på nivået for *kunnskap* og *kontroll*, ettersom dette er prosesser som eksisterer uavhengig av definisjoner av kompetanse. ”*The notion of processing control allows us to separate out those factors of language ability that we want to call ‘knowledge’ and those factors that have to do with the more or less skillful deployment of knowledge in actual performance millisecond per millisecond*” (Sharwood Smith 1994:16) Spørsmålet blir derfor ikke om det er et kompetansefenomen eller et performansefenomen, men om det er *kunnskapsrelatert* eller *kontrollrelatert*.

3.3.2 Kunnskap, kontroll og variasjon

Tilegnelse av kunnskap om målspråket er grunnleggende i læringsprosessen, og kan defineres som kunnskap eller bevissthet rundt språkstrukturer og når de skal brukes både rent lingvistisk og sosialt; ”... *a systemised body of mental representations underlying the learner’s language use, irrespective of whether those mental representations coincide with those of a mature native speaker of the language*” (Sharwood Smith 1994:14). Det er altså ikke bare den kunnskapen som stemmer overens med målspråket som regnes med, det er hele den mentale representasjonen til innlæreren, innlærerens eget bilde av målspråket. Det omfatter derfor også overgeneraliseringer og annen kreativ bruk av målspråksstrukturer, selv om de er avvikende i forhold til målspråket.

Oppbyggingen av kunnskap om strukturer kan beskrives som en tredelt prosess¹³ som består av en *introduksjon* av nye former, en *utvidelse* av kunnskap om bruken av formene, og en *restrukturering* av mellomspråkssystemet på grunnlag av de to første stegene (Ellis 1989:38/ jf. McLaughlin 1987).

Parallelt med oppbyggingen av kunnskap om strukturene, utvikler innlæreren *kontroll* over disse i faktisk bruk. Slik kontroll opparbeides gjennom en automatiseringsprosess, hvor strukturer krever mindre og mindre oppmerksomhetsressurser for å gjennomføres.

¹³ Innovation, Elaboration og revisjon

De ulike nivåene av prosesseringen av strukturene - kunnskap og kontroll - kan begge gi opphav til variasjon på ulike måter; ”Hence, hesitant linguistic behaviour may be attributed to a lack of relevant knowledge, but it also might be attributed to the actual possession of the relevant knowledge without fluent control over that knowledge ...” (Sharwood Smith 1994: 16).

Når det gjelder kunnskapsrelatert variasjon, hevdes det at mellomspråk kan være variabelt i forhold til hvor *analysert* det er, definert som ”the extent to which the learner has formed a mental representation of the structures which have been internalised.”(Ellis 1989:37).

Variasjon kan derfor være relatert til at man har en ufullstendig representasjon av målspråket. Men i forhold til definisjonen av kunnskap over, er det ikke helt rett å snakke om ”ufullstendig” kunnskap, det er bedre å si at det er en ikke-målspråkslik kunnskap, siden den ikke bare er ufullstendig, men også ”overfullstendig” i noen tilfeller.

Kontrollrelatert variasjon oppstår når ulike strukturer har oppnådd ulik grad av *automatisering*. Dette går altså ikke på opparbeiding av kunnskap, men på opparbeiding av kontroll i reell språkbruk. Trekk som ikke er fullt ut automatisert, vil være sårbare, og dette fører til en variabel anvendelse, som altså er avvikende til tross for kunnskapen innlæreren sitter inne med.

3.3.3 To hypoteser om variasjon i anvendelse av invertert ordstilling

Det første forslaget til forklaring i 3.2.2, som sier at variasjonen er resultat av tolkningsproblemer, forutsetter altså at variasjonen er *kunnskapsrelatert*. Hvis inversjonsregelen ikke er fullt ut analysert/avtolket, og innlærerens kunnskap om når det skal inverteres ikke stemmer overens med målspråket, kan dette føre til variasjon med veksling mellom invertert, underinvertert og overinvertert ordstilling.

Når en innlærer på en eller annen måte har blitt introdusert for inversjonsstrukturen, enten ved å oppdage den på egen hånd i innputt, eller ved å ha blitt fortalt i en undervisningssituasjon at subjekt og verb skal bytte plass når noe annet innleder setninga, kommer neste steg som er oppbygging av kunnskap om *når* denne skal brukes (utvidelse/elaboration). Innlæreren må da få kunnskap både om forskjellen mellom *subjekt og andre ledd*, og mellom *ikke-subjekt som krever invertert ordstilling og ikke-subjekt som ikke krever det*.

Det er kanskje først og fremst det siste som er problemet, ettersom setninger som innledes av subjekt ikke vil ha noe tredje ledd som kan skape V3-ordstilling. Ufullstendig kunnskap om når det skal inverteres og når det ikke skal det, er derimot en mer sannsynlig kilde til variasjon. Det hele blir et spørsmål om identifisering av hva som er inversjonskontekster og hva som ikke er det, og derav følger min første hypotese:

Identifiseringshypotesen

Variasjonen er kunnskapsrelatert, og den variable anvendelsen av invertert ordstilling er styrt av en ufullstendig identifisering av når det skal være invertert ordstilling og når det ikke skal være det.

Det andre forslaget til forklaring knytter variasjonen til kontroll og automatisering. Hvis inversjonsstrukturen ikke er 100% automatisert, og innlæreren har *lav kontroll* over den, kan dette føre til variasjon, enten i tråd med den lingvistiske kontekstens kapasitetsbelastning, eller forhold ved selve produksjonssituasjonen. Dette tilsvarer Hagens kapasitetsbetingede variasjon. Ulik grad av kontroll over regelen kan ha effekt på om en regel anvendes eller ikke, og neste hypotese følger av dette:

Kapasitetshypotesen

Variasjonen er kontrollrelatert, og den variable anvendelsen av invertert ordstilling er resultat av varierende kontroll over kunnskapen og varierende kapasitet til å anvende kunnskapen om invertert ordstilling.

Jeg ønsker å teste ut begge disse hypotesene for å se hvilken av dem som best kan predikere variasjonsmønstrene i materialet mitt best. I kapittel 4 tar jeg for meg identifiseringshypotesen, og i kapittel 5 kapasitetshypotesen. Hva jeg legger i kunnskap om og kontroll over inversjonsregelen mer spesifikt, vil jeg komme tilbake til der. Den vesentlige forskjellen mellom disse to hypotesene, er at mens kapasitetshypotesen forutsetter at innlærerne vet at og når regelen skal anvendes, så vil identifiseringshypotesen forutsetter at de ikke har en målspråklig kunnskap om alle kontekstene som krever inversjon.

3.4 Metode

Jeg vil nå gå over til undersøkelsen av de to hypotesene. Dette vil være en avgrenset undersøkelse, og kan selvfølgelig ikke si alt om fenomenet, men jeg synes det er en viktig vinkling som kan komme i forlengelsen av for eksempel Lunds funksjonelle forklaring på hvorfor det læres inversjon som faller sent på plass i deklorative setninger. Jeg kan selvfølgelig ikke si noe direkte om det læres sent eller ikke, ettersom dette er en tverrsnittstudie og man skal være forsiktig med å konkludere for mye om utvikling over tid på grunnlag av det, noe jeg vil si mer om i kapittel 6.

3.1.2 Datamateriale for variasjonsundersøkelsen

Undersøkelsene av variasjonsmønster vil begrense seg til å være av de variasjonstekstene som har minst to kontekster uten invertert ordstilling. Jeg ønsket i utgangspunktet å ha en jevn fordeling av antall tekster per morsmål, men dette lot seg ikke gjøre med det tekstene som er tilgjengelige i ASK-test, så det har ikke latt seg gjøre å kontrollere for morsmålsbakgrunn, noe som er en svakhet ved denne undersøkelsen.

Følgende tekster har minst to underinverterte, angitt med dokumentnavn i ASK, morsmål og antall inversjonskontekster.

Ref.ASK	Morsmål	Ant. oblig. kontekster	Inverterte kontekster	Ikke-inverterte kontekster	Overinverterte kontekster
0049	albansk	5	1	4	
0050	albansk	9	7	2	
0056	albansk	8	3	5	
0064	albansk	9	5	4	1
0054	albansk	6	3	3	
0057	polsk	8	5	3	
0061	polsk	6	2	4	
0037	russisk	8	6	2	
0068	russisk	5	2	3	
0070	russisk	9	5	4	3
0032	Serbokr	9	6	3	3
0082	Serbokr	6	2	2	
0085	Serbokr	12	6	6	5
0044	Serbokr	11	2	9	1
0088	Somali	7	2	5	
0089	Somali	6	4	2	
0098	Somali	4	1	3	
0044	Spansk	11	9	2	
0002	Spansk	5	3	2	1
0006	spansk	3	1	2	

Inversjonskontekstene i disse tekstene har blitt skrevet inn i tabeller i Excel, og den videre behandlingen har vært forskjellig for de to hypotesene, så den vil jeg komme nærmere inn på i kapittel 4 og 5.

3.2.3 Individuelle variasjonsmønstre

Felles for de to undersøkelsene er at jeg fokuserer på enkelttekster. Jeg ser på *individuelle variasjonsmønstre*, hvor hver inversjonskontekst kun blir sammenlignet med de andre i samme tekst, ikke med setningsmaterialet som helhet. Dette fordi utgangspunktet for hypotesene er at variasjonen er styrt av forhold innenfor den individuelle innlærerens kunnskap og kontroll, og i og med at innlærerne kan være på ulike stadier vil en kollektivundersøkelse kunne gi et feil bilde; *”Since we know that each individual may have their own idiosyncratic pathway to developing language competence, case studies are also able to show that the development of individual language acquirers may be different of that described of groups.”* (Seliger & Shoamy 1989:125)

I praksis har jeg identifisert hver kontekst etter kriterier som blir gitt for hver av hypotesene, og kontekstene blir sammenlignet med de andre i samme tekst. Tekstenes profiler blir regnet som enten positive eller negative i forhold til den hypotesen. I drøftingskapitlet foretar jeg en sammenligning av de ulike tekstenes profiler, både i forhold til hverandre, og i forhold til om de er positive i forhold til en, begge eller ingen av hypotesene.

Kapittel 4: IDENTIFISERINGSHYPOTESEN

Jeg vil i dette kapittelet undersøke den først av hypotesene mine, identifiseringshypotesen, som tar utgangspunkt i variasjon som et kunnskapsfenomen med opphav i avkodinga av målspåket. Den variable anvendelsen av invertert ordstilling blir da et resultat av en ikke-målspåkslik kunnskap om *når* det skal inverteres.

4.1 Opparbeiding av kunnskap om inversjonens distribusjon

Kunnskap om en struktur eller en regel dreier seg, som jeg var inne på i kapittel 3, ikke bare om å vite om den, og å være i stand til å utføre den rent mekanisk, det handler også om en dypere form for kunnskap om strukturens *distribusjon*. Når en innlærer får presentert, eller selv blir bevisst på, en struktur i målspåket, gjenstår fremdeles den krevende oppgaven med å finne ut *når* den skal anvendes. Den mest avgjørende kilden til dette er målspåksinnputt. Som Carroll skriver: *"Although SLA researchers agree about very little when it comes to explaining the how and why of L2 acquisition, one point on which it is consensus is that SLA requires exposure to the second language."* (Carroll: 2)

I en beregning gjort av Faarlund (1992), er så mye som 40 % av norske hovedsetninger innledet av et annet ledd enn subjektet (Faarlund 1992: 91), så det skulle ikke være noen tvil om at innlærerne faktisk møter inversjonsstrukturen i innputt. Men norsk innputt er allikevel ikke en så enkelt skog å kartlegge som det kan virke i en ren formell beskrivelse av inversjonsregelen. Innlæreren kan for eksempel møte begge de følgende strukturene i norsk innputt.

1. **Da gikk jeg** en tur
2. **Da jeg gikk** på tur, så jeg en skog.
3. **Hvor var jeg?**
4. Stedet **hvor jeg var**, var mørkt og uoversiktlig.

Her ser vi at ett og samme ord, *da/hvor*, i noen typer setninger (1,3) følges av invertert ordstilling, mens det i andre tilfeller (2,4) ikke gjør det. Når det gjelder setningsparet 3 og 4, har Karen Lund foreslått at spørsmålsinversjonen fanges opp tidligere fordi dens distribusjon er styrt av pragmatiske forhold, men hva kan innlæreren støtte seg på når det gjelder

forskjellen på 1 og 2? I tillegg til disse kan innlæreren komme over følgende motstridende konstruksjoner:

5. Heldigvis, **læreren min** bodde like ved.
6. Heldigvis **bodde læreren** min like ved.

Her står også det samme ordet både med og uten invertert ordstilling, denne gangen uten at betydningen endres. I Norsk Referansegrammatikk (NRG), behandles setninger som 5 som ekstraposisjon, hvor det foranstilte leddet regnes for å være utenfor setningen, og det er derfor ikke noen egentlig foranstilling som krever invertert ordstilling, slik som i setning 6. Det står skrevet i NRG: ”*I mange høve kan språkbrukaren velje om ein vil la det fyrste leddet komme i forfeltet eller i laust forfelt. Det er da ofte eit spørsmål om stil eller register.*” (Faarlund, Lie og Vannebo 1997: 907). Konsekvensen av dette er at man ender opp med enkelte ledd som kan etterfølges både av invertert og ikke-invertert ordstilling. I møte med innputt kan det tenkes at det er vanskelig for en innlærer å vite forskjellen på disse to setningskonstruksjonene, og det kan fremstå som uklart om ord som *heldigvis* og *dessverre* skal stå med eller uten invertert ordstilling. Å skille mellom når det skal være det ene og det andre, blir derfor et spørsmål om å skjønne når det er vanlig å bruke den ene og den andre stilvarianten.

Disse eksemplene på relativt vanlige norske setningskonstruksjoner, er ment for å illustrere hvor ugjennomtrengelig målspråksinnputten kan fremstå for en innlærer som bevisst eller ubevisst har satt ut for å søke kunnskap om inversjonsstrukturens distribusjon. Skal man invertere eller ikke etter *da*? Etter hvilke andre ledd skal man det og når skal man det ikke? Det er ikke lett å skulle danne seg en representasjon av en regel, når kilden kan være såpass tvetydig som disse setningene viser, dessuten finnes det en stor andel kontekster som krever inversjon, som innlæreren sjelden eller aldri støter på i innputt, for eksempel mer eller mindre unike leddsetningskonstruksjoner og ord og uttrykk som ikke er så frekvente som *da*, *i dag*, *så* osv.

Det er denne utfordringen som ligger til grunn for identifiseringshypotesen. I en prosess med oppbygging av kunnskap om distribusjon, vil innlæreren veksle mellom å produsere kontekster som de *har identifisert* som inversjonskrevende, og kontekster som *ikke er identifiserte*. I tillegg kan noen kontekster som ikke skal ha invertert ordstilling i norsk, bli *feilaktig identifisert* og dermed overinvertert.

Implisitt i identifiseringshypotesen ligger en påstand om at noen kontekster *lettere* lar seg identifisere som inversjonskrevende, enn andre kontekster. Jeg vil nå se nærmere på hva som kan være årsaken til dette, og hvilke signaler innlæreren kan tenkes å støtte seg på i identifiseringsprosessen.

4.2 Foranstilt ledd som signal om invertert ordstilling

Det som skiller hovedsetninger som krever invertert ordstilling, fra dem som ikke gjør det, er at de *ikke* innledes av subjekt. Det er på grunn av foranstillingen av et ikke-subjekt at man får et krav om invertert ordstilling, og en kan derfor si at et foranstilt ikke-subjekt fungerer som *signal* om at det skal være invertert ordstilling. Dette vil si at å identifisere en inversjonskontekst, er det samme som å identifisere et ledd som inversjonskrevende. Identifiseringshypotesen kan derfor spissformuleres til det følgende:

Den variable anvendelsen av invertert ordstilling i obligatoriske kontekster er et resultat av en ufullstendig identifisering av foranstilte ledd som inversjonskrevende.

Dette gjør undersøkelsen av identifiseringshypotesen til en undersøkelse av hva som skiller de foranstilte leddene i setninger med inversjon, fra setningene uten inversjon, og i tillegg hva som gjør at noen blir *feilaktig* identifisert som inversjonskrevende.

Et vesentlig spørsmål i denne sammenhengen blir hvilke egenskaper ved de foranstilte leddene som kan tenkes å være opphavet til variasjonen, hvilke egenskaper som er knyttet til identifiseringen. Alle ord og konstituenten har både strukturelle, syntaktiske, pragmatiske, semantiske og fonologiske egenskaper, men det vil ikke være uvesentlig hvilke egenskaper man tar utgangspunkt i når man skal undersøke identifisering som påvirkningsfaktor. Tidligere undersøkelser av det foranstilte leddets påvirkning på inversjonsmestring og variasjon, har så vidt jeg kan se hovedsaklig fokusert på *frasetype/syntaktisk funksjon* (Bolander 1988), og *strukturell kompleksitet* (Brautaset 1996, Vikøy 2004 og Hagen 1992). Bolanders inndeling er i objekt, adverb, preposisjonsfrase, predikativ og leddsetning, altså den type ledd en kan forvente å finne i foranstilt posisjon beskrevet med velkjent grammatisk terminologi. Både Brautaset og Vikøy tar utgangspunkt i leddets lengde og kompleksitet, for å teste ut om det er prosesseringstyngde som styrer variasjonen. Hagen vektlegger også prosesseringstyngde, men ikke bare med hensyn til kompleksitet. Han samstiller både syntaktiske, strukturelle og pragmatiske egenskaper; kort/lang, tema/rema, frekvens og

tekstkoherens, og hans undersøkelse fremstår derfor som mer nyansert og innsiktsgivende, selv om linken mellom disse egenskapene og prosesseringstygde forblir noe uklar, som nevnt i kapittel 2.

Det som kan være problematisk med undersøkelsene til Bolander, Brautaset og Vikøy, er at de benytter velkjente grammatiske kategorier for å beskrive hvilke ledd som ofte og sjelden forekommer sammen med invertert og ikke-invertert ordstilling hos innlærerne, men de sier egentlig lite om *hvorfor* noen ledd står med og noen uten invertert ordstilling. De beskriver altså konteksten rundt variasjon, men sier ikke noe om hvorfor det er variasjon. Selv om vi som fagpersoner er vant til å bruke disse kategoriene, er det, som blant annet Lund har poengtert, ikke nødvendigvis de samme kategoriene innlærerne opererer med (Lund 1997: 157). De kan være gode til å beskrive, men mangler en *forklarende* dimensjon. Derfor kan det være viktig å søke egenskaper ved leddene som er kognitivt/psykologisk reelle for innlærerne, med mindre man kan knytte de tradisjonelle grammatiske kategoriene direkte til slike forhold.

4.2.1 Prosessering av invertert innputt

Mens de fleste andre undersøkelser har tatt utgangspunkt i det foranstilte leddets kapasitetsbelastning, noe jeg selv vil komme tilbake til i kapittel 5 når jeg ser nærmere på kapasitetshypotesen, er jeg i denne omgang mer opptatt av leddets rolle i oppbyggingen av kunnskap om distribusjon av invertert ordstilling. Det som da er viktig, er å finne ut hvilke(n) egenskap(er) som kan tenkes å ha påvirkning på i hvilken grad uttrykk lar seg identifisere som inversjonskrevende. Som sagt er den mest tilgjengelige kilden til inversjonsstrukturens distribusjon innputt. Det er derfor mer interessant for meg å fokusere på de foranstilte leddenes innputtrelaterte egenskaper, fremfor de mer språkformelle egenskapene. I denne sammenhengen er det avgjørende hvorvidt det finnes positiv evidens for en struktur i det språket innlærerne møter. Det vil si at det finnes eksempler, eller bevis, på en struktur i innputt, som innlæreren kan bruke til å bygge opp sin representasjon av målspråket. Bergreen og Tenfjord skriver :

”I den innputten språkinnlæreren får fra omgivelsene ved å delta i ordinær kommunikasjon, fins ingen direkte tilgjengelige regler, generaliseringer eller abstraksjoner som forteller hvordan målspråket er konstruert. Reglene må rekonstrueres av det språklærende individet på basis av ”bevis”, indikasjoner, nøkler som fins i innputt – ...” (Bergreen og Tenfjord 1999: 163)

Videre skriver de at ”*Positiv evidens regnes som det basale slutningsgrunnlaget ved alle former for språklæring.*”, og jeg vil gå ut fra at innlæreren i identifiseringsprosessen benytter seg av slik positiv evidens. Men det er ikke så enkelt som at alt som forekommer i innputt registreres og tas opp som kunnskap av innlæreren, derav Chaudrons skille mellom *innputt* og *inntak* (jf. 3.3).

Enkelte teorier hevder at positiv evidens ikke er nok, men at man også trenger negativ evidens, og det finnes dessuten teorier som hevder at positiv evidens underdeterminerer den kompetansen som språkinnlærere oppnår, de kan umulig oppnå en fullverdig førstespråkskompetanse uten noe mer enn bare innputt, noe medfødt (Bergreen og Tenfjord 1999: 164). Denne problemstillingen omtales ofte som *The poverty of the stimulus*, et hovedargument for at mennesker må ha en egenspråkmodul, en universell grammatikk (UG), uten at jeg skal gå nærmere inn på dette her.

I teorier om språklig prosessering er det vanlig å trekke frem to faktorer som letter den reseptive prosesseringen av innputt, om høyner sjansen for at det blir til faktisk inntak. *Saliens* letter prosesseringen ved at noen trekk er mer tydelige og iørefallende enn andre, enten på grunn av plassering i setninga eller på grunn av form, og de blir dermed lagt lettere merke til (Bergreen og Tenfjord 1999: 95). Spesielt enheter initialt i en setning blir lettere lagt merke til, noe man mener har å gjøre med at initiale enheter lettere overføres til langtidsmindet fordi de prosesseres på et tidspunkt der innlæreren har overskudd av kapasitet (Bergreen og Tenfjord 1999: 95). I tillegg til *saliens* spiller *frekvens* i innputt en viktig rolle, ved at ord som forekommer ofte lettere blir lagt merke til. Men allikevel er det bare når frekvensen kommer over et visst nivå at det har positiv effekt på prosesseringen; ”*Innputteneheter må forekomme med en viss hyppighet for å bli registrert, utover dette har relativ frekvens liten betydning.*” (Bergreen og Tenfjord 1999:96). Dette betyr at det ikke lar seg gjøre å basere noen som helst form for prediksjon om relative læringsvansker basert på den relative frekvensen til ord, man kan bare predikere et enten/eller-forhold, hvor frekvensen påvirker læring av de aller hyppigste ordene, men ikke andre. En ulempe med frekvente ord, er at de ofte er mindre saliente, noe som til en viss grad kan begrense den positive effekten frekvensen har. Funksjonsord er for eksempel veldig frekvente, men korte og lite saliente (Bergreen og Tenfjord 1999: 96).

Sammenhengen mellom frekvens og læring gjelder ikke bare enkeltord, men kan også gjelde større enheter og uttrykk. Innen kognitiv grammatikk ser en gjerne på hvordan deler av innputt læres som helhet, som hele konstruksjoner, og ikke som gjennomanalyserte enheter. Ideen er at en stor andel av språklige enheter forekommer i relativt frekvente kollokasjoner, og at det ofte gir mer mening å fokusere på hele kollokasjoner, *chunks/constructions*, enn på mindre bestanddeler (Ellis 2003: 66).

Om man skal se på frekvens i forhold til inversjonskontekster, er det altså ikke den generelle frekvensen til det enkelte foranstilte ledd som er avgjørende, men frekvensen av **kollokasjonen** 'foranstilt ledd + invertert ordstilling' i innputt. Hvis identifiseringen av ledd som inversjonskrevende påvirkes av frekvens, kan vi forvente at frekvensen av et bestemt ledd sammen med invertert ordstilling kan ha betydning for prosesseringen og identifiseringen av leddet som inversjonskrevende. Dette vil si at ledd som er *frekvente* i foranstilt posisjon, som altså ofte opptrer sammen med invertert ordstilling, *lettere* vil la seg identifisere som inversjonskrevende, enn ledd som sjelden eller aldri blir observert sammen med invertert ordstilling.

Slike kollokasjoner trenger ikke være helt "faste" (hvor alle leddene er de eksakt samme) deler av kollokasjonen kan være variabel (Ellis 2003:70), og i dette tilfellet er det altså ikke snakk om *spesifikke* finite verb og subjekter, men en kollokasjon av et *spesifikt* foranstilt ledd og *generelle* finite verb og subjekter.

4.2.2 Identifiseringsgrunnlag

I denne sammenhengen kan vi skape et hypotetisk skille mellom identifisering som kan skje på grunnlag av at innlæreren *ofte nok* har registrert et ord/uttrykk i foranstilt posisjon sammen med invertert ordstilling; *innputtbasert identifisering*, og identifisering som må skje som rene slutninger, i tilfeller hvor innlæreren ikke kan støtte seg på innputt; *slutningsbasert identifisering*. Det siste gjelder både ord/uttrykk som ikke er særlig frekvente i foranstilt posisjon, eller som er direkte ugrammatiske, og dessuten lengre fraser som kan sies å være egenproduserte og *unike* for den gitte sammenhengen. Å identifisere disse som inversjonskrevende, er av åpenbare grunner en mer krevende prosess, fordi man ikke har frekvent nok evidens i innputt, eller ikke evidens i det hele tatt.

I kartleggingen av inversjonsstrukturens distribusjon, og spesielt i en produksjonssammenheng, vil innlæreren måtte basere seg på ett av disse to identifiseringsgrunnlagene. Mens de mest frekvente kollokasjonene kan hentes ut av innputt på en mer helfraseaktig måte, kanskje direkte uanalysert, krever de mindre frekvente og unike kollokasjonene en dypere innsikt og dypere prosesseringskapasitet for å bli identifisert. Det krever overføring av mer abstrakte kunnskaper om når invertert ordstilling skal brukes.

På dette tidspunkt kan jeg altså sette frem følgende påstand:

- **Identifiseringen av foranstilte ledd som inversjonskrevende henger sammen med leddets innputtfrekvens i initial posisjon sammen med invertert ordstilling.**
- **Språklige uttrykk som er høyfrekvente i initial posisjon sammen med invertert ordstilling, vil kunne være gjenstand for innputtbasert identifisering, og er derfor lettere å identifisere som inversjonskrevende, enn de som er mindre frekvente eller aldri forekommer med invertert ordstilling i innputt.**

Skillet mellom innputtbasert og slutningsbasert er selvfølgelig ikke klart og direkte påvisbart. Det vil være en gråsoner mellom de aller mest frekvente uttrykkene, og de som man med rimelig stor sikkerhet kan vite at innlærerne aldri har møtt i innputt. I denne gråsonen ligger vanlige, men mindre frekvente uttrykk, som innlæreren sannsynligvis har evidens for, men kanskje ikke frekvent nok til å kunne dra nytte av innputtbasert identifisering. Det er allikevel først og fremst en metodisk gråsoner, som omfatter de uttrykkene jeg i analysen ikke kan være like sikker på. Det vil alltid være individuelle grenseopptrekninger mellom innputtbasert og slutningsbasert, styrt både av nivå og personlige erfaringer med innputt. Dette gir videre opphav til en abstrakt frekvensskala som uttrykker enhver type uttrykks status i innputt, og i hvor stor grad dette er positivt eller negativt for identifiseringen.



4.2.3 Syntaktisk uklarhet

Det er allikevel ikke så enkelt som dette. Mens frekvens er en faktor som kan lette prosesseringen av strukturer, finnes det også forhold som kan gjøre den vanskeligere. Som jeg skrev i 4.1 kan innputt fremstå som forvirrende hvis man ikke har et tydelig bilde av hvilke ledd som krever inversjon og hvilke som ikke skal ha det. Et forhold som flere andre studier av inversjon i mellomspråk har brakt frem, er homonymien mellom enkelte subjunksjoner og ledd i foranstilt posisjon, spesielt når det gjelder 'så' som kan være både adverbial og subjunksjon.

7. Så **gikk jeg** inn i skogen.

8. Men det var mørkt, **så jeg gikk** ut igjen.

Andre eksempler på slike homonymer er 'når' og andre spørreord, som står med invertert ordstilling i spørsmål, men uten i interrogative leddsetninger. I tillegg kan noen leddsetninger uten konjunksjon gi tvetydig innputt.

9. Da **gikk jeg** inn.

10. Da **jeg gikk** inn, var jeg redd.

Tvetydig innputt av denne typen kan gjøre ordene syntaktisk uklare for innlæreren; ved at de får en innputt som tilsier både at de krever invertert ordstilling og at de ikke gjør det, noe som kan tenkes å vanskeliggjøre identifiseringen. Man kan vente tre typer av ledd mht syntaktisk klarhet.

Syntaktisk klare ledd:

- Subjunksjoner og konjunksjoner som alltid står uten invertert ordstilling
- Foranstilte ledd som alltid står med invertert ordstilling

Syntaktisk uklare ledd:

- Homonyme subjunksjoner og adverbialer/spørreord (*så, da, når, hv-ord*)
- Foranstilte ledd som kan stå i ekstraposisjon (*heldigvis, dessverre etc.*)
- Ledd som er modifiserte med leddsetning uten subjunksjon ('*mannen jeg kjenner*', '*året jeg kom til Norge*')

Det er riktignok ikke en fullstendig tvetydighet mellom disse ordene, når en tar ytringsforhold i betraktning. I en ytring vil disse ordene ha et ulikt trykk. Som enkle foranstilte ledd vil de ha mer trykk enn når de er denotative eller er en subjunksjon. Jeg vil allikevel anta at likheten er stor nok, eventuelt at forskjellene er for små, til at det allikevel kan påvirke identifiseringen.

På grunn av dette vil jeg nyansere mine tidligere påstander om frekvens. Disse gjelder altså bare hvis uttrykkene ellers er syntaktisk (eller semantisk) entydige og klare. Hvis de ikke er det, er ikke høy frekvens nødvendigvis tilstrekkelig for identifisering.

- **Ledd som er høyfrekvente i initial posisjon sammen med invertert ordstilling, men som kan oppfattes som syntaktisk (eller semantisk) uklare, vil være vanskeligere å identifisere enn ledd som har en mer gjennomskuelig bruk.**

4.2.4 Prediksjoner for materialet

Hovedpåstanden til identifiseringshypotesen er altså at frekvente, syntaktisk klare ledd vil være lettere å identifisere som inversjonskrevende, enn mindre frekvente og syntaktisk uklare ledd. Hvordan disse forholdene er operasjonalisert for å muliggjøre uttesting av hypotesen kommer jeg tilbake til i neste kapittel – både operasjonaliseringer av innputtbasert og slutningsbasert, og en undersøkelse av hva som er de mest frekvente foranstilte ordene i norsk.

4.3 Operasjonalisering og gjennomføring

Jeg vil nå komme inn på hvordan undersøkelsen har blitt gjennomført, en operasjonalisering av frekvenskategoriene og hvilke tendenser materialet må vise for at hypotesen ikke skal bli avkreftet.

4.3.1 Gjennomføring

Hoveddelen av undersøkelsen har gått på å identifisere alle foranstilte ledd i variasjonstekstene etter hvor frekvent de regnes for å være i innputt i kollokasjon med invertert ordstilling, i kategoriene *høyfrekvent*, *mellomfrekvent* eller *unik*. Denne inndelingen skjer på grunnlag av en empirisk undersøkelse som jeg kommer tilbake til lenger ned. Det er

altså bare de obligatoriske kontekstene som er med i denne delen av undersøkelsen, de overinverterte kontekstene undersøkes for seg (se 4.3.4).

Tildelingen av frekvenskategorier har blitt ført inn i et skjema som er laget for å vise hver enkelt teksts fordeling av invertert/ikke-invertert i forhold til frekvenskategoriene.

	Høyfrekvente	Mellomfrekvente	Unike
Invertert			
Ikke-invertert			

Figur 4.1: frekvensfordelingskjema.

Hver kontekst har blitt markert med et kryss, **X**, mens de tvetydige *da* og *så* har understrekning, **X**, og de som kan tolkes som ekstraposisjoner er markert med (**X**). Når alle kontekstene er ført inn, kan det leses av om tekstens helhetlige *profil* støtter eller ikke støtter hypotesen.

4.3.2 Prediksjoner og profiler

For at identifiseringshypotesen skal støttes av materialet, må de inverterte kontekstene i hver enkelt tekst være *mer frekvente* enn de som ikke er inverterte. Dette vil si at om de unike kontekstene er inverterte, så skal også de mellomfrekvente og de høyfrekvente være det, og om de høyfrekvente ikke er inverterte, så skal ikke de unike være det heller. Rent visuelt i skjemaet kommer dette tilsyne med at de øverste kryssene, de inverterte kontekstene, ligger til venstre for de nederste kryssene, som er de ikke-inverterte. Unntak her er tvetydige kontekster, som kan tenkes å ikke ha invertert ordstilling, til tross for høy frekvens.

Hver enkelt teksts skjema utgjør en *variasjonsprofil*, som enten støtter eller ikke støtter hypotesen. Hver tekst vil ha én av tre profiler:

1. **Positiv profil:** Fordelingen stemmer overens med prediksjonene, de inverterte *er* mer frekvente enn de som ikke er inverterte.
2. **Negativ profil:** Fordelingen er motsatt av prediksjonene; de inverterte er *mindre* frekvente enn de som ikke er inverterte.
3. **Jevn profil:** fordelingen av inverterte og ikke-inverterte kontekster er jevn over alle frekvenskategorier.

En mer spesifikk operasjonalisering av de ulike profiltypene:

a) Positiv profil (+)

Dette er tekster hvor materialet innfrir forventningene. De inverterte setningene (øverste rad) har foranstilte ledd som ligger til venstre for de ikke-inverterte (nederste rad) i skjemaet.

	Høyfrekvent	mellomfrekvent	Unik
Invertert	XX	XX	
Ikke-invertert			XX

Figur 4.2: eksempel på positiv profil som stemmer med en streng fortolkning av prediksjonene.

I strengeste fortolkning av hypotesene skal det ikke finnes noen variasjon på kategorinivå her; altså både inverterte og ikke-inverterte innen én og samme *frekvenskategori*. Men da dette er lite sannsynlig å finne, opererer jeg også med en mildere fortolkning, som også regner profiler som har variasjon innen én kategori som positive.

	Høyfrekvent	mellomfrekvent	Unik
Invertert	XX	XX	XX
Ikke-invertert			XX

Figur 4.3: eksempel på en positiv profil som stemmer overens med en mild fortolkning av prediksjonene.

Når denne variasjonen er blant de høyfrekvente eller de unike, er dette greit, så lenge de øvrige kategoriene ikke har variasjon og ellers stemmer overens med hypotesen, altså at de inverterte ligger til venstre for de ikke-inverterte. Det vil si at variasjon innen de høyfrekvente bare godtas dersom de mellomfrekvente og de unike *ikke er inverterte*. Variasjon innen de unike godtas bare om de mellomfrekvente og de høyfrekvente er *inverterte*. Om variasjonen er innen de mellomfrekvente, kan det være nødvendig å differensiere disse innad i forhold til frekvens, om dette lar seg gjøre.

b) Negative profil (--):

Dette er tekster hvor materialet viser seg motsatt av prediksjonene; de inverterte ligger til høyre og er dermed mindre frekvente, enn de som ikke er inverterte. Slike profiler vil antyde at implikasjonsretningen går motsatt vei av det identifiseringshypotesen predikerer, at det er de unike frasene som identifiseres lettest, og slike profiler gir derfor ingen støtte til hypotesen, men utfordrer den. Også her kan det være variasjon innen én kategori.

	Høyfrekvent	mellomfrekvent	Unik
Invertert		XX	XX
Ikke-invertert	XX		

Figur 4.4: eksempel på negativ profil

c) Jevne profiler (Ø)

Dette er tekster som har en jevn spredning av inverterte og ikke-inverterte setninger over to eller alle frekvenskategoriene.

	Høyfrekvent	mellomfrekvent	Unik
Invertert	XX	XX	XX
Ikke-invertert	XX		XX

Figur 4.5: eksempel på jevn profil

d) Oppgradering av profiler pga tvetydige kontekster

Selv om profiler viser seg å være negative eller jevne etter frekvenskategoriseringen, kan det tenkes at noen av dem stiller seg annerledes i forhold til hypotesen om en tar hensyn til de tvetydige leddene. På grunn av dette har jeg i neste omgang undersøkt om noen av motprofilene allikevel ser ut til å være positive til hypotesen, om en tar hensyn til at tvetydige ledd kan være vanskeligere å identifisere enn entydige ledd.

4.3.3 Frekvenskategorier

Hva som regnes som høyfrekvente, mellomfrekvente og unike foranstilte ledd, må være klart definert, så her følger en operasjonalisering av de tre frekvenskategoriene.

a) Høyfrekvente ledd

For å finne ut hva som er de mest frekvente foranstilte leddene i norsk innputt, har jeg gjort en empirisk undersøkelse av norsk skriftspråk ved hjelp av *Oslo-korpuset av taggede norske teksters*¹⁴ utvalg av avistekster og skjønnlitteratur¹⁵. Målet har vært å kartlegge frekvensen til ulike typer ord/uttrykk som står i kollokasjon med invertert ordstilling, helt ned på

¹⁴ <http://www.tekstlab.uio.no/norsk/bokmaal/>

¹⁵ Det er også mulig å inkludere sakprosa i søket, men jeg har ønsket å utelate disse tekstene da språkbruken ligger et stykke unna den innputten en kan forvente at innlærerne kommer over.

leksemnivå, og jeg har derfor gjort et søk etter alle setninger der subjektet står etter det finite verbet. Jeg har altså spurt hva som står foran, når subjektet står etter det finite verbet¹⁶?

En opptelling blant ca. 1400 tilfeldig utvalgte kontekster viste at noen ord/uttrykk har merkbart høyere frekvens i foranstilt posisjon enn andre. For å få bekreftet at frekvensforholdene stemmer også i et større materiale, har jeg kontrollsjekket frekvensen i hele korpuset, ved å søke på det bestemte ordet foran finitt verb + subjekt.

De mest frekvente ordene i foranstilt posisjon viste seg å være: *nå, så, da, her, der* og *derfor*, og derfor er det disse ordene jeg regner som høyfrekvente i min undersøkelse. To av disse, *da* og *så*, er tvetydige kontekster.

En kan selvfølgelig ikke vite helt sikkert at dette avspeiler hver enkelt innlærers erfaring med innputt. Et problem er at jeg baserer meg på et korpus av skriftlige tekster, mens en innputtundersøkelse kanskje heller skulle tatt utgangspunkt i et muntlig materiale. Allikevel, en kjensgjerning når det gjelder ordfrekvens er at de aller mest frekvente ord og uttrykk, da også i foranstilt posisjon, har en tendens til å være de samme uavhengig av om det er muntlig eller skriftlig språk. Dette er også grunnen til at jeg begrenser høyfrekventkategorien til bare 6 ord, fordi disse har såpass mye høyere frekvens enn de neste ordene på listen.

b) Unike ledd

Hva som skal regnes som unike ledd, kan ikke fastslås empirisk, men det er nødvendig med klare kriterier. En generell betingelse for denne kategorien, er at leddene skal være konstruksjoner som man med stor sannsynlighet kan vite at innlæreren *ikke* har møtt i innputt sammen med invertert ordstilling. Dette vil si at det er grunn til å anta at det er snakk om *egenproduserte* konstruksjoner, tydelig konstruert i forhold til sammenhengen de står i, eller ledd som rett og slett er ugrammatiske i foranstilt posisjon, for eksempel utenlandske ord/uttrykk og norske ord/uttrykk som ikke kan foranstilles.

c) Mellomfrekvente ledd

I denne kategorien ligger alt fra relativt frekvente ledd, til mer uvanlige konstruksjoner, som allikevel ikke kan sies å være unike. Denne kategorien er derfor definert som *alt som ikke er*

¹⁶ I dette søket ble jeg nødt å plukke ut spørresetninger og vilkårssetninger med invertert ordstilling manuelt.

høyfrekvent eller unikt. Det vil altså være alt fra relativt vanlige ord eller faste uttrykk, til infrekvente, komplekse fraser som allikevel ikke er unike, for eksempel *tre måneder etter*. Om det blir nødvendig, i tilfelle en tekst mangler høyfrekvente eller unike kollokasjoner, vil jeg gjøre en videre differensiering av leddene i denne kategorien, ved å sjekke opp mot frekvensen i Oslo-korpuset.

En type ledd som også regnes med her, til tross for at de kan fremstå som unike, er holofrastiske leddsetningskonstruksjoner. Selv om leddsetninger kan regnes som en unik konstruksjon, siden de som regel har et unikt innhold, kan det finnes enkelte slike setninger som allikevel ikke nødvendigvis er helt egenproduserte. Noen typer leddsetninger kan være holofrastiske, kanskje fordi de er relativt frekvente både generelt eller i en innlærersituasjon. Dette kan være setninger som *'Da jeg kom til Norge', 'da jeg var liten', 'da jeg bodde i Russland'* etc. Det kan tenkes at slike setninger er inverterte, til tross for at de ikke er så frekvente i innputt, rett og slett fordi det er setninger de med stor sannsynlighet har vært borti på kurs eller ellers. Det er ikke analyserte enheter, og derfor er anvendelsen av invertert ordstilling mer innputtbasert enn slutningsbasert. De er allikevel ikke høyfrekvente, og derfor blir de regnet som mellomfrekvente.

d) Tvetydige ledd

Jeg har regnet to typer tvetydige ledd. De som er tvetydige fordi de kan stå både med og uten invertert ordstilling, men med ulik setningsbetydning, som *så* og *da*. Spesielt *så* er både semantisk og syntaktisk komplisert ord å bruke. Den andre gruppen er de leddene som kan stå i ekstraposisjon, og altså kan stå både med og uten invertert ordstilling, uten at betydningen endres. Denne siste typen er tvetydig i innputt, i tillegg til at den skaper tolkningsproblemer når jeg skal lese av materialet, fordi jeg ikke vet helt sikkert om det er en underinversjon eller en bevisst setningskonstruksjon fra innlærerens side.

4.3.4 Andre sjekkpunkt

a) Fri variasjon på leksemnivå

Mens variasjon på kategorinivå lett kan la seg forklare ved at kategoriene er vide og at identifiseringsprosessen går gradvis, vil variasjon på leksemnivå, altså både invertert og ikkeinvertert etter ett og samme ord/uttrykk, være en utfordring for hypotesen. Om variasjon skyldes identifisering av kontekster, skal det strengt tatt ikke finnes enkeltord eller uttrykk som både er identifisert og ikke-identifisert i samme tekst. De positive profilene blir derfor

svekket av en slik variasjon, med mindre leddet er tvetydig, noe som kan gjøre variasjonen til et resultat av bevisst valg (ekstraponering) eller usikkerhet på grunn tvetydig innputt.

b) Overinversjon

En undersøkelse av obligatoriske kontekster er, som jeg sa tidligere, metodisk svak om den ikke også tar med overgeneraliseringer. Men som jeg har vært inne på, ønsker jeg å holde obligatoriske kontekster og overgeneraliserte kontekster fra hverandre. Grunnen til at de overgeneraliserte kontekstene ikke er tatt med i frekvenstildelingen, er at denne tildelingen kun går på frekvensen av *foranstilte ledd + invertert ordstilling* i innputt, og dette er selvfølgelig ikke relevant for disse kontekstene ettersom de aldri har invertert ordstilling.

I en identifiseringsammenheng vil overinversjonene først og fremst bare fungere som et tegn på feilidentifisering. Innlærerens utfordring er å identifisere obligatoriske kontekster på grunnlag av innputt, og overgeneralisering kan regnes som et bieffekt av dette, når identifiseringen er avvikende. Det hadde selvfølgelig vært interessant å måle frekvensen av de tvetydige ordene som henholdsvis foranstilte adverbial og subjunksjon, men dette har jeg valgt å se bort fra denne gangen.

Overgeneralisering skjer gjerne i leddsetninger, etter subjunksjon, og i hovedsetninger etter konjunksjon. Det tredje alternativ kan være overgeneralisering etter venstre-diksløkering av et lengre subjekt, som i *'*Det å være utlending i Norge, er det vanskelig'*, men jeg har ikke funnet noen slike i materialet mitt.

Overgeneraliserte kontekster blir først og fremst interessante på to punkter: Det første er om det bare er overgeneralisering til leddsetninger som innledes av tvetydige subjunksjoner, eller om det også forekommer etter ledd som ikke er tvetydige, som aldri står med inversjon, for eksempel *hvis* og *fordi*. Dette kan si noe om hvorvidt det bare er en type innputtbasert overgeneralisering, eventuelt en ren overføring av konstruksjoner som forekommer i innputt, eller om det også er resultat av en slutningsprosess.

Den andre måten overinversjon er interessant på, er om de overgeneraliserte kontekstene som er tvetydige/innputtbaserte, er i fri variasjon med enten obligatoriske kontekster eller spørsmål. Dette kan vise om det er *fri variasjon* mellom ulike *så* og *da*-kontekster. Er det slik at innlæreren bruker én ordstilling, enten invertert eller ikke-invertert, etter samtlige *da* og/eller *så*? Eller er det en eller annen form for fri variasjon, en variabel innlærerregel, hvor

noen er inverterte og noen ikke er det? Her er det selvfølgelig vekslingen mellom tvetydige overgeneraliserte kontekster og obligatoriske kontekster som er mest interessant.

c) Utelukking av kompleksitet

Et grunnleggende skille som jeg bør ha kontroll over, er skillet mellom påvirkningsfaktorer som har med *identifisering/kunnskap* å gjøre, og faktorer som har med *kontroll/kapasitet* å gjøre. Det er viktig for å formulere betingelsene til hypotesen slik at funnene mine faktisk tilsier at det er identifisering som er årsaken til variasjonen, og ikke andre faktorer som kan gi samme utslag. Det dreier seg altså om kontroll av andre variabler, i dette tilfellet *det foranstilte leddets kompleksitet*. Kompleksitetsforhold utgjør et problem for undersøkelsen fordi frekvensskalaen jeg har brukt også til en stor grad avspeiler leddenes lengde/kompleksitet (Berggreen og Tenfjord 1999: 235). Frekvente ord er ofte korte, og de unike, egenproduserte konstruksjonene må nødvendigvis komme over en viss lengde for å kunne kalles unike. Spesifikke leddsetninger er sjelden høyfrekvente, og kortere ord og faste uttrykk er ikke unike. Det er derfor viktig å kunne finne eksempler på foranstilte ledd som bryter med denne ikonisiteten, for å støtte og styrke de positive profilene. Først da blir identifiseringsforhold isolert som årsak, og kompleksitet som påvirkningsfaktor kan utelukkes. En utelukking som dette er strengt tatt bare nødvendig for de positive profilene. I kapittel 5 vil jeg uansett foreta en undersøkelse av kompleksitet.

4.4 Resultater

4.4.1 Resultat etter frekvensfordeling

Implikasjonstabellen under viser fordelingen av kontekster for hver tekst. Plusstegn (+) betyr bare inverterte kontekster i den gitte kategorien, minustegn (--) bare ikke-inverterte, og V betyr variasjon innen kategorien. Tomt felt betyr at det ikke er noen forekomster innen kategorien.

Hovedkravet for en positiv profil er at de inverterte kontekstene (+) ligger til *venstre* for de ikke-inverterte (-). I tillegg har jeg åpnet for variasjon innen en kategori, så lenge kravet om at + ligger til venstre for - er oppfylt. De typiske positive profilene er da: + + V , + V - , eller V - - . De typiske negative profilene er : - - V , - V + , eller V + + , og profilen er jevn (Ø) om det er variasjon innen to eller tre kategorier.

Tekst	HF	MF	UNIK	Profil
2	V		V	Ø
6	o	V		Ø
32	+	+	V	+
37	V	+	+	--
44	+	V		+
49	V		-	+
50	+	+	V	+
54	V	V		Ø
56	+	-	V	Ø
57		V	+	--
61		V	-	+
64	V	V	-	Ø
68	V	V	-	Ø
70	+	V	-	+
82	+	V	-	+
85		V	V	Ø
88	-	V	-	Ø
89		V	V	Ø
98		-	+	--

Profiler	antall
Positive	7
Jevne	8
Negative	4

Det er altså 7 profiler som støtter hypotesen om en bare tar frekvensforhold i betraktning, mens 3 er negative, og 8 jevne. De jevne profilene har en jevn fordeling, med variasjon innen flere kategorier og derfor ingen tydelige implikasjonsmønster som tyder på at det er frekvens som styrer vekslingen mellom inverterte og ikke-inverterte kontekster. Disse avkrefter ikke hypotesen, men kan antyde at det ikke er frekvens som er den avgjørende faktorene. Tekst 56 og 88 har ikke det typiske jevne fordelingen, men regnes som jevne fordi de verken har positiv eller negativ fordeling av kontekster.

4.4.2 Eksempler

a) Jevn profil

Tekst 2 har følgende frekvensprofil:

TEKST 2	Høyfrekvent	Mellomfrekvent	Unik
Inverterte	X		XX
Ikke-inverterte	X		X

Her ser vi at det er en jevn fordeling av frekvente og unike foranstilte ledd i inverterte og ikke-inverterte setninger. Det er variasjon innen mer enn én frekvenskategori, og er derfor en jevn profil.

De foranstilte leddene i denne teksten er disse:

Ikke-invertert

Also
Nå

Invertert

. etter vi flytter, fra hjemlang
; fordi hvis man mistet tradjsoner
har (=her)

De ikke-inverterte utgjør både det unike leddet *also*, som jeg med sikkerhet kan vite aldri har forekommet med inversjon i innputt siden det ikke er et norsk ord; og *nå*, som er det aller mest frekvente foranstilte leddet i norsk. De inverterte setningene har tilsvarende fordeling; to unike foranstilte leddsetninger, og det høyfrekvente *her* (i denne teksten med ortografisk feil: *har*.) For denne innlæreren er det altså ingenting som tilsier at det er inversjonskollokasjonens frekvens i innputt som styrer variasjonen.

b) Negativ profil

Tekst 98 går motsatt vei av hypotesens prediksjoner, da den ene inverterte setningen er *mindre* frekvent enn de ikke-inverterte. Denne profilen utgjør derfor en utfordring for hypotesen.

Tekst 98	Høyfrekvent	Mellomfrekvent	Unik
Inverterte			X
Ikke-inverterte		XXX	

De foranstilte leddene er:

Ikke-invertert

Først
... etterpå
I denne forbindelse

Invertert

;På grunn av dårlig være i Malmø

c) Positiv profil

En tekst som stemmer overens med prediksjonene er tekst 32, som har følgende profil:

Tekst 32	Høyfrekvent	Mellomfrekvent	Unik
Inverterte		XXX	XXX
Ikke-inverterte			XXX

Her ser vi at alle kontekstene som ikke er inverterte er *unike*, mens de mer frekvente kontekstene er inverterte. Det er variasjon innen den unike kategorien.

De ulike foranstilte leddene er:

Ikke-invertert

At bli en sjåføren på tunge kjøretøy
, også
I STED om sommer når er jazz festivalen

Invertert

For meg selv
Da kjørte jeg første gang lastebil
Da fikk jeg min første jobb
I dag
I hoved sak
I dag

4.4.3 Profiler etter å ha tatt tvetydighet i betraktning

Frekvens var altså ikke den eneste avgjørende faktoren for identifiseringen av inversjonskrevende kontekster. *Syntaktisk tvetydighet* spiller også en viktig rolle, og det kan være en rimelig forklaring på hvorfor ledd som er frekvente og ellers forventes å være inverterte, allikevel ikke er det. I neste omgang har jeg altså undersøkt om noen av profilene allikevel er positive når en tar hensyn til at tvetydige ledd kan gjøre identifiseringen vanskeligere. To av tekstene, 68 og 89, endrer profil etter å ha tatt hensyn til tvetydige kontekster. (Leddene det gjelder er uthevet i skjemaet. X = da/så, (X) = kan tolkes om bevisst ekstraposisjon)

Tekst 68	Høyfrekvent	Mellomfrekvent	Unik
Inverterte	X		X
Ikke-inverterte	<u>X</u>	(X)	X

Tekst 68 viser jevn fordeling når man bare ser på frekvens, med variasjon både innen de høyfrekvente og de unike leddene. Ser man nærmere på setningene, ser man allikevel at den høyfrekvente konteksten som ikke er invertert, er 'da'¹⁷, som altså er en tvetydig kontekst. I tillegg er det mellomfrekvente leddet 'selvfulgelig' etterfulgt av et komma, noe som viser at innlæreren bevisst har valgt ekstraposisjon og dermed ikke-invertert ordstilling. Tar man hensyn til dette, støtter også denne teksten hypotesen, ettersom variasjonen innen de

¹⁷ Og dublering

høyfrekvente kan forklares med at det inverterte leddet er tvetydig og derfor kan være vanskeligere å identifisere.

Tekst 89 sier også men mer positiv tendens etter å ha sett på tvetydige ledd:

Tekst 89	Høyfrekvent	Mellomfrekvent	Unik
Inverterte		XXX	X
Ikke-inverterte		(X)	X

Her er det mellomfrekvente leddet 'F.eks', som jo kan stå både med og uten inversjon. Ser man bort fra denne ene konteksten, vil profilen bli positiv, fordi variasjonen innen de mellomfrekvente kan være resultat av en bevisst ekstraponering.

Antallet av de ulike profilene etter denne omgangen er :

Profiler	Antall
Positive	9
Jevne	6
Negative	4

Dette gir resultater med overvekt av positive profiler (9).. Fem av tekstene har fremdeles en jevn fordeling som ikke kan forklares ved at noen kontekster er tvetydige ledd, og i disse tekstene er det altså ikke noe tydelig mønster som kan knyttes til det foranstilte leddets frekvens. Det betyr derimot ikke at disse tekstene er motprofiler.

Det kan tolkes som et positivt resultat at så mye halvparten av variasjonstekstene som har blitt undersøkt her, kan predikeres av identifiseringshypotesen.

4.4.4 Andre sjekkpunkter

Jeg har også hatt en rekke ekstra sjekkpunkt for de positive profilene (se 4.3.4): Fri variasjon på lekse nivå, overinversjoner og utelukkning av kompleksitet som påvirkningsfaktor.

a) Fri variasjon på lekse nivå

Om vekslingen mellom invertert og ikke-invertert er styrt av identifisering, vil det innebære at ett og samme ledd ikke kan være både identifisert og ikke-identifisert av samme innlærer. Fri

variasjon, hvor samme foranstilte ledd står både med og uten inversjon, vil derfor være et negativt funn i positive profiler.

Det er bare tre av de positive profilene som har fri variasjon på lekse nivå, tabellen gir en oversikt over disse.

32	
44	
49	Da
50	
61	Om sommeren
68	
70	
82	I hvert fall
89	

Vi ser at både 49 og 82, har fri variasjon etter tvetydige ledd, noe som kan fungere som en forklaring på at ledd er både identifiserte og ikke-identifiserte.

Ellis sier om slik variasjon at det kan være en innlærerstrategi for å teste ut hypoteser (Ellis 1989: 37), og med tvetydige ledd skulle det være ekstra god grunn til å prøve seg frem på en måte som kan gi fri variasjon. I tillegg er det klart at en ren glipp er mulig, og at det er det som er grunnen til at noe fremstår som både identifisert og ikke-identifisert.

b) Utelukking av kompleksitet

Når det gjelder utelukking av kompleksitet som påvirkningsfaktor, er det ikke alle de positive profilene som gjør det. Det vil si at når man ser på de foranstilte leddene, ser man at det er en fordeling hvor de inverterte er kortere og mindre komplekse enn de ikke-inverterte. Dette betyr at variasjonen like godt kan være resultat av det foranstilte leddets kompleksitet.

32	X
44	-
49	X
50	X
61	-
68	-
70	-
82	-
89	X

Tekst 32, 49, 50 og 89 har eksempler som bryter med ikonisiteten mellom kompleksitet og frekvens. De andre kunne like godt vært predikert av hypotesen om at det er kompleksitet som styrer variasjonen. Her har jeg bare sett på lengden, altså sett på det med et forenklet syn på

kompleksitet, og det kan hende at de mer utviklede operasjonaliseringene av kompleksitet som jeg bruker i neste kapittel, vil være mer fruktbare.

4.5 Konklusjon

Undersøkelsen av identifiseringshypotesen har fokusert på frekvensen og den syntaktiske entydigheten til det foranstilte leddet, med tanke på hvordan dette påvirker identifiseringen av en kontekst som inversjonskrevende. Prediksjonen var at de inverterte kontekstene skulle være mer frekvente og enn de underinverterte.

Resultatene av undersøkelsen viste at når bare frekvens ble tatt i betraktning, var det overvekt av positive profiler, altså profiler som stemmer med prediksjonen. 7 av 19 tekster var positive. Når syntaktisk tvetydighet ble brakt inn, viste flere av tekstene seg å ha positive profiler. Det ble da 9 positive profiler, altså halvparten.

I forbindelse med dette kan man være kritisk til min tolkning av jevne kontekster. Det kan tenkes at man med en mindre streng tolkning av disse kontekstene, vil få en mer positiv profil, om man for eksempel tillater én glipp, eller et moteksempel som ikke kan forklares.

Profilene ble videre undersøkt for en rekke ekstra sjekkpunkt; om profilen utelukket kompleksitet som årsak, overinversjoner, og om der var fri variasjon.

Tre av de positive profilene hadde slik fri variasjon. Hvorfor disse leddene forekom både med og uten inversjon må forklares på andre måter enn identifisering, men to av den kan jo forklares ved at de er tvetydige.

Jeg sjekket også om tekstene hadde eksempler som utelukket lengde og isolerte frekvens, noe som bare forekom i 4 av tekstene. De øvrige 5 har ikke utelukket lengde som påvirkningsfaktor. Dette betyr allikevel ikke at det ikke er identifisering som er årsaken, bare at det like godt kan være kompleksitet.

Det helhetlige inntrykket er at en del av tekstene støtter hypotesen, men langt fra alle. I de positive profilene finnes det også motfaktorer som fri variasjon. og bare fire av tekstene utelukker lengde som årsak til variasjonen. Hvor mye vekt som skal tillegges disse faktorene, kan diskuteres.

En interessant problemstilling er *grunnene* til fordelingen mellom positive, negative og jevne profiler. At noen tekster ser ut til å innfri prediksjonene kan være en tilfeldighet, men det kan også være et resultat av at innlærerne har ulike strategier. Det kan hende at variasjonen i noen av tekstene skyldes identifisering, mens den i andre tekster er styrt av andre forhold, for eksempel kapasitetsbelastning. Resultatene av denne undersøkelsen vil drøftes videre i kapittel 6, når de kan sammenlignes med resultatene fra undersøkelsen av kapasitetshypotesen.

4.6 Metodiske forbehold

4.6.1 Identifiseringsfaktoren

I undersøkelsen av identifiseringshypotesen har jeg knyttet kunnskap om når invertert ordstilling skal brukes, til identifisering av det foranstilte leddet som inversjonskrevende. At det foranstilte leddet signaliserer inversjon, trenger ikke diskuteres, men hvorvidt det bare er *frekvensen* i kollokasjon med invertert ordstilling som er avgjørende, er mer diskutabelt. Det er sannsynlig at identifiseringen er styrt av flere faktorer enn frekvens, og en svakhet med undersøkelsen min er det ensidige fokuset på dette. Min tilnærming tar for eksempel ikke hensyn til morsmålsbakgrunn, en faktor som med all sannsynlighet er viktig, slik som for eksempel Lund viser (Lund 1997:163). Også Saariks studie viser at estlenderne overfører kunnskap om invertert ordstilling fra sitt eget morsmål. Morsmål *uten* en V2-regel kan også tenkes å påvirke identifiseringen, siden den generelle prosesseringen av ord og uttrykk påvirkes av forholdene i morsmålet.

Tilnærmingen min tar heller ikke hensyn til evnen til å trekke induktive slutninger fra enkeltleksemer til typer av ledd. Det er godt mulig at når først ett tidsrefererende ledd er identifisert som inversjonskrevende, så vil denne kunnskapen kunne overføres til andre ledd med samme funksjon, til tross for at de ikke er særlig frekvente. Også synet på leddsetninger som unike, og ikke som klasser av når-setninger, hvis-setninger osv, kan kritiseres for å være for strengt. En undersøkelse som tar mer hensyn til denne typen slutninger, hadde vært bra.

Jeg har heller ikke tatt med mer grammatiske forhold, som hvilke syntaktiske funksjoner som er markert og umarkert ved fremflytting. Her kan forhold som har med syntaktisk funksjon å

gjøre, og hvorvidt et foranstilt ledd er komplementært eller fritt, kanskje vise seg å ha stor betydning.

En annen sak som er viktig i forhold til opparbeidinga av kunnskap, er at innlærerne er forskjellige, ikke bare mht morsmål, men også når det kommer til utdanningsbakgrunn og læringsstil. Noen er godt vant til grammatikk, og vil kanskje være bevisst på hvilke kontekster som krever inversjon fra første stund, mens andre vil være mer usikre, og må kartlegge distribusjonsregelen på den måten som identifiseringshypotesen tar utgangspunkt i.

Det er i det hele tatt ganske mange faktorer som spiller inn i identifiseringen av en inversjonskontekst, og en undersøkelse som tar alt med i betraktningen vil måtte være betydelig mer omfattende enn jeg har mulighet til her. At halvparten av tekstene lot seg forutse av hypotesen slik den ble testet ut, må regnes som et positivt funn.

4.6.2 Operasjonaliseringene

Identifiseringsgrunnlaget jeg har regnet som viktigst, er frekvens, og jeg har delt inn i høyfrekvente, mellomfrekvente og unike ledd. Dette er en noe grov inndeling, og en mer nyansert inndeling kunne sikkert vært på sin plass, men dette kan selvfølgelig være vanskelig ettersom det ikke er lett å få innblikk i hva slags innputt innlærerne faktisk møter.

Tvetydighetskategorien min er noe diskutabel. Det er en legitim påstand at tvetydige ledd kan føre til usikkerhet om hvorvidt det skal være eller ikke være invertert ordstilling, men det kan diskuteres om dette er en holdbar faktor for å oppgradere profiler til å være positive, på den måten jeg har gjort.

4.6.3 Den praktiske gjennomføringen

I undersøkelsen sammenlignet jeg inverterte og underinverterte kontekster i en form for implikasjonsskjema. Det viste seg at det ikke var noen tekster med positive profiler i strengeste forstand, og at jeg måtte åpne opp for variasjon innen én kategori, noe som ga flere positive profiler. Det kan hende at denne tolkningen er for streng, og at jeg kanskje kunne åpnet opp for at ikke alle kontekstene skal stemme overens med prediksjonene, men at i hvert fall én skal kunne stå uforklart, så lenge resten av kontekstene innfrir prediksjonene. Dette fordi det skal godt gjøres å finne tekster som stemmer 100% med hypotesen.

Jeg vil komme tilbake til mer generelle metodiske forbehold, som gjelder både identifiseringshypotesen og kapasitetshypotesen, i kapittel 6.

Kapittel 5: KAPASITESHYPOTEESEN

I dette kapittelet vil jeg undersøke den andre av hypotesene mine, kapasitetshypotesen, som altså tar utgangspunkt i at variasjon som et kontrollrelatert fenomen, først og fremst er knyttet til produksjon. Den variable anvendelsen av invertert ordstilling er resultat av at fremhenting av kunnskap bli vanskeligere av kapasitetsårsaker. Variasjonen er altså ikke nødvendigvis knyttet til kunnskap om *når* det skal være invertert ordstilling, men til hvorvidt innlæreren klarer å anvende kunnskapen i en faktisk kommunikasjonssituasjon.

5.1 Kontroll og kapasitetsbelastning

At variasjonen er relatert til kontroll, innebærer at innlæreren vet når det skal inverteres, men at denne kunnskapen ikke alltid hentes frem. Sharwood-Smith skriver om kontroll at ”... *control involves the exploitation of knowledge in actual performance and the degree to which the learner is able to switch attention to and fro the relevant bits and pieces of knowledge available*”(Sharwood Smith 1994:113)

MacLaughlin tar utgangspunkt i generell kognitiv psykologi, hvor språklæring sidestilles med andre kognitive læringsutfordringer, og beskriver opparbeiding av kontroll som bevegelse fra *kontrollert prosessering* til *automatisert prosessering*. Hovedtankegangen er at man i begynnelsen bruker mye oppmerksomhetsressurser på nye oppgaver (kontrollert prosessering), mens man etter hvert automatiserer oppgaven slik at den ikke lenger krever så mye av innlæreren kapasitet. Når en innlærer mestrer et trekk i enkle oppgaver, men ikke i faktisk kommunikasjon, forklarer MacLaughlin dette med at innlæreren benytter seg av kontrollerte prosesser, som det altså bare er tid og plass til når forholdene er mindre pressende. Når innlæreren begynner å mestre trekket i mer spontan kommunikasjon, har trekket blitt automatisert (Tarone 1988: 34). På grunn av dette vil en struktur spre seg fra de helt enkle til de mer kompliserte kontekster; ”*In this framework, a form would always appear first under conditions appropriate to controlled processes, and only later appear under conditions appropriate to automatic processes.*” (Tarone 1988: 34)

Dette gjelder altså som en forklaring både på variasjon på tvers av situasjoner som legger ulikt kapasitetspress på innlæreren, men også når det gjelder variasjon i én og samme situasjon, fordi noen typer språklige kontekster krever mer oppmerksomhet og kapasitet enn andre;

”Dersom informasjonsinnhenting ikke kan foregå nokså automatisert, blir det avgjørende for prosesseringen at den språklige konteksten ellers er enkel og lite krevende. Bruk av ikke-automatisert viten vil bli variabel etter språklig kontekst.” (Bergreen og Tenfjord 1999:235)

Det er dette som gjerne blir kalt kapasitetsbetinget variasjon (se kapittel 2). Når trekket ikke er automatisert, når man har lav kontroll, vil kunnskapen forbli ubrukt om konteksten rundt blir for krevende. Konteksten stjeler av oppmerksomhetsressursene, og det blir lite igjen til å hente frem kunnskapen om inversjon.

Dette kan også knyttes opp til Lunds funksjonelle forklaring på at inversjon læres sent; dens manglende kommunikative tyngde blir nedprioritert i kampen mot språklige elementer som krever oppmerksomhet for å uttrykke innhold med mer kommunikativ tyngde, for eksempel tempus, bestemthet, referanse osv.

Jeg vil derfor undersøke om variasjon er styrt av de språklige omgivelsenes *kapasitetsbelastning*, og i hvor stor grad de kan tenkes å stjele oppmerksomheten fra anvendelsen av den inverterte ordstillingen. Når omgivelsene er enkle og lettprosesserte, er det større sjanse for at innlæreren får benyttet kunnskapen han har, og når omgivelsene er tunge og tar mye oppmerksomhet, er det stor sjanse for at kunnskapen forblir ubrukt.

Kontekster som medfører lav kapasitetsbelastning er med større sannsynlighet inverterte enn kontekster som er mer belastende for innlæreren.

5.2 Operasjonalisering: Lingvistisk kontekst og kapasitetsbelastning

Før jeg kan undersøke denne hypotesen nærmere er det to spørsmål må avklares: Hva utgjør den lingvistiske konteksten, og på hvilken måte er den kapasitetsbelastende?

5.2.1 Avgrensning av den lingvistiske kontekst

I forrige kapittel var det det foranstilte leddet som ble regnet som den delen av den språklige omgivelsen som styrte variasjonen. Det gjorde det i kraft av frekvens og syntaktisk entydighet i innputt, det var altså en bestemt grunn til at påvirkningen var avgrenset til bare det foranstilte leddet. Men i forbindelse med kapasitetshypotesen, vil jeg utvide påvirkningskilden til også å inkludere subjektet og det finite verbet. På setningsnivå er det først og fremst disse tre leddene som konkurrerer om oppmerksomhet med anvendelsen av invertert ordstilling. En kan også tenke seg andre kapasitetsbelastende faktorer, både i den språklige konteksten og utenfor, men i denne omgang konsentrerer jeg meg om disse tre leddene, da disse er den nærmeste språklige konteksten og derfor skal produseres i samme omgang med inversjonen.

Undersøkelsen av kapasitetshypotesen blir derfor en undersøkelse av hvor kapasitetsbelastende forleddet, subjektet og det finite verbet er, og hvorvidt dette er avgjørende for om invertert ordstilling blir anvendt eller ikke. Selv om det er disse tre som er regnet som hovedpåvirkningskilden til variasjonen, vil jeg også trekke inn andre faktorer, for eksempel setningsadverbialer og venstre-dislokerte leddsetninger. Jeg skal komme tilbake til dette.

Kapasitetshypotesen kan spissformuleres på følgende måte:

Variasjonen er resultat av at kunnskapen om invertert ordstilling ikke er helt automatisert, og at den omkringliggende språklige konteksten i ulik grad belaster innlærerens kapasitet og påvirker innlærerens muligheter til å nytte seg kunnskapene om inversjon.

Variasjonen avspeiler altså ikke i første omgang et eget system hos innlæreren, ulikt målpråkets, men en *systematisk påvirkning* av et system som forutsettes å være målpråkslikt.

5.2.2 Overinverterte kontekster

I denne undersøkelsen er også de overinverterte kontekstene regnet med på lik linje med de obligatoriske kontekstene. Dersom det er kapasitetsbelastning som avgjør om setningen har invertert ordstilling eller ikke, så skal dette gjelde for overinverterte setninger også.

Overinverterte setninger skal derfor ikke være mer belastende enn de underinverterte. Som regel vil en overinvertert setning ha lette setningsinnledere, derfor bør subjektet og verbet være relativt lette, alt avhengig av hvor mye kapasitetsbelastning den enkelte innlærer kan tåle.

De overinverterte setningene skal ikke være mer belastende enn de underinverterte. Om det finnes overinverterte kontekster i variasjon med ikke-overinverterte setninger etter samme type ledd, er det forventet at den ikke-overinverterte er mer belastende, og at det altså er kapasitetsforhold som styrer også denne variasjonen.

5.2.3 Operasjonalisering av kapasitetsbelastning

Graden av kapasitetsbelastning er altså den graden av oppmerksomhet inversjonen krever for å kunne utføres. Jo mer automatisert invertert ordstilling er, jo mindre oppmerksomhet kreves det for å anvende den. De begrensede ressursene innlæreren har til rådighet, må fordeles over hele ytringen, og om andre deler krever mer oppmerksomhet, vil det gå ut over inversjonsanvendelsen.

Kapasitetsbelastning knyttes gjerne til *prosesseringstyngde* eller *kontekstuell vekt* (Ellis 1989: 37), men som jeg har vært inne på, kan man lett ende opp med en sirkeldefinisjon av tyngde ved å kategorisere kontekster på denne måten, I denne sammenheng vil det si at man stempler de inverterte kontekstene som lette, mens de ikke-inverterte regnes som tunge. Dette ønsker jeg selvsagt å unngå, og ved å utlede kategoriene fra en hypotese om hva som er årsaken til variasjonen – kapasitetsbelastning - kan jeg undersøke materialet uten et sirkeldefinert tyngdebegrep. Det største problemet blir heller å sikre båndet mellom den påståtte forklaringen og de operasjonaliserte kategoriene. Tar jeg utgangspunkt i en bestemt operasjonalisering av kapasitetsbelastning, vet jeg jo aldri om det er akkurat dette som utgjør den faktiske, eller totale, kapasitetsbelastningen. Dette metodiske forbeholdet kommer jeg tilbake til i slutte av dette kapitlet og i kapittel 6.

5.2.4 Mål for kapasitetsbelastning

Som mål for kapasitetsbelastning har jeg først og fremst valgt **strukturell kompleksitet**. Det er altså ikke bare den rene *lengden* på leddene som er interessant, men den indre strukturen.

Kategoriene jeg har valgt til å representere belastning, avspeiler først og fremst en kompleksitetsskala over leddstrukturer som når de produseres utgjør ulike grader av kapasitetsbelastning, fra **enkeltord**, via en **sammensatt frase**, til en hel **setning**. Jo mer komplekst, jo mer kapasitetsbelastende. På ordnivå skiller jeg mellom kortere og lengre ord, og på frasenivå mellom kortere og lengre fraser. Dette utgjør til sammen 5 kategorier, hvor **1** (korte ord) er minst belastende, og **5** (setninger) mest belastende. Tallangivelsene er kun representasjoner for rang, de utgjør ikke noen som helst form for numerisk verdi utover det.

Kapasitetsbelastning	lav	—————▶			høy
kompleksitet	1	2	3	4	5
	ord		fraser		setning

5.2.5 Operasjonalisering av hvert enkelt ledds relative kapasitetsbelastning

Jeg vil nå utdype de ulike kategoriene for de ulike leddenes belastning i den språklige konteksten. Når to eller flere ledd i samme tekst er på samme kompleksitetsnivå, vil *lengden* på leddene avgjøre hvilket som er mest belastende.

a) Foranstilte ledds kapasitetsbelastning

For foranstilte ledd har jeg gjort følgende inndeling i kompleksitetskategorier (nominale foranstilte ledd vurderes etter kriteriene for subjekter lenger nede):

Kategori 1: Korte enkeltord.

Her regnes de korteste ordene som forekommer som foranstilte ledd - ord med én stavelse.

Kategori 2: Lengre enkeltord/faste uttrykk.

Ord med mer enn to stavelser, og faste uttrykk/idiomer. Grunnen til at faste uttrykk sidestilles med enkeltord, er at disse kan sees på som uanalyserte fraser, og derfor ikke har noen

hierarkisk frasestruktur, for eksempel i 'hvert fall', 'hver gang' og 'for eksempel'. Uttrykkene forekommer bare i én form, og har ingen indre bevegelighet. Nominale foranstilte ledd tilsvarer subjekter i kategori 2.

Kategori 3: Enkle fraser.

Enkle preposisjonsfraser, altså preposisjon pluss substantiv. Nominale ledd tilsvarer subjekter i kategori 3.

Kategori 4: Lengre fraser.

Mer komplekse preposisjonsfraser; preposisjon pluss utbygd substantivfrase, eller nominale ledd tilsvarende subjekter i kategori 4.

Kategori 5: Leddsetninger

Foranstilte ledd som er leddsetninger, eller inneholder leddsetninger (denotasjon).

b) Subjektets kapasitetsbelastning

Subjektets belastning kategoriseres i tråd med det foranstilte leddets, men de er ikke umiddelbart sammenlignbare, selv om de tilhører tilsvarende rangkategorier.

For substantivfraser har jeg basert meg på Maria Axelssons avhandling, hvor hun har kommet frem til en læringsløype på tre nivåer for substantivfraser (Axelsson 1994:99). Denne læringsløypen tar utgangspunkt i frekvens og korrekthet hos innlærerne, og dette går parallelt med frasenes kompleksitet. Kompleksiteten ble beregnet etter det Axelsson kaller ”*antal formella svårigheter*” som frasen krever, med hovedvekt på bestemthetsmarkering. (Flertallsbøyninger er ikke regnet som en formell utfordring hos Axelson.) Hos meg har hvert nivå hos Axelsson fått en kategori (2, 3 og 4):

Kategori 1: Korte enkeltord

Pronomen, det (som formelt subjekt).

Kategori 2: Lengre enkeltord/substantivfraser 1

Tilsvarende Axelssons nivå 1: navn, ubøyde substantiver, og premodifiserte ubøyde substantiv.

Kategori 3: Substantivfraser 2

Tilsvarende Axelssons nivå 3; substantiv med ubestemt artikkel, substantiv med bestemthetsendelse, og ubestemte substantiv med possessiv.

Kategori 4: Substantivfraser 3

Tilsvarende Axelssons nivå 3; *den grønne bilen* og *en grønn bil*, eller enda mer utbygd.

Kategori 5: Leddsetninger

Leddsetninger, denotasjon, infinitivkonstruksjoner.

c) Finitte verbs kapasitetsbelastning

Verbet vil bare være ett ord, og derfor skiller jeg kun mellom korte og lange verb.

Kategori 1: korte verb

Kategori 2: lengre verb

d) Sammenlignbarheten mellom kategori 1-5 for foranstilte ledd og subjekter:

Selv om kategoriene for de ulike leddene skal representere omtrent samme kompleksitet, er de ikke umiddelbart sammenlignbare på tvers av leddtype. Det vil si at et foranstilt ledd i kategori 3 ikke nødvendigvis tilsvarer et subjekt i kategori 3. Nivå 1 og 5 kan sammenlignes direkte, mens 2, 3 og 4 må vurderes mer nøye.

e) Ekstra belastningsfaktorer

Det er selvfølgelig ikke bare strukturell kompleksitet og produksjon av ren form som belaster kapasiteten i produksjonen av en setning. Jeg har derfor operert med noen ekstra belastningsfaktorer:

- Anaforiske foranstilte ledd eller subjekter
- Kontrasterte (trykksterke) foranstilte ledd og subjekter
- Verb med sterk knytning til en partikkel eller et annet komplement
- Frekvente foranstilte ledd og subjekter og verb

Ord som er anaforiske kan være mer belastende enn ord som ikke er det. Dette fordi selve prosessen rundt anafori krever en del oppmerksomhet. Det at et pro-ords innhold ikke er direkte tilgjengelig gjennom selve ordet, men må etableres med referanse til konteksten, gjør det mer krevende å bruke for eksempel anaforiske pronomen, som han, hun og de, enn pronomen som har en mer deiktisk og ubestemt referanse, som jeg, vi, man og en. Disse er ikke anaforiske, og referansen kan hentes direkte ut av den generelle situasjonen.. Også foranstilte ledd kan være anaforiske, spesielt pro-adverbene da og derfor.

Man kan også skille mellom trykk tungt og trykk lett pronomen. De fleste pronomen vil være trykk lette, men noen pronomen trykk legges for å kontrasteres eller fremheves. Hagen viser til at trykk lette pronomen virker positivt på inversjonsmestringen, fordi de lettere danner en enklise med verbet. Siden materialet mitt består av skriftlige tekster, blir det ikke helt rett å snakke om *trykk*, og jeg velger derfor heller ordet *kontrastering*. Trykk markering av ord i muntlig språk vil ofte være uttrykk for kontrastering, et behov for å oppløse en tvetydighet i referansen, men det kan også være emfatisk trykk, som fremhever ny informasjon.

For verb vil dette om de på en eller annen måte er bundet til en verbpartikkel eller et ord i den etterfølgende konteksten, være vesentlig for hvor belastende det er. Verb som har verbpartikkel kan være mer belastende, siden verbet er nært knyttet til partikkelen, som i de ”*giftet seg*”, eller som idiomatiske uttrykk; ”*å ha det bra*”.

f) Belastende faktorer utenfor den avgrensede lingvistiske konteksten:

Selv om jeg har avgrenset den lingvistiske konteksten til foranstilt ledd, subjekt og det finite verbet, vil jeg trekke inn noen få andre faktorer som ikke har direkte med disse leddene å gjøre:

- Setningsadverbialer i midtfeltet
- Foranstilling i leddsetninger
- Venstre-dislokerte ledd

Setninger med negasjon eller andre adverbialer i midtfeltet, kan være belastende fordi de konkurrerer med subjektet om plassen bak det finite verbet, men også fordi plasseringen av

adverbialet kan være en utfordring i seg selv. Derfor kan det tenkes at tilstedeværelse av negasjoner og andre midtadverbialer kan være belastende.

Siden leddsetninger regnes som en mer komplisert struktur enn hovedsetninger (Lund 1997:167), regner jeg foranstilling av ledd i leddsetninger som generelt mer krevende enn i vanlige hovedsetninger.

e) Positive faktorer

Det finnes noen forhold som kan gjøre en ellers belastende inversjonskontekst lettere å prosessere.

- Holofrastiske foranstilte leddsetninger
- ”Tags” (som *sier jeg*)
- Ekstraponering

Som jeg var inne på i kapittel 4, kan noen foranstilte leddsetninger være *holofrastiske*, eller lært som en uanalysert enhet. Slike setninger vil rent overfladisk ser ut til å tilhøre den høyeste kompleksitetskategorien, men utgjør kanskje ikke utgjør den store belastningen fordi innlæreren ikke konstruerer dem fra grunnen selv - de har en ”flat struktur”. Dette vil si at kontekster som har slike foranstilte ledd, i realiteten kan ha en lavere belastning enn kompleksitetskategorien tilsier.

Vedheng til helsetninger, som *'sier jeg'*, *'synes jeg'*, *'mener jeg'*, kalles gjerne for en *tag*. Slike tagger er ofte helfraser, og dermed lite belastende til tross for at de kommer etter en lang helsetning. Kontekster med tagger kan derfor regnes som mindre belastende enn kompleksiteten tilsier (jf. Bolander 1988: 209).

Som i kapittel 4 vil foranstilte ledd som kan stå som ekstraposisjon i norsk, kunne stå uten invertert ordstilling uten at det er ugrammatisk. Om kontekster uten invertert ordstilling kan tolkes som et bevisst valg, er ikke mangelen på inversjon et resultat av kapasitetsbelastning. Dette kan være grunnen til at kontekster som ellers har en relativt *lav* belastning, allikevel ikke er invertert. Dette gjelder den typen ekstraposisjon som ikke gjentar innholdet med et pro-ord i forfeltet.

f) Om bruk av disse ekstrarfaktorene.

Disse ekstra belastningsfaktorene og positive faktorene vil bli brukt til å forklare hvorfor ledd som kompleksitetsmessig har lav belastning, allikevel *ikke* er invertert, eventuelt hvorfor ledd som har kompleksitetsmessig høy belastning *er* invertert. Det er viktig å kvalitetssikre bruken av disse ekstrarfaktorene, slik at de tas til inntekt for en ting, men ikke for andre. Jeg har derfor noen retningslinjer for bruk av disse.

- 1) Hvor mye påvirkning disse faktorene har, må vurderes i forhold til om de kan anses som belastende *nok* til å gi utslag?
- 2) Faktorene skal være legitimt brukt som forklaring. Om de brukes som forklaring på en underinvertert setning, skal også eventuelle korrekte setninger i samme tekst som også har denne faktoren vurderes. Om tekster viser seg positive til hypotesen uten at slike faktorer trekkes inn, skal forekomsten av disse allikevel tas stilling til, og eventuelt vurderes som en motfaktor.

5.3 Gjennomføring av undersøkelsen

Undersøkelsen av kapasitetshypotesen vil først og fremst se på om det er de språklige omgivelsenes kompleksitet som avgjør om en setning har invertert ordstilling eller ikke. I følge hypotesen skal de setningene som ikke er inverterte, ha en lingvistisk kontekst som er mer kompleks og belastende for kapasiteten, enn de setningene som er inverterte.

I hver enkelt tekst vil setningene med invertert ordstilling ha mindre belastende språklige omgivelser enn setningene uten invertert ordstilling.

5.3.1 Tildeling av kategorier og ekstrarfaktorer

Hvert ledd i hver inversjonskontekst blir kategorisert etter kompleksitet utfra kriteriene over, og eventuelle andre belastningsfaktorer som er tilstede blir registrert, slik at kontekstene med invertert ordstilling kan sammenlignes med dem uten i hver enkelt tekst.

Undersøkelsen har to runder, i første om gang ser jeg bare på kompleksitet, og i neste omgang bringer jeg inn de andre belastningsfaktorene. Dette er blant annet for å kunne skille ut hvor

mye av belastningen som kan tilskrives kompleksitet og hvor mye som kan tilskrives andre faktorer.

Etter å ha identifisert kompleksiteten til hvert ledd i inversjonskonteksten etter kategoriene i 5.2.4, får jeg en *tall-representasjon* av den lingvistiske kontekstens samlede kapasitetsbelastning. Om det foranstilte leddet tilhører kategori 1, subjektet 5 og verbet 2, vil kapasiteten kunne representeres som **[1-5-2]**. Om hypotesen stemmer, skal de inverterte setningene ha *lavere* kapasitetsbelastning enn de underinverterte.

Det er et viktig poeng at lav og høy belastning er et *individuell* forhold. Innlærerne kan ha opparbeidet ulik grad av automatisering av inversjonsstrukturen, så hva som er belastende og ikke-belastende, må defineres ut fra den enkelte innlærerens nivå. På grunn av dette har jeg tatt utgangspunkt i en antatt belastningstoleransegrense for hver innlærer/tekst, som jeg har kalt *den maksimale belastningen*. Dette er *den mest krevende inverterte konteksten i hver tekst*. Siden denne konteksten er invertert, må en regne med at kapasitetsbelastningen er innenfor grensene for det innlæreren kan tåle før inversjonen utelates. Med den maksimale kapasitetsbelastningen som referansepunkt, kan jeg undersøke om kontekstene som *ikke* er inverterte, er *mer* eller *mindre* belastende enn den.

Denne metoden for å definere en maksimal toleransegrense, grensen mellom tungt og lett for hver innlærer, må ikke forveksles med den sirkeldefineringsen av tungt og lett som jeg har vært inne på. Den er kun et metodisk grep, et målepunkt; kapasitetsbelastningen i seg selv er definert uavhengig, i kompleksitet. Det er altså ikke en definisjon av kapasitetsbelastning, men en angivelse av grensen mellom tungt og lett for den enkelte innlærer.

5.3.2 Sammenligning av kontekster

De underinverterte setningene i hver tekst blir sammenlignet med den maksimale belastningen i hver tekst på følgende måte:

Finnes det underinverterte setninger hvor <foranstilt ledd + subjekt + finitt verb> er:

- 1) *mindre* belastende enn den maksimale belastningen? (= minuskontekst)
- 2) *mer* belastende enn den maksimale belastningen? (= plusskontekst)
- 3) *like* belastende som den maksimale belastningen? (= lik kontekst)

I en *plusskontekst* er enten alle ledd i kollokasjonen *mindre* belastende enn leddene i den maksimale belastningen, eller så er ett eller to ledd like belastende, mens resten er lettere. I en *minuskontekst* er det motsatt; alle ledd er *mer* belastende, eller ett eller to ledd er like, og resten tyngre. I en *lik kontekst* er tallrepresentasjonene like, eller de vurderes som tilsvarende om direkte sammenligning ikke er mulig. Som jeg var inne på før, kan kategori 1 og 5 sammenlignes direkte på tvers av foranstilt ledd og subjekt, mens det må vurderes mer nøye for kategori 2, 3 og 4. Dette vil si at om noen kontekster ikke umiddelbart lar seg sammenligne med den maksimale belastningen, må jeg vurdere om de er mer eller mindre belastende ved å se nærmere på teksten.

5.3.3 Kompleksitetsprofiler

Etter å ha sammenlignet underinverterte kontekster med den maksimale belastningen, sitter jeg igjen med en *kompleksitetsprofil* for hver tekst. Det er tre mulige profiler:

Negativ profil: Bare minuskontekster (alle underinverterte kontekster i teksten er *mindre* belastende enn den maksimale belastningen).

Positiv profil: Bare plusskontekster eller like kontekster (alle underinverterte kontekster i teksten er *mer* belastende enn maksimal belastning).

Blandet profil: Både plusskontekster/like kontekster og minuskontekster.

Denne profilen sier kun noe om hvordan teksten fremstår når det kommer til kompleksitet. Videre er det ønskelig å se om noen av ekstraraktorene kan være grunnen til at de underinverterte mot forventning er mindre belastende enn den maksimale belastningen, eventuelt om noen av de positive faktorene gjør at ellers belastende kontekster allikevel er invertert.

5.4 Resultater

5.4.1 Oversikt over tekstenes profiler når kun kompleksitet tas i betraktning

Følgende tabell viser de ulike tekstenes kompleksitetsprofil, etter å ha vurdert dem som positive (+), negative (--) eller blandete profiler (B):

	2	6	32	37	44	49	50	54	56	57	61	64	68	70	82	85	88	89	98	sum
+																				0
-	X					X	X	X										X	X	6
B			X	X					X			X	X	X	X	X	X			9

(tekst 6, 44, 57 og 61 lot seg ikke undersøke med metoden jeg har utarbeidet, så for disse tekstene har jeg ingen profilvurdering)

Det er altså ingen av tekstene som viser noen umiddelbar positiv tendens når bare kompleksitet tas i betraktning. Dette kan tyde på at det ikke bare er ren kompleksitet som er avgjørende, noe som er en interessant observasjon i seg selv. Det er en liten overvekt av blandede profiler (9), og negative (6), og jeg har sett nærmere på disse for å se om minuskontekstene har noen ekstra belastningsfaktorer som kan forklare hvorfor de ikke er invertert, til tross for at kompleksiteten ligger innenfor det innlæreren kan tåle av belastning. Det kan også være grunn til å revurdere den maksimale belastningen i noen av tekstene, om den konteksten som er regnet som maksimalt belastende har noen positive faktorer som gjør den mindre belastende enn kompleksiteten tilsier.

I tillegg må minuskontekstene vurderes i forhold til om de er mindre belastende enn samtlige inverterte setninger, eller om de bare er mindre belastende enn de maksimale belastningen. De står sterkere som minuskontekst dersom de er lavere enn samtlige inverterte, enn om de bare er lavere enn den maksimale belastningen, men ellers mer belastende enn de andre kontekstene.

5.4.2 Eksempler

I de følgende tabellene over kontekster er de underinverterte *over* den røde streken og den maksimale belastningen er uthevet med **fete** typer. De inverterte er sortert etter den innbyrdes kompleksiteten. Kolonne **K** viser om de underinverterte er pluss- eller minuskontekster. **O** foran kontekstnummeret betyr overinvertert kontekst.

a) Negative profiler

De tekstene som har en negativ profil, stemmer altså ikke overens med prediksjonene om at de underinverterte kontekstene skal være mer belastende enn de inverterte. Utfordringen med

disse tekstene er å se om noen av minuskontekstene kan vise seg å ha ekstra belastningsfaktorer som ikke avspeiles i kompleksitetskategoriene. Tre av seks negative profiler lar seg oppgradere til positive, for eksempel tekst 50. I Denne teksten er det kontekster uten invertert ordstilling, og begge er mindre belastende enn den maksimale belastningen på **5-3-1**, som er setningen: *'så hvis det skulle være noe galt med dem, må lærerne kontakte...'*

	FL	SU	FV	K	
50.1	5-	1	1	-	Når barna er små de har det
50.2	5-	1	2	-	Fordi barna føler seg bra, de føler seg
50.3	5+	3	1		Når barna begynner å gå på barneskolen, for det første må lærerne
50.4	5	3	1		, så hvis det skulle være noe galt med dem, må lærerne
50.5	5	1	1		Da barna er ferdig med ungdomsskolen, kan man....
50.6	4	3	2		I den alderen begynner puberterten
50.7	4+	1	2		For fremmede folk kaller jeg
50.8	1	3	2		(Etter at barna har blitt 3 eller 4 år,) da trenger foreldrene
50.9	1	1	2		(Når jeg sier foreldre,) da mener jeg

Begge kontekstene uten invertert ordstilling har foranstilte ledd på samme kompleksitetsnivå som den maksimale belastningen, men disse er lengre i de inverterte kontekstene, som også har mer belastende subjekter. De underinverterte kontekstene har allikevel en del andre belastende faktorer. Begge setningene har, til forskjell fra alle de inverterte, anaforisk subjekt – 'de'. I tillegg har begge kontekstene verb med knytning til etterfølgende ord; 'føler seg' og det idiomatiske uttrykket 'har det bra'.

Denne profilen er altså kompleksitetsmessig negativ, men begge minuskontekstene har ekstra belastninger - anaforisk subjekt og verb med knytning til partikkel/komplement - som kan være årsaken til at de allikevel ikke er inverterte.

Noen av de negative profilene lar seg imidlertid ikke forklare ved ekstrarfaktorer, disse tekstene forblir negative, for eksempel tekst 2. Denne teksten har to kontekster uten invertert ordstilling, og begge er mindre belastende enn den maksimale belastningen på **5-1-2**, setningen *'... fordi hvis man mistet tradisjoner vi mister...'*

k.nr.	FL	SU	FV	K	
2.1	2	1	2	-	Also man trenger
2.2	1	1	2	-	, nå han begynner
2.3	5	1	2		; fordi hvis man mistet tradisjoner mister vi
2.4	5	1	2		.etter vi flytter, fra hjemlang begynte vi
2.5	1	4	1		(fordi) har er den beste sted
O2.5	1	1	1		Og må være man høy

Subjektene er mer belastende i de underinverterte kontekstene, siden de inverterte har *vi* som subjekt, som er mer frekvent og har en mer tilgjengelig deiktisk referanse enn *man* og *han*. Dette veier allikevel ikke opp for leddsetningene som innleder de inverterte kontekstene. Det ene underinverterte subjektet er anaforisk, men utover dette er det lite som stemmer overens med prediksjonene, verken kompleksitetsmessig eller med hensyn til andre belastningsfaktorer.

Det finnes også andre forhold som taler imot denne teksten som positiv: To av de inverterte kontekstene er i en leddsetning. I tillegg er det slik at alle minuskontekstene har lavere belastning enn samtlige inverterte. Det er altså sterke minuskontekster, og siden ingen har andre belastningsfaktorer, er denne tekstens profil negativ i forhold til hypotesen. Overinversjonen i denne teksten har relativt lav belastning, men både hjelpeverbet og hovedverbet er invertert.

b) Blandede profiler

De blandede profilene har altså både underinverterte minuskontekster og plusskontekster. Plusskontekstene stemmer overens med prediksjonene, mens minuskontekstene må sjekkes for ekstrarfaktorer. En tekst som viser positive tendenser når jeg trekker inne ekstrarfaktorer er tekst 68. Den maksimale belastningen i denne teksten er **5-1-1**; '*For eksempel, hvis jeg trenger noe, kan jeg spørre*', og det er tre kontekster uten invertert ordstilling.

	FL	SU	FV	Kompl	
68.1	5+	1	1	+	Hvis du har noen venner og har forhold med noen det er
68.2	2	1	1	-	Selvfølgelig , de kan også spør meg
68.3	1	2	2	?	(Hvis jeg er i nød,) da kamerater hjelper meg .
68.4	5	1	1		For example, hvis jeg trenger noe, kan jeg
68.5	1	1	1		...og nå har jeg
O68.6	1	1	1		Så har jeg

Den første konteksten, 68.1, har *lik* kompleksitet som den maksimale belastningen, men den har en mye lengre leddsetning i forfeltet, og er derfor mer belastende. Den eneste minuskonteksten, 68.2, har et tydelig markert ekstraposisjonert forledd, så jeg velger å tolke manglende invertert ordstilling som et bevisst valg. Konteksten har uansett flere belastende ekstrarfaktorer: et anaforisk trykksterkt subjekt, og adverb i midtfeltet. Den siste konteksten, 68.3, er ikke umiddelbart sammenlignbar med den maksimale belastningen. Men med belastende faktorer som venstre-dislokert leddsetning, inversjonskontekstenes eneste

substantivsubjekt, og et verb med verbpartikkel, er det mye som tyder på at denne setningen er mer belastende.

Det er også en sjanse for at konteksten som er regnet som den maksimale belastningen, har en holofrastisk foranstilt leddsetning, *hvis jeg trenger noe*, som i så fall betyr at den maksimale belastningen kan være ”feilberegnet”. Når den andre inverterte konteksten i tillegg er 1-1-1, er det verdt å stille spørsmålsteget ved den valgte maksimale belastningen. Minuskontekstene er nemlig alle mer belastende enn den andre inverterte konteksten.

En annen positiv tendens i denne teksten er at alle de belastende ekstraraktorene utelukkende er tilstede i de underinverterte kontekstene. Den overinverterte konteksten passer også inn, da den har belastningen 1-1-1.

Tekst 32 har også en blandet kompleksitetsprofil som viser seg å være positiv. Teksten har en maksimal belastning på **5-1-2**, i to av kontekstene. Det er tre kontekster uten inversjon, og disse er litt vanskelige å vurdere i forhold til den maksimale belastningen, fordi de ikke er direkte sammelignbare.

	FL	SU	FV	Kompl	
32.1	5+	1	1	=	At blir en sjåføren på tunge kjøretøy han må gå
32.2	2	2	2	-	, også lærer snakket ofte
32.3	5+	1	1	=	I STED om sommer når er jazz festivalen det er
32.4	5	1	2		Da kjørte jeg første gang lastebil trodde jeg...
32.5	5	1	2		Da fikk jeg min første jobbe skjønte jeg...
32.6	3	1	1		For meg selv var det...
32.7	2	1	2		I dag kjør er jeg
32.8	2	1	2		I dag trives jeg
32.9	2	1	1		I hoved sak må det ...
O32.10	1	1	2		Da kjør te jeg...
O32.11	1	1	1		Da fikk jeg..
O32.12	1	1	2		Hva mente han

To av kontekstene, 32.1 og 32.3, har et lengre forfelt enn begge de maksimale inverterte, men de har kortere verb, så kompleksitetsmessig er disse fire setningene omtrent *like*. De inverterte har imidlertid *jeg* som subjekt, mens de underinverterte har *han* og *det*, så de inverterte kontekstene er allikevel mindre belastende. Kontekst 32.2 er ikke umiddelbart sammenlignbar med den maksimale kompleksiteten - verbet er likt, forfeltet lettere og subjektet litt lengre. Allikevel er det tydelig at kompleksitetsforskjellene mellom subjektene og verbene i disse kontekstene er så små, at jeg lar det foranstilte leddet veie tyngst, noe som gjør den underinverterte konteksten til en minuskontekst. Men denne minuskonteksten har et adverbial

i midtfeltet, og den er derfor mer belastende enn den maksimale. I tillegg er den svak som minuskontekst, fordi den er mer belastende enn de andre inverterte setningene.

De overinverterte kontekstene i denne teksten er ikke mer belastende enn de underinverterte.

En siste blandet profilsom kan vise seg positiv er tekst 82. Denne teksten har relativt lav maksimal belastning, 2-1-2. Kompleksitetsprofilen er blandet, med én kontekst som er mer belastende, og én som er mindre.

k.nr	FL	SU	FV	K	
82.1	5	1	2	+	men fra det hva hørte jeg av folk som har sitt eget barn jeg kunne..
82.2	2	1	1	-	..., men i hvert fall vi kan prøve
82.3	2	1	2		I hvert fall jobber jeg
82.4	1	1	2		(, hvis jeg var), da beklager jeg
O82.5	1	1	2		hva hørte jeg
O82.5	2	1	1		Hvor mye er de

Minuskonteksten kan være en ekstraposisjon, men den er ikke markert med komma, og uttrykket finnes også med inversjon i den maksimalt inverterte konteksten.

..., men **i hvert fall** vi kan prøve

I hvert fall jobber jeg

En mulig utslagsgivende forskjell mellom disse er at den inverterte konteksten har *jeg* som subjekt, mens den underinverterte har *vi*, og på den måten er minuskonteksten allikevel hakket mer belastende enn den maksimale belastningen.

c) Blandete profiler med kun ett moteksempel

Andre blandete profiler har minuskontekster som ikke lar seg forklare, og i denne sammenhengen kommer spørsmålet om hvorvidt alle minuskontekstene må være mer belastende enn de underinverterte for at profilen skal være positiv. Det er rimelig å forvente at ikke all variasjonen er styrt av belastningsforhold, og at noe av variasjonen er fri. At noen kontekster er i fri variasjon, er enten resultat av for dårlig leting mht belastning, eller av at det er andre faktorer enn det jeg har regnet som belastende som spiller inn. Det er umulig å ha kontroll over alle faktorer som kan belaste mestringsen, derfor har jeg sjekket om det er noen tekster som bare har én minuskontekst som ikke lar seg forklare. Om denne ene forekomsten er en svak minuskontekst, en som kun er mindre belastende enn den maksimale belastningen,

men mer belastende enn de øvrige inverterte, vil jeg si at konteksten generelt sett er mer belastende enn de inverterte, og at profilen er positiv.

Tekst 56 har for eksempel en maksimal belastning på **4-1-1**: ”..., og etter ti dager på åpent plass, eller i tenda, jeg kom....”. Det er fem kontekster uten invertert ordstilling, tre minuskontekster og to plusskontekster.

	FL	SU	FV	K	
56.1	5	1	2	+	, men de som jeg har jeg holder kontakt med
56.2	3	1	1	-	, med henne jeg var alltid
56.3	5	1	2	+	Når vi var ferdige med videregående skolen og høyskolen etterpå, vi funnet
56.4	2+	1	2	-	.., etter kort tid hun giftet seg
56.5	4-	1	1	-	Etter tre måneder i Norge jeg fikk vite
56.6	4+	1	1		Og etter ti dager på åpent plass, eller i tenda, kom jeg
56.7	1	4	1		Men da kom en darlig samvittighet
56.8	1	3	1		Så kom krigen

En av de underinverterte kontekstene (56.5) har lik kompleksitet som den maksimale belastningen, men siden det foranstilte leddet i denne er betraktelig kortere, er den altså mindre belastende. Den har ingen ekstra belastningsfaktorer, noe som gjør at den alt i alt ikke kan sies å være mer belastende enn den maksimale belastningen.

Begge de to andre minuskontekstene har andre belastende faktorer. Kontekst 56.2 har både et anaforisk foranstilt ledd, og et adverbial i midtfeltet, mens kontekst 56.4 har et kontrastert anaforisk subjekt, i tillegg til et verb med partikkel. Dette vil si at denne teksten bare er én kontekst ”unna” en positiv profil. Dens moteksempel, 56.5, har allikevel bare så vidt lavere belastning enn den maksimale, og den har ellers høyere belastning enn alle de andre inverterte, derfor viser denne teksten en generell positiv tendens.

Det som også er interessant med denne teksten, er at alle de inverterte kontekstene har samme korte verb, *kom*, og kanskje kan tilstedeværelsen av dette verbet være en annen faktor som gjør anvendelsen av invertert ordstilling lettere i disse setningene.

5.4.3 Oversikt over tekstenes profil med både kompleksitet og ekstra faktorer

Når 15 tekster er analysert, ser en at hele 11 stemmer overens med prediksjonene. Fire tekster er negative.

	2	6	32	37	44	49	50	54	56	57	61	64	68	70	82	85	88	89	98	
P:	-		+	+		+	+	+	+			+	+	+	+	-	+	-	-	

5.4.4 Konklusjon til undersøkelsen av kapasitetshypotesen

Undersøkelsen av kapasitetshypotesen med vekt på den strukturelle kompleksiteten til den språklige konteksten og en rekke andre belastende faktorer i området rundt inversjonskonteksten, viste seg å gi positive resultater. Det viste seg riktignok at fokus *kun* på kompleksitet ikke ga noen positive profiler, men at ekstra belastningsfaktorer måtte trekkes inn. Dette betyr allikevel ikke at kompleksiteten ikke er avgjørende, bare at det ikke er tilstrekkelig for å forklare *alle* kontekstene. Mange kontekster var mer belastende bare i kraft av kompleksitet. Det at kompleksiteten ikke var tilstrekkelig, tyder bare på at det er flere faktorer som påvirker, og kanskje kan enda flere faktorer som ikke er tatt med her, avdekke positive mønstre også i de negative profilene.

Å knytte variasjonen til kapasitetsbelastning har altså gitt mange positive resultater. Det er dermed nærliggende å si at i mitt materiale ser det ut til at variasjonen kan være kapasitetsbetinget, selv om ikke alle tekstene støtter like godt opp under denne påstanden. Dette er bare en eksplorativ undersøkelse, så jeg vil ikke se på hvorvidt resultatene er statistisk signifikante, men nøye meg med å antyde at det er positive tendenser i materialet.

5.6 Vurdering av metoder og operasjonalisering (og positive)

Jeg vil nå se på en del metodiske svakheter med undersøkelsen.

5.6.1 Avgrensningen av den lingvistiske konteksten:

Jeg har i undersøkelsen avgrenset påvirkningskilden til å være foranstilt ledd, subjekt og finitt verbal, og i tillegg inkludert andre faktorer i området rundt konteksten, som har med setningen generelt å gjøre. Det er mulig at avgrensningen i utgangspunktet også skulle tatt med eventuelle setningsadverbialer, da dette viste seg å være en viktig utslagsgivende faktor. Dette er allikevel ikke noe problem, ettersom denne faktoren har blitt regnet med som ekstrarfaktor. Andre forhold som kunne vært regnet med, er binding til foregående kontekst med konjunksjoner.

Det er også sannsynlig at morsmålsfaktoren er sentral. Morsmålet kan ha innvirkning på hvor belastende en kontekst er. Karen Lund påpeker at de tyske og nederlandske elevene allerede kan ha automatisert inversjonen gjennom morsmålet, og at alt som skal til er å oppdage strukturen i målspråket. I mitt materiale var det ingen variasjonstekster fra morsmål som også har V2-regel, de nederlandske og tyske hadde alle kontekstene på plass. Noe som kan tyde på at disse har automatisert sine kontekster.

5.6.2 Operasjonalisering av kapasitetsbelastning

Jeg har vært nødt til å begrense antall kategorier for kompleksitet, og det kan hende at det hadde vært bedre med flere, og mer nyanserte, kategorier. Blant annet kunne jeg innlemmet anaforske korte ord i selve hovedkategoriseringen, og spesielt for kategori 5 (setninger) oppdaget jeg at det kunne vært greit å hatt flere underkategorier. I tekster med mer enn én leddsetning, rangerte jeg dem etter lengde, mens det kanskje hadde vært mer relevant å se litt på hvilken *type* leddsetning det var, og være litt mer nyansert i forhold til de ulike leddene i leddsetningen, for eksempel substantivfraser og tilstedeværelse av setningsadverbialer. Men jeg vil tro lengden et godt stykke på vei avspeiler slike forhold uansett.

Andre faktorer som skulle vært regnet med er frekvens, og kanskje en ytterligere nyansering av bøyningssendelser både for substantiv og verb.

Jeg mener allikevel at min noe grove inndeling, som tross alt avspeiler en skala fra ett ord til setning, har fungert tilfredsstillende og gitt gode resultater. Om det derimot er tilstrekkelig i en annen type materiale er usikkert.

5.6.3 Undersøkellesmetoden

Undersøkellesmetoden jeg har brukt, å sammenligne de ulike kontekstenes kompleksitetsrepresentasjoner, vil kanskje også være vanskelig å bruke i andre typer materialer og undersøkelser. I de fleste tekstene i mitt materiale fungerte denne metoden bra, fordi de fleste kontekstene lot seg sammenligne. Men jeg støtte også på problemer i noen av tekstene, hvor representasjonene var for ulike til at sammenligning var mulig. I et materiale med større variasjon av språklige uttrykk, spesielt for subjektet, som er lite variert i mitt materiale, vil kanskje metoden komme til kort og kreve for det første at kategoriene til

subjektet og det foranstilte leddet tilsvarer hverandre i hver kategori, og dessuten mer raffinerte retningslinjer, for å gjøre sammenligningen lettere.

Videre diskusjon av resultatene og metoden kommer i neste kapittel, hvor jeg sammenligner kapasitetshypotesen og identifiseringshypotesen.

Kapittel 6: DRØFTING

6.1 Identifisering eller kapasitet

6.1.1 Oppsummering av hypotesene

Målet med undersøkelsene jeg har gjort, har vært å finne ut hva som kan være årsaken til den variable anvendelsen av invertert ordstilling i deklarativsetninger. Jeg har satt fram to hypoteser og disse bygger på ulike forutsetninger om hva variasjon i mellomspråk skyldes, og hvor i prosesseringen av andrespråket variasjonen er lokalisert. Identifiseringshypotesen bygger på at variasjon er kunnskapsrelatert, knyttet til oppbyggingen av kunnskap om regelens distribusjon, mens kapasitetshypotesen bygger på at variasjon er kontrollrelatert, knyttet til innlærerens kapasitet til å anvende kunnskapen.

Identifiseringshypotesen tok utgangspunkt i at variasjon er resultat av en variabel kunnskap om inversjonsregelen, og en ufullstendig identifisering av foranstilte ledd som inversjonskrevende. Undersøkelsen fokuserte på det foranstilte leddets frekvens i kollokasjon med invertert ordstilling, med den hypotesen at jo mer frekvent leddet er, jo lettere er det å identifisere det som inversjonskrevende. Jeg har operert med et grunnleggende skille mellom *innputtbasert* identifisering og *slutningsbasert* identifisering, som går på hvorvidt innlæreren kan støtte seg på evidens i innputt, eller om det må slutninger til for å vite om det skal være invertert ordstilling eller ikke.

Kapasitetshypotesen tok utgangspunkt i at variasjon er resultat av varierende kapasitetsbelastning. Inversjonsstrukturen risikerer å bli utelatt om konteksten rundt krever for mye oppmerksomhet. Undersøkelsen fokuserte på den strukturelle kompleksiteten til det foranstilte leddet, subjektet og det finite verbet, samt en rekke ekstra belastningsfaktorer, med en hypotese om at jo mindre belastende konteksten rundt er, jo lettere er det å anvende invertert ordstilling.

Langt på vei avspeiler forskjellen på disse to Corders skille mellom *postsystematiske* feil og *presystematiske/ idiosynkratiske* feil, som han beskriver på følgende måte:

”... *such sentences could be accounted for as being cases of failure (for whatever reason) to follow a known rule, in contradiction to what I'm calling idiosyncratic sentences, which*

involve no failure in performance and cannot be corrected by the learner precisely because they follow the only rules known to them, those of his transitional dialect.”(Corder 1982: 18).

Det er en grunnleggende forskjell mellom disse to vinklene til variasjonen, som går på om variasjonen avspeiler et *eget system*, hvor innlærerne har andre regler enn målspråket, eller om variasjonen avspeiler en *systematisk påvirkning av et system* som er målspråkligt. Dette kan også knyttes til diskusjonen om variasjon er kompetanserelatert eller performanselatert, som jeg diskuterte i kapittel 2.

Følgende tabell illustrerer noen av forskjellene mellom hypotesene:

	Identifiseringshypotesen	Kapasitetshypotesen
Lokalisering i prosessering	<i>Resepsjon</i>	<i>Produksjon</i>
Utfordring	<i>Tolkning</i>	<i>Kapasitet</i>
Lingvistisk kontekst/påvirkningskilde	<i>Foranstilt ledd</i>	<i>Foranstilt ledd, subjekt, finitt verb, ektrafaktorer</i>
Avgjørende egenskaper ved den lingvistisk konteksten	<i>Frekvens og syntaktisk gjennomskiktighet</i>	<i>Strukturell kompleksitet, produksjon av form</i>
Variasjonsårsak	<i>Kunnskapsrelatert</i>	<i>Kontrollrelatert</i>
Underinversjonene er	<i>Presystematisk/ideosynkratiske</i>	<i>Postsystematiske</i>

6.1.2 Oppsummering av resultatene

Undersøkelsen av identifiseringshypotesen, viste at **ni** av tekstene i materialet mitt, innfridde prediksjonene om at de inverterte skulle være mer frekvent enn de ikke-inverterte. Noen flere profiler så mer positive ut når jeg dro inn tvetydighetsfaktoren. Bare fire profiler viste et motsatt mønster enn predikert, og fem var jevne profiler.

Undersøkelsen av kapasitetshypotesen ga enda tydeligere positive resultater, hele 11 av de 15 tekstene som lot seg undersøke, innfridde prediksjonene. Det var ingen entydige positive profiler når jeg bare så på kompleksitet, men en god del av minuskontekstene lot seg forklare av andre belastningsfaktorer.

6.1.3 De ulike tekstenes stilling til de ulike hypotesene

Det viser seg altså at de ulike profilene stiller ulikt forhold til hypotesene. Noen støtter begge hypotesene, noen støtter bare en av dem, og noen få støtter ingen. Tabellen under gir en oversikt over hvilke tekster som støtter hvilke hypoteser, og en helhetlig vurdering av profilen. **K** (støtter kapasitetshypotesen), **I** (støtter identifiseringshypotesen), **B** (støtter begge hypotesene) og '-' (støtter ingen av hypotesene):

	2	6	32	37	44	49	50	54	56	57	61	64	68	70	82	85	88	89	98
I	Ø	-	+	-	+	+	+	Ø	Ø	-	+	Ø	+	+	+	Ø	Ø	+	-
K	-		+	+		+	+	+	+			+	+	+	+	-	+	-	-
	--	--	B	K	I	B	B	K	K	--	--	K	B	B	B	-	K	I	-

Seks av tekstene ser ut til å støtte begge hypotesene, og det er to som er motprofiler til begge, to støtter bare identifiseringshypotesen, mens fem støttes bare av kapasitetshypotesen. At noen av tekstene bare kan predikeres av en av hypotesene, kan vitne om ulike strategier eller tilnærminger til inversjonsregelen, eventuelt at de er på ulike stadier i innlæringa. Dette vil jeg komme tilbake til lenger ned. De tekstene som predikeres av begge hypotesene, er interessante, fordi det tyder på et bestemt tendens i den lingvistiske konteksten, hvor frekvensen av det foranstilte leddet går parallelt med hele kontekstens kapasitetsbelastning, kanskje på den måten at subjektet og verbalet har en jevn belastning i alle kontekstene, mens det er det foranstilte leddets belastning som varierer.

6.1.4 Identifisering eller kapasitet? En diskusjon om problemer med hypotesene

Jeg har i kapittel 4 og 5 vært inne på mulige svakheter med hver av hypotesene på det metodiske plan, men det finnes også en rekke svakheter med dem på et mer overordna nivå, knyttet til selve oppdelingen i to ulike lokaliseringer av og årsaker til variasjon som ble omtalt i kapittel 2.

Begge hypotesene har den svakheten at de ser på årsakene de er knyttet til *isolert*.

Kapasitetshypotesen forutsetter en målspråkslik kompetanse, og tar egentlig ikke hensyn til at innlæreren ikke nødvendigvis har en målspråkslik kunnskap over regelen. Riktignok har jeg tatt med i betraktning eventuelle overgeneraliseringer, men ikke det faktum at noen kontekster

kan være uidentifiserte som inversjonskrevende. Dette er til en viss grad et problem i forhold til *the comparative fallacy*, fordi det forutsetter at kunnskapen om regelen er målspråkslik.

Identifiseringshypotesen har sine begrensinger fordi den ikke tar forbehold om at andre faktorer enn identifisering styrer anvendelsen av invertert ordstilling, faktorer knyttet til hvorvidt den etablerte kunnskapen faktisk blir brukt, som for eksempel automatisering¹⁸.

Det er vel egentlig mer realistisk at disse faktorene – identifisering og automatisering – kunnskap og kontroll - virker sammen, og det er vel heller ingen ny tanke innen språklæringsteori. Grunnen til at skillet ble gjort fra min side, var at det eksisterer ulike syn på hvor i prosesseringen av målspråket variasjonen oppstår, og det ble derfor viktig å skille dem fra hverandre av metodiske grunner. Robert Bayley, som er kjent innen sosiolingvistisk variasjonsteori, har satt frem '*The principle of multiple causes*': "... *it is unlikely that any single contextual factor can explain the variability observed in natural language data.*" (Bayley 2002: 118). Dette gjelder med stor sannsynlighet for mellomspråksvariasjon også. Enhver kontekst i et variasjonsmønster kan skyldes *uidentifiserte* faktorer, og i tillegg kan flere faktorer kan styre *samtidig*. En kan sikkert gå så langt som å si at flere faktorer VIL styre, at det heller er regelen enn et unntak, jf. Kapittel 2.2.1c om Ellis' frie variasjon.

Når det gjelder de antatte påvirkningskildene som mine hypoteser forutsetter (kunnskap og kontroll), er det derfor mer sannsynlig at disse virker sammen, enn hver for seg. Innlæringsprosessen består *både* av opparbeiding av kunnskap og opparbeiding av kontroll, og derfor er det mulig at en mer treffsikker hypotese vil ta utgangspunkt i et samspillet mellom disse to som årsaker til variasjon. Det er altså ikke *enten* et resultat av et idiosykkratisk mellomspråkssystem, *eller* en systematisk påvirkning på et utviklet system, det kan heller være snakk om en ***systematisk påvirkning av et eget system***. Noen kontekster er underinvertert fordi de ikke er identifiserte, andre fordi de er for belastende. En slik dobbelpåvirkning vil kanskje kunne forklare at lettidentifiserte kontekster allikevel ikke er inverterte, fordi de har for stor kapasitetsbelastning, mens kontekster som er minimalt belastende kan være underinverterte på grunn av manglende identifisering. Selv om et ledd er identifisert som inversjonskrevende av innlæreren, vil det kunne lide under høy belastning om det ikke er automatisert, og da er det lite hjelp en lettidentifisert kontekst.

¹⁸ I tillegg til at min operasjonalisering av identifiseringsfaktorer kan være for snever, som nevnt i kapittel 4.

Jeg vil på tampen av denne oppgaven forslå en mulig ny hypotese, *SPIS-hypotesen*, som ser på variasjon som et resultat av en *Systematisk Påvirkning av et Idiosykratisk System*, som kombinerer faktorene kunnskap og kontroll, og dermed ivaretar det symbiotiske, overlappende forholdet som eksisterer mellom disse to, og frir dem fra eventuelle svakheter knyttet til å isolere disse faktorene. En nærmere undersøkelse av denne hypotesen er det ikke tid og plass til i denne avhandlingen, men jeg ønsker å antyde at det kan være interessant å se nærmere på.

SPIS-hypotesen

Variabel anvendelse av invertert ordstilling er både kunnskaps- og kontrollrelatert, og derfor både resultat av variabel identifisering og variabel kapasitetsbelastning.

Det er allikevel ikke slik at denne prosessen med opparbeiding av kunnskap og kontroll går rett frem, lineært, først kunnskap og så kontroll. Det er heller en overlappende prosess, hvor trekk læres og automatiseres vekselvis. Én kontekst identifiseres og automatiseres, og neste kontekst læres og automatiseres. På denne måten er kanskje det første inversjonskonteksten automatisert lenge før en annen i det hele tatt er identifisert. Forskjellen mellom innputtbasert og slutningsbasert identifisering er interessant her, spesielt med henvisning til Krashens teori om at det som er tilegnet gjennom slutninger vanskeligere lar seg automatisere enn det som er lært mer ubevisst. Dette er ikke en tråd jeg skal følge opp, men det kan være et viktig innspill om en skal se på samspillet mellom identifisering og automatisering, ettersom automatiseringen kanskje ikke vil gå like lett med de slutningsbaserte kontekstene som med de innputtbaserte.

I tillegg til dette kan innlærere være ulikt orientert når det kommer til opparbeiding av kunnskap og kontroll, noe Ellis viser i sin "*Bidimensional model of second language proficiency*" (Ellis 1989:40). Ulike innlærere har ulik fokus når det kommer til kunnskap og kontroll, og Ellis deler i 3 typer innlærere:

"Type A learner develops a high level of control but fails to advance sociolinguistic knowledge to any extent. Type B learner is the opposite; considerable sociolinguistic competence is developed but there is little control over this (i.e. the learner can act on this competence only in planned discourse). The type C learner is more balanced, progressing equally on both dimensions." (Ellis: 41)

Sett i forhold til disse innlærertypene, kan det i denne konteksten tenkes at innlærerne har ulike orienteringer mht inversjonsregelen. Noen vil fokusere på kunnskapen om distribusjon, og dermed identifisere flere kontekster, men ikke tåle så høy kapasitetsbelastning, mens andre er mer orienterte mot automatisering, og mindre mot identifisering. En tredje gruppe igjen kan ha en mer balansert orientering. Slike forhold er selvfølgelig også interessante å finne ut mer om.

6.1.5 En oppsummering

Jeg har nå drøftet om resultatene mine kan være et produkt av et for dramatisk skille mellom antatte årsaker til variasjon, og at det kanskje er enda mer fruktbart å se på samspillet mellom de faktorene jeg har tatt utgangspunkt i. Den oppdelingen i to hypoteser som jeg har gjort i denne avhandlingen, har allikevel vært metodisk viktig, fordi de tok utgangspunkt i to ulike paradigmers holdinger til variasjonsfenomenet, og antatt lokalisering av variasjonens opphav i prosesseringen – om det avspeiler en variabel kompetanse, eller en systematisk påvirkning på en homogen kompetanse.

Det var også positiv støtte i materialet til begge hypotesene, da det bare var noen få tekster som ikke støttet en eller begge hypotesene. Håpet er altså at andelen positive profiler blir større i en undersøkelse som slår hypotesene sammen på en mer utarbeidet måte en jeg har gjort her. Her kan man også bruke materiale fra ulike oppgavetyper, ulike kommunikative situasjoner, for å bedre isolere hva som er kunnskapsrelatert og hva som er kontrollrelatert.

Det er ikke påvist at svaret ligger i en forening av de to faktorene, dette er bare et forslag, en tolkning av profilene og resultatene, og det er godt mulig at det heller er individuelle forskjeller hos innlærerne som er grunnen til at noen støtter den ene eller den andre hypotesen, enten individuelle faktorer, som morsmål og utdanning, eller innlærernes ulike orienteringer i kunnskaps- og kontrolldimensjonene. Det er klart at slik faktorer må bringes inn i en mer omfattende undersøkelse.

6.3 Synkroni og diakroni

En stor del av variasjonsteoriene innen mellomspråksforskningen er påvirket av sosiolingvistiske teorier om variasjon. Her er det en allmenn oppfatning at synkron variasjon avspeiler diakron variasjon (historisk utvikling) (Bayley 2002: 117). Dette knyttes ofte til Baileys *bølgeteori*, hvor variasjon i bruk av trekk hos en gruppe mennesker skyldes at trekket sprer seg som ringer i vann, fra en gruppe til en annen.

Som jeg var inne på i kapittel 2, hevder også noen at mellomspråksvariasjon er en direkte avspeiling av utvikling over tid, og Ellis skriver:

"The changes that take place over time as the learner revises some interlanguage rules in the direction of the chosen target language norms. The principle thesis which will be explored is that horizontal variability is not only a window for viewing the process of L2 development but is concomitant to this process." (Ellis 1989: 36)

Ellis ser bort fra et skille mellom kompetanse og performanse, og han hevder at den variasjonen som er tilstede på ett tidspunkt i utviklinga er et fingeravtrykk av den diakrone utviklingsmønsteret. I dette ligger det at de forholdene som styrer bruk og ikke bruk av en form på ett tidspunkt i utviklinga, er de samme som styrer en spredning fra kontekst til kontekst over tid. Fra det tidspunktet trekket ikke er i bruk, går man over til et variasjonsperiode, frem til det tidspunktet trekket anvendes målspråklikt, om dette skjer. Variasjonsmønsteret vil endre seg, i takt med at nye kontekster faller på plass. Ikke nødvendigvis fordi flere og flere kontekster blir korrekte, men fordi mer og mer krevende kontekster automatiseres. Det er snakk om endringer i et system, et system som hele tiden restruktureres. Om en kartlegger endringene i dette systemet, vil en kanskje finne at det vokser, både i retning av målspråkssystemet og med egne kreative konstruksjoner.

I identifiseringshypotesen ligger det en slik implikasjon. Invertert ordstilling er lettere å lære i de høyfrekvente kontekstene, og de unike kontekstene krever slutninger på et høyere plan for at invertert ordstilling skal anvendes. Det dreier seg om opparbeiding av kunnskap, og dette vil si at i tillegg til å predikere synkron variasjonsmønster, vil det også ha påstander om utviklingen over tid. Hele forutsetningen for kunnskapsrelatert variasjon, er at det avspeiler en utvikling over tid, en gradvis utvikling av kunnskap om distribusjon, som utvider seg på en

forutsigbar måte. Et annet forhold med kunnskapsrelatert variasjon, er at den er *stabil*, på den måten at variasjonen vil være den samme i ulike situasjoner i samme periode av utviklinga.

Det er kanskje lettere å legge merke til forholdet mellom synkron og diakron variasjon når det gjelder kunnskapsrelatert variasjon, enn når det gjelder kapasitetsbetinget variasjon.

Sammenhengen mellom utvikling over tid og kontrollvariasjon da heller ikke like anerkjent, ettersom det av flere regnes utelukkende som resultat av *performanseforhold*; regelen er lært, men den glipper på grunn av forholdene rundt produksjonen, kun relatert til situasjonen og mangel på kapasitet der og da. Ulike situasjoner med ulikt press på innlæreren vil kanskje føre til at den maksimale belastning varierer, og at innlæreren får flere kontekster på plass jo mindre belastende situasjonen er, og jo mer tid til selvkorrigerende som er tilgjengelig. Men allikevel er det ikke slik at anvendelsen av regelen bare er et spørsmål om tid til rådighet. Det er også et spørsmål om automatiseringsgrad, og hvor mye oppmerksomhet et trekk og den nærmeste konteksten krever for å anvendes.

Det vil være slik at i hver fall to forhold styrer hvor stor belastning en kan tåle: For det første hvor automatisert trekkene er, hvor god kontroll man har, og for det andre tiden en har til rådighet. Kapasitetsbelastningen kan derfor regnes over 2 dimensjoner; graden av press som ligger i situasjonen, i hvor stor grad trekket og konteksten rundt er automatisert. Jo mer automatisert trekket er, jo mindre betydning har tiden en har til rådighet å si. Jo mindre automatisert og jo mer oppmerksomhet noen krever, jo mer avgjørende er det at situasjonen er enkelt og at man har mer tid til rådighet.

Den lingvistiske konteksten til et trekk kan derfor være lite belastende fordi 1) den er kort enkelt og umarkert, eller 2) den er automatisert, eller 2) det er god tid til å utøve kontrollert prosessering.

6.3.1 Metodiske problemer med dette synet på variasjon

Tarone (1988: 116) sier at skal man være forsiktig med å trekke konklusjoner om utvikling over tid på grunnlag av tverrsnittundersøkelser. Det er allikevel ikke usannsynlig at synkron variasjon kan regnes som en avtrykk av en utviklingsprosess. Det krever imidlertid mer omfattende, undersøkelser for å kunne fastslå dette, både av diakron variasjon, og av variasjon i ulike kommunikasjonssituasjoner.

En kunnskapsrelatert variasjon, som jo vil avspeile et system på et bestemt tidspunkt i utviklinga, vil for eksempel være tilnærmet lik fra situasjon til situasjon, og bare endre seg over tid ettersom systemet restruktureres. Ved en sammenligning av variasjonsmønstrene hos én og samme innlærer i to ulike situasjoner, vil variasjonsmønstrene være relativt like, uavhengig av hvor mye oppmerksomhetspress situasjonen tilfører innlæreren.

Det er selvfølgelig et viktig poeng at forskjellene mellom ulike situasjoner ikke bare dreier seg om tid og oppmerksomhet, men også at ulike kommunikative kontekster fører til bruk av ulike språklige uttrykk og formuleringer. Variasjonsmønstrene blir derfor ikke nødvendigvis like i de ulike situasjonene med hensyn til hvor stor del som er invertert og ikke invertert, men med hensyn til hvilke *typer* kontekster som er inverterte. Variasjonsmønstrene vil derfor bare være like i den grad de har de samme typer kontekster.

6.3.2 diakroni og synkroni?

Jeg ønsker ikke å legge frem noen konklusjon om hvorvidt de variasjonsmønstrene jeg har funnet i mitt materiale er en avspeiling av en diakron mellomspråksvariasjon, eller noe generelt utviklingsmønster, på den måten at den er et synkront uttrykk for at trekket faller på plass gradvis i tyngre og tyngre kontekster. Min undersøkelse er bare av variasjonen i en bestemt situasjon, og derfor vet jeg ingenting om hvilke variasjonsmønstre innlærerne har vist før, eller senere. Det er allikevel en implisitt påstand i hvert fall i identifiseringshypotesen om at inversjon faller på plass tidligere i noen kontekster enn andre, selv om undersøkelsen er begrenset.

Det som ville vært interessant er en undersøkelse som ivaretar samspillet mellom opparbeiding av kunnskap og kontroll, og samtidig sørger for å kontrollere for faktorer som ligger i kommunikasjonssituasjoner, slik at man eventuelt kan sortere ut hvilke deler av variasjonen som har sitt opphav i et idiosykratisk mellomspråkssystem, og hvilke som kan tilskrives møtet mellom grad av kontroll og forholdene med den kommunikative situasjonen.

6.4 Metodiske forbehold

Jeg har i slutten av kapittel 4 og 5 vært inne på noen metodiske svakheter og forbehold med undersøkelsen min. Jeg vil nå se på en rekke andre metodiske forbehold.

6.4.1 Feilkilde: bruk av skriftlige tekster

Materialet i denne oppgaven har vært skriftlige tekster, som er langvarsoppgaver fra Språkprøven i norsk for voksne innvandrere. Å bruke et skriftlig materiale har sine svakheter, fordi denne typen språkproduksjon ikke har høy grad av spontanitet. Innlæreren har tid å korrigere språkbruken, i tillegg til at det er tid til å tenke over hver setning, på en måte som ikke er mulig i for eksempel muntlig språkbruk.

Dette har konsekvenser for generaliserbarheten til resultatene mine. Spesielt for kapasitetshypotesen må en ta forbehold om at innlærerne har hatt tid til å se over og korrigere språkbruken sin, noe som kan ha resultert i at innlæreren for eksempler har anvendt mer belastende kontekster enn i han ville gjort i en annen mer spontan situasjon. Men det faktum at det faktisk finnes underinverterte setninger i tekstene, kan tyde på at det ikke har vært nok tid til eventuell selvkorrigerings og anvendelse av ikke-automatisert kunnskap. Det er derfor ikke sikkert at resultatene og de konklusjonene de har ført til lar seg overføre til også å gjelde muntlig språkbruk, eller andre typer skriftlige kontekster.

En annen faktor er at tekstene er skrevet i en generell *testsituasjon*, noe som har sine fordeler og ulemper. Fordelen er med testsituasjonen er at jeg unngår at kandidatene er spesielt fokusert på akkurat på det trekket jeg skal undersøke. Hadde jeg brukt et elisert materiale med fokus på inversjonsstrukturen, er det større sjanse for at kandidatene hadde hatt en fokus på dette trekket som kunne virket negativt inn på validiteten. Det er likevel en ulempe med testsituasjonen, nemlig at det er en situasjon hvor fokuset på form vil være høyt. Testen blir tatt for å dokumentere norskkunnskaper, og innlærerne vil sannsynligvis være spesielt nøye med selvkorrigerings o.l. Som jeg var inne på over, vil det også innebære lav grad av spontanitet i språkbruken.

Dette er altså en undersøkelse av et skriftlig materiale, og generaliserbarheten begrenses til skriftlige tekster. Det ligger allikevel en rekke påstander i hypotesene og operasjonaliseringene som også kan gjelde for muntlig, spontan kommunikativ språkbruk. Identifiseringsprosesser og kapasitetsbelastning vil være generelle prosesser som også gjelder her, selv om forholdene rundt kommunikasjonen, spesielt med tanke på kapasitetsbelastning, vil variere mellom oppgavetyper.

Et annet viktig spørsmål er om undersøkelsen faktisk undersøker det den påstår at den undersøker. Dette har jeg blant annet forsøkt å sørge for i identifiseringshypotesen, hvor jeg ser etter eksempler som kan utelukke at det er det foranstilte leddets kompleksitet som styrer. Dette var nødvendig på grunn av den gjennomgående ikonisiteten som finnes mellom språklige uttrykks frekvens og lengde. Slike eksempler fant jeg dessverre ikke i alle de positive profilene, som vi så at noen av de positive profilene ikke har utelukket kompleksitet som årsak. Men som Tarone skriver i sin vurdering av ulike teorier (Tarone 1988:24): om to teorier ellers gir like resultater, vinner den mest estetiske av dem. Jeg mener identifiseringshypotesen er en klar estetisk og intuitiv vinner, fordi den knyttes til en realistisk forklaring på hvordan det foranstilte leddet *alene* kan ha konsekvenser for anvendelse av invertert ordstilling.

6.4.2 Hypotesenes begrensninger

Eventuelle begrensninger med hver hypotese i forhold til operasjonalisering og øvrig metode, er tatt opp i kapittel 4 og 5 etter hver av undersøkelsene, men jeg vil gjenta noen hovedpunkter her.

For det første er identifisering og kapasitetsbelastning bare to av flere operasjonaliseringer av kunnskapsrelatert og kontrollrelatert variasjon. For det andre finnes det alternative operasjonaliseringer av identifisering og kapasitetsbelastning enn de jeg har tatt utgangspunkt i (frekvens og strukturell kompleksitet), og disse kunne med fordel ha vært mer nyanserte.

Det må også nevnes at det finnes mange andre mulige styrende årsaker til variasjonen, både lokalt på setningsnivå, på et mer globalt *tekstuel* nivå, og ikke minst på det ikke-språklige nivået (for eksempel distraksjoner i testlokalet), men i denne oppgaven har jeg altså valgt å fokusere på den språklige konteksten. En faktor som kunne vært ønskelig å trekke inn i operasjonaliseringen av hypotesene, er innlærervariabler som morsmålsbakgrunn og

utdanning. Som jeg har vært inne på tidligere er det stor sannsynlighet for at morsmålet kan spille en sentral rolle i både identifiseringen og i hvor belastende en kontekst oppleves som. Av relevante trekk her er for eksempel om morsmålet har fast eller fri ordstilling, og i hvilken grad de lar pragmatikk styre ordstillingen. I tillegg er det flere ganger vist at det spiller en rolle om morsmålet har en egen V2-regel.

Det var i starten ønskelig å skulle ha en jevn fordeling av morsmålsbakgrunner i materialet mitt, slik at jeg kunne kontrollere for morsmålsfaktoren, og i tillegg undersøke om det var noen tendenser som kunne føres tilbake til morsmålsforskjeller. ASK-materialet i sin helhet, som jeg hadde planlagt å bruke, ble dessverre ikke ferdig i tide, så jeg måtte nøye meg med det som var tilgjengelig i ASK-test. Der fantes det to tekster fra hver morsmålsbakgrunn, men når jeg skulle plukke ut variasjonstekster var ikke fordelingen lenger jevn, blant annet var det *ingen* nederlandske, tyske, engelske og vietnamesiske tekster som hadde variasjon slik jeg hadde definert det. Dette vil si at det i denne undersøkelsen ikke er kontrollert for morsmålsbakgrunn, og jeg har av samme grunn ikke ønsket å gjøre noen videre undersøkelser av morsmål og variasjonsmønstre.

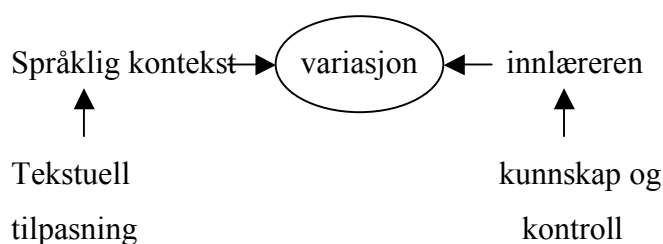
6.4.3 *Comparative fallacy*

Et stort metodisk varsku en må passe seg for er den før omtalte *comparative fallacy*. Jeg skrev tidligere at man ved en ekstern definisjon av variasjon, med å bare fokusere på obligatoriske kontekster, sto i fare for ikke å ta hensyn til mellomspråkets egne system.

Jeg har i min undersøkelse definert variasjon i forhold til en målspråksnorm: som : ”*en veksling i én og samme tekst mellom invertert og ikke-invertert ordstilling der en i målspråket har et grammatisk krav om invertert.*” (Johansen 2007: 26). Jeg har tidligere sagt at dette ikke er en sammenligningsbrist, fordi jeg skiller mellom fenomenet jeg ønsker å undersøke, og metoden for å gjøre det. I undersøkelsen har jeg tatt med overgeneraliseringer, for å også ta med i betraktning elementer som er særegent for et idiosynkratisk mellomspråkssystem. En annen grunn til at jeg har mitt på det rene, er rett og slett at jeg ikke fokuserer på feil og avvik, men heller skiller mellom anvendelse av en struktur, som er tilstede noen steder, og andre ikke.

6.5 Noen sluttbemerkninger om variasjon og koherens i tekst

Den variable anvendelsen av inversjon kan sies å være noe som oppstår i møtet mellom den språklige konteksten og innlærerens kompetanse i form av kunnskap og kontroll over inversjonsregelen. Når det gjelder dette har jeg gjort meg noen tanker om noe som på sett og vis kan være en parallell årsak til variasjonen. Variasjon og endringer i den språklige konteksten skaper variasjon i mellomspråket, men det som i sin tur skaper variasjonen i den språklige konteksten er den *tekstuelle tilpasningen* innlæreren gjør med budskapet, for å passe setningene inn i den tekstlige og situasjonelle konteksten. Det formelle innholdet i den språklige konteksten er resultat av tekstuell tilpasning hvor de setningene tilpasses til resten av teksten. Årsakene til variasjonsmønstrene kan derfor ligge på to nivåer, i den tekstuelle tilpasningen som skaper variasjon i den språklige konteksten, som igjen fører til mellomspråksvariasjon i møte med innlærerens kompetanse.



Ettersom både det foranstilte leddet og subjektet er nedslagsfelt for to av de mest sentrale kohesjonsmekanismer i språket – *tematisering og referentkobling* – synes jeg det hadde vært interessant å se på om variasjon kan knyttes til et overordna ønske fra innlæreren om å skape koherens, flyt, i teksten eller samtalen. Spesielt interessant er det at selve kravet om inversjon er resultat av at man tematiserer et annet ledd enn subjektet. Det vil være en spenning mellom den pragmatiske, kommunikativt viktige foranstillingen, og den grammatiske, kommunikativt tomme inversjonen, og kanskje er dette årsaken til at inversjon nedprioriteres. For subjektets del, vil valg av referensielt uttrykk til subjektets referanse føre til at det kan være veldig mange ulike språklige uttrykk her, alt fra pronomen til leddsetninger.

En videre tankegang rundt dette kan være fruktbar fordi det kan belyse noe av forskjellene mellom ulike ytringssituasjoner og ulike oppgavetyper, Fordi disse kan legge ulike føringer

for den tekstuelle tilpasningen. Dette ligger imidlertid utenfor denne oppgavens rammer, så jeg vil nøye meg noen fabulerende antydninger her avslutningsvis.

Kapittel 7: UTLEDNING

Jeg har i denne avhandlingen undersøkt variasjonsmønstre i anvendelsen av invertert ordstilling i mellomspråk, i håp om å finne en sammenheng mellom vekslingen av invertert og ikke-invertert ordstilling og den umiddelbare språklige konteksten. Hovedmålet har vært å finne ut om den språklige konteksten enten på grunn av dens rolle i identifiseringen av når det skal være inversjon, eller i kraft av sin kapasitetsbelastning, fører til en systematisk veksling mellom invertert og ikke-invertert ordstilling i målspråkets obligatoriske kontekster.

I undersøkelsen har jeg tatt utgangspunkt i to hypoteser som skulle representere to ulike syn på hva som forårsaker variasjon: For det første *identifiseringshypotesen* hevder at variasjon er et *kunnskapsrelatert* fenomen, og for det andre *kapasitetshypotesen* som hevder at variasjon er kontrollrelatert.

Undersøkelsene viste at 9 av 19 tekster lot seg predikere av identifiseringshypotesen, mens 11 av 15 tekster støttet kapasitetshypotesen. Seks av disse tekstene var de samme, så de støttet altså begge hypotesene. Selv om jeg ikke har fokusert på annet enn det rent deskriptive i fremleggingen av resultatene, har jeg tolket disse funnene positivt, og som et tegn på at det *er* en sammenheng mellom variasjonen og den språklige konteksten.

Egenskapene med den språklige konteksten som har vært brukt i undersøkelsen, er altså for identifisering det foranstilte leddets inntreffrekvens i foranstilt posisjon etterfulgt av invertert ordstilling. For kapasitetbelastning er det strukturell kompleksitet og en rekke ekstrarfaktorer, hvor spesielt midtadverbialer så ut til å spille en viktig rolle.

En variasjonsstudie som ser på mellomspråkets veksling mellom to former, uten nødvendigvis fokusere på om den ene er korrekt eller avvikende, kan stå i kontrast til studier som fokuserer på trekk som lært eller ikke lært, definert etter et bestemt tilegnelseskriterium (*development as sequence*), men variasjonsstudier kan også utfylle slike studier. I det *dynamiske paradigmet*, blant annet representert av Huebner (1983), ønsker man å avskrive denne typen studier¹⁹; men som Ellis sier, en bør både ha studier som fokuserer på det sekvensielle med innlæringen av

¹⁹ Hos Huebner omtales dette som the Quantitativ paradigm, og tilsvarer så vidt jeg kan se Ellis' development as sequence, og eksterne tilnærminger til variasjon.

trekk, og studier som ser på veksten til et trekk, kontekst for kontekst, *development as growth* (Ellis 1989: 37).

Om en kan trekke ut hvilke faktorer som styrer et variasjonsmønster, for eksempel den språklige konteksten, vil kunnskapen om disse faktorene kunne gi en større dybde i kunnskapen om hvordan trekk læres, gradvis. I en tilnærming som ser på om trekk er tilegnet eller ikke, kan det være vesentlig å også se på hvor stor variasjon det er i den språklige konteksten; forekommer trekket bare i de aller enkleste kontekstene, eller finnes også eksempler på de mest komplekse kontekstene? Fenomenet underrepresentasjon er interessant i denne sammenhengen, fordi lav frekvens av et trekk i innlærerspråk kan være like mye tegn på vansker som en ren feil, som Schachter (1974) har påpekt, og drar man dette videre vil det kanskje også være vesentlig å ta forbehold om underrepresentasjon av *typer* av kontekster i undersøkelser med fokus på tilegnelsespunkt. Om en skal se på om et trekk er tilegnet eller ikke, bør en også ta hensyn til hvilke kontekster som er representert, og man kan kanskje ikke si at trekket er tilegnet før det er anvendt både i de mest vanskelige kontekstene og i de enkleste. Undersøkelser av variasjon, kan gi oss innsikt i dette, i hvilke kontekster som er det lette og de tunge for et gitt trekk.

*

Tittelen '*Ja takk, begge deler*' har flere lag og oppsummerer denne avhandlingen på en kort og grei måte. For det første viser den til *studieobjektet*, at innlærerne anvender to former, både invertert og ikke-invertert, istedenfor én som i målpråket (Ja takk begge deler!). Den viser også til mitt *metodiske valg* om foreta undersøkelsen med utgangspunkt to hypoteser (Ja, takk begge deler!), og mitt avslutningsvise forslag om sammenslåing av disse to hypotesene for bedre å kunne avdekke det som ligger bak variasjonen, både kunnskap og kontroll (Ja takk begge deler!). Det viser til at variasjonen kanskje ikke bare er synkron, men også kan si noe om det diakrone ved innlæringa (Ja takk, begge deler!). Til sist understreker tittelen litt av poenget med et studie som dette, det å skulle fokusere på vekstdimensjonen i mellomspråksutviklinga, og ikke bare på om trekket er tilegnet eller ikke (men, ja takk, begge deler!).

LITTERATURKISTE

- Anderson, R. 1989. The Teoretical Status of Variation in Interlanguage Development. I S. Gass, C. Madden, D. Preston & L. Selinker *Variation in Second Language Acquisition Volum II: Psycholinguistic issues*, Side 46-64.
Clevedon: Multilingual Matters Ltd.
- Axelsson, M. 1994. *Noun phrase development in Swedish as a second language: a study of adult learners acquiring definiteness and the semantics an and morphology of adjectives*.
Stockholm University
- Bayley, R. 2002. The quantitative paradigm. I J.K. Chambers, P. Trudgill & N. Schilling-Estes *The Handbook of Language Variation and Change*, Side 117-141.
Blackwell Publishing Ltd.
- Bergreen, H. & Tenfjord, K. 1999. *Andrespråkslæring*
Oslo Gyldendal akademisk
- Bolander, M. 1988. Nu ja hoppas inte så mycke. Om inversion och placering av negation och adverb i svenska som andraspråk. I K. Hyltenstam & I. Lindberg *Första symposiet om svenska som andrespråk, Vol.1*, side 203-214
Stockholms Universitet
- Bley-Wroman, R. 1983. The Comparative Fallacy in Interlanguage Studies: The case og systematicity. *Language Learning* 33, 1-17.
- Brautaset, A. 1996. *Inversjon i norsk mellomspråk: En undersøkelse av inversjon i stiler skrevet av innlærere med norsk som andrespråk*.
Oslo: Novus forlag
- Carroll, S.E. 2001. *Input and Evidence. The raw material of second language acquisition*.
Amsterdam: Benjamins
- Corder, S. P. 1982. *Error Analysis and Interlanguage* (2. utgave.)
Oxford University Press,
- Ellis, R. 1989. Sources of Intra-Learner Variablity in Language Use and Their Relationship to Second Language Acquisition. I S. Gass, C. Madden, D. Preston & L. Selinker *Variation in Second Language Acquisition Volum II: Psycholinguistic issues*, Side 22-45.
Clevedon: Mulitlingual Matters Ltd.

Ellis, N.C. 2003. Constructions, Chunking and Connectionism: The Emergence of Second Language Structure. I C.J.Doughty, & M.H. Long, *The Handbook of Second Language Acquisition*, side 63-103
Blackwell Publishing

Faarlund, J.T., Lie, S. Og Vannebo, K.I. 1997. *Norsk referansegrammatikk*.
Oslo: Universitetsforlaget

Golden, A. 1994. *Ordforråd, ordbruk og ordlæring*
Oslo: Ad Notam Gyldendal

Hagen, J. E. 1992. Feilinvertering, overinvertering og underinvertering.
I NOA 15, s. 27-38

Hagen, J. E. 1998. *Norsk grammatikk for andrespråklærere*.
Oslo: Ad Notam

Huebner, T. 1983. *A longitudinal analysis of the acquisition of English*.

Lund, K. 1997. *Lærer alle dansk på samme måde? En længdeundersøgelse af voksnes tilegnelse af dansk som andetsprog*.
Special-pædagogiskforlag

McLaughlin, B. 1987. *Theories of Second-language Learning*.
London: Arnold

Meisel, clahse peineman2983

Meisel, J.M., Clahsen, H. Og Pienemann 1981. On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition* 3, 109-135.

Mitchell, R.& Myles, F. 1998. *Second Language Learning Theories*
London: Arnold

Pienemann, M. 1988. *Language Processing and Second Language Development. Processability theory*.
Amsterdam:Benjamins

Rickford, J. R. 2002. Implicational Scales
I J.K. Chambers, P. Trudgill & N. Schilling-Estes *The Handbook of Language Variation and Change*, Side 142-167.
Blackwell Publishing

Romaine, S. 2003. Variation. I C. J. Doughty & M.H. Long *The Handbook of Second Language Acquisition*, side 409 – 435
Blackwell Publishing

Saarik, K. 2004. *Iblant jeg må vente svare - En typologisk basert undersøkelse av tilegnelsen av V2-regelen i indikativsetninger hos estisk- og finsktalende innlærere av norsk som andrespråk*. Hovedfagsoppgave, nordisk institutt, Universitetet i Bergen.

Schachter, J. 1974. An Error in Error Analysis. *Language Learning* 24, 567-59.

Seliger, H. W. & Shohamy, E. 1989. *Second Language Research Methods*
Oxford University Press

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics* 10, 209-231

Sharwood Smith, M. 1994. *Second Language Learning: Theoretical Foundations*.
New York: Longman

Tarone, E. 1988. *Variation in Interlanguage*
London: Arnold

Vikøy, Aa. 2004. *Med italiensk som morsmål og norsk som målspråk : ein typologisk basert studie av italienske informantar sitt tileigningsmønster av norsk subjektstvang og inversjon*. Hovedfagsoppgave, nordisk institutt, Universitetet i Bergen.

Sammendrag

Tittel: *Ja takk, begge deler – en undersøkelse av variasjonsmønstre i anvendelse av invertert ordstilling i norsk mellomspråk.*

Forfatter: Hilde Johansen

Jeg har i denne avhandlingen prøvd å undersøke variasjonsmønstrene i anvendelse av invertert ordstilling i deklarativer hovedsetninger i norsk mellomspråk, i håp om å kunne knytte den til variasjon i de umiddelbare språklige omgivelsene til inversjonsstrukturen.

Jeg har hatt to hypoteser, *identifiseringshypotesen* og *kapasitetshypotesen*, som skilles ved at den ene legger til grunn at variasjon er knyttet til innlærerens kunnskap om inversjonsregelen, mens den andre hevder at det er knyttet til i hvor stor grad denne kunnskapen er automatisert.

Identifiseringshypotesen

Variasjonen er kunnskapsrelatert, og den variable anvendelsen av invertert ordstilling er styrt av en ufullstendig identifisering av når det skal være invertert ordstilling og når det ikke skal være det.

Kapasitetshypotesen

Variasjonen er kontrollrelatert, og den variable anvendelsen av invertert ordstilling er resultat av varierende kontroll over kunnskapen og varierende kapasitet til å anvende kunnskapen om invertert ordstilling.

Materiale

Jeg har bruk 19 skriftlige langvarsoppgaver fra Norsk andrespråkskorpus, som alle har variasjonsmønstre hvor de veksler mellom invertert og ikke-invertert ordstilling, der det er obligatorisk i norsk. Studien er en flerkasusstudie, med vekt på individuelle variasjonsmønstre. Jeg har ikke fremstilt resultatene statistisk, bare rent deskriptivt, da dette først og fremst er en eksplorativ studie.

Resultat

Undersøkelsene viste at 9 av 19 tekster lot seg predikere av identifiseringshypotesen, mens 11 av 15 tekster støttet kapasitetshypotesen. Seks av disse tekstene var de samme, så de støttet altså begge hypotesene. Selv om jeg ikke har fokusert på annet enn det rent deskriptive i fremleggingen av resultatene, har jeg tolket disse funnene positivt, og som et tegn på at det *er* en sammenheng mellom variasjonen og den språklige konteksten.

Jeg antyder i drøftingen min at en mer omfattende undersøkelse som slår sammen disse to faktorene – kunnskap og kontroll – som årsak til variasjon, og tar hensyn til det kontinuerlige samspillet mellom disse to dimensjonene i språkinnlæringen, muligens kan være enda mer fruktbar.

Primærkilder

Følgende litteratur har vært mest sentral i arbeidet med oppgaven:

- Bergreen & Tenfjord (1999) *Andrespråklæring*
- Ellis (1989) *Sources of Intra-Learner Variability in Language Use and Their Relationship to Second Language Acquisition.*
- Hagen (1992) *Feilinvetering, overinvetering og underinvetering.*
- Lund (1997). *Lærer alle dansk på samme måte? En længdeundersøgelse af voksnes tilegnelse af dansk som andetsprog.*
- Sharwood Smith (1994.) *Second Language Learning: Theoretical Foundations.*
- Tarone (1988) *Variation in Interlanguage*