

*Vanlige måter å gjøre feil

Forholdsvis hyppig forekommende, ugrammatiske konstruksjoner i norsk – at, når og hvorfor de brukes

Eirik Tjervåg
Mastergradsoppgave i lingvistikk
Institutt for lingvistiske, litterære og estetiske studier
Universitetet i Bergen
Høsten 2011

Forord

Da jeg begynte arbeidet med denne oppgaven så jeg for meg at den ville ende opp med å først og fremst dreie seg om forekomsten av én type hyppig forekommende, ugrammatiske konstruksjoner (måteNP), og den empiriske undersøkelsen av forekomsten slike. Slik har det ikke blitt. I oppgaven som foreligger nå, utgjør presentasjonen av og diskusjonen omkring disse konstruksjonene en forholdsvis liten del. Det er to grunner til dette: Den første grunnen er at det empiriske materialet ikke gav meg like store muligheter som jeg hadde håpet til å vise hvilken rolle spesifikke grammatiske konstruksjoner i måteNP spiller for antallet ugrammatiske forekomster av måteNP som er å finne i korpuset. Den andre grunnen er at jeg under arbeidet med oppgaven ble oppmerksom på flere konstruksjoner det kan stilles spørsmål om grammatikaliteten til enn jeg var klar over i utgangspunktet, som resumptive pronomen og langdistanseavhengigheter (og da spesielt øyrestriksjonene knyttet til disse). Særlig slike konstruksjoners stilling i de skandinaviske språkene har jeg vurdert som relevante i denne oppgaven. Selv om den delen av oppgaven som dreier seg om måteNP og den empiriske undersøkelsen har endt opp langt mindre enn først tenkt, kan resultatene fra de empiriske undersøkelsene som er foretatt i denne oppgaven være med på å vise egenskaper ved parseren som kan være relevant i diskusjonen omkring andre problematiske konstruksjoner enn måteNP.

Jeg vil takke veilederen min, professor Helge Dyvik, for verdifulle tilbakemeldinger på utkast til oppgaven. Jeg er også takk skyldig alle som under arbeidet har lånt meg et øre og en oppfatning når jeg har behøvd andre parsere enn min egen til å vurdere akseptabiliteten til setninger med vanskelige konstruksjoner.

Innhold

Del 1

1	Grammatikalitet og akseptabilitet	5
1.1	Kompetanse, performans og grammatikk	5
1.2	Grammatiske og akseptable ytringer	7
1.3	To akseptabilitetsbegreper	9
1.4	Forståelig/uforståelig	11
1.4.1	Akseptable, ugrammatiske ytringer	12
1.4.2	–Forståelig	13
1.4.3	Hvorfor er flertydige setninger problematiske?	14
1.4.4	+Forståelig	16
1.5	Oppsummering	17
2	Rekursjon	19
2.1	Iterativ og ikkeiterativ rekursjon	19
2.2	Hvorfor er dyp egeninnføyning vanskelig?	20
3	Langdistanseavhengighet	23
3.1	Setningsknuter	23
3.2	Er bruken av RP grammatisk eller akseptabel?	24
3.3	Svensk RP: kompetanse, performans og normering	25
3.4	RP som 'last resort'	27
4	Referansebegrensninger	32
4.1	Begrensninger for personlige pronomen	32
4.2	Begrensninger for gap	33
4.2.1	Parasitic gaps	33
4.2.2	Er PG pronominal?	35
4.3	RPs mulige antesedenter	37
4.3.1	Er RP fonologisk uttrykk for en gap?	37
5	Øyrestriksjoner	39
5.1	Øyrestriksjoner og akseptabilitet	39
5.1.1	'The Complex NP Constraint'	40
5.2	Norske øyrestriksjoner	41
5.2.1	Uenighet om norske øyrestriksjoner	43

5.2.2	Hvorfor er norsk unntatt øyrestriksjonene?	44
5.3	Øyrestriksjoner og grammatikalitet	47
6	Mulige årsaker til feil i parsingen	49
6.1	Parserens egenskaper	49
6.2	Parserens begrensninger	50
6.3	Parsing av garden path-konstruksjoner	51
7	Preposisjoner som forsvinner	52
7.1	Tilgrensende preposisjoner	52
7.2	Strandede preposisjoner	53
Del 2		
8	Grammatiske feil i måteNP	56
8.1	Rollen til <i>måte</i> i måteNP	56
8.2	Mulige årsaker til utelatt måteP	56
8.3	Tidlig lukking av måteNP	57
8.4	Fraværet av måteNP i engelsk	58
9	Metode	61
9.1	Materiale	61
9.2	Skriftlig språkproduksjon og -persepsjon	61
9.3	Utvalg	62
10	Resultater	64
10.1	Tallene	64
10.1.1	Lengden på måteF	64
10.1.2	Andel grammatiske måteNP	65
10.1.3	Andel måteNP med verb	67
10.1.4	Andel grammatiske måteNP med og uten verb	68
10.1.5	Grammatiske måteNP med og uten preposisjoner	69
10.1.6	Konkrete preposisjoner	69
10.2	En ekstra <i>på</i> i måteNP	71
10.3	RP og <i>på</i> som 'lukkere'	72
11	Konklusjon	76

0 Innledning

Denne oppgaven vil ta for seg konstruksjonstyper i norsk som har en tendens til å ende opp som ugrammatiske ved bruk. Oppgaven er delt inn i to deler: Del 1 (kapittel 1-7) gjør rede for begrepsbruken i oppgaven, og gir eksempler på typer konstruksjoner det kan være vanskelige å bruke og å denotere betydningen til. Del 2 (kapittel 8-10) tar for seg én av disse konstruksjonstypene (som jeg kaller 'måteNP')¹, og legger frem en empirisk undersøkelse av forekomster av måteNP i norske avistekster.

Kapittel 1 forklarer hvordan begreper som grammatikalitet, akseptabilitet, kompetanse og performans vil brukes i denne oppgaven. Kapittel 2 viser hva det er som kan gjøre noen typer ikkeiterative konstruksjoner vanskelige å bruke. Kapittel 3 tar for seg langdistanseavhengigheter, og vurderer hvorvidt konstruksjoner med slike som inneholder resumptive pronomen er grammatiske eller ikke. Kapittel 4 sammenligner bruken av – og referansemulighetene til – gap og resumptive pronomen i langdistanseavhengigheter, og viser hvorfor det ikke er likegyldig om det finnes et pronomen eller ikke i en 'parasitic gap'. Kapittel 5 diskuterer hvorvidt visse øyrestriksjoner som gjelder for f.eks. engelsk gjelder for norsk eller ikke, og i hvilken grad brudd på øyrestriksjonene kan 'reddes' av resumptive pronomen. Kapittel 6 beskriver parserens egenskaper, og viser på hvilken måte visse konstruksjoner kan forårsake feil under parsingen. Kapittel 7 presenterer eksempler på realiserte ytringer som av ulike, men lignende grunner ikke følger de grammatiske reglene, og gir mulige forklaringer på hvorfor slike konstruksjoner kan gi parseren problemer produsere en grammatisk setning. Kapittel 8 gjør rede for mulige problemer parseren kan få ved produksjon av måteNP, og viser hvorfor tilsvarende konstruksjoner ikke finnes i engelsk. Kapittel 9 beskriver materialet og metoden for den empiriske studien i denne oppgaven, og vurderer hva man kan forvente av en empirisk undersøkelse med avistekster som materiale i forhold til en tilsvarende undersøkelse av talt språk. Kapittel 10 tar for seg funnene i studien, gir noen eksempler og tall fra søkene i NA, og viser fellestrekk mellom måteNP og eksemplene på 'vanskelige' konstruksjoner i del 1. Kapittel 11 runder av oppgaven, og diskuterer hvordan det kan ha seg at ugrammatiske setninger i visse situasjoner kan være bedre egnet til kommunikasjon enn grammatiske.

¹'måteNP' er nominalfraser med *måte* som kjerne, som er objekt til en preposisjonsfrase med en strandet *på* (måteP) som kjerne: *Per fant en bedre måte å spise fisk på* (her er måteNP understreket og måteP uthevet).

DEL 1

1 Grammatikalitet og akseptabilitet

I dette kapitlet vil jeg gjøre rede for begreper jeg kommer til å bruke i behandlingen av det teoretiske og empiriske materialet i løpet av oppgaven. De fleste av disse begrepene er velkjente, men jeg vil her spesifisere nøyaktig hva det er jeg mener når jeg bruker begreper som *grammatisk*, *akseptabel*, *kompetanse* og *performans*.

1.1 Kompetanse, performans og grammatikk

Denne oppgaven bruker Chomskys (1965) distinksjon kompetanse/performans: *We thus make a fundamental distinction between competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations)*. (s. 4). Kompetansen er språket som det er tenkt brukt av det Chomsky kaller en 'ideell taler-lytter':

(...) an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. (Chomsky, 1965, s. 3)

Grammatikken i et språk utgjøres ifølge Chomsky av denne tenkte taler-lytterens kompetanse: *A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence*. (s. 4). I denne oppgaven vil jeg ikke forsøke å gi en beskrivelse av denne kompetansen, men vise eksempler på tilfeller der avvik fra grammatikken (eller det som synes å være grammatikken) kan skyldes det Chomsky refererer til som *memory limitation* (her: parserens begrensninger). Jeg vil også gi mulige forklaringer på hvorfor enkelte setninger er vanskeligere å bruke enn andre, hvorfor enkelte grammatiske setninger ikke regnes som akseptable av et stort antall språkbrukere, og hvorfor visse grammatiske, akseptable setningers proposisjon ikke så lett kan forstås av lytteren.

Med 'performans' menes det språkbrukeren rapporterer å vite om språket. Performansen ville for en ideell taler-lytter utgjøres av kompetansen alene, men det er flere elementer som påvirker hvordan språket faktisk brukes: Performansen for vanlige, ikke-eideelle språkbrukere i et heterogent språksamfunn (som er situasjonen for alle eksisterende språkbrukerne) utgjøres av (i) grammatisk relevante forhold (kompetansen) og (ii)

grammatisk irrelevante forhold (minnebegrensninger, distraksjoner, kontekst osv.). I denne oppgaven vil det hovedsaklig være minnebegrensningene det fokuseres på. Figur 1 er en enkel², skjematisk fremstilling av hvordan kompetansen virker sammen med parseren for taleren og lytteren under ytringen av en setning.

FIGUR 1:



Kompetansen er språkbrukerens kunnskap om språket. Parseren prosesserer språk (både ved produksjon og persepsjon) ved bruk av kompetansen. Både produksjon og persepsjon av språk (performans) er avhengig av to elementer for å fungere: Kompetansen og parseren. Når vi produserer setninger, (1) behandler talerens parser proposisjonen taleren vil formidle (ved bruk av kompetansen) på en måte som produserer en grammatisk setning, mens (2) lytterens parser behandler setningen (ved bruk av kompetansen) på en måte som avleder proposisjonen taleren forsøker å formidle. Chomskys tenkte ideelle taler-lytter er uavhengig av parseren, og en prosess som den mellom taleren og lytteren som vist i figur 1 – der både taleren og lytteren er ideelle – vil alltid ende opp med vellykket, språklig kommunikasjon. Selv om ingen virkelig språkbruker er av denne typen, ender som oftest en slik prosess (også mellom virkelige språkbrukere) med vellykket kommunikasjon. Dette fordrer at ingenting skjærer seg underveis, verken for talerens produksjon eller lytterens prosessering. I grove trekk er det tre ting som kan gå galt under prosessen som er beskrevet i figur 1:

- *Kompetansen* i det gjeldende språket er ikke til stede hos taleren eller lytteren (ved at taleren produserer en setning som ikke kan denotere proposisjonen, eller at lytteren ikke forstår hva setningen betyr).
- *Parseren* klarer ikke å håndtere relasjonene mellom elementene i setningen.³
- *Setningen* kan være flertydig, og avlede flere enn én proposisjon for lytteren.⁴

²Med 'enkel' menes at den kun tar hensyn til minnebegrensninger i språkbruk, og ikke inkluderer kontekst eller andre elementer som kan virke inn på språkbruken.

³Hvilke årsaker det kan være til at parseren 'svikter' vil jeg komme tilbake til i kapittel 6

⁴Dette vil jeg komme nærmere inn på i 1.3

Denne oppgaven vil kun ta for seg de *språklige* aspektene ved kommunikasjon, og jeg vil kun gjøre rede for situasjoner der det *prøves* å formidle en proposisjon. Jeg vil bruke betegnelsen 'ugrammatisk' om ytringer jeg mener det er grunner til å regne for å bryte med de grammatiske reglene av årsaker det er mulig å forklare. Med begreper som 'grammatiske regler' og 'grammatikalitet' mener jeg ikke objektive, absolutte regler for språket (noe det neppe er mulig å definere), men noe som skiller seg fra språkbruken og akseptabilitetsbegrepet ved at den (grammatikken) er uavhengig av parseren, og er noe parseren (som en ubevisst språkprosessor) bruker som retningslinjer under prosesseringen av språk. Når det ytres uventede konstruksjoner (f.eks. som inneholder resumptive pronomen), er det ikke gitt at disse utelukkende er et resultat av at grammatikken tillater dem; de kan like godt være et resultat av at parseren 'svikter' og velger løsninger som gjør det lettere for lytteren å forstå hva det er taleren ønsker å formidle. Denne bruken av ugrammatiske, kommunikative 'nødløsninger' vil jeg komme tilbake til senere, særlig i 3.4.

1.2 Grammatiske og akseptable ytringer

I 'Syntactic Structures' (1957) omtaler Chomsky forholdet mellom grammatikalitet og akseptabilitet på følgende måte (min understreking):

The fundamental aim in the linguistic analysis of a language L is to separate the grammatical sequences which are the sentences of L from the ungrammatical sequences which are not sentences of L and to study the structure of the grammatical sequences. The grammar of L will thus be a device that generates all of the grammatical sequences of L and non of the ungrammatical ones. One way to test the adequacy of a grammar proposed for L is to determine whether or not the sequences that it generates are actually grammatical, i.e., acceptable to a native speaker, etc. (Chomsky, 1957, s. 13)

I denne definisjonen er grammatikalitet og akseptabilitet det samme: En akseptabel sekvens er en grammatisk sekvens. En forutsigbar grammatikk forutsetter dermed at 'akseptabilitet' er like kontekstuavhenig som 'grammatikalitet', og at alle språkbrukere alltid er like godt i stand til å vurdere hvorvidt en sekvens tilhører språket eller ikke (som en *ideal speaker-listener*. Slik er det ikke i naturlige språk.

I 'Aspects of the Theory of Syntax' (1965) bruker Chomsky 'grammatikalitet' og 'akseptabilitet' som to separate begreper, og introduserer dikotomien competence/performance.

The term 'acceptable' is not to be confused with 'grammatical'. Acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence. (...) the scales of grammaticalness and acceptability do not coincide. Grammaticalness is only one of many factors that interact to determine acceptability. Correspondingly, although one might propose various operational tests for acceptability, it is unlikely that a necessary and sufficient operational criterion might be invented for the much more abstract and far more important notion of grammaticalness. (Chomsky, 1965, s. 11)

Et annet begrep for akseptabilitet i denne definisjonen kunne vært 'brukbarhet'. Det er ikke alle grammatiske sekvenser brukere av et språk L er like godt i stand til å bruke. Uakseptable, grammatiske sekvenser blir sjelden eller aldri brukt (og vanskeligere forstått når en bruker av L får dem presentert for seg). En akseptabel setning er ifølge Chomsky *immediatly comprehensible without paper-and-pencil analysis* (1965, s. 10). Det er mulig å forstå uakseptable setninger, men når de brukes kreves det mer innsats (og kanskje mer kunnskap om grammatikk) fra lytteren enn når det brukes akseptable setninger.

Blant faktorene som spiller en rolle for hvorvidt en sekvens er akseptabel eller ikke nevnes intonasjon, stilistikk, diskurs og begrensninger i minnet. Sistnevnte må regnes for å være det som spiller den viktigste rollen, siden bruken av uakseptable setninger nesten alltid forekommer når det brukes konstruksjoner som åpenbart krever mye av minnet. Eksempler på slike konstruksjoner er (Chomsky, 1965, s. 12):⁵

- (1) a. nested constructions ('called [the man who wrote the book that you told me about] up')
- b. self-embedded constructions ('who the boy [who the students recognized] pointed out')
- c. multiple-branching constructions ('[John], [Bill], [Tom], and [several of their friends] visited us last night')
- d. left-branching constructions ('[[[[John]'s brother]'s father]'s uncle]')
- e. right-branching constructions ('[this is [the cat [that caught [the rat that stole the cheese]]]')

⁵(1-b) betegnes av Chomsky (1965) som den konstruksjonen som bidrar mest til inakseptabilitet, etterfulgt av (1-a).

Lyons (1968, s. 137) påpeker at begrepet 'akseptabilitet' *is a more primitive term than either 'grammatical' or 'meaningful' in the sense that, unlike these terms, it does not depend upon any technical definitions or theoretical concepts of linguistics*. Nettopp dette (for ikke å snakke om muligheten for variasjon i akseptabilitetsvurderingene fra språkbruker til språkbruker) gjør det vanskelig å behandle akseptabilitet på en objektiv måte, selv om det sjelden er stor variasjon mellom språkbrukerne i vurderingen av hva som er en akseptabel setning.⁶ I det følgende vil jeg selv gi vurderinger av akseptabilitet der det er grunn til å tro at det store flertallet av språkbrukere vil være enige med min vurdering. Noen steder vil jeg gi bevis for påstandene om akseptabilitet.

1.3 To akseptabilitetsbegreper

Definisjonen Chomsky (1965) gir av 'akseptabel' lar begrepet referere til ytringer som er (i) *perfectly natural and* (ii) *immediately comprehensible without paper-and-pencil analysis, and* (iii) *in no way bizarre and outlandish* (s. 10) (min nummerering). Denne definisjonen 'siler ut' ytringer som er mer eller mindre 'ubrukbare' for språkbrukerne, som dype setningsinnføyninger (*Mannen hunden katten fryktet jaget falt*), som for mange kan virke som kun en oppramsing av ord, og som selv for lingvister krever en grad av analyse for å gi mening. Slike setninger kan ikke betegnes som 'perfectly natural' (i). Et annet eksempel på en grammatisk, uakseptabel setning med denne definisjonen vil være *Colorless green ideas sleep furiously* (Chomsky, 1957, s. 15), der (i motsetning til i *Mannen hunden...*) det syntaktiske forholdet mellom konstituentene lar seg forstå, men proposisjonen til ytringen er logisk umulig. Disse setningene kan betegnes som 'outlandish' (iii). En type setninger som både havner innenfor og utenfor Chomskys definisjon av en akseptabel ytring er strukturelt flertydige setninger som *Flying planes can be dangerous* (Chomsky, 1965, s. 21). Selv om ingen med engelsk som morsmål ville trekke i tvil om denne setningen er mulig på engelsk, er den ikke 'immediately comprehensible without paper-and-pencil analysis'. En slik analyse viser at den har to mulige lesninger (som kan skilles fra hverandre nettopp ved en slik analyse):

- (2) a. [S [VP Flying [NP planes]] [VP can be dangerous]]
 b. [S [NP [AP Flying [NP planes]]] [VP can be dangerous]]

⁶Dette ser vi blant annet i tilfellene der ugrammatiske konstruksjoner brukes oftere enn sine 'grammatiske motsvar', noe jeg vil komme tilbake til.

Ytringer som (2) kan riktignok forekomme i en kontekst som kan røpe hvilken tolkning som er den riktige, og intonasjonen når den ytres muntlig vil ofte avsløre hva som menes. Men dette er *språklige* ytringer, og muligheten for at de kan forekomme i en sammenheng som ikke hjelper mottakeren i tolkningen av dem er til stede. Riktignok er begge alternativene man har i tolkningen umiddelbart åpenbare, men den strukturelle flertydigheten gjør slike setninger mindre brukbare, på lik linje med (om enn på en annen måte enn) setninger som *Mannen hunden...* og *Colorless green...* Flertydige setninger har altså fellestrekk med uakseptable setninger (både grammatiske og ugrammatiske), samtidig som de flertydige setningene utvilsomt er å regne som både akseptable og grammatiske. I del (ii) av Chomskys definisjon av 'akseptabel ytring' heter det at akseptable ytringer er *immediately comprehensible without paper-and-pencil analysis* (min understrekning). Det er etter denne definisjonen flertydige setninger ikke regnes som akseptable. 'Forståelig' (*comprehensible*) og 'akseptabel' er åpenbart ikke det samme. Akseptable setninger er ikke nødvendigvis umiddelbart forståelige, som (2), og forståelige setninger kan være uakseptable, som (3) (Chomsky, 1956, s. 15):

(3) Read you a book on modern music?

I tillegg til å være ugrammatisk, bryter (3) med (i) i Chomskys akseptabilitetsdefinisjon: Den er ikke *perfectly natural*. En lytter med engelsk som morsmål ville reagere dersom en sekvens som (3) forekom i en samtale, selv om det er tydelig hva taleren mener. (3) er med andre ord uakseptabel (etter (i)), men forståelig. En syntaktisk analyse av denne ytringen vil ikke bidra til større forståelse – noe den vil for (4-c) – og den er ikke mindre forståelig før en slik analyse enn etter, i motsetning til (4-c).

Så langt har vi sett fire eksempler på setninger det kan stilles spørsmål ved akseptabiliteten eller grammatikaliteten til (og som gjentas her under (4) for oversiktens skyld):

- (4)
- a. Mannen hunden katten fryktet jaget falt
 - b. Colorless green ideas sleep furiously
 - c. Flying planes can be dangerous
 - d. Read you a book on modern music?

Av eksemplene er (4-a) og (4-b) like vanskelige å oppfatte proposisjonen til umiddelbart, men av ulike grunner. (4-b) er uforståelig av semantiske årsaker: NPen *colorless green ideas* har en umulig referanse, lik som S-en *ideas sleep* og VPen *sleep furiously*. Likevel er ikke dette en setning som krever skriftlig analyse for å avgjøre det syntaktiske forholdet

mellom elementene; Det er dette som gjør at lytteren umiddelbart kan gi denne setningen betegnelsen 'uakseptabel'. Den oppfyller ikke (iii) i Chomskys akseptabilitetskrav (om å ikke være *bizarre and outlandish*).

(4-a) er uforståelig (i betydningen 'ikke *immediately comprehensible*') av syntaktiske årsaker: Den kan synes som en oppramsing av ord (en rekke substantiver (*mannen hunden jaget*) etterfulgt av en rekke verb (*fryktet jaget falt*)), men er egentlig en dobbel 'self-embedding' (som i (1-b)) (heretter 'egeninnføyning'). Denne typen konstruksjoner regnes av Chomsky (1965, s. 13) for å bidra radikalt til inakseptabilitet. Slike setninger vil kun kunne forstås ved skriftlig analyse, og faller derfor ikke inn under Chomskys akseptabilitetsdefinisjon (ii).

(4-c) er utvilsomt grammatisk, og vil av brukere av engelsk regnes som akseptabel. Det som skiller den fra de fleste andre ytringer er muligheten for flere tolkninger (2). Den har to ulike (om enn lignende) proposisjoner, og som en isolert språklig ytring er det umulig å forstå hvilken av dem taleren prøver å formidle. Det er kun ved å foreta en analyse av strukturen i setningen man kan forstå hvilken tolkning som er den riktige. Dermed havner også (4-c) i kategorien 'uakseptabel', med Chomskys definisjon (ii).⁷

1.4 Forståelig/uforståelig

Introduksjonen av dikotomien akseptabel/uakseptabel som et tillegg til grammatisk/ugrammatisk i behandlingen av setninger det ikke er uproblematisk å bruke i dagligtalen har gjort det lettere å vurdere på hvilken måte mulige ytringer er relevante for lingvistikken. Siden grammatiske setninger med en syntaks som gjør det uklart hva som menes med dem kan være både uakseptable (4-a) og akseptable (4-c), kan det behøves en presisering av begrepet 'akseptabel', gjerne med enda en dikotomi, med utgangspunkt i Chomskys betegnelse av akseptable setninger som *immediately comprehensible* (ii).⁸ Samtidig vil definisjonen av 'akseptabel' slik begrepet brukes her reduseres til å favne setninger som er *perfectly natural and in no way bizarre and outlandish* (i). Dette vil blant annet gjøre det mulig å

⁷Igjen, det synes forholdsvis klart at Chomsky mener *immediately comprehensible* i en kontekst, og i de fleste tilfeller fremgår det av konteksten hvilken lesning av en strukturelt flertydig setning som er den riktige. Men det er setningen isolert som er interessant. Visse typer setninger (som (4-a) og (4-c)) er syntaktisk mer interessante enn andre.

⁸På norsk vil jeg kalle denne dikotomien forståelig/uforståelig. En 'uforståelig' setning i denne sammenhengen er ikke nødvendigvis umulig å forstå. 'Uforståelig' betyr snarere 'ikke umiddelbart forståelig', og betegner setninger med en syntaks som gjør at de skiller seg ut ved å være strukturelt flertydig (4-c) eller dypt innføyd (4-a), eller ved å mangle denotasjon (4-b)

kategorisere flertydige setninger som akseptable. Setninger som (4-a) og (4-c) har slik den samme verdien i grammatikalitet (+G)⁹, men ulik verdi i akseptabilitet (-A og +A). Når det gjelder 'forståelighet' har de samme verdi (-F).

Ytringer som (4-b) regnes etter Chomskys (1965) definisjon av akseptabilitet som uakseptable, og er verken 'perfectly natural' eller 'immediately comprehensible'. Verdiene er -A og -F, men +G.

(4-d) er verken flertydig eller har en uklar denotasjon, men den bryter med de grammatiske reglene for engelsk (siden den innledes av hoved verbet) (-G), og vil ikke regnes som akseptabel av språkbrukere med engelsk som morsmål (-A). Likevel er det ingen tvil om hva taleren mener (*Do you read a book on modern music?*). Altså er ytringen +F.

1.4.1 Akseptable, ugrammatiskse ytringer

Som nevnt har Chomsky (1965) betegnet grammatikalitet og akseptabilitet som to svært ulike trekk i språket. Akseptabilitet kan testes empirisk, men det kan ikke grammatikalitet. En akseptabel setning i L kan defineres som (1) En setning som betegnes som akseptabel (dvs. tilhørende L) av en lytter med L som morsmål, eller (2) En setning som produseres av en taler med L som morsmål. Ofte er (1) og (2) sammenfallende, og det er dette som gjør det mulig å formidle en proposisjon ved hjelp av ugrammatiskse (men akseptable) setninger. Hvis en grammatisk feil blir vanlig å bruke ser den ut til å ha mulighet for å bli akseptert. Dette skal jeg komme tilbake til senere.

Det de fleste typene grammatiskse, uakseptable setninger har til felles er at de sjelden eller aldri brukes i dagligtale. Dette er en av grunnene til at et korpus av ytrede setninger i et språk L ikke kan favne alle mulighetene grammatikken i L gir. Den andre grunnen er forekomsten av ugrammatiskse setninger i et slikt korpus.

If we rank the sequences of a given length in order of statistical approximation to English, we will find both grammatical and ungrammatical sequences scattered throughout the list; there appears to be no particular relation between order of approximation and grammaticality. (Chomsky, 1965, s. 17)

I 1.4.2 og 1.4.4 vil jeg ta for meg ugrammatiskse setninger (-G) som kan tenkes å forekomme i dagligtalen (+A).

⁹For enkelthets skyld vil jeg av og til forkorte begrepene grammatisk, akseptabel og forståelig til G, A og F, og markere positive og negative verdier ved dem med et tilgrensende +/-

1.4.2 –Forståelig

Det finnes mer eller mindre faste uttrykk som fungerer godt til å formidle en proposisjon, men som likevel ikke følger de grammatiske reglene for språket. Følgende er hentet fra Lakoff (1970, s. 157):

(5) One more beer, and I'll leave

Ved første øyekast kan en ytring som (5) virke grammatisk, kanskje først og fremst fordi den er så anvendelig til sitt bruk, og fordi den neppe er til å misforstå i en gitt kontekst. Men igjen, grammatikalitet tar ikke hensyn til konteksten, og (5) bryter med de grammatiske reglene for engelsk (siden *one more beer* ikke er tilknyttet et verb).¹⁰ Lakoff (1970) gir tre mulige lesninger av (5):

- (6) a. If I drink one more beer then I'll leave
b. If you drink one more beer then I'll leave
c. If you pour one more beer down my back, then I'll leave

Videre antyder Lakoff at det finnes flere mulige lesninger av (5). Denne flertydigheten er ikke strukturell, som i (4-c): Det finnes kun én mulig lesning av syntaksen i (5). Flertydigheten skyldes den manglende konteksten. Chomsky (2004, s. 321) sier følgende om lytterens tolkning av ugrammatiske ytringer:

Given a grammatically deviant utterance, we attempt to impose an interpretation on it, exploiting whatever features of grammatical structure it preserves and whatever analogies we can construct with perfectly well-formed utterances (it is precisely for this reason that a well-formed deviant utterance may be richer and more effective. (...))

Det ser ut til å være effektivitet som først og fremst motiverer bruk av setninger som (5), og forekomsten i en kontekst som rettferdiggjør dem. Bruken av flertydige setninger som (4-c) kan også være motivert av det samme, siden en slik ytring er mer effektiv enn et mulig, entydig men mer omstendelig alternativ (f.eks. *Planes that fly can be dangerous*). Effektiviseringer som dem som er eksemplifisert i (4-c) og (5) har til felles at de både er akseptable (+A) og 'uforståelige' (–F), men de har ulik grammatikalitet. Setninger som (4-c) er +G. Setninger som (5) er –G.

¹⁰Det kan ut fra (5) se ut som akseptabilitet også er uavhengig av kontekst, siden få (om noen) engelsktalende ville betegne en setning som (5) som uakseptabel.

Det er i grove trekk to ulike typer konstruksjoner som kan gjøre grammatiske setninger 'uforståelige' (-F). Eksempler på disse er allerede vist (i (4-b) og (4-c)), men gjengis her i (7):

- (7) a. Colorless green ideas sleep furiously
b. Flying planes can be dangerous

Det er svært ulike grunner til at setninger som disse er å regne som -F. En semantisk problematisk setninger som (7-a) er -F fordi den ikke denoterer en proposisjon på tross av at ordstillingen er nøytral (SVO), og det ikke er noe uvanlig ved syntaksen som skulle gi problemer med lesning av den, også uten papir-og-penn-analyse. Det er ordenes betydning som hindrer en fornuftig proposisjon å kunne avledes fra slike setninger. Derfor regnes de for å være 'uforståelige', som definert her. En syntaktisk flertydig setning som (7-b) er -F fordi den ikke denoterer én proposisjon, selv om også den har nøytral ordstilling og ellers ikke har i seg noen uvanlige syntaktiske konstruksjoner. I slike tilfeller er det ikke semantiske forhold som hindrer proposisjonen, men muligheten for tolkning av gitte elementer som tilhørende ulike ordklasser.

1.4.3 Hvorfor er flertydige setninger problematiske?

En effektivisering (som dem vi så eksempler på i (4-c) og (5)) som en generell forenkling av kommunikasjonen åpner muligheten for at setningen ikke fungerer like godt til kommunikasjon som en entydig setning. Fodor (1981, s. 286) påpeker hvordan forekomsten av flertydige setninger – og den allmenne oppfatningen av dem som akseptable – viser at kompetansen avleder konstruksjoner som ikke er perfekte for kommunikasjon.

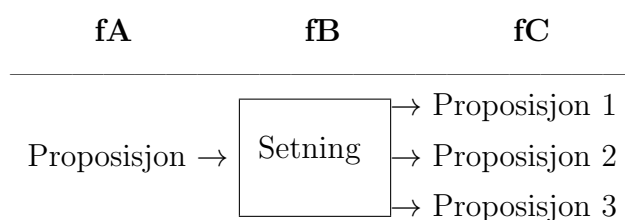
One very obvious problem that the sentence processing routines face is ambiguity – both full ambiguity and temporary ambiguity – that is resolved by subsequent words in the sentence. If, as seems likely, our brains lack the capacity to compute all possible analyses of an ambiguous word string, then we can only guess what the intended structure is; sometimes, inevitably, we shall guess wrongly. Ambiguity is not a problem in sentence production, since the producer knows which interpretation he intends and has only to find some well-formed realization of it.

Som nevnt gjør ofte kontekst eller intonasjon den tilsiktede proposisjonen bak en flertydig setning mulig å avlede. Fodors poeng er at man vanskelig kan finne en forklaring på hvorfor kompetansen tillater konstruksjoner som ikke lar taleren ytre seg presist, eller snarere,

hvordan kompetansen kan ha utviklet seg til å godta slike konstruksjoner: At én proposisjon gjennom en grammatisk, akseptabel setning kan denotere flere proposisjoner er ikke den beste løsningen for kommunikasjon.

Figur 2 (en forenklet utgave av figur 1 (i 1.1)) viser hvordan denoteringen av en proposisjon (fase A (fA)) kan uttrykkes ved en (flertydig) setning (fB), og resultere i flere mulige proposisjoner hos lytteren (fC).

FIGUR 2:



At kompetansen tillater produksjonen av setninger som gir flere mulige proposisjoner enn den taleren ønsker å formidle, mener Fodor tyder på at begrensningen grammatikken setter for hva som er en mulig setning (f.eks. øyrestriksjoner)¹¹ ikke først og fremst skyldes vansker med prosesseringen (ved at de er for 'tunge' for parseren): *I suggest that the parsing routines play a crucial but less specific role in determining the [island] constraints.* (Fodor, 1981, s. 285). Det er ikke begrensninger ved parserens kapasitet som gjør setninger som (4-c) 'uforståelige' (i motsetning til (4-a)): Det er et resultat av performansen natur (dvs. at den er endelig og må virke lineært)¹². Den er (i motsetning til kompetansen) ikke i stand til å skille mellom proposisjon 1, 2 og 3 (se figur 2). Så snart en flertydig setning er ytret, har lytteren alle de mulige tolkningene tilgjengelig, men vet ikke hvilken av dem som gir proposisjonen taleren forsøker å formidle. Problemene i kommunikasjonen ved bruk av flertydige setninger oppstår altså senere enn ved bruk av dypt innføyde setninger (som (4-a)). Problemene ved dype, egeninnføyde setninger oppstår allerede i løpet av ytringen (fB), siden lytterens parser ikke klarer å koble predikatene med argumentene. Det er av samme grunn slike setninger så sjelden brukes: De krever like mye av taleren som av lytteren, siden produksjonen av dem er like krevende for parseren som prosesseringen. Taleren vil ved bruk av dype, egeninnføyde setninger få problemer med å kode proposisjonen (fA) til en setning (fB), mens problemene ved bruk av flertydige setninger oppstår først *etter* selve produksjonen av ytringen (fC), og utgjør derfor kun et problem for lytterens parser.

¹¹Øyrestriksjoner vil jeg komme grundigere inn på i kapittel 5.

¹²Jf. Chomsky (1963, s. 390): (...) *the performance of the speaker or hearer must be representable by a finite automaton of some sort.*

1.4.4 +Forståelig

Følgende eksempler har blitt ytret i NRK radio:

- (8) a. Det_i kom vi i høyesterett i plenum til at det ville stride mot en uskyldspresumpsjon i den europeiske menneskerettighetskonvensjonen å opprettholde $_{-i}$ *det_i. (Resumptivt pronomener)
- b. Japan opplever sin verste krise siden andre verdenskrig, der_i jordskjelv og tsunami har ført til store ødeleggelser $_{-i}$ *der_i. (Resumptiv adjunkt)
- c. Det er mange ting jeg *_ skulle *gjørne ha spurt ham om. (Manglende adverbialskift i leddsetninger)
- d. Min forventning som stor eier i Statoil er selvsagt at de følger lover og regler i de landene de opptrer *_ . (Utelatte strandede preposisjoner)

Proposisjonen til samtlige av disse ytringene er lett å forstå, og ved første øyekast kan de se ut til å være grammatiske. Likevel bryter de med ulike regler for grammatikken:

- Den siste forekomsten av *det* i (8-a) er et resumptivt pronomener (som forekommer i tillegg til en gap i leddsetningen den forekommer i (med den første forekomsten av *det* som filler)). (8-a) bryter med prinsippet om koherens (Dalrymple, 1990, s. 39), som sier at et predikat ikke kan ta flere enn ett av hvert argument: *opprettholde* har to objekter (*det* og $_{-i}$ (*det_i*)).
- Leddsetningen i (8-b) har to like stedsadverbialer (*der* og *der*). De ugrammatiske tilleggene i (8-a) og (8-b) ser ut til å ha samme pragmatiske 'funksjon': De letter forståelsen av setninger med lang avstand mellom filler og gap, og at behovet for å bruke resumptive pronomener er sterkt nok til at taleren (muligens ubevisst), bryter de grammatiske reglene. Konstruksjoner med resumptive elementer kommer jeg tilbake til i 3.1.
- I (8-c) er det brudd på reglene for plassering av adverbialet (*gjørne*) i leddsetninger (*jeg skulle gjørne ha spurt ham om*) (Faarlund, Lie & Vannebo, 1997, s. 891)¹³ som gjør ytringen ugrammatisk.¹⁴

¹³Norsk referansegrammatikk (Heretter NRG).

¹⁴Reglene som er definert i NRG baserer seg på setningsskjema. I skjema B (leddsetningsskjemaet) plasseres setningsadverbialer mellom subjektet og det finite verbet, og ikke etter det finite verbet, som i (8-c).

- I (8-d) står *opptre* som et transitivt verb som tar objekt (*landene*). Det er nærliggende å tro at *landene* er ment å være utfylling i en adjunktPP med *i* som kjerne. Denne typen grammatisk feil, der en strandet preposisjon er utelatt, skal det kommes tilbake til i 7.2.

Felles for setningene i (8) er at de til tross for å være ugrammatiske ($-G$) har proposisjoner som ikke er til å misforstå ($+F$). Konstruksjoner som disse vil være ytret uten at verken ytreren eller lytteren legger merke til at de bryter med de grammatiske reglene (dvs., de er $+A$).

1.5 Oppsummering

Tabell 1 viser mulige kombinasjoner av parametrene grammatisk (G), akseptabel (A) og (umiddelbart) forståelig (F).¹⁵

TABELL 1:

G	A	F	Eksempel
-	-	-	Fjor i Roma til flyttet Per
-	-	+	Read you a book on modern music?
-	+	-	One more beer, and I'll leave
-	+	+	Per har en rar måte å sitte på bussen
+	-	-	Mannen hunden katten fryktet jaget falt
+	-	+	Per har en rar måte å sitte på bussen på
+	+	-	Flying planes can be dangerous
+	+	+	Per flyttet til Roma i fjor

Det er grunnleggende forskjeller i hva som ligger til grunn for vurderingen av grammatikalitet, akseptabilitet og forståelighet. Grammatikalitet baserer seg på språkbrukerens kunnskap om språket (kompetanse), akseptabilitet baserer seg på språkbrukerens språkfølelse (ved parserens virke med utgangspunkt kompetansen)¹⁶ (performans), og forståelighet baserer seg på språkbrukerens evne til å oppfatte hva (den tilsiktede) proposisjonen det gis uttrykk for ved en ytring er.

¹⁵Det er ikke alle setningstypene i tabellen jeg har tatt for meg så langt. Setninger som eksemplene i rad 4 og 6 skal jeg komme tilbake til i del 2. Ytringer med $-$ eller $+$ for alle parametrene (rad 1 og 8) er ikke interessante her.

¹⁶Med dette menes prosessen som er beskrevet i figur 1 i dette kapitlet.

At jeg anerkjenner at språkbrukere har lingvistisk kompetanse betyr ikke at jeg regner denne for å være en medfødt evne, eller at de grammatiske reglene utgjør objektive, absolute sannheter. Likevel mener jeg at måten vi bruker språk på i seg selv ikke kan definere hva som er grammatisk. Dette er fordi det er flere faktorer enn kompetansen som påvirker språkbruken vår. Selv om det er vanskelig å si nøyaktig hva forskjellen mellom kompetanse og performans er, tror jeg det lar seg gjøre å sannsynliggjøre hvilke begrensninger parseren er belemret med, og på bakgrunn av dette resonnerer jeg frem til noen av trekkene ved språkbrukerens kompetanse. Et mulig problem i redegjørelsen for når det er parserens begrensninger som er årsaken til at ugrammatiske konstruksjoner brukes i språket, er selve påstanden om at slike konstruksjoner er ugrammatiske. I løpet av oppgaven vil jeg gi uttrykk for oppfatninger om grammatikalitet som neppe deles av alle språkbrukere. I slike tilfeller vil jeg forsøke å gi eksempler og resonnementer som kan støtte grammatikalitetspåstandene. I de tilfellene jeg hevder at visse konstruksjoner er ugrammatiske som vanligvis ikke regnes for å være det (f.eks. konstruksjoner med resumptive pronomen – særlig i svensk) er dette et uttrykk for at en årsak til at de forekommer *kan* være at parseren bruker visse elementer som ikke er en del av kompetansen for å gi best mulig uttrykk for en proposisjon, snarere enn en påstand om at de er absolutt ugrammatiske.

Når jeg underveis i oppgaven hevder at ytringer eller konstruksjoner er ugrammatiske, mener jeg altså ikke at de bryter med objektive, grammatiske regler for språket, men at det er grunn til å anta at parserens utilstrekkelighet i håndteringen av vanskelige, uakseptable konstruksjoner kan få oss til å regne dem som ugrammatiske når de i virkeligheten bare er mindre forståelige enn tilsvarende¹⁷ ugrammatiske, men akseptable konstruksjoner.

¹⁷Med 'tilsvarende' menes det mulige alternative, foretrukne konstruksjoner (som bruk av resumptivt pronomen fremfor en gap (kapittel 4), eller å utelate fremfor å uttrykke måteP i måteNP (kapittel 8)).

2 Rekursjon

I dette kapitlet vil jeg gjøre rede for mulige typer rekursive konstruksjoner og presentere forklaringer som er gitt for hvorfor enkelte ikkeiterative konstruksjoner er vanskelige å bruke. Jeg vil også referere til empiriske undersøkelser som viser hvordan små forskjeller i en setnings struktur kan utgjøre forholdsvis store forskjeller for arbeidet som kreves av parseren under prosesseringen av dem.

2.1 Iterativ og ikkeiterativ rekursjon

Samtlige av setningene i (9) inneholder tre nivåer av rekursjon.

- (9) a. Mannens venns sønns bils motor stoppet.
- b. Katten i stolen i stuen i huset i skogen sov.
- c. Mannen (som) hunden (som) katten (som) jenten eide fryktet jaget snublet.
- d. At at at at Lars dro gledet Lise plaget Kari moret Jens overrasket Svein.

Setningene i (9) viser tilfeller av iterativ rekursjon (venstreforgreining: (9-a) og høyreforgreining: (9-b)) og ikkeiterativ rekursjon (egeninnføyning: (9-c) og (9-d)). Det setningene i (9) har til felles er at de sjelden eller aldri brukes. En åpenbar forskjell mellom (9-a)/(9-b) og (9-c)/(9-d) er hvor vanskelig det er å forstå proposisjonen til de to sistnevnte. Få, om noen språkbrukere vil være i stand til å bruke setninger som (9-c) og (9-d), og ved bruk av dem i dagligtale vil de av mange lyttere regnes som uakseptable, i motsetning til (9-a) og (9-b).

Det som skiller (9-c) og (9-d) fra de to øvrige setningene er at de inneholder egeninnføyde konstruksjoner: Den ytterste av de innføyde setningene lukkes ikke før den neste innføyde setningen åpnes, som igjen ikke lukkes før den neste åpnes, osv. Det som skiller (9-c) fra (9-d), er at det i (9-d) ikke nevnes noen predikater eller argumenter før det har blitt gitt et 'varsel' om hvor mange innføyde ledd som er i vente (i antallet *at*). Her finner ikke predikatene og argumentene som styres av dem seg langt fra hverandre, men innledningen til hvert COMP (*at*) kan befinne seg langt fra leddsetningen den innleder. Setninger som (9-c) og (9-d) har sine mer akseptable, forståelige ekvivalenter:¹⁸

¹⁸Det kan diskuteres hvorvidt (9-c) og (10-a) har samme denotasjon, siden rekkefølgen til elementene kan ha noe å si for rekkefølgen til hendelsene i proposisjonen.

- (10) a. Jenten eide katten som fryktet hunden som jaget mannen som snublet
b. Det overrasket Svein at det moret Jens at det plaget Kari at det gledet Lise
at Lars dro

Her lukkes hver predikat-argument-relasjon underveis: Hvert predikat får oppfylt kravet sitt om argumenter før det neste verbet introduseres. Begge setningene i (10) er tilfeller av høyreforgreining.

2.2 Hvorfor er dyp egeninnføyning vanskelig?

At setninger som (9-c) og (9-d) kan være uforståelige når de presenteres for en lytter, men på den andre siden lar seg papir-og-penn-analysere og omformulere til de mer forståelige (10-a) og (10-b), viser at det er forskjell mellom kompetanse og performans. Chomsky (1963) gjør rede forholdet mellom kompetansen som en generativ evne, og performansen som en lineær øvelse:

(...) we must conclude that the competence of the native speaker cannot be characterized by a finite automaton. The grammar stored in his brain cannot be a one-sided linear grammar, a fact that is not in the least surprising. Nevertheless, the performance of the speaker or hearer must be representable by a finite state automaton of some sort. The speaker-hearer has only a finite memory, a part of which he uses to store the rules of his grammar (a set of rules for a device with unbounded memory), and a part of which he uses for computation in actually producing a sentence or 'perceiving' its structure and understanding it. (Chomsky, 1963, s. 390)

Mens performansen (som avfeier dype egeninnføyde setninger som uakseptable) fungerer lineært, er kompetansen (som aksepterer egeninnføyde setninger som grammatiske) ikke-lineær. Performansen er en 'minneøvelse', som dreier seg om å holde rede på forholdet mellom de ulike elementene som ytres/høres i en setning, noe som gjøres av parseren ved bruk av språkbrukerens ubevisste kunnskap om grammatikken (kompetansen). Årsaken til at det finnes grammatiske setninger som er vanskelige eller umulige å bruke og forstå i en vanlig taler-lytter-situasjon er altså at kompetansen er for 'kraftig' for performansen. Minnet vårt er rett og slett ikke av en slik natur at det er i stand til å utnytte mulighetene som finnes i grammatikken fullt ut. Ytringer forløper lineært, mens grammatikken er generativ.

Det har blitt gjort flere empiriske studier av setninger med egeninnføyning. I et forsøk

viser Miller & Isard (1963) sammenhengen mellom antallet egeninnføyninger og evnen språkbrukere har til å gjengi setninger. Resultatene er ikke overraskende: Jo dypere egeninnføyningen er, desto vanskeligere er det å gjengi setningen. Som forklaring foreslås det at egeninnføyninger håndteres på en annen måte enn vanlige setninger, at det finnes en form for underkategori av setningsprosessering, og at kapasiteten til å bruke slike uvanlige metoder for parsing er svært begrenset. Dette er en mulig forklaring, men det vanskelig å si noe sikkert om årsaken til at egeninnføyde setninger er så vanskelige å håndtere.

It is difficult to do more than speculate about these matters, for we are here faced by a type of cognitive task – the handling of interruptions – about which very little is known. (Miller & Isard, 1963, s. 302)

Senere studier (Just, Carpenter & Keller, 1996) har vist at en leddsetning med objekt-relativisering (11-b) (der objektet kommer før subjektet og verbet) tar langt lengre tid å lese enn en leddsetning med subjektrelativisering (med nøytral SVO-stilling) (11-a) (mine nedstreker (gap) og klammer (filler)):

- (11) a. [The reporter]_i that —_i attacked the senator admitted the error
b. [The reporter]_i that the senator attacked —_i admitted the error

Resultatene til Just et al. (1996) kan vise en sammenheng mellom parserens virke og Behaghels (1923) første lov om at elementer i språket som hører sammen er mer troende å befinne seg nær hverandre i setningen enn elementer som ikke hører sammen. I (11-a) er filleren (*the reporter*) både strukturelt og lineært nærmere gappen enn i (11-b).

Undersøkelser av blodgjennomstrømningen i Brocas område (foretatt av Stromswold, Caplan, Alpert & Rauch, 1996) viser mer aktivitet ved eksponering for setninger med egeninnføyning (12-b) enn setninger med høyreforgreining (12-a) (mine nedstreker og klammer):

- (12) a. The child spilled the juice that stained the rug
b. [The juice]_i that the child spilled —_i stained the rug

Også her gir større strukturell og lineær avstand mellom elementene som hører sammen parseren mer å gjøre. Mens *spilled* i (12-a) grenser til både subjektet og objektet sitt, er objektet topikalisert i (12-b). Verd å merke seg er det også at hoved verbet ikke er det samme i de to. I (12-a) er hoved verbet *spilled*, mens *stained* inngår i modifikatoren til *juice*. I (12-b) er *stained* hovedverb, mens *spilled* inngår i modifikatoren til *juice*.

Selv om de nevnte studiene på ulike måter viser at det kreves mer av parseren i prosesseringen av setninger med lang avstand mellom elementer som hører sammen, er det vanskelig å trekke konklusjoner om hvordan parseren håndterer egeninnføyninger på bakgrunn av dette. Siden b-eksemplene i (11) og (12) (de mest krevende for parseren) viser både større *strukturell* og større *lineær* avstand enn a-eksemplene, kan det heller ikke avgjøres på bakgrunn av dette hvilken av avstandene som er mest krevende for parseren å håndtere. Om ikke annet viser dette at selv små forskjeller i kompleksitet kan utgjøre en forskjell i anstrengelsene parseren utsettes for, selv når de mer komplekse konstruksjonene (b-eksemplene) intuitivt ikke er det minste mindre akseptable eller forståelige enn de mindre komplekse konstruksjonene (a-eksemplene).

3 Langdistanseavhengighet

Dette kapitlet gjør rede for – og viser eksempler på – bruk av personlige pronomen som erstatning for gap (resumptive pronomen). Det diskuteres hvorvidt konstruksjoner med resumptive pronomen er grammatiske eller ikke, særlig i svensk, der bruken av resumptive pronomen i langt større grad enn i f.eks. norsk regnes som akseptabel. Det gis også mulige forklaringer på hvorfor resumptive pronomen i visse konstruksjoner – også i språk der de er forholdsvis uvanlige – foretrekkes fremfor gap.

3.1 Setningsknuter

På norsk kan filleren i en setningsknute¹⁹ være et tematisert setningsledd eller et spørreord (NRG, s. 1095). Filleren er ledd i en matrisesetning, og har felles referanse med gappen, som befinner seg i en underordnet setning.

(13) Det_i vet jeg ikke [hvem som har sagt —_i] (NRG)

Resumptive pronomen (RP) er et alternativ til langdistanseavhengigheten i setningsknuter, som tilsynelatende gir anaforisk referanse (anafor – antesedent) snarere en setningsknute (filler – gap).²⁰ I NRG betegnes ikke bruken av RP som ugrammatisk: Det slås fast at de forekommer, særlig i dialekter fra Trøndelag og Nord-Norge. Det gis – i tillegg til eksempler på resumptive pronomen (14-a) – eksempler på resumptive determinativer (14-b) (NRG, s. 1104). I tillegg til disse forekommer det resumptive adjunkter (14-c). Uthevingene (og grammatikalitetsvurderingene²¹) er mine egne:

- (14) a. *(...) erklæringer som det er helt klart hvem *de* er rettet mot. (NRG/NRK Radio 1978)
b. *(...) en begivenhet som man vet at *dette* er verdenshistorie. (NRG/NRK TV 1979)
c. *De siste fire søndagene har jeg vært i kirker her i Oslo, hvor det har vært slående hvor mange barnefamilier og ungdom som har vært *der*. (NRK Radio 2011)

¹⁹'Setningsknute' er den norske betegnelsen på det som i den engelskspråklige litteraturen kalles 'filler-gap-dependency'.

²⁰Hvorvidt det virkelig er snakk om anaforisk referanse ved bruk av RP vil jeg komme tilbake til i (18).

²¹I NRG regnes (14-a) og (14-b) som grammatiske. I løpet av denne oppgaven vil jeg vise hvorfor jeg mener bruken av RP gjør ytringer ugrammatiske.

Ingen av setningene i (14) hadde endret betydning om det ikke var for de resumptive elementene. Språkbrukeren som skal ytre setninger som disse, med en gap det er mulig å plassere et resumptivt element i (eller ved), stilles overfor et dilemma: Mens forekomsten av RP gjør setningen ugrammatisk, kan utelatelsen av det resumptive elementet skje på bekostning av forståeligheten. Med andre ord må taleren velge mellom forståelighet (resumptivt element) og grammatikalitet.

Konstruksjoner med resumptive elementer kan ikke kategoriseres som grammatiske, siden de bryter med prinsipper for grammatikken: Resumptive pronomen og determinativer (som refererer til nominale ledd) kan ikke tildeles en syntaktisk rolle av verbet, siden denne rollen allerede er fylt av antesedenten deres, ved en gap. Resumptive adjunkter (som refererer til adverbiale ledd) har koreferanse med en adjunkt som allerede har fylt rollen som adjunkt for det gjeldende forholdet (tid/sted).

(15-a) viser et forsøk på å bøte på uforståeligheten til en dyp, egeninnføydd setning (gjengitt her (noe endret) som (15-a)) ved bruk av RP (15-b):

- (15) a. Mannen hunden katten jenten eide fryktet jaget snublet.
b. *Mannen_k hunden_j katten_i jenten eide den_i fryktet den_j jaget ham_k snublet.

Intuitivt virker det verken lettere eller vanskeligere å lese (15-b) enn (15-a). Her skiller den dype egeninnføyningen seg fra setningsknuten, i at den ikke lar seg redde ved tilføyning av et element. Det kan tenkes at dette er årsaken til at dype egeninnføyninger sjelden eller aldri tas i bruk: Det finnes ingen syntaktisk 'medisin' for å gjøre slike konstruksjoner forståelige, akseptable, eller i det hele tatt mulig å bruke i dagligtalen.

3.2 Er bruken av RP grammatisk eller akseptabel?

Falk (2002) påpeker at bruken av RP bryter med prinsippet om 'Economy of Expression' (fra Bresnan (2001)), ved at man tillegger c-strukturen elementer som ikke er nødvendig for f-strukturens velformethet, eller for det semantiske uttrykket. For å rettferdiggjøre bruken av RP som 'hjelpere' for parseren introduserer Falk 'Sufficiency of Expression': *Syntactic elements which provide clues to parsing are exceptions to Economy of Expression. Such elements include markers of clause boundaries and extraction sites.* (Falk, 2002, s. 164f) Som eksempel på 'markers of clause boundaries' nevnes *that* i setninger som *I believe (that) the world is flat*. Som eksempel på 'extraction sites' (dvs. RP) refereres det til Erteschik-Shir (1992, s. 89) (Erteschik-Shirs grammatikalitetsvurderinger):

- (16) a. This is the girl that John likes $_/*$ her.
 b. This is the girl that Peter said that John likes $_/??$ her.
 c. This is the girl that Peter said that John thinks that Bob likes $_/?$ her.
 d. This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to $?_/$ her.

Grammatikalitetsvurderingene (*, ? og ??) i (16) er egentlig akseptabilitetsvurderinger. I den grad RP brukes for å lette forståelsen av en ytring, er det ikke først og fremst for å gjøre setningen mer akseptabel (A) eller grammatisk (G) (siden målet med tale er kommunikasjon), men å gjøre den mer forståelig (F). Mens versjonene av setningene i (16) som inneholder RP er ugrammatiske, og (i økende grad med økende avstand mellom filleren og pronomenet) akseptable, er samtlige forståelige. Måten Erteschik-Shir bruker 'grammatisk' på i sin vurdering av setningene i (16), sammenfaller godt med Chomskys (1965) begrep 'akseptabel' når han beskriver forskjellen mellom akseptabilitet og grammatikalitet (som gjengitt i 1.2).

3.3 Svensk RP: kompetanse, performans og normering

NRG betegner RP som helt vanlig i svensk (s. 1104), men ifølge Teleman, Hellberg, Andersson & Christensen (1999)²² er ikke bruken av RP helt uproblematisk.

*Många språkvårdare bedömer bruket av resumti vt pronomen som talspråkligt och avråder från satsflätekonstruktionen i skrift i de fall då denna skulle kräva resumti vt pronomen i den underordnade satsen.*²³ (SAG, bind 4, s. 427)

Det er med andre ord uenighet om hvorvidt bruken RP er 'riktig svensk' i skrift. De setningsknutene som *krever* RP i de fleste svenske dialekter er konstruksjoner der *[d]et resumptiva pronomenet är subjekt i en underordnad sats med inledare (subjunksjon eller satsbas)* (SAG, s. 427).

Teleman (1999, s. 297) gir blant andre følgende eksempler på grammatikaliserte RP (17-a) og mer tvilsom bruk av dem (f.eks. som objekt i den underordnede setningen) (17-b):

- (17) a. Din bror har jag inte hört om *han* kommer.
 b. Din bror har jag inte hört om hon har träffat *honom*.

²²Svenska Akademiens Grammatik (Heretter SAG)

²³En 'satsfläta' er en *satsförbindelse där en satsbas i en överordnad sats har satsledsfunktion i en underordnad sats: 'Henne tror jag inte att du har talat med'* (SAG, bind 1, s. 222) (m.a.o. en setningsknote).

RP som en måte å gjøre kompliserte konstruksjoner mer forståelige og lettere å bruke på er ifølge Teleman (1999) noe som først og fremst (skal) finnes i tale: *I skrift har den vane skribenten redan redigerat bort såväl icke-standardiserade uttryck som performansberoande egenheter (vilka i tal alltid blir ståande kvar på bandet)* (s. 297). Han vedgår m.a.o. at bruken av RP i svensk skriver seg fra et behov for å kunne uttrykke seg bedre, og at performansen 'overstyrer' kompetansen ved bruken av dem. I samme artikkel skriver Teleman:

Distinktionen mellan kommunikativa lösningar och grammatiska regler eller – om man så vill – skillnaden mellan performans och kompetens är avgörande för grammatiken. Att kommunikativa lösningar kan generaliseras och konventionaliseras eller grammatikaliseras, så att de t.o.m. kan användas i andra situationer än dem som ursprungligen lett till deras användning är inget argument mot distinktionen. Grammatikaliseringsprocessen kan leda till att gränsen mellan kompetens och performans ofta blir svår att dra, t.ex. vid olika slag av omtagningar, självrettelser eller ellipser. Men den svårigheten är något vi måste leva med. (Teleman, 1999, s. 296)

Det synes som om den utbredte bruken av RP-subjekt som en *kommunikativ* løsning (dvs. for å lette forståelsen) har ført til grammatikalisering av visse typer RP i svensk. Teleman skriver at performansen styrer kompetansen gjennom grammatikalisering, og at grammatikken bør ta hensyn til kommunikative løsninger. Det er dette siste prinsippet forfatterne av SAG, som er en *deskriptiv* grammatikk, følger. At normeringen SAG presenterer er deskriptiv gjør reglene påvirkelige fra performansen, som er både lingvistisk og ekstralingvistisk. Derfor kan ikke normeringen beskrive lingvistisk kompetanse. Vurderingen SAG og Teleman (1999) gjør av RP i subjektsposisjon som grammatiske synes å være en anerkjennelse av at det forekommer utstrakt bruk av dem. Dette betyr ikke nødvendigvis at de er en del av kompetansen, men kun brukbare som kommunikative løsninger.

Hyltenstam (1984) (gjengitt i Hawkins, 2001) gjør rede for bruken av RP i svenske restriktive relativsetninger hos personer med svensk som andrespråk. Forsøkspersonene hadde som morsmål språk som tillater RP i slike setninger (gresk og persisk), og språk som ikke tillater dem (finsk og spansk). Resultatene viser at personer med gresk og persisk som morsmål i gjennomsnitt bruker flere resumptive pronomen (17,1 og 20) enn personer med finsk og spansk som morsmål (3 og 12) (Hawkins, 2001, s. 160). Dette tyder på at forsøkspersonene har 'tatt med seg' vanen med RP inn i svensk. Kompetansen kan ikke defineres av hvordan andrespråksbrukere av L bruker L, men språkbrukere med L som

andrespråk har en tendens til å benytte seg av ekstralingvistiske kommunikative løsninger når språkkunnskapen kommer til kort. Slike løsninger kan være hentet fra morsmålet, og det synes som det er dette som er tilfellet i Hyltenstams undersøkelse. Dette er som når personer med norsk som andrespråk bryter V2-regelen, og plasserer subjektet foran verbet i f.eks. *hv*-setninger (**Hvor du kommer fra?*)²⁴ eller setninger med topikalisert adjunkt (**I går Per var hjemme*).²⁵ At slike feil er forholds vanlige gjør dem ikke mindre ugrammatiske, på samme måte som tendensen personer med svensk som andrespråk (spesielt dem med morsmål som tillater RP) har til å bruke RP i svensk, ikke gjør bruken av dem mer grammatisk.

3.4 RP som 'last resort'

Chomsky (1977) omtaler bruken av RP²⁶ som en 'redning' for brudd på øyrestriksjonene i Engelsk kort i en fotnote (s. 126, fn13): *I suppose that this is an ancillary process, not to be incorporated, strictly speaking, within the grammar*. Med andre ord betegner han ikke RP som grammatiske, men snarere som ekstralingvistiske verktøy for kommunikasjon.²⁷ De 'viser til' et element tidligere i samme setning, men ikke som anaforer. Chomsky tar ikke for seg RP i konstruksjoner som *ikke* bryter med øyrestriksjonene i engelsk, men viser eksempler på slike i hebraisk (som han vedgår å ikke kjenne detaljene om reglene for bruk av resumptive pronomen til).²⁸ RP omtales altså ikke i Chomsky (1977) som fullverdige alternativer til filler-gap-relasjonen, men snarere som et tillegg.

I sin omtale av obligatoriske 'parasitic gaps' (dvs. gaps som står foran den ekte gappen) som ikke kan byttes ut med pronomen. blant annet på grunn av brudd med Postals (1971) prinsipp om 'cross-over-effects'.²⁹ viser Engdahl (1983) hvordan det at bruk av pronomen

²⁴Riktignok finnes det flere norske dialekter der *hv*-spørsmål er unntatt V2-regelen (Vangsnes, 2004), men dette gjelder ikke for bokmål eller nynorsk.

²⁵I bruken av hovedsetninger blant personer med norsk som andrespråk er den vanligste feilen at man plasserer for mange ledd foran verbalet i setningen. (Golden, Mac Donald, Michalsen & Ryen, 1990, s. 20).

²⁶Chomsky bruker ikke betegnelsen 'resumptive pronoun', men beskriver fenomenet som *a construction with a pronoun where an island constraint would block relativization*.

²⁷Chomsky (2004) har også påpekt at en ugrammatisk setning i visse tilfeller kan være å foretrekke fremfor en grammatisk: (...) *a well-formed deviant utterance may be richer and more effective* (...). Dette skal jeg komme tilbake til i kapittel (6).

²⁸At RP kan brukes som et alternativ til gap i visse konstruksjoner i hebraisk bekreftes av Rizzi (1991, s. 123, fn25): (...) *in headed relatives in Modern Hebrew the gap strategy and the resumptive strategy alternate freely in all the contexts in which specific principles do not block one strategy* (...)

²⁹Kort fortalt hindrer dette prinsippet anaforisk referanse til et antesedent som domineres av anaforen – noe Engdahl (1983, s. 17-18) viser at hindrer koreferanse mellom f.eks. et *hv*-element og et pronomen

i slike konstruksjoner reduserer grammatikaliteten kan kontrasteres med hvordan RP kan øke grammatikaliteten (I følgende sekvens er det snakk om bruk av obligatoriske 'parasitic gaps'):

We can contrast this way of avoiding the utterance of an ungrammatical sentence with the use of resumptive pronouns. It is often assumed that inserting a resumptive pronoun may 'save' a sentence which would otherwise violate an extraction island. (...) A sentence that would violate a constraint is 'saved' or at least considerably improved, if the pronoun is replaced with a gap. (Engdahl, 1983, s. 18)

Her er det snakk om mulig bruk av RP som et grammatisk, mer forståelig alternativ til en grammatisk, mindre forståelig bruk av gap, på samme måte som en erstatning av et obligatorisk 'parasitic gap' med en pronomen fører til redusert grammatikalitet (som i (18)).

(18) Räkna upp de filmer_i som alla som sett $_i$ /?dem_i tyckte bra om $_i$

Hvis Engdahl har rett i at utelatelsen av et pronomen i tilfeller som (18) er motivert av det samme som bruken av RP er, er ikke RP først og fremst et verktøy for å lette kommunikasjonen, men for å bedre grammatikaliteten. Dette er et standpunkt som går imot mye av den refererte litteraturen i denne oppgaven (f.eks. Chomsky (1977) og Shlonsky (1992)). Forekomsten av pronomen i konstruksjoner med 'parasitic gaps' skal jeg komme nærmere inn på i 4.2.1.

Shlonskys (1992) hypotese (etter studier av hebraiske og palestinsk-arabiske data) er at RP er en siste utvei for å redde en ugrammatisk setning: (...) *resumptive pronouns only occur as a last resort, when wh-movement fails to yield a grammatical output* (s. 443). Dvs. at RP brukes når taleren oppdager at setningen som produseres ikke følger de grammatiske reglene, og legger inn et RP for i det minste å redde forståeligheten, siden den mislykkede grammatiske strukturen står i fare for å ikke kunne denotere proposisjonen taleren prøver å gi uttrykk for. Det Shlonsky ikke forklarer med dette, er forekomsten av RP i ellers grammatiske setninger (f.eks. *Dette er katten_i som naboens hund løp rundt hele kvartalet for å fange $_i$ /*den_i*). Det synes som det er språkbrukerens vurdering av akseptabiliteten – og ikke grammatikaliteten – som avgjør om det settes inn et RP. Når f.eks. parseren oppdager at det ekstraheres fra en setningsknode i setningen som produseres, regnes setningen av ytreren å være ikke bare ugrammatisk, men også uakseptabel. Det er som befinner seg mellom *hv*-elementet og dets gap: *Who_i did he*_i/j claim $_i$ had won?* Se også eksempel (18).

ikke det *grammatikalske* utfallet som bestemmer forekomsten av RP (som Shlonsky (1992, s. 443) hevder), men snarere hvorvidt ytringen er akseptabel eller ikke. Hvis avstanden mellom filler og gapen blir for lang til at den enkelte språkbrukeren vil vurdere setningen som akseptabel, legges RP inn for i det minste å bøte på forståeligheten (og kanskje også akseptabiliteten³⁰ (Erteschik-Shir, 1992)).

Siden det ikke er grammatikk, men akseptabilitet det dreier seg om, er spørsmålet om hvorvidt konstruksjoner med RP er grammatiske et spørsmål om performans snarere enn kompetanse; RP er et ekstralingvistisk 'verktøy' for å lette kommunikasjonen (jf. Chomsky, 1977). Dersom dette stemmer, er det problematisk å forsvare vurderinger av setninger med RP som grammatiske, selv om de forekommer i (og øker akseptabiliteten til) ekstraksjon fra setningknuter. Dette gjelder i det minste for språk med filler-gap-avhengighet (ved hv-spørsmål og topikalisering) som engelsk og norsk, og også for svensk hvis ikke bruken av RP kan vises å forekomme *i stedet for* en gap, og ikke i tillegg (som bl.a. Chomsky (1977) hevder). Slik situasjonen er for norsk og engelsk, er ikke RP bundet til filleren, siden en gap allerede, ved regel, er har opptatt denne plassen.³¹ Det er denne mangelen på binding mellom RP og filleren som gjør RP til et overflødig, nominalt ledd, uten tilhørighet til noe predikat, og derfor årsaken til at setningen bryter med prinsippet om koherens.

I *Resumptive pronouns can be syntactically bound* argumenterer Zaenen, Engdahl & Maling (1981) nettopp for at RP (i noen språk) kan bindes til filleren på samme måte som en gap, ved syntaktisk binding.³² De setter seg fore å vise at en syntaktisk binding kan foregå fra en setningsknote i svensk, i tillegg til at *syntactic binding can occur between a wh-element and a resumptive pronoun, as well as into positions that seem to be islands for the binding between a wh-phrase and a phonologically null element* (s. 680). Argumentasjonen i Zaenen et al. (1981) er todelt. Den første delen forklarer hvorfor bindingen mellom filler og RP er syntaktisk snarere enn anaforisk. Det vises hvordan det svenske refleksive pronomenet *sin* kan befinne seg før antesedenten hvis antesedenten er subjekt i en underordnet setning (19-a). Videre vurderes (19-b) som grammatisk.

³⁰Igjen, Erteschik-Shir (1992) bruker akseptabilitetsvurderinger som et mål for grammatikalitet, så hans definisjon av *grammatisk* er den samme som den som brukes for *akseptabel* her.

³¹I tilfelle brudd på øyrestriksjonene vil det ikke være en gap i den underordnede setningen, siden filler-gap-forholdet er avhengig av at disse restriksjonene overholdes. Ved et slikt brudd spiller det ingen rolle for grammatikaliteten om det forekommer RP eller ikke: Setningen vil uansett være ugrammatisk. Likevel kan en setning *med* brudd på øyrestriksjonene være mer forståelig med RP enn uten. Det kan synes som det er dette som får slike setninger til å virke mer grammatiske enn tilsvarende uten RP.

³²Det understrekes i artikkelen at syntaktisk binding (gap – filler) ikke er det samme som anaforisk binding (anafor – antesedent).

- (19) a. Vilken av sina_i flickvänner tror du att Kalle_i inte längre träffar?
 b. [Vilken av sina_i flickvänner]_j undrade du om det att Kalle_i inte längre fick träffa henne_j kunde ligga bakom hans dårliga humör?

I den andre delen argumenteres det for det samme – at det ikke er anaforisk, men syntaktisk binding mellom en filler og en RP – ved å vise til 'across-the-board'-betingelsene³³, og Ross' (1967) observasjon om at en setningsknuteregulering som gjelder for én av konjunktene i en koordinasjon også må gjelde for den andre konjunkten. Følgende eksempel (som av Zaenen et al. regnes som grammatisk) brukes til å vise at begge disse betingelsene angår RP.

- (20) Där borta går en man_i som jag ofta träffar men inte minns vad han_i heter.

Videre begrunnes bruken av 'across-the-board'-betingelsene for setninger som (20) på følgende måte:

We assume that Wh Movement or Relative Deletion must apply across-the-board to an identical element in both conjuncts. Then the only way that [(20)] can be derived is to assume that Wh Movement or Relative Deletion has applied to both conjuncts, and that the trace that is left behind is 'spelled out' as a (resumptive) pronoun in one case, but as a gap in the other. Assuming further that this spelling-out cannot change the syntactic category of the trace, both gaps and resumptive pronouns must be of the same syntactic category. The requirement of across-the-board-rule application means that the same rule must have applied to both conjuncts; hence, we cannot assume that one conjunct exhibits syntactic binding established by the application of, say, Wh Movement, and that the other conjunct simply exhibits anaphoric binding. (Zaenen et al., 1981, s. 681f)

Både Ross' (1967) og Williams' (1978) prinsipp kan altså benyttes for å vise at RPen *han* i (20) ikke kan ha anaforisk referanse med *man* som antesedent. RP regnes av Zaenen et al. (1981) for å være tilsynelatende anaforiske, men i virkeligheten kun fonologiske uttrykk for en gap.

Diskusjonen i dette kapittelet viser at RP vanskelig kan betegnes som ekte syntaktiske

³³Blant 'across-the-board'-betingelsene (Williams, 1978) er kravet om at *the affected elements in each conjunct be 'identical' in syntactic category* (Zaenen et al. (1981)).

elementer, men at de snarere er (tidvis svært nyttig) ekstralingvistisk 'støtte' for kommunikasjon. Hvis dette er tilfellet, vil det i (20) fremdeles være riktig å regne det for å finnes en syntaktisk binding mellom *man* og en gap på samme plass som RPen *han*. Vurderinger av grammatikaliteten til mer og mindre komplekse setninger med og uten RP baserer seg på språkbrukernes vurdering av setninger som brukbare til kommunikasjon,³⁴ med andre ord akseptabilitet og/eller forståelighet.

³⁴Jf. Erteschik-Shirs vurdering av setningene i (16). De empiriske studiene presentert i 2.2 viser at 'brukbarheten' til setninger er gradert, helt uavhengig av både akseptabilitet og forståelighet. Vi som språkbrukere betegner ikke a-setningene i (11) og (12) som mer eller mindre grammatiske enn b-setningene. Heller ikke i grammatikalitet er det den minste forskjell mellom dem. Likevel viser de seg å variere i 'umiddelbar forståelighet' (ved forskjell i responstid og hjerneaktivitet): Litt mer komplekse setninger er litt mindre forståelige. Det er nærliggende å anta at mye mer komplekse setninger er mye mindre forståelige (og kan behøve ekstralingvistisk 'støtte').

4 Referansebegrensninger

I dette kapitlet sammenlignes mulighetene og begrensningene i referanse for pronomener og gap. I tillegg blir det tatt opp hvorvidt det utgjør noen forskjell for potensielle fillere om en 'parasitic gap' opptas av et pronomen, og vist hvordan resumptive pronomener ikke kan ha *syntaktisk* binding til en filler

4.1 Begrensninger for personlige pronomener

I norsk er den eneste umulige antesedenten til et personlig pronomen i objektsposisjon subjektet i den samme setningen (gitt at antesedenten samsvarer semantisk) (21-a). Den eneste mulige antesedenten til et refleksivt pronomen er subjektet i setningen den forekommer i (21-b).³⁵

- (21) a. Lars_i vasker ham_{*i}
b. Lars_i vasker seg_i

Det samme gjelder for leddsetninger. Personlige pronomener i objektsposisjon i en leddsetning kan referere til et element i matricesetningen (*Per* i (22-a)) eller en hvilken som helst enhet utenfor setningen, som samsvarer i tall, kjønn og person med pronomenet (bortsett fra objektet i leddsetningen) (22-a). Refleksivt pronomen kan kun referere til subjektet (22-b).

- (22) a. Lars_i sier at Per_j vasker ham_{i/*j}
b. Lars_i sier at Per_j vasker seg_{*i/j}

I (22-a) begrenses mulige antesedenter for *ham* av at *seg* er det eneste pronomenet som kan referere til subjektet. Subjektet kan ikke være antesedenten til et personlig pronomen i objektsposisjon.³⁶

³⁵I de følgende eksemplene vil jeg for enkelthets skyld kun indeksere anaforene etter hvilke antesedenter de kan ha i den gjeldende setningen. Koindexering av f.eks. *Lars* og *ham* i (22-a) viser mulig anaforisk referanse i denne setningen. *ham* kan også referere til enheter utenfor setningen, selv om indexeringen ikke viser det.

³⁶*hverandre* har riktignok de samme syntaktiske egenskapene som *seg*, men er ikke relevant her (siden *hverandre* krever koordinert subjekt).

4.2 Begrensninger for gap

Setninger som (23) er både grammatiske og akseptable.

(23) Lars_i sier at Per_j ikke vet om Frank_k tror at Åge_m liker ham_{i/j/k}

Årsaken til at slike setninger aksepteres på tross av muligheten for binding mellom elementer som befinner seg langt fra hverandre (f.eks. *ham* og *Lars*) kan skyldes at anaforisk referanse ikke er begrenset av like strenge regler som f.eks. langdistanseavhengigheter.

Når langdistanseavhengighet på samme måte krysser flere setningsgrenser kan det være vanskelig å holde tritt med hvilket element gappen representerer. I (24) må referansen til filleren (akkurat som i 'verste fall' i (23)) krysse tre setningsgrenser før det blir klart hvilket element gappen ikke kan referere til (*Per*). Denne setningen bryter ikke med øyrestriksjonene, men kan på grunn av dybden til gappen av mange regnes som uakseptabel.

(24) Lars_i sier Per_j at han_j ikke vet om Frank_k tror at Åge_m liker $_i$

4.2.1 Parasitic gaps

På spørsmålet om hvorfor fenomenet 'parasitic gaps'³⁷ eksisterer, svarer Chomsky (1982) følgende:

The phenomenon arises from the null hypothesis, the optionality of phonological content for pronominals, the same property that accounts for control and the possibility of 'missing' subjects in pro-drop languages.

Det er altså PG (nullelement) som skriver seg fra bruken av pronomen (fonologisk element), og ikke omvendt. Hvis det skal kunne være snakk om ekte valgfrihet i bruken av et uttrykk for et gitt innhold, må alle de alternative uttrykkene representere det samme innholdet (dvs. ha den samme referansen. I det følgende vil jeg vise at dette ikke er tilfellet for PG (nulluttrykk) og pronomen (fonologisk uttrykk).

NRG (s. 1105) viser hvordan 'parasitic gaps' kan brukes i norsk. Engdahl (1983, s. 5) sier følgende om PG: *Sentences with parasitic gaps show that one filler may control more than one gap.* Dette er illustrert i (25) (der '...' markerer PG).

(25) Lars_i sier Per_j at $_i$ ikke vet om Frank_k tror at Åge_m liker $_i$

³⁷NRG bruker termen 'snyltarhol' for 'parasitic gaps'. Heretter vil jeg bruke forkortelsen 'PG'.

I (25) er PG avhengig av den 'ekte' gappen for å kunne referere til *Lars*. Noe av det som skiller PG fra en 'vanlig' gap er at det er det ikke nødvendigvis endrer grammatikaliteten å erstatte PG med et pronomen:

(26) $Lars_i$ sier Per_j at ${}_i$ ikke vet om $Frank_k$ tror at $\overset{\circ}{A}ge_m$ liker $ham_{i/j/k}$

De to foregående eksemplene *kan* ha den samme betydningen, men har det ikke nødvendigvis. I (25) har PGen kun én mulig referent (filler), mens pronomenet den er erstattet av i (26) har tre mulige referenter (antesedenter). Pronomenet i (26) har akkurat de samme referansemulighetene som *ham* har i (23): Selv om det er plassert i posisjonen til en PG har det ikke samme funksjon som en gap (Det er den syntaktiske strukturen som bestemmer hvilket element gappen representerer). Et slikt pronomen er (i motsetning til PG) ikke avhengig av at det finnes en gap for å kunne referere nøyaktig, selv om den *kan* ha gappen som antesedent. Følgende eksempler viser en setning tilsvarende (25) uten en ekte gap (27-a), og tilsvarende (26) uten en ekte gap (27-b):

(27) a. $*Lars_i$ vet ikke om $Frank_k$ tror at $\overset{\circ}{A}ge_m$ liker $..._{*i/*k/*m}$
 b. $Lars_i$ vet ikke om $Frank_k$ tror at $\overset{\circ}{A}ge_m$ liker $ham_{i/k}$

Her blir forskjellen mellom PG og et pronomen på PGs plass tydelig: PG kan ikke referere uten en gap, mens *ham* fremdeles kan referere til alle nominale elementer det samsvarer med, og som referanse med ikke fører til at det brytes noen restriksjoner. Årsaken til at (27-a) er ugrammatisk er at (forsøket på en) PG mangler en gap å referere til – noe som skyldes at det mangler et topikalisert subjekt. Dette fører igjen til at *liker* står uten objekt. (27) viser også at en gap er mer 'treffsikker' i referansen enn personlige pronomen. Mangelen på den 'ekte' gappen i (27-a) setter 'PG' ute av stand til å referere til noe (noe som strengt tatt gjør at den ikke lenger er PG).

Det finnes altså fordeler både ved PG og ved resumptive pronomen: PG er et mer nøyaktig 'redskap' for referanse enn et pronomen, mens pronomenet har en mer fleksibel distribusjon enn PG; det er ikke avhengig av en 'ekte' gap å bindes til. I tillegg kan et personlig pronomen på PG-plass – i setninger med flere nominale ledd med ulike attributter for kjønn, tall og person – minne lytteren om hvilket nominalt element fra tidligere i setningen (eller diskursen) det refereres til på den pronomenbesatte PG-plassen. Intuitivt synes det vanskeligere å holde rede på hvilket element som hører til PG når den 'ekte' gappen (som PG jo er bundet til) ikke inneholder noen informasjon om filleren. Dermed kan det være lang lineær avstand mellom den semantiske informasjonen i filleren, og mangelen

på denne i en PG. Denne semantiske informasjonen er noe et pronomen på PGs plass i det minste til en viss grad kan formidle.

4.2.2 Er PG pronominale?

For oversikten skyld samles eksemplene (24), (25) og (26) her, som (28):

- (28)
- a. Lars_i sier Per_j at han_j ikke vet om Frank_k tror at Åge_m liker _i
 - b. Lars_i sier Per_j at _i ikke vet om Frank_k tror at Åge_m liker _i
 - c. Lars_i sier Per_j at _i ikke vet om Frank_k tror at Åge_m liker ham_{i/j/k}

Samtlige av disse setningene er grammatiske, men det skjer til dels drastiske endringer i betydningen av dem når man veksler mellom bruk av gap og pronomen:

- I (28-a) må *han* tolkes som anaforisk, utvetydig refererende til *Per*, siden referering til et topikalisert objekt (som *Lars*) i slike konstruksjoner kun kan bindes til en gap. Slik er også gapen utvetydig bundet til *Lars*.
- I (28-b) må den ekte gapen (som i (28-a)) være bundet til *Lars*, og PG være indirekte bundet til *Lars*, via den ekte gapen.
- I (28-c) er PG (lovlig) byttet ut med et pronomen. Dette eksempelet viser at avhengighetsforholdet mellom PG og ekte gap ikke er gjensidig: PG er avhengig av den ekte gapen for å referere,³⁸ mens den ekte gapen fungerer når den forekommer i setninger uten PG (28-c) på akkurat samme måte som når PG er til stede (28-b). Det kommer med andre ord ikke noe forvarsel i den ekte gapen om at den skal bindes til en PG.

Vekslingen mellom gap og pronomen viser også variasjonen i mulige referanser for de to elementene. Det eneste tilfellet der gap og pronomen *kan* referere til det samme elementet er der et pronomen er satt inn for PG i (28-c). Plasseringen av *ham* på PGs plass i denne setningen reduserer ikke grammatikaliteten, men kan – som dette eksempelet viser – redusere forståeligheten (ved å gjøre setningen flertydig).³⁹

Den siste mulige kombinasjon av gap og pronomen i setninger som dem i (28) er (29).

³⁸ Dette kommer til syne i (28-a), der det samme elementet som er en PG i (28-b) er en uavhengig gap (siden det ikke finnes noen andre gapper i setningen).

³⁹ Riktignok er dette et eksempel som er konstruert for å vise muligheten for flertydighet ved utskiftning av PG med et pronomen. I de fleste tilfeller vil det være mulig å forstå hva som er antesedenten til en pronomen på PGs plass ved samsvar i person, tall og kjønn. Likevel viser (28-c) at setningen kan bli

(29) *Lars_i sier Per_j at han_j ikke vet om Frank_k tror at Åge_m liker ham_{i/j/k}

Her hindres en tolkning av *Åge* som antesedenten til *ham* av forbudet ikkerefleksive objektspronomen har mot å referere til sitt eget subjekt. Da gjenstår *Lars*, *Per* og *Frank* som mulige antesedenter til *ham*. En tolkning av *Per* eller *Frank* som antesedent byr på et åpenbart problem: *Lars* står uten tilhørighet til et predikat. Dermed tvinges *ham* til å ta *Lars* som antesedent, i mangel på andre verb enn *liker* som *Lars* kan være argument til. Det er ikke altså ikke syntaktisk *binding* som binder *Lars* til *ham*, men snarere syntaktisk *nødvendighet*, siden en tolkning av andre elementer enn *Lars* som antesedenten til *ham* fører til brudd på prinsippet om koherens. Det er altså en ekstralingvistisk nødløsning, eller det som i 3.4 har blitt omtalt som 'last resort'. Slik kan også pronomen få 'obligatorisk' referanse på samme måte som gappen i (28-a), og i motsetning til pronomenet i (28-c). Denne 'obligatoriske' referansen kreves ikke av grammatikken (snarere tvert imot), men den kan øke akseptabiliteten.

Chomsky (1982, s. 38) mener som nevnt at PG er en form for pronomen – en valgfri nullvariant av det pronomenet som kan ta PGs plass. Dette skulle tilsi at den eneste forskjellen mellom '...' i (28-b) og *ham* i (28-c) er formen. Det har jeg over vist at ikke er tilfellet. Mens PG har syntaktisk binding til den 'ekte' gappen, på samme måte som den 'ekte' gappen har syntaktisk binding til filleren (som '...' har til *Lars* i (28-a)), har *ham* på PG sin plass i (28-c) de samme mulige referentene som et 'vanlig' pronomen, nemlig alle nominale elementer med attributtene 'tredje person+entall+hankjønn' som ikke er subjektet til verbet pronomenet står som objekt til (på akkurat samme måte som *ham* har i (23)). Muligheten et pronomen på PGs plass ofte har til å referere på samme måte som PG selv kan forlede oss til å tro at et slikt pronomen i praksis er et fonologisk uttrykt PG. På samme måte kan kravet filleren til et RP har til å tildeles en syntaktisk rolle av et verb (som vist ved (29)) få oss til å tro at bindingen mellom RP og filleren er av samme art som den mellom en gap og filleren, mens det i virkeligheten er ekstralingvistiske årsaker (forståelighet og/eller akseptabilitet) som motiverer RP.

Det eneste en PG og et pronomen har til felles er distribusjonen i konstruksjoner som tillater PG. Det eneste RP har til felles med en gap er distribusjonen i konstruksjoner med filler-gap-avhengigheter.

'uforståelig' (-F) (pga. flertydighet) ved bruk av pronomen i slike tilfeller, heller enn PG (som i (28-b), som ikke er flertydig, og dermed 'forståelig' (+F)).

4.3 RPs mulige antesedenter

Følgende eksempler viser gap som subjekt (30-a) og som objekt (30-b).

- (30) a. Lars_i sier Per_j at _i vasker ham_j
b. Lars_i sier Per_j at han_j vasker _i

Her er det *gappens* plassering som avgjør anaforenes referanse, siden *ham/han* ikke kan referere til et koargument: Personlige pronomener i objektsposisjon kan ikke ta subjektet som antesedent (det er det det refleksive pronomener som gjør), og personlige pronomener i subjektsposisjon kan ikke ta objektet som antesedent. Dermed *må* det vedkommende pronomener referere til det argumentet i matrisesetningen som ikke er filler.

I (31) vises to mulige tolkninger av den flertydige (og problematiske) ytringen *#Lars sier Per at han vasker ham*⁴⁰. De to tolkningene er RP-versjoner av setningene i (30): Gappen er byttet ut med en RP (uthevet). Indekseringen viser ikke *faktiske* bindinger, men bindingene i den gjeldende tolkningen.

- (31) a. #Lars_i sier Per_j at **han**_i vasker ham_j
b. #Lars_i sier Per_j at han_j vasker **ham**_i

I disse ytringene gis det ikke noe klar indikasjon på hvilken anafor som refererer til hvilken antesedent, siden RP er 'forkledd' som et vanlig pronomener. At RP har flere mulige antesedenter (selv om det i virkeligheten er fonologisk uttrykk for en gap) gjør at også det 'ekte' pronomeners antesedent er ubestemmelig.

4.3.1 Er RP fonologisk uttrykk for en gap?

Eksempel (22-a) gjengis for enkelhets skyld her som (32):

- (32) Lars_i sier at Per_j vasker ham_{i/*j}

Forskjellen mellom eksemplene i (32) og (31) er illustrerende for en viktig forskjell mellom langdistanseavhengighet og personlige pronomeners anaforiske referanse. Mens det er strengt syntaktisk bestemt i hva som kan være filleren til en bestemt gap, er personlige pronomeners mulige antesedenter begrenset av *semantiske* krav (i (31) hankjønn, entall), og i liten grad av syntaktiske krav (antesedenten *må* være ikkesubjekt, som motiverer asterisken i (32)). Slik minner langdistanseavhengighet mer om forholdet mellom et refleksive

⁴⁰Her bruker jeg '#' til å merke setninger som kan være akseptable, men likevel er problematiske.

pronomen og en antesedent, siden både en filler og antesedenten til et refleksivt pronomen er syntaktisk bestemt, og i at refleksive pronomen i mindre grad enn personlige pronomen er avhengig av semantiske forhold (i og med at *seg* er kjønn- og talluavhengig (om enn ikke personuavhengig)). Engdahl (1983, s. 29) betegner PG som *null anaphors* snarere enn et fonologisk uttrykt gap. I så fall er det snakk om en nullanafor med mer avgrenset referanse enn f.eks. personlige pronomen, siden 'nullanaforen' PG *må* ha en gap som 'antesedent'.

I motsetning til et personlig pronomen kan verken en gap eller et refleksivt pronomen ha flere mulige referenter (dvs. filler/antesedent). Riktignok er RPs anafor entydig dersom eventuelle andre anaforer har andre attributter (f.eks. #*Lise_i sier Per_j at hun_i vasker ham_j*), men dette gir en *semantisk* avgrensning av RPs mulige referenter (i motsetning til langdistanseavhengighetens *syntaktiske* avgrensning). Mens en gap har syntaktisk (entydig) referanse, har et personlig pronomen semantisk (potensielt flertydig) referanse. Det er derfor personlige pronomener er uegnet som fonologiske erstatninger for gap.⁴¹

⁴¹At semantisk referanse er potensielt flertydig betyr at hvorvidt det kan refereres til flere enn ett element avhenger av semantikken (i dette tilfellet kjønn-, tall- og personattributtene til eventuelle andre nominale elementer i setningen). At syntaktisk referanse er entydig betyr at slik referanse er den samme, uavhengig av andre elementer i setningen. Riktignok kan syntaktiske avhengigheter være bundne (og ha absolutt referanse til ett bestemt element) eller ubundne (og ha relativ referanse til en posisjon i setningen), men syntaktisk referanse er noe annet enn – og uavhengig av – semantisk samsvar.

5 Øyrestriksjoner

I dette kapittelet vil det diskuteres om øyrestriksjonene er grammatiske begrensninger, eller om de i virkeligheten baserer seg på akseptabilitet. Noe som kan tyde på at konstruksjoner som bryter med øyrestriksjonene ikke nødvendigvis er ugrammatiske, er den varierende graden av akseptabilitet slike brudd har mellom forholdsvis nært beslektede språk (som engelsk, norsk og svensk). Det vil også refereres til en mulig forklaring for hvorfor øyrestriksjonene som gjelder for engelsk ikke gjelder for svensk – en forklaring som også er gyldig for norsk.

5.1 Øyrestriksjoner og akseptabilitet

I prinsippet er det ingen grenser for hvor lang avstanden mellom en filler og en gap kan være i en setningsknote. NRG (s. 1096f) viser to typer avstand mellom filler og gap: Strukturell (33-a) (NRGs klammer) og lineær (33-b) (fra Saltveit, 1994):

- (33) a. Denne bilen_i trudde eg [du meinte [at ho ville bli glad [om eg kjøpte —_i]]]
b. Den personen som kom til å oppta meg noe, var en nordmann_i, som Engen Haugland, mest kjent som idrettsmann – og dessuten både klassekamerat og turvenn, men nå i politiet, bad meg være til stede ved forhøret av —_i

Mens filler-gap-relasjonen i eksempel (33-a) krysser tre setningsgrenser, skyldes den lange avstanden i (33-b) en lang subjekts-NP i setningen gappen befinner seg i (*Engen Haugland, mest kjent som idrettsmann – og dessuten både klassekamerat og turvenn, men nå i politiet*). Eksempelene i (33) bryter ikke med noen grammatiske regler. Likevel kan den lange strukturelle og lineære avstanden mellom filler og gap redusere akseptabiliteten i slike setninger. Intuitivt (ut fra eksemplene i (33)) reduseres forståeligheten i slike setninger mer med stor lineær avstand (33-b) enn der den strukturelle avstanden er stor (33-a).

Det finnes også eksempler på setningsknuter som regnes som ugrammatiske på grunn av brudd på øyrestriksjonene, men som regnes som akseptable. NRG (s. 1103) viser noen av setningene Creider (1987) i en undersøkelse fant at ble regnet som akseptable av flertallet av forsøkspersonene:

- (34) a. Her er den bakken som vi aldri et mat før vi går opp
b. Kva melodi puttarr du bomull i øyra når han spelar?

Jeg har tidligere vist eksempler på setninger som er ugrammatiske, men er likevel akseptable. Setningene i (34) er på grunn av brudd på øyrestriksjonene regnet som ugrammatiske, men de er både akseptable og forståelige. Eksempelet jeg tidligere har brukt for setninger som er $-G +A +F$ er å finne i tabell 1, her gjengitt som (35).

(35) Per har en rar måte å sitte på bussen

Forskjellen mellom setningene i (34) og den i (35) er årsaken til at de er ugrammatiske: (35) er ugrammatisk fordi den bryter med det generelle prinsippet om koherens, mens (34) er ugrammatiske fordi de bryter med øyrestriksjonene for adjunkter (ingenting kan ekstraheres fra en adjunkt). På samme måte som inkohrens gjør det, fører brudd på øyrestriksjonene ofte til inakseptabilitet.⁴² I det følgende vil jeg prøve å finne ut om det er hensynet til akseptabilitet eller grammatikalitet øyrestriksjonene er definert ut fra.

5.1.1 'The Complex NP Constraint'

Jeg vil nå vise noe av resonnementet som ligger bak en av øyrestriksjonene (*The Complex NP Constraint*), som ble presentert av Ross (1967).⁴³ Eksemplene i (36) er hentet fra Ross (1967, s. 27). (37) er hentet fra Chomsky (1964), og gjengitt av Ross.

- (36) a. He told me about a book which I can't figure out $\left\{ \begin{array}{l} \text{whether to buy or not.} \\ \text{how to read.} \\ \text{where to obtain.} \\ \text{what to do about.} \end{array} \right\}$
- b. He told me about a book which I can't figure out $\left\{ \begin{array}{l} \text{why he read.} \\ \text{?whether I should read.} \\ \text{??when I should read.} \end{array} \right\}$
- c. Which books did he tell you $\left\{ \begin{array}{l} \text{why} \\ \text{?whether} \\ \text{??when} \end{array} \right\}$ he wanted to read?

(37) *what did he wonder where John put?

⁴²Senere vil jeg vise at norsk ikke er underlagt de samme øyrestriksjonene som f.eks. engelsk.

⁴³I avsnittene som følger vil jeg også vise til andre øyrestriksjoner, men går kun til en viss grad inn på ett av dem her. De øvrige restriksjonene virker på samme måte, i at de forbyr ekstrahering av et element fra visse konstruksjoner.

Ross vedgår – i resonnementet som leder til en justering av Chomskys (1964, s. 43f) betingelser⁴⁴ for reglene som styrer dannelsen av spørsmål og relativfraser (og i sin tur til bl.a. øyrestriksjonen for komplekse NPer)⁴⁵ – at han baserer seg på akseptabilitet⁴⁶:

For some reason that is obscure to me, I find sentences like those in [(36-a)], where the embedded question consists of a wh-word followed by an infinitive, by and large more acceptable than corresponding sentences, like those in [(36-b)], where the wh-word is followed by a clause with a finite verb. And yet there are many sentences which differ in no way which I can discern from those in [(36-b)-(36-c)] but which I find totally unacceptable. (Chomsky's example, [(37)] is a good case in point). (Ross, 1967, s. 27)

Det er altså Ross' egne akseptabilitetsvurderinger behovet for en justering av Chomskys betingelser for reglene som styrer dannelsen av spørsmål baserer seg på, siden disse betingelsene regner setninger som dem i (36) som ugrammatiske.

5.2 Norske øyrestriksjoner

Mens engelske setninger som bryter med øyrestriksjonene som oftest synes å være ugrammatiske, er dette ikke alltid tilfellet for norsk. Se eksempel (38-a)-(38-b) Creider (1987, s. 7) og deres norske motsvar (38-c)-(38-d) (Creiders asterisker):

⁴⁴Betingelsen går ut på at transformasjonen som får som resultat et spørsmål eller en relativsetning *must be limited to a single application to each underlying terminal string*, dvs. at en *wh*-setning ikke selv kan inneholde en *wh*-setning.

⁴⁵*The Complex NP Constraint: No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation* (Ross, 1967, s. 127)

⁴⁶Ross (1967) spesifiserer ikke om asterisk og 'graderte' asterisk (*?. ? og ??) markerer grammatikalitet eller akseptabilitet. Det er setninger som på et nivå regnes som problematiske som merkes med asterisk. Ved minst ett tilfelle markeres en setning som er 'unintelligible' (uforståelig) med asterisk **10 dollars was cost by the parking ticket*, mot en setning som er uttalt ugrammatisk, uten asterisk (*they forced me for me to wash myself*). (Ross' begrep 'intelligible' er åpenbart et annet enn 'forståelig' som det er definert i 1.4.) Dersom asterisk betyr 'uakseptabel', vil det være Ross' egen intuisjon (snarere enn faktisk grammatikk) som gir grunnlaget for argumentasjonen – noe som ikke er helt uproblematisk: Hvis asterisk betyr 'ugrammatisk' vil argumentet gå i sirkel: En setning er ugrammatisk fordi det har blitt definert en regel (basert på intuisjon) som sier at den er ugrammatisk.

- (38) a. *[This road]_i I don't know where _i goes
 b. *[Which book]_i do you deny the possibility that Kendall wrote _i?
 c. [Denne veien]_i vet jeg ikke hvor _i går
 d. [Hvilken bok]_i benekter du muligheten for at Kendall skrev _i?

Eksemplene viser ekstraksjon fra en *wh*-konstruksjon ((38-a) og (38-c)), og en kompleks NP ((38-b) og (38-d)). Det er åpenbart ulike regler for engelske og norske øyrestriksjoner. Det er konstruksjoner som dem i (38-c)-(38-d) Helland & Christensen (1986) påpeker som skandinaviske unntak fra øyrestriksjonene for germanske språk:

Scandinavian differs from other Germanic languages, including English, in that the rule of Wh-movement is only to a limited subject to such restrictions as Ross' (1967) "Wh-island Constraint" and the "Complex NP Constraint". (s. 4)

Videre vises det hvordan de øvrige av Ross' restriksjoner gjelder for norsk. Også mellom skandinaviske språk varierer graden av akseptabilitet til ulike typer langdistanseavhengigheter. Helland & Christensen (1986) viser til det Engdahl (1984a,b) rapporterer om akseptabiliteten til svenske, danske og finlandsvenske versjoner av setningene i (39).

- (39) a. Petter_i vet jeg at _i skal komme
 b. Petter_i vet jeg ikke om _i skal komme
 c. Petter_i vet jeg ikke hva_j _i sa _j

Av disse typene konstruksjoner er kun (39-c) uakseptabel på dansk, og (39-a) uakseptabel på finlandssvensk. Ifølge Engdahl er ingen av dem tillatt⁴⁷ på svensk, i motsetning til tilsvarende setninger med et resumptivt pronomen i subjektsposisjon (for _i) (f.eks. *Petter_i vet jag inte vad_j han_i sa _j*). Den store variasjonen mellom så nært beslektede språk som de skandinaviske når det gjelder denne typen konstruksjoner (og den antatte variasjonen i akseptabilitet for langdistansekonstruksjoner mellom dialekter innad i språkene) gir grunn til mistanke om at øyrestriksjonene som tilsynelatende gjelder disse språkene ikke er grammatiske begrensninger. Det synes som om avvisningen av konstruksjoner som dem i (39) i skandinaviske språk er så sterkt befestet i språket at brudd på 'øyrestriksjonene' gir inntrykk av å bryte med grammatikken uten virkelig å gjøre det. I de 5.3 vil jeg vise hvordan dette kan være en årsak til at øyrestriksjonene regnes som en del av grammatikken, uten nødvendigvis å være det.

⁴⁷Det kommer ikke frem om det angår grammatikalitet eller akseptabilitet når det snakkes om 'tillatt' om slike setninger i svensk.

5.2.1 Uenighet om norske øyrestriksjoner

Det foregikk på 70- og 80-tallet en diskusjon i litteraturen om hvorvidt enkelte av de norske avvikene fra øyrestriksjoner som gjelder f.eks. engelsk egentlig var avvik (dvs. at visse konstruksjoner var uakseptable), eller om norsk virkelig var 'fritatt' fra slike restriksjoner. Blant annet ble det diskutert hvorvidt et *hv*-element kan ekstraheres fra en interrogativ setning. Taraldsen (1978) gjengir Malings (1977) eksempel på en slik konstruksjon (40-b) og en konstruksjon der filleren ekstraheres fra en *som*-relativ (40-a) (Malings asterisk). Taraldsen understreker at det forskjell på akseptabiliteten til disse ytringene i hans egen tale (riksmål). Samtlige av de norske språkbrukerne han har spurt har også sagt seg enige med ham. *Both are acceptable, although somewhat marginal.* (Taraldsen, 1978, s. 635):

- (40) a. [en sang]_i som ingen vet hvem som har skrevet —_i
b. *[Hvilken sang]_i vet ingen hvem som har skrevet —_i

Videre forklarer han at samtlige norsktalende han har konsultert om forskjellen i akseptabiliteten til disse ytringene, har ment det samme som ham: (40-b) er ikke mindre akseptabel enn (40-a).

Maling (1978) forklarer forskjellen mellom sine funn og Taraldsen og hans informan- ters vurdering av akseptabiliteten til setninger som (40-b) med dialektforskjeller. Hennes asterisk i (40-b) baserer seg på et bredere utvalg informanter: (...) *since my informants so consistently report that there is a difference, I feel that the asymmetry (...) represents a linguistic reality for many dialects.* (Maling, 1978, s. 81) Videre forklarer hun forskjellen mellom hennes brede og Taraldsens smale utvalg informan- ters akseptabilitetsvurderinger med at en syntaktisk endring er i emning. Muligheten for ekstraksjon fra 'som-konstruksjoner' har smittet over på *hv*-konstruksjoner, som dermed aksepteres analogt med de akseptable ekstraksjonene fra *som*-konstruksjoner.

Lie (1982) har et standpunkt mellom Malings og Taraldsens. Han mener det er en forskjell i grammatikaliteten⁴⁸ til setninger som (40-a) og (40-b). Han mener denne asymmetrien er tilstede, men ikke i like stor grad som Maling hevder, og at den er av en annen art: *I will claim that this difference has nothing to do with relativization and wh-movement as such, but is due to other facts* (Lie, 1982, s. 200). Lies påstand er at Maling og hennes informan- ters vurdering av grammatikaliteten skyldes at (40-b) har en uklar denotasjon, siden *hvilken sang* ikke refererer til noen bestemt enhet, samtidig som den er bundet til

⁴⁸Lie ser ut til å bruke begrepet *grammatical* her på samme måte som Taraldsen og Maling bruker *acceptable*.

en gap. Setningen blir for uklar til at den kan gi skikkelig referanse. Lie viser hvordan clefting kan hjelpe setninger tilsvarende (40-b), og konkluderer med at det finnes analoge, grammatiske setninger (Lie, 1982, s. 202) (min indeksering):

(41) Hva_i veit alle hva $_i$ er (for noe)?

Lie hevder med andre ord at Malings eksempler er ugrammatiske av andre årsaker enn brudd på øyrestriksjonene, noe som skulle tilsi at også norske *hv*-konstruksjoner er unntatt øyrestriksjonene.

5.2.2 Hvorfor er norsk unntatt øyrestriksjonene?

Taraldsen (1978) mener variasjonen i tilfeller der norske setninger *kan* (men ikke nødvendigvis) underlegges øyrestriksjoner (f.eks. 'Complex NP constraint') skyldes andre forhold enn selve den syntaktiske strukturen:

- (42) a. Hvilket krydder_i visste du om en butikk som solgte $_i$ billig?
b. *Hvilket krydder_i visste du om den butikken som solgte $_i$ billig?

Forskjellen mellom eksemplene i (42) viser et eksempel på det Taraldsen kaller *certain narrow conditions* som han mener brudd på 'the Complex NP constraint' kan godtas på grunn av: I (42) er det bestemtheten til kjernen i den komplekse NPen gappen befinner seg i som avgjør om bruddet på øyrestriksjonen har et akseptabelt utfall eller ikke. Taraldsen viser ikke til annet enn sin egen intuisjon i akseptabilitetsvurderingen av disse setningene. Likevel – enten (42-b) er akseptabel eller ikke – forklarer ikke dette hvordan (42-a) kan være akseptabel selv om den bryter med øyrestriksjonene som gjelder for engelsk og mange andre språk. Disse restriksjonene fører til at de engelske oversettelsene av begge eksemplene i (42) er ugrammatiske:

- (43) a. *What spice_i did you know of a shop that sold $_i$ cheap?
b. *What spice_i did you know of the shop that sold $_i$ cheap?

Min egen intuisjon samsvarer med Taraldsens når han vurderer (42-a) som mindre akseptabel enn (42-b), men det synes ikke å være noen større grunn til å regne sistnevnte som uakseptabel eller mer av et brudd på en øyrestriksjon enn (42-a), siden det ikke er noen *strukturell* forskjell mellom de to. Taraldsens eksempler viser at bestemthet kan forstyrre parseren, og dermed få setninger til å virke 'tyngre', men ikke at det er noe ved reglene for språklig struktur som hindrer ekstraksjon fra *som*-konstruksjoner med bestemt kjerne.

Engdahl (1982) mener også at at brudd med visse øyrestriksjoner aksepteres i skandinaviske språk skyldes noe annet enn syntaktiske forhold:

(...) I still seriously question the assumption that restrictions on extractions should be explained in structural terms. I believe that the types of restrictions on unbounded dependencies that we encounter in Swedish are not primarily structural in nature. Consequently they need not be stated in the syntax, which thus can be kept maximally simple. (s. 171)

Taraldsens (1978) syn på øyrestriksjoner er noe av det samme: Slike 'brudd' med øyrestriksjonene er ikke egentlige brudd på øyrestriksjonene. Restriksjoner som 'Complex NP Constraint' finnes egentlig ikke i norsk, og forekomsten av setninger som tilsynelatende overholder disse restriksjonene skyldes snarere at *certain narrow conditions* er til stede i setningen. Engdahls beskrivelse av hva disse tilsynelatende overholdelsene av restriksjoner i skandinaviske språk egentlig består i er *facts having to do with how these sentences can be used, i.e. facts about listeners' limited processing capabilities and speakers' desire to make themselves understood*. Unntakene fra øyrestriksjoner i skandinaviske språk skyldes altså ikkelingvistiske forhold. Derfor verken kan eller bør disse øyrestriksjonene beskrives strukturelt, siden 'ulovlig' ekstrahering kun er et verktøy for kommunikasjonen. Siden restriksjonene kun er til hjelp for kommunikasjonen (dvs. uavhengig av kompetansen), er ikke en strukturell forklaring (som dreier seg om kompetansen) nødvendig. I så fall er det vanskelig å begrunne eksistensen av øyrestriksjoner overhodet, også for engelsk – noe jeg ikke (fra verken det ene eller det andre standpunktet) skal begi meg ut på i denne oppgaven.

Som nevnt tidligere er performansen mer enn et rent uttrykk for kompetansen. Vi kan ta i bruk konteksten (f.eks. ved bruk av pronomen som refererer til ikkespråklige elementer) for å forenkle språket, eller ugrammatiske grep kan tas, som bruk av resumptive pronomen. Slike kommunikative 'støtter' brukes for å appellere til mottakerens språk eller språkfølelse. Det er ikke den ubevisste parseren som håndterer ikkespråklige elementer som disse, men en bevisst 'kanal' for kommunikasjon som tillegg til språket. Det kan tenkes at øyrestriksjonene gjør seg gjeldende fordi de er upraktiske og ikke egnet til bruk ved kommunikasjon. I så fall er årsaken til at norsk i så stor grad bryter med øyrestriksjonene vanskelig å begripe, for ikke å snakke om den tilsynelatende tilfeldige distribusjonen av aksept for brudd på ulike restriksjoner for dansk og finlandsvensk. På samme måte er det vanskelig å forstå hvorfor resumptive pronomen ofte foretrekkes fremfor gap i – også i de tilfellene gap er helt uproblematisk på norsk.

Engdahl (1982, s. 168) presenterer en norsk setning med to fillere i posisjoner som gjør det vanskelig å avgjøre hvilken gap som er bundet til hvilken filler, og den svenske (utvetydige) oversettelsen av setningen (Mine indekseringer av gappene i (44-a)):

- (44) a. Det var Lisa_i læreren spurte hvilken gutt_j vi trodde —_j var sint på —_i
b. Det var Lisa_i läraren frågade vilken kille_j vi trodde hon_i var arg på —_j

Engdahl indekserer ikke selv denne setningen, for å vise hvordan refereringen i den mest sannsynlige konteksten foregår mellom de samme elementene som i den svenske. Men hun sier selv at *[o]ut of context, a sentence like [(44-a)] receives a nested interpretation* (dvs. som i min indeksering).

Intuitivt (med norsk intuisjon) er (44-b) flertydig. De syntaktiske og anaforiske bindingene kan like godt som i Engdahls tolkning være som i (45), der *Lisa* blir tolket som filler til gappen i relativsetningen *Lisa* selv modifieres av, lik min tolkning av gappen som *hvilken gutt* er filler til i (44-a).

- (45) Det var Lisa_i läraren_k frågade —_i vilken kille_j vi trodde hon_k var arg på —_j

En tolkning som denne fordrer riktignok en forståelse av det resumptive pronomenet *hon* (som et fonologisk uttrykt gap) som et ekte pronomen, som refererer *anaforisk* til *läraren*⁴⁹. En tilsvarende tolkning av gappen som følger *trodde* i den norske versjonen (44-a) er ikke mulig: Det er bare *Lisa* (som del av et topikalisert ledd) og *gutt* (som del av en *hv*-nominal) som kan fungere som filler til gappen. Dette viser igjen (som vist i 4.2.1) både hvordan filler-gap-konstruksjoner refererer mer nøyaktig enn pronomenen, og at resumptive pronomen refererer anaforisk, og har derfor samme funksjon som (og *er* i virkeligheten) vanlige pronomen.

Som nevnt i 5.2 er samtlige av de norske setningene som bryter med (de engelske) øyrestriksjonene akseptable på svensk, men gappen uttrykkes ved RP. I tilfeller med én gap kan det synes marginalt fordelaktig å ha fonologisk uttrykk fremfor et nullelement i posisjonen til gappen, men bruken av det fonologiske uttrykket (pronomenet) tar med seg andre egenskaper uttrykket har, som muligheten til å referere friere enn en gap.

En mulig forklaring på hvorfor svensk står friere til å bryte øyrestriksjonene som gjelder for f.eks. engelsk, gis av Allwood (1982, s. 29):

⁴⁹En kontekst med en mannlig lærer ville ikke tillatt en tolkning som (45). Det er ikke usannsynlig at læreren i eksempelet er ment å være mannlig (siden det ikke er snakk om en *lärarinna*), men dette er ikke relevant her. Poenget er å vise muligheten også resumptive pronomen har til å referere til flere enn ett element i setningen.

Topicalization is a much more extensive rule of sentence formation in Swedish than it is in English (...). Since Swedish speakers are so thoroughly accustomed to topicalization, it may be that this conventionally permits us to perceive as acceptable (and certainly as interpretable) many satsflåtor which a language with less common topicalization could not tolerate. In this sense even our syntactic perception, at least in some respects, might be to some extent conventionally determined by the language we speak.

Her pekes det på vane som en mulig årsak til at konstruksjoner som i noen språk regnes for å være akseptable er uakseptable for andre – også nært beslektede – språk (som 'Complex NP Constraint'). Allwood bruker også begrepet 'interpretable' om slike setninger, en betegnelse som synes å sammenfalle med 'forståelig' som definert i 1.4.⁵⁰ Allwood vegrer seg for å si at kompetansen kan endres av performansen gjennom vane, men antyder at måten vi bruker språk på og hvorvidt vi aksepterer visse konstruksjoner kan påvirkes av selve språkbruken.

5.3 Øyrestriksjoner og grammatikalitet

Ross viser ved relativisering (47) av to ulike setninger (46) (som denoterer den samme proposisjonen på litt ulike måter) nøyaktig hva det er som hindrer ekstraksjon av f.eks. *wh*-fraser som befinner seg i en relativsetning (Ross' asterisk).

- (46) a. I read a statement which was about a man.
 b. I read a statement about a man.
- (47) a. *The man who I read a statement which was about is sick.
 b. The man who I read a statement about is sick.

At (47-a) er ugrammatisk mens (47-b) ikke er det, forklarer Ross med Klimas (1964) begrensning for nominalfraser,⁵¹ og sitt eget prinsipp om 'S-Pruning'⁵². Konsekvensen av dette er, ifølge Ross, at (46-a) må transformeres til (47-b) (istedenfor (47-a), siden den

⁵⁰Også de fleste engelske setninger som bryter med øyrestriksjonene er +F. Det lar seg gjøre å forstå hva de er ment å bety fra ytreren – gjennom å se hvilket element gappen er bundet til – selv om de regnes som uakseptable.

⁵¹Klima (1964): *Elements dominated by a sentence which is dominated by a noun phrase cannot be questioned or relativized.*

⁵²S-Pruning: *[D]elete any embedded node S which does not branch (i.e., which does not immediately dominate at least two nodes).* Dette prinsippet er en konsekvens av at det har vært vanlig å regne f.eks.

er ugrammatisk). Dette tillates av 'the Relative Clause Reduction Rule', siden *which was about a man* 'kuttet' av prinsippet om 'S-Pruning'. Et prinsipp som involverer en form for 'S-Pruning' finnes også i LFG, og får de samme konsekvensene: Prinsippet om 'Economy of Expression' (EoE)⁵³.

Det synes som om tilleggingen av prinsippet om S-Pruning er motivert av intuisjonen om at bl.a. (47-a) er ugrammatisk, og at dette begrunnes med at setningen har et mulig kortere uttrykk som gir samme denotasjon i (47-b). Intuitivt vil både (46-b) og (47-b) være både et mer naturlig uttrykk for denne proposisjonen og mer akseptable enn a-alternativene, men å kalle en tilstrekkelig uakseptabel setning *ugrammatisk* er et skritt på vei mot å viske ut skillet mellom dem:

(...) it is quite apparent that a speaker's reports and viewpoints about his behaviour and his competence may be in error. Thus a generative grammar attempts to specify what the speaker actually knows, not what he may report about his knowledge. (Chomsky, 1965, s. 8)

Proposisjonen til setninger som (47-a) er ikke til å ta feil av (m.a.o. er de +F), og at det finnes mer effektive måter å gi uttrykk for den samme denotasjonen på er ikke nok til å kunne regne slike setninger som ugrammatiske. Øyrestriksjonene kan i beste fall hindre setninger som er uakseptable å produseres. Med andre ord synes øyrestriksjonene å være restriksjoner for performansen. Brudd på øyrestriksjonene reduserer akseptabiliteten på samme måte som dype egeninnføyninger (som vist i (4-a)) gjør det.

I seg selv medfører altså brudd på øyrestriksjonene (i) *ikke* redusert grammatikalitet (i motsetning til resumptive pronomer og utelatelse av måteP i måteNP)⁵⁴, (ii) redusert akseptabilitet, og (iii) redusert forståelighet. Et annet eksempel på en setning med disse parametrene (+G, -A, -F) er den dypt egeninnføyde setningen i (4-a): *Mannen hunden katten fryktet jaget falt*. Årsaken til at øyrestriksjonene regnes som uakseptable synes å være den samme som for slike egeninnføyninger: Parseren er ikke i stand til å prosessere et så stort antall predikat-argument-relasjoner.

en NP bestående av en kjerne og en possessor for å være en setning, som har etterlatt seg en S-node med én NP-datter.

⁵³*Economy of Expression: All syntactic phrase structure nodes are optional and are not used unless required by independent principles (completeness, coherence, semantic expressivity)* (Bresnan, 2001, s. 91). EoE er slik mer generelt enn S-Pruning, i at det gjelder alle noder, og ikke bare S-noder.

⁵⁴Definisjonene av måteP og måteNP er gitt i innledningen.

6 Mulige årsaker til feil i parsingen

Uavhengig av om brudd på øyrestriksjonene og bruk av RP reduserer akseptabiliteten, grammatikaliteten eller både akseptabiliteten og grammatikaliteten, er det trolig feil under parsingen (ved produksjon og/eller persepsjon) av slike konstruksjoner som gjør at ugrammatiske konstruksjoner produseres og at enkelte grammatiske konstruksjoner regnes som uakseptable. I dette kapitlet vil jeg vise mulige årsaker til – og følger av – at parseren svikter.

6.1 Parserens egenskaper

Det er vanskelig å si nøyaktig hvordan mekanismen som prosesserer språk i en språkbrukers hjerne opererer, men det lar seg gjøre å si noe om egenskapene til parseren. Milne (1982, s. 349) betegner parseren som (i) *unconscious*, (ii) *deterministic and* (iii) *fast, but* (iv) *limited* (min nummerering).

At parseren er ubevisst (i) betyr at den er adskilt fra generell menneskelig tankevirksomhet. Språkbrukeren behøver ikke å tenke seg om før det ytres en setning, og det behøves ikke (eller sjelden, som vi snart skal se) en bevisst analyse av sammensetningen av ordene i en setning man får presentert for seg. At parseren er deterministisk⁵⁵ (ii) betyr at den i seg selv ikke kan forkaste en struktur som blir prosessert (Marcus, 1975, s. 6). Dette er en konsekvens av at parseren virker ubevisst. Det er først når parseren viser seg å ha feiltolket strukturen i en setning (for eksempel når man støter på en garden path-konstruksjon (GP)) at språkbrukeren må ty til bevisst prosessering. At parseren er rask (iii) synes ved hvordan språkbrukeren kan bruke og forstå språklige ytringer som består av mange morfosyntaktiske elementer i løpet av kort tid. At parseren er begrenset (iv) viser seg i den begrensede muligheten språkbrukere har til å utnytte mulighetene som finnes i språket til å ytre 'vanskelige' setninger (med konstruksjoner som dyp egeninnføyning, pakking, osv.).

Om forholdet mellom bevisst og ubevisst prosessering av språk sier Milne (1982, s. 353):

*It is assumed that the syntactic and semantic processes of the HSPM rarely fail. When they do, conscious effort is used to recover from the error and the error component takes control.*⁵⁶

⁵⁵Milne (1982) presiserer at det ikke er hele språkprosesseringsapparatet som er deterministisk, men kun den syntaktiske parseren, som virker når en ytring kan prosesseres ubevisst.

⁵⁶HSPM = the Human Sentence Parsing Mechanism.

Det er altså først når den syntaktiske parseren ikke klarer å prosessere en ytring at språkbrukeren må analysere ytringen bevisst, med en 'feilkomponent'. GP-setninger betegnes som tilfeller som krever bevisst prosessering. Dette er fordi slike setninger skiller seg fra normalen: *As an armchair definition; any grammatical sentence that seems abnormal to read, requires conscious effort.* (Milne, 1982, s. 353). Videre antar Milne at parseren består av en syntaktisk prosessor, og en semantisk prosessor. Det er den semantiske, bevisste prosessoren som 'slår inn' når den syntaktiske, ubevisste prosessoren feiltolker en setning (f. eks. en GP-konstruksjon).

6.2 Parserens begrensninger

Chomsky (1965, s. 13) viser eksempler på to typer konstruksjoner som bidrar til inakseptabilitet: 'Nesting' (pakking) (48-a) og 'self-embedding' (egeninnføyning) (48-b).

- (48) a. I called the man who wrote the book that you told me about up
 b. The man who the boy who the students recognized pointed out is a friend of mine

Felles for disse konstruksjonene er at de krever mye av minnet. Pakking involverer elementer som hører sammen men befinner seg langt fra hverandre i setningen (*called* og *up*), mens egeninnføyning krever at parseren 'husker' flere predikat-argument-relasjoner før setningen lukkes (ved at *the man* tildeles en rolle som subjekt til *is a friend of mine*): *the man is a friend of mine; the boy pointed out the man og the students recognized the boy*. I dagligtale er setninger som (48) uvanlige, om brukt i det hele tatt. Det finnes alternative måter å uttrykke disse proposisjonene på ((49-a) ved heavy NP shift, og (49-b) ved passivisering):

- (49) a. I called up the man who wrote the book that you told me about
 b. The man who is a friend of mine was pointed out by the boy who was recognized by the students

Siden det finnes alternativer til pakking og egeninnføyning, og disse konstruksjonene for de fleste språkbrukere krever for mye av minnet til å kunne forstås, foretrekkes ofte alternativene.

Selv om setningene i (48) og (49) denoterer de samme proposisjonene, er måten parseren behandler de to på svært forskjellig, ifølge Milne (1982). Setningene i (49) prosesseres ubevisst, mens setningene i (48) krever bevisst innsats fra lytteren.

6.3 Parsing av garden path-konstruksjoner

Det som skjer når prosessoren feiltolker en GP-konstruksjon, er at parseren 'antar' en tolkning av en konstruksjon som den riktige. I (50-a) feiltolker parseren *man* som en del av subjektsNPen, med det resultat at frasestrukturregelen $S \rightarrow NP VP$ ikke følges:

- (50) a. *[S [NP The old man] [NP the boat]]
b. [S [NP The old] [VP man the boat]]

Tolkningen av *the old man* som en NP foregår automatisk i den syntaktiske prosessoren, og lukkes automatisk (jf. Kimballs femte prinsipp⁵⁷, om at en frase lukkes så snart det er mulig (med mindre den påfølgende noden er en umiddelbar konstituent i den gjeldende frasen)). Ifølge Kimballs sjette og syvende prinsipp vil en lukket frase skyves ned i et syntaktisk (og kanskje semantisk) prosesseringsnivå, og være tung for parseren å hente tilbake og flytte til eventuelle andre funksjoner. Det er altså denne prosessen parseren ikke klarer på egenhånd når lesningen (50-a) skal omtolkes til (50-b) ved at substantivet *man* (som allerede er 'pakket ned' til neste nivå som en NP sammen med *the old*) flyttes til VPen verben *man* utgjør kjernen av i (50-b).

Kimballs (1973) femte prinsipp brytes i den første foretrukne lesningen (50-a): Lukking av subjektsNPen så snart som mulig ville bety lukking etter *the old*, siden *man* ikke er en del av subjektsNPen i den riktige tolkningen (50-b). Årsaken til at den ikke lukkes 'i tide' er at parseren (som virker lineært) ikke vet hva som følger etter subjektet, og dermed ikke kan forutsi at *man* (som forekommer langt oftere som et substantiv enn som et verb) skal leses som et verb.

At garden path-konstruksjoner er mer krevende å prosessere enn andre konstruksjoner har blitt demonstrert ved reaksjonstidsforsøk foretatt av Milne (1982), som viste at setninger med GP-konstruksjoner tok betydelig lengre tid å lese enn normale setninger. Tiden det tar å lese setninger med dype setningsinnyngninger har Just et al. (1996) vist er lengre enn tiden det tar å lese normale setninger (som nevnt i 2.2). En viktig forskjell mellom GP-konstruksjoner og dype setningsinnyngninger er at førstnevnte lar seg reanalyse og forstå forholdsvis enkelt, mens en setningsinnyngning (avhengig av dybden) kan være umulig å forstå uten papir-og-penn-analyse.

⁵⁷Kimball presenterer i *Seven principles of surface structure parsing in natural language* (1973) syv hypoteser om hvordan parseren virker. Disse vil jeg referere til som 'Kimballs prinsipper'.

7 Preposisjoner som forsvinner

I dette kapittelet vil jeg vise eksempler på to typer konstruksjoner som mangler en preposisjon for å kunne regnes som grammatisk. Slike mangler forekommer ved bruk av ulike preposisjoner, men ofte forekommer en preposisjon lik den som mangler i slike konstruksjoner. Det vil også gis en mulig forklaring på hvorfor preposisjoner i slike tilfeller 'forsvinner'.

7.1 Tilgrensende preposisjoner

Følgende har blitt ytret i NRK radio (mine kursiveringer):

- (51)
- *Vi skal snart snakke *om* det er sjefen eller arbeidstakeren som er verst når det gjelder mobbing på jobb
 - *Det ble ikke noe av samarbeidet, men en indisk konsulent fikk en million dollar, uten at Yara kan redegjøre *for* hva.
 - *Det er jo en helhetsvurdering, og det kan man jo veldig vanskelig svare *på* direkten.

Felles for disse ytringene er at de inneholder en preposisjon (kursivert) med ubestemmelig funksjon, og at de mangler en preposisjon for å oppfylle prinsippet om koherens. Selv om det er tydelig hva som menes i alle ytringene, kan det ikke bestemmes hvilken funksjon de uthevede preposisjonene har i ytringen.

- I (51-a) kan *om* tolkes som (1) en preposisjon (som kjernen i adjunkten til *snakke*: *Vi skal snart snakke om X*, eller (2) som en subjunksjon med den påfølgende leddsetningen som utfylling: *om det er sjefen eller arbeidstakeren (...)*. En tolkning av *om* som både en preposisjon og en subjunksjon lar seg ikke gjøre.
- I (51-b) kan *for* enten tolkes som (1) kjernen i det oblike leddet til 'redegjøre': *Yara kan redegjøre for X*, eller (2) som kjernen i adjunkten til *en indisk konsulent fikk en million dollar* (med *hva* som objekt).
- I (51-c) kan *på* enten tolkes som (1) den strandede kjernen i adjunkten til *man kan jo veldig vanskelig svare* (med *det* som et topikalisert objekt), eller (2) kjernen i adjunkten til *det kan man jo veldig vanskelig svare på* (med *direkten* som objekt).

De grammatiske (om enn uakseptable) ekvivalentene til (51) er (52).⁵⁸

⁵⁸I NRG betegnes setninger med sekvensen *om om* (som (52-a)) som ugrammatiske, uten noen begrunnelse utover at de inneholder 'to *om* etter hverandre'. Som et alternativ foreslås det å bruke *hvorvidt*

- (52) a. Vi skal snart snakke om om det er sjefen eller arbeidstakeren som er verst når det gjelder mobbing på jobb.
- b. Det ble ikke noe av samarbeidet, men en indisk konsulent fikk en million dollar, uten at Yara kan redegjøre for for hva.
- c. Det er jo en helhetsvurdering, og det kan man jo veldig vanskelig svare på på direkten.

Årsaken til at alternativene i (51) ble ytret heller enn de i (52) kan synes åpenbar for de fleste som har norsk som morsmål: Det kjennes unaturlig å bruke sekvenser som *om om*, *på på* og *for for*. Dette er sekvenser vi sjelden eller aldri bruker i dagligtale. For parseren sin del kan det tenkes at den tolker preposisjonen som en som 'lukker'⁵⁹ eventuelle åpne fraser med langdistanseavhengigheter (som i (51-b) (*en million dollar – for hva*) og (51-c) (*svare på – det*)), og at det er vanskelig å forholde seg til langdistanserelasjonen samtidig som man skal håndtere en lokal preposisjonsrelasjon.

7.2 Strandede preposisjoner

På spørsmål om hva han mente om Statoils utvinning av olje fra oljesand i Canada gav olje- og energiministeren følgende to kommentarer i NRK radio, med få dagers mellomrom ((53-a) ble ytret muntlig av statsråden i radiostudio, og (53-b) er et sitat fra en epost fra ham, lest opp av programlederen):

- (53) a. *Min forventning som stor eier i Statoil er selvsagt at de følger lover og regler i de landene de opptrer.
- b. Generelt forventer jeg at Statoil følger de lover og regler som gjelder i landene selskapet opererer i.

Ytringene har flere likhetstrekk både i form og i innhold, men (53-a) mangler, i motsetning til (53-b), en strandet *i*. Disse eksemplene illustrerer hvor vanskelig det kan være å håndtere langdistanseavhengigheter når man må parse en setning samtidig som man ytrer den.

Følgende har blitt ytret i NRK radio:

- (54) a. *De bor ennå i den kommunen de er folkeregistrert __.
- b. *(...) om ikke veldig mange år så er jo og Europa i den situasjonen Norge er __.

istedenfor subjunksjonen *om* (*om hvorvidt*).

⁵⁹Jf. Kimballs (1973) femte prinsipp om at parseren lukker frasen så snart som mulig (50).

- c. *Det hun gjorde på de konferansene jeg var —, det synes jeg hun gjorde på en god måte.
- d. *(...) som gjør oss i stand til å være svært direkte på mål i de operasjonene vi deltar —.

Felles for disse ytringene er at de mangler en strandet preposisjon, at plassen til den strandede preposisjonen befinner seg i XCOMP, og at objektet til denne preposisjonen befinner seg i den overordnede setningen. (54-c) er den eneste av ytringene som mangler en strandet *på*. Resten mangler en strandet *i*. (54-d) skiller seg fra de øvrige ytringene ved at den manglende preposisjonen innleder en adjunkt til et intransitivt verb (*delta*), snarere enn en utfylling til et transitivt verb (*være*).

Det kan synes som den strandede preposisjonen i disse tilfellene 'glemmes' fordi parseren anser den åpne predikat-argument-relasjonen (for preposisjonen) for å være lukket ved den første forekomsten av en slik preposisjon (jf. Kimballs (1973) femte prinsipp). Spesielt med dette er det at preposisjonen både har et lokalt objekt og befinner seg på feil sted. Mens det i (51) og (52) kun er preposisjonene selv som er de samme, har i tillegg objektet til preposisjonen som forekommer og den manglende, strandede preposisjonen, samme referent i setningene i (54). En strandet preposisjon har et gap-objekt, med objektet til preposisjonen som forekommer som filler.

Det finnes også eksempler på ytringer der en strandet preposisjon 'glemmes' selv om NPen den befinner seg i styres av det samme verbet som den strandede preposisjonens objekt. I slike setninger deler ikke den forekommende og den utelatte preposisjonen objektsreferent:

(55) *Hvilket lys setter dette da til syvende og sist krigen i Afghanistan —?

I (55) grenser den manglende strandede preposisjonen (*i*) til en PP med en kjerne med samme form. Den strandede preposisjonen har også her (som i (54)) en gap som objekt, men her er den kjerne i et oblikt ledd, noe som ikke er tilfellet i noen av eksemplene i (54). En mulig årsak til at den strandede preposisjonen utelates her kan være en kombinasjon av det som antas å være årsaken til utelatelsen i (51) (lineær nærhet til en lik preposisjon) og (54) (høy belastning for parseren på grunn av langdistanseavhengigheter). Avstanden mellom preposisjonen og plassen til den strandede preposisjonen er større i (55) enn i (51), og det er færre innføyninger mellom plassene i (55) enn i (54), men kombinasjonen av disse svakere effektene kan føre til at den strandede preposisjonen glemmes og fører til ugrammatiske ytringer som (55).

En tilsvarende elisjon kan oppstå når to ulike strandede preposisjoner forekommer samtidig. Eksempel (56-a) er fra NRK radio. (Her er (56-b) den grammatiske varianten av (56-a)).

- (56) a. *Var det denne måten du ble snakket til — som barn?
b. Var det denne måten du ble snakket til på som barn?

I resultatene fra søket i NA (i del 2) vil jeg blant annet gjøre rede for andelen ugrammatiske forekomster blant setninger der *på* og andre preposisjoner forekommer i før måteP i måteNP⁶⁰. Siden det kan være en mulighet for at det er vanskelige filler-gap-relasjoner (som det er mellom *måte* og måteP) som er årsaken til at måteP utelates, vil jeg også gjøre rede for betydningen forekomsten av flere enn ett verb i måteNP har.

⁶⁰En definisjon av måteNP er gitt i innledningen.

DEL 2

8 Grammatiske feil i måteNP

I dette kapitlet vil jeg gjøre rede for noe av det teoretiske grunnlaget jeg vil behandle resultatene fra de empiriske undersøkelsene i denne oppgaven med bakgrunn i, og vise hvorfor konstruksjoner tilsvarende måteNP ikke finnes i engelsk. Eksemplene som brukes i dette kapitlet er hentet fra Norsk Aviskorpus der ikke annet er oppgitt.

8.1 Rollen til *måte* i måteNP

Forskjellen mellom en måteNP der måteP inkluderes og en der måteP ikke inkluderes, er rollen *måte* får. Når måteP inkluderes, er *måte* kjernen i måteNP, som modifierer objektet i matrisetningen måteNP inngår i (som objekt til måteP). Når måteP utelates, deltar *måte* i strukturen som om det skulle være objektet til hovedverbet i måteNP. En slik struktur (med en infinitivkonstruksjon uten subjekt) antyder at det er en kontrollrelasjon mellom kjernen i objektet til hovedsetningen (*måte*) og verbet i infinitivkonstruksjonen: *måte* blir objektet til verbet i infinitivkonstruksjonen. Følgende eksempler er mine egne:

- (57) a. Per_i har en ny bok_j å —_i lese —_j
b. *Per_i har en ny måte_j å —_i løpe —_j

Den syntaktiske strukturen i (57-a) krever at *gå* tar *en ny måte* som objekt på samme måte som *lese* tar *en ny bok* som objekt i (57-a), men *gå* kan ikke ta objekt. Når måteP inkluderes, er det måteP som 'overtar' objektet *løpe* etter de syntaktiske reglene måteNP ble påpraktet i (57-b):

- (58) Per_i har en ny måte_j å —_i løpe på —_j

MåteNP er altså helt avhengig av at måteP inkluderes for å være grammatisk.

8.2 Mulige årsaker til utelatt måteP

Umiddelbart kan årsaken til at måteP utelates i konstruksjoner måteNP inngår i se ut til å være den store avstanden mellom *måte* og plassen til måteP:⁶¹

⁶¹Heretter vil jeg kalle sekvensen mellom *måte* å og (plassen til) måteP for måteF ('måtefrase').

- (59) a. Å herme etter andre er en fin måte å [starte] på
 b. *Jeg leter etter en måte å [gi kompensasjon til de som har lidd under denne alvorlige og brutale mishandlingen utført av noen få i USAs bevæpnede styrker] —.

Mens det i (59-a) kun er ett ord i måteF, er det tilsvarende antallet ord i (59-b) 22. Intuitivt er det ikke vanskelig å forstå at måteP utelates i (59-b), og den lange avstanden kan være en av årsakene til utelatelsen av måteP. Likevel forekommer det eksempler på måteNP med lengre avstand mellom *måte å* og plassen til måteP der måteP forekommer (60-a), og kortere eksempler der måteP utelates (60-b), (60-c).

- (60) a. Dette er en fin måte å [skape litt positiv PR rundt idretten] på.
 b. *(...) drømmene var Guds måte å [forberede meg på det] —.
 c. *(...) finne seg en alternativ måte å [komme på jobb fredag] —.

Det ser av disse eksemplene (og lignende eksempler) ut som det er mer enn bare antallet ord i måteF som spiller en rolle for hvorvidt måteP utelates fra måteNP. Disse eksemplene er ikke tilfeldig valgt: Det viser seg at forekomsten av *på* i måteF (som i (60-b) og (60-c)) spiller en avgjørende rolle for om måteNP i bruk inneholder måteP eller ikke.

8.3 Tidlig lukking av måteNP

I Kimballs (1973) femte prinsipp heter det at en frase skal lukkes dersom den neste noden etter en mulig lukking *er* en umiddelbar konstituent av den gjeldende frasen. Det synes som om parseren ofte virker etter dette prinsippet når *på* forekommer i måteF, og lukker måteNP ved første og beste anledning. Siden en *på* i en plassering som tillater lukking av måteNP er en mulighet til å parse ferdig frasen, lukkes ofte frasen når en slik *på* forekommer.⁶²

- (61) a. *(...) er en veldig fin måte å komme seg ned på bakken igjen, sier kongen.
 b. Dette er en måte å legge press på Telenor på.

Ettersom sekvensen *en veldig fin måte å komme seg ned på* (i (61-a)) er en komplett NP (med *måte* som kjerne), tolker parseren *på* som måteP. I (61-b) er den tilsvarende sekvensen *en måte å legge press på*. I denne sekvensen mangler objektet til det oblike leddet (*på ...*) som kreves av *legge*, og en tolkning av den første forekomsten av *på* som måteP er ikke

⁶²Dette skal jeg gå grundigere inn på i kapittel 10.

mulig. Hvis vi tar bort elementene som forekommer mellom den første forekomsten av *på*, og (plassen til) måteP, endres grammatikaliteten i begge ytringene i (61):

- (62) a. (...) er en veldig fin måte å komme seg ned på, sier kongen.
b. *Dette er en måte å legge press på.

Årsaken til at (61-a) mangler måteP er som nevnt en feiltolkning fra parseren sin side. Kimballs syvende prinsipp sier at en lukket frase skyves ut fra korttidsminnet, og inn i et syntaktisk/semantisk prosesseringstrinn. MåteNP skyves altså videre ut av korttidsminnet så snart måteP forekommer, eller snarere, når det parseren *tror* er måteP, forekommer. Ifølge Kimballs sjette prinsipp er det svært kostbart å hente tilbake konstituenten fra lukkede fraser, i den forstand at det krever mye av talerens parser. Når *en veldig fin måte å komme seg ned på* tolkes som en fullstendig måteNP, og måteNPen lukkes, er det krevende for parseren å åpne den 'falske' måteNPen for å inkludere resten av den 'egentlige' måteNPen (*bakken igjen*). Dette får konsekvenser for grammatikaliteten, siden sekvensen *bakken igjen* ikke har tilhørighet til noe predikat i ytringen.⁶³

Bruken av resumptive pronomeren har flere likhetstrekk med utelattelsen av måteP, selv om det tilsynelatende er det motsatte som gjøres i de to tilfellene (et RP er 'ulovlig' tillagt, mens måteP er 'ulovlig' utelatt): Motivasjonen for å bruke dem er å øke akseptabiliteten (RP og måteP) og øke forståeligheten (RP). Konsekvensen av begge er brudd på prinsippet om koherens: Et resumptivt pronomeren er et argument som mangler et predikat som kan gi det en syntaktisk rolle, på samme måte som objektet til måteP mangler et slikt predikat når måteP er utelatt.

8.4 Fraværet av måteNP i engelsk

I norsk og engelsk finnes det (minst) to måter å uttrykke fremtid på: Ved et modalt hjelpeverb, og ved en infinitivskonstruksjon. Fremtid ved modalt hjelpeverb i de to språkene har åpenbare likhetstrekk. (Eksemplene som involverer Peter er mine egne.)

- (63) a. Peter vil spise fisk
b. Peter will eat fish

Forskjellen mellom disse er flertydigheten til *vil*, som skyldes muligheten for flere enn én lesning: Epistemisk lesning gir (63-a) en proposisjon tilsvarende den for (63-b) (fremtids-

⁶³Setninger med måteNP som inneholder en *på* med en annen funksjon enn som måteP vil jeg komme tilbake til i 10.2.

aspekt), mens ved deontisk lesning av *vil* denoterer (63-a) at Peter har et ønske om å spise fisk. (63-b) har bare én mulig lesning: den epistemiske.

En tilsvarende proposisjon uttrykt uten *vil/will* er:

- (64) a. Peter kommer til å spise fisk
b. Peter is going to eat fish

Her benyttes den andre muligheten man har for å uttrykke fremtid i norsk og engelsk: ved et fraseverb i presens. Eksemplene i (64) viser at norsk og engelsk har motsatt utgangspunkt for å uttrykke fremtid ved infinitivskonstruksjoner. Norsk tar utgangspunkt i tidspunktet for den fremtidige handlingen: agensreferenten 'kommer til' handlingen 'å spise fisk', mens engelsk tar utgangspunkt i ytringstidspunktet: agensreferenten 'is going to' handlingen 'eat fish', som vil foregå i fremtiden. Forskjellen i uttrykkene i blir tydeligere om infinitivskonstruksjonene i (64) byttes ut med en NP:

- (65) a. Peter kommer til Roma
b. Peter is going to Rome

Verd å merke seg er det at *to* i uttrykk som (65-b) og (64-b) er en preposisjon, og ikke et infinitivmerke: Infinitivskonstruksjonen i (64-b) inkluderer ikke *to* som infinitivmerke (i motsetning til det norske motsvaret (64-a), som inneholder både preposisjon og infinitivmerke).

Den engelske ekvivalenten til *på en måte* er *in a way*.

- (66) a. Peter spiser fisk på en uvanlig måte.
b. Peter eats fish in an unusual way.

Forskjellen mellom de to uttrykkene er – bortsett fra de åpenbare: det brukes ulike ord (om enn av lik ordklasse): *på/in*, *måte/way* – flertydigheten det engelske substantivet har: Den vanligste måten dette ordet brukes på er i betydningen *vei*.⁶⁴ Denne flertydigheten (og årsaken til at det er *way* som brukes i slike uttrykk) er tydeligere i den engelske ekvivalenten til måteNP:

- (67) a. Peter fant en bedre måte å spise fisk på
b. Peter found a better way to eat fish

⁶⁴Norsk *måte* har riktignok også en annen betydning, som i uttrykket *å drikke med måte*, men denne betydningen er ikke mulig (og derfor ikke relevant) i denne sammenheng, siden det er et ikke tellbart substantiv, og derfor ikke kan ha ubestemt artikkel.

Den engelske måten å uttrykke dette på (67-b) har til felles med uttrykk med fremtidsuttrykk med *be going to* (64-b) at den fokuserer på 'veien' frem til handlingen hovedverbet denoterer (*eat fish*). Et slikt uttrykk kan også brukes om man bytter ut VPen i infinitivkonstruksjonen med en NP (analogt med (65)). På norsk går ikke dette like bra:

- (68) a. *Peter fant en bedre måte å Roma på
b. Peter found a better way to Rome

Strukturlikheten mellom (67-b) og (68-b) viser at engelske *in a way* som en ekvivalent til *på en måte* skriver seg fra betegnelsen for en fysisk vei (som i (68-b)), overført til et strekke av tid fra ytringstidspunktet til en situasjon oppstår (i eksemplene over: at fisken er spist). Dette er ikke tilfellet for det norske *måte*, som ikke fokuserer på handlingen som veien frem til situasjonen den forårsaker, men snarere handlingen i seg selv.

Årsaken til at setninger som (67-b) ikke har måteP (som (67-a)) eller en annen strandet preposisjon (*in*, som i (66-b)) er mest nærliggende, er at alle de nominale elementene allerede står i et predikat-argument-forhold. Dette er ikke tilfellet for *måte* i (67-a). Uten måteP vil *spise* formelt ha to direkte objekter: *måte* (gjennom en kontrollrelasjon, som kjernen i objektet til matriseverbet *ha*, og som objektet til *spise*) og *fisk*. Men med måteP blir *måte* objektet til *på*, og *spise* har kun ett objekt.

9 Metode

I dette kapitlet forklares fremgangsmåten under innsamlingen og vurderingen av materialet den empiriske undersøkelsen baserer seg på. Her diskuteres også kort hvilken forskjell det antas å utgjøre for resultatene (i forhold til talespråk) at materialet er hentet fra avistekster.

9.1 Materiale

Jeg har brukt Norsk Aviskorpus (NA) til å finne forekomster av måteNP. Målet med søkene er å finne ut hvilke konstruksjoner som gjør det vanskelig å bruke måteNP på riktig måte, dvs. å inkludere måteP i setningen.

Når det er aviser en empirisk undersøkelse av språket baserer seg på, er det rimelig å tro at funnene blir noe annerledes enn om man tok for seg tilfeldige setninger ytret av et tilfeldig utvalg av språkbrukerne. For det første er det journalisters språk man tar for seg. Journalister (som produserer mer tekst enn de fleste andre språkbrukere) er troende til å gjøre færre feil enn andre språkbrukere. For det andre er det tekst man henter materialet fra. Siden språkbrukeren har mulighet til å redigere teksten i ettertid (noe enten en journalist, korrekturleser eller deskmedarbeider som regel gjør før teksten publiseres) er det naturlig å anta at det forekommer færre grammatiske feil enn ved tale. Likevel forekommer altså disse feilene, og det kan gjøres rede for årsaken til dem, på samme måte som når grammatiske feil forekommer i tale.

9.2 Skriftlig språkproduksjon og -persepsjon

Det som skiller de empiriske studiene i denne oppgaven fra f.eks. de refererte studiene til Just et al. (1996) og Stromswold et al. (1996) (som vist i 2.2) er at empirien i denne oppgaven er språkproduksjon snarere enn -persepsjon. Likevel – som det har blitt antydnet gjennom hele oppgaven – det kan ikke sees bort fra at det kan være de samme språklige elementene som skaper problemer for parseren hos lytteren som hos taleren. At materialet for denne oppgavens empiriske studie er avistekster gjør at det ikke nødvendigvis *bare* er språkproduksjon som gir utslag i tekstene som foreligger. Som nevnt vil ofte (om ikke alltid) deskmedarbeidere lese igjennom tekstene før de kommer på trykk, noe som betyr at eventuelle grammatiske feil (som utelatt måteP) i de fleste tilfellene har blitt prosessert og godtatt av minst én annen språkbrukers parser enn parseren som produserte konstruksjonen. Jeg har ikke vist – og skal ikke vise – forskjellen i prosessering av skriftlig og muntlig

språk, men fordelene i skrift fremfor tale synes åpenbare når det gjelder f.eks. produksjon av setninger ved hjelp av et flyktig medium som tale mot et redigerbart medium som tekst. Dette gjør de mest oppsiktsvekkende resultatene (som den lave frekvensen av grammatiske måteNP der en *på* befinner seg i måteF (tabell 8 i kapittel 10)) om mulig enda mer oppsiktsvekkende, siden de grammatiske feilene har 'overlevd' prosessering av flere enn én parser.

De empiriske studiene som er foretatt her skiller seg også fra Miller & Isards (1963) studier av språkbrukeres evne til å gjengi egeninnføyde setninger (også gjengitt i 2.2), i at materialet jeg har tatt for meg ikke har blitt til i et labforsøk, men ved naturlig språkbruk. Riktignok kan det i slikt materiale finnes feilkilder det er mulig å luke ut i et laboratorium (som stilvariasjon mellom språkbrukerne), men til gjengjeld fås språkmateriale slik det brukes naturlig, og relevante setninger slik de forekommer blant tusener av andre, irrelevante setninger, uten at språkbrukeren som produserer dem i større grad enn ellers anstrenger seg for å unngå å gjøre feil.

9.3 Utvalg

Søkemotoren i NA gir mulighet for søk på strenger av ord. Jeg vil finne materialet for den empiriske undersøkelsen ved å gjøre søk på sekvensen *måte å*. Et slikt søk gir relevante treff for denne undersøkelsen, men det vil både utelate relevante treff (69-a), og inkludere treff med andre konstruksjoner enn måteNP (69-b).

- (69) a. Relevante, utelatte treff:
- (i) Andre bøyninger av *måte* (*måten, måter, måtene*) (*Per har lært den vanligste måten å fiske på*)
 - (ii) måteNP der *måte* står like foran en NP (*Her er én måte dyrene rømmer på*)
- b. Irrelevante, inkluderte treff: AdjunktPP med *måte* som objekt, der *på* forekommer før *måte* (*Og humor behøver jo på ingen måte å være delikat*)

De relevante, utelatte treffene i (69-a) er utelatt for enkelhetens skyld. De kunne ha vist variasjon ved ulike bøyninger, eller når f.eks. subjektet til kjernen i hovedverbet i måteNP står *i* måteNP (dvs. i COMP) (og ikke refereres til ved en kontrollrelasjon) (og verbet er finitt (som i (69-a-ii) (XCOMP))), men slik variasjon skal jeg ikke ta for meg her. Jeg vil finne hvilke konstruksjoner som typisk fører til utelatelse av måteP i en måteNP med formen *måte*. Jeg har valgt formen *måte* helt enkelt fordi måteNP med denne formen er

hyppigst forekommende i NA,⁶⁵ og jeg har valgt å søke på strengen *måte å* fordi forekomsten en bøyning av *måte* etterfulgt av N ved søk har vist seg å variere med bøyningen av *måte* i så stor grad at det kompliserer arbeidet med funnene unødig. Treffene ved søkene jeg vil gjøre vil i alle tilfeller vise det jeg vil finne: Hva slags konstruksjoner er det måteP oftest utelates fra?

Avistekstene jeg har funnet forekomstene av måteNP som brukes i denne undersøkelsen i er fra 2010. Materialet jeg satt igjen med etter søket som beskrevet over har jeg sjekket for duplikater maskinelt og fjernet disse.⁶⁶ Deretter har jeg ved manuell gjennomgang funnet antall ord i måteF og eventuelle forekomster av verb og preposisjoner, samt hvilke av dem måteP forekommer i. Jeg har tatt for meg 1000 forekomster av måteNP.

⁶⁵Formen *måte-Ø/-n/-r/-ne* utgjorde henholdsvis 70,6; 27,6; 0,9 og 0,9 prosent av tilfellene der en form av lemmaet 'måte' forekom i Norsk Aviskorpus (2010), umiddelbart etterfulgt av *å*.

⁶⁶Enkelte setninger viste seg etter den maskinelle gjennomgangen fremdeles å forekomme flere enn én gang, da de ble gjengitt i flere aviser. Disse har jeg fjernet når jeg har oppdaget dem i løpet av den manuelle gjennomgangen av funnene.

10 Resultater

I dette kapitlet presenteres resultatene fra søkene i NA. Tallene vil kommenteres underveis, og det vil gis eksempler på setninger slik de forekom i korpuset. Jeg vil også vise hvor avgjørende det er for hvorvidt måteP forekommer i måteNP om en ekstra forekomst av *på* befinner seg på et sted i ytringen som gir parseren syntaktiske grunner til å tro at denne ekstra forekomsten av *på*. Til slutt vil jeg sammenligne betydningen bruken av resumptive pronomen og utelatelsen av måteP har for grammatikaliteten i setningene de forekommer i eller kan forekomme i.

10.1 Tallene

I tabellene angir kolonnen 'Lengde' lengden på måteF (i radene viser 1F tallene for måteF bestående av ett ord (*(...) en fin måte å [starte] på*), 2F er måteF bestående av to ord (*(...) en annen måte å [låne penger] på*) osv.) Jeg har bare tatt for meg måteNP med inntil ni ord i F, siden frekvensen av lengre måteFer enn dette er svært lav. (Ved de innledende søkene fant jeg måteNP med opptil 22F (som er gjengitt i eksempel (59-b)).

10.1.1 Lengden på måteF

Antallet ord i måteF blant de 1000 setningene søkene i NA gav, fordelte seg som vist i tabell 2. Antallet ord i xF er omvendt proporsjonal med antallet forekomster av xF, med ett unntak: 1F (151) er betydelig lavere enn 2F (413), noe som trolig skyldes at det finnes flere transitive enn intransitive verb i norsk. Når hovedverbet er intransitivt er det laveste mulige antallet ord i måteF 1 (hovedverbet). Når måteNP inneholder et transitivt verb vil det laveste mulige antallet ord i måteF være 2 (hovedverbet+objekt).

TABELL 2:

Lengde	Antall
1F	151
2F	413
3F	152
4F	100
5F	76
6F	47
7F	33
8F	18
9F	10
Tot	1000

10.1.2 Andel grammatiske måteNP

I tabell 3 vises antallet måteNP som inkluderer måteP, og dermed er grammatiske (G), samt andelen disse utgjør av det totale antallet måteNP for hver xF (%G). Det er ingen overraskelse at lengden på måteF er tilnærmet omvendt proporsjonal med andelen grammatiske måteF for forekomstene av xF. Av tabellen kan det se ut som det er antallet ord i måteF som avgjør om en ytret måteNP er grammatisk eller ugrammatisk.

Ved første øyekast kan det synes som det er dette Erteschik-Shir (1992) viser med eksemplene som er gjengitt i (16) i 3.2. Men lengden på eksemplene til Erteschik-Shir er proporsjonal med hvor komplekse de er, noe som også trolig er tilfellet for for måteNP, siden lengre setninger er troende til å inneholde flere predikater.

Likevel viser tallene i tabell 3 at grammatikaliteten til måteNP følger det samme mønsteret som Erteschik-Shirs grammatikalitetsvurderinger: Lengre setninger får det til å virke mindre naturlig å bruke måteP i måteNP (på samme måte som lange avstander mellom filler og gap i konstruksjoner med langdistanseavhengigheter gjør bruken av RP mer akseptabel (som i (16-d) – gjengitt her som (70-b)), mens kortere måteNPer svært sjelden forekommer uten måteP (jf. inakseptabiliteten til korte setninger som inkluderer RP, som i (16-a) – gjengitt her som (70-a)).

- (70) a. This is the girl that John likes $_/*$ her.
 b. This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to $?_/$ her.

Følgende eksempler viser én av to forekomster fra søkene i NA av måteNP *uten* måteP der det befant seg ett ord i måteF (71-a), og den eneste forekomsten av måteNP *med* måteP der det befant seg ni ord i måteF (71-b). I disse eksemplene er det alternativet (av måteP og '–') som er minst sannsynlig å forekomme⁶⁷ markert med '??'. Det alternativet setningen faktisk forekom med i NA er uthevet:

⁶⁷Med 'minst sannsynlig å forekomme' mener jeg det alternativet (på/–, henholdsvis +/- grammatisk) med lavest frekvens i det gjeldende antallet måteF (basert på tallene i tabell 3)).

TABELL 3:

Lengde	G	%G
1F	149	98,7
2F	388	94,2
3F	132	86,8
4F	77	77,0
5F	43	56,6
6F	17	36,2
7F	12	36,4
8F	4	22,2
9F	1	10,0
Tot	823	82,3

- (71) a. 'Å skrive er en måte å leve ?**—**/på', sa Flaubert.
 b. (...) en morsom måte å åpne øynene til folk for at de skal stemme **—**?**på**, (...)

Selv om (71-b) etter regelen skal inneholde måteP, kjennes det unektelig langt mindre naturlig å plassere måteP i slutten av måteNP i (71-b) enn å utelate det. I (71-a) er det omvendt: Det kjennes unaturlig å utelate måteP, som også forårsaker at setningen blir ugrammatisk. Dette stemmer overens med det som ble funnet i NA, der det kun var to av de 149 forekommende måteNPene (1,3%) med ett ord i måteF som manglet måteP (se tabell 3).

Hvis vi sammenligner eksemplene i (70) og (71) ser vi at forholdet mellom lengde på setningen og akseptabiliteten til forekomsten til eller fraværet av det siste elementet (RP og måteP) er omvendt for de to konstruksjonene. I (70) er akseptabiliteten for bruk av RP størst i lange sekvenser, mens i (71) er akseptabiliteten for bruk av måteP størst i korte sekvenser. Denne motsetningen i forholdet mellom lengde og akseptabilitet for inkludering av et element sammenfaller med motsetningen i årsaken til brudd på prinsippet om koherens for de to (som vist i 5.1): RP utgjør et argument som ikke tildeles en syntaktisk rolle av et predikat, og utelatelse av måteP forårsaker at et argument (NPen *måte* er kjerne i) mangler et predikat å tildeles en syntaktisk rolle til.

Samsvaret mellom avvik fra normalen (enten ved at måteP utelates fra måteNP eller ved at RP tillegges filler-gap-konstruksjoner) og lengden/kompleksiteten til konstruksjonen de forekommer i, gjør det rimelig å tro at årsaken til de to avvikene er den samme: Setningen er for komplisert til at parseren kan ta i bruk språkkunnskapen kompetansen stiller til rådighet (se figur 1 i 1.1). Hvis dette stemmer, ser det ut som det ikke bare er for å lette kommunikasjonen parseren bryter med reglene, men også simpelthen fordi de 'åpne' konstruksjonene parseren er i ferd med å behandle (som f.eks. måteNP før måteP forekommer) 'glemmes' av parseren på grunn av overbelastning som følge av lengden på konstruksjonen som parses (som i seg selv – som empirien jeg har vist så langt tyder på – kan få parseren til å glemme hvilke elementer konstruksjonen må inneholde for å være grammatisk). Slik kan utelatelsen av måteP jammføres med den lite utstrakte bruken av for eksempel dype, egeninnføyde setninger (som vist i 2.2): Parseren har ikke kapasitet til å håndtere så tunge konstruksjoner.

10.1.3 Andel måteNP med verb

Arbeidene med langdistanseavhengigheter viser at enkle strukturer – uten 'avbrudd' fra f.eks. predikatrelasjoner – sjelden er nok til å forstyrre parseren. Det er også grunn til å tro at det forekommer flere verb i lange strenger enn i korte. Tabell 4 viser hvor mange av hver xF som inneholder andre verb (FV) enn hovedverbet, og andelen måteNPer med FV utgjør av det totale antallet måteNP (%FV) (1F og 2F kan ikke inneholde flere verb enn hovedverbet, siden det ikke gir plass til argumenter i tillegg til hovedverbet og det ekstra verbet). Jo lenger måteNP er, desto flere verb inngår i den. Av tabellen synes det som det like godt som antallet ord i måteF kan være forekomsten av verb som fører til at måteP utelates.

TABELL 4:

Lengde	FV	%FV
1F	–	–
2F	–	–
3F	3	2,0
4F	5	4,9
5F	8	10,5
6F	13	27,7
7F	10	29,5
8F	10	55,6
9F	7	70,0
Tot	56	5,6

Eksemplene i (72) viser forekomster av måteNP i NA som inneholder verb (her brukes det kun eksempler på funn der måteP mangler):

- (72) a. *(...) en ny måte å gjøre noe vi alltid har gjort (...) (*som-konstruksjon*)
b. *(...) min måte å vise at vi ikke tolererer det (...) (*at-konstruksjon*)
c. *(...) en måte å fortelle dere hvordan jeg har det. (*hv-konstruksjon*)
d. *(...) en måte å prøve å bekjempe seksuell trakassering (...) (*infinitivskonstruksjon*)

I tillegg inneholdt noen få av måteNPene blant funnene i NA flere enn ett verb. Eksempel (73) viser en av disse (en som-konstruksjon pakket i en at-konstruksjon). Intuitivt er det forståelig at parseren utelater måteP i disse tilfellene.

- (73) (...) en måte å sikre at alle søk nettleseren vår sender er kryptert.

Siden de typene verbkonstruksjoner som er vist i (72) hver for seg forekommer svært sjelden blant de 1000 funnene av måteNP i NA, har jeg ikke gjort noen sammenligning av hvor ofte hver av dem forekommer med og uten måteP. Jeg behandler dem samlet i tabell 5.

10.1.4 Andel grammatiske måteNP med og uten verb

Andelen måteNPer *uten* andre verb enn hovedverbet som er grammatiske (u/V), og andelen grammatiske måteNPer *med* slike (m/V) vises i tabell 5.⁶⁸ Av tabellen kan det leses at differansen (Diff) mellom andelen måteNP som er grammatisk når et verb forekommer i måteNP og når det ikke gjør det, er størst når det er få ord i måteNP. Forekomsten av verb er altså en av årsakene til at måteP utelates, og tilbøyeligheten verb har til å forårsake at måteP utelates er størst når måteNP er kort.

TABELL 5:

Lengde	u/V	m/V	Diff
3F	87,2	66,7	20,5
4F-6F	61,0	43,3	17,7
7F-9F	35,8	22,2	13,6
Tot	61,4	43,7	17,6

Den lave frekvensen til måteNP med et høyt antall ord i måteF, og den lave frekvensen til måteNP med de ulike typene verb som (72) gir eksempler på, gjør det vanskelig å trekke noen konklusjoner om hvilken innvirkning de ulike verbene har på hvorvidt måteP utelates eller ikke. Om ikke annet kan andelen grammatiske måteNP med flere verb enn hovedverbet sammenlignes med tilsvarende tall for måteNP som inneholder preposisjoner i tillegg til en eventuelt måteP (se tabell 6 på neste side). Generelt synes det som differansen i grammatikalitet for måteNP med og uten det gjeldende ekstra elementet (verb/preposisjon) er omtrent dobbelt så stor for verb som for preposisjoner, i det minste av det som kan leses av de korteste (og mest frekvente) måteNPene (3F: 20,5/10,3; 4F-6F: 17,7/8,2). Dette betyr at verb som forekommer i måteNP medfører større sannsynlighet for at måteP utelates enn det preposisjoner gjør. En mulig forklaring på dette er at verb kan forekomme som transitive, og kreve to argumenter (subjekt og objekt), mens preposisjoner ikke kan kreve mer enn ett argument (objekt). På grunn av dette er behandlingen av verb i slike konstruksjoner generelt trolig tyngre for parseren enn behandlingen av preposisjoner.

⁶⁸Her er gjennomsnittet av tre og tre xF trukket sammen på grunn av den lave frekvensen de lengste setningene forekommer i. (måteNPer som verb ikke kan forekomme i (1F og 2F) er også utelatt her).

10.1.5 Grammatiske måteNP med og uten preposisjoner

Tabell 6 viser andelen måteNP som er grammatisk når måteF inneholder en preposisjon (m/P) og når den ikke gjør det (u/P). At differansen mellom måteF med og uten preposisjoner utjevnes og blir negativ⁶⁹ (F7-F9) for lengre setninger kan tyde på at tilstedeværelsen til en preposisjon spiller mindre rolle jo lengre setningen er.

TABELL 6:

Lengde	u/P	m/P	Diff
3F	88,1	77,8	10,3
4F-6F	60,8	52,6	8,2
7F-9F	21,0	28,6	-7,6
Tot	56,6	53,0	3,6

10.1.6 Konkrete preposisjoner

Toppen av Korrekturavdelingens frekvensordliste (fra 4,7 millioner ord fra norske aviser og romaner) er gjengitt i tabell 7. De mest frekvente preposisjonene er *i*, *på* og *til*. Av de 1000 måteNPene som ble funnet i NA forekommer *i* i 46, *til* i 36 og *på* i 32. Til sammen utgjør disse henholdsvis 24,6, 19,3 og 17,1 prosent av de totalt 187 forekomstene av måteP som inneholder preposisjoner. Den lave forekomsten av preposisjoner for hver Fx gjør det hensiktsmessig å sammenligne alle forekomstene av dem under ett.

TABELL 7:

1	og
2	i
3	det
4	på
5	som
6	er
7	en
8	til

⁶⁹At måteNP tilsynelatende oftere er grammatisk ved forekomsten av en preposisjon (jf. 7F-9F m/P sammenlignet med u/P i tabell 6) skyldes trolig den lave forekomsten av lange sekvenser uten preposisjoner i materialet. Selv blant 1000 forekomster av måteNP vil det være et fåtall som er så lange som 7F-94 (totalt 61). Av disse vil det igjen være svært få som ikke inneholder preposisjoner (som jo er hyppig forekommende i språket). Det er altså sjelden det formuleres så lange måteNPer (kanskje nettopp for å unngå problemer med parseren), og resultatet blir et lavt antall måteNP som inkluderer måteP i korpus (NA). Blant annet inneholder ingen av de 9 forekomstene av måteF som består av 9 ord, måteP. Til tross for at frekvensen for så lange måteNP er så lav, har jeg inkludert disse tallene for å vise forskjellen mellom grammatikaliteten til måteNP med og uten preposisjoner i forhold til grammatikaliteten til måteNP med og uten verb.

De eneste tilfellene av måteNP fra funnene i NA som inneholder preposisjoner er de med en lengde på 3F-9F, siden preposisjoner ikke kan forekomme når måteF er kortere enn F3. Derfor viser raden 'Alle' i tabell 8 antallet måteNP med 3F-9F. Eksempelene i (74) viser parvis (a-b, c-d, e-f) eksempler på grammatiske (med måteP) og ugrammatiske (uten måteP) måteNP der henholdsvis *i*, *til* og *på* forekommer i måteF.

- (74) a. (...) en ganske elegant måte å komme videre i saken på.
 b. *En måte å ta gevinst i skjelvende markeder (...)
 c. Berlusconis måte å ønske sin kinesiske kollega velkommen til byen på.
 d. *(...) en superbillig måte å skaffe seg adgang til skjærgården.
 e. En tredje måte å tjene penger på uroen på (...)
 f. *En annen måte å tjene penger på svingningene (...)

I tabell 8 viser kolonnen 'Antall' det totale antallet forekomster av de angitte elementene i måteNP med F3-F9. Kolonnen 'Andel' viser prosentandelen hver av dem utgjør av det totale antallet setninger som kan inneholde preposisjoner, og kolonnen '%G' viser antall grammatiske måteNP med de gjeldende elementene. Raden 'Alle' viser antallet måteNP totalt for F3-F9, og raden 'P' viser antallet måteNP som inneholder en preposisjon.

TABELL 8:

	Antall	Andel	%G
Alle	436	100,0	65,6
P	187	42,8	52,4
'i'	46	10,6	63,0
'til'	36	8,3	52,8
'på'	32	7,3	21,9

Mens *i* i mindre grad enn preposisjone samlet bidrar til en grammatisk måteNP (63% grammatisk, mot preposisjonene samlet: 52,4% grammatisk), bidrar *til* (52,8% grammatisk) og *på* (21,9% grammatisk) i større grad til at måteNP mangler måteP. Særlig i måteNP som inneholder *på* glemmes måteP. Dette kan skyldes at parseren tolker en *på* som forekommer før plassen til måteP som måteP, og regner måteNP for å være 'lukket' ved den første forekomsten av *på*. Slik 'tidlig lukking' minner om setninger der en strandet preposisjon utelates ved forekomsten av en lik preposisjon med en annen funksjon tidligere i setningen, og der en lik, tilgrensende preposisjon utelates (som vist i 7.1).

10.2 En ekstra *på* i måteNP

Blant de 1000 funnene av måteNP i NA inneholdt 32 en *på* i måteF.⁷⁰ Av disse inneholder kun 21,9% måteP (tabell 8), mot 82,3% blant alle funnene (tabell 3). Setningene som inneholder annP kan deles inn i to kategorier.

- (75) a. *(...) en humoristisk måte å sette søkelyset på et problem.
b. *(...) en måte å tjene penger på Jacksons minne.

MåteF i (75-a) og (75-b) har to ting til felles: De mangler måteP, og de inneholder en annP. Dette endrer seg om man i begge eksemplene fjerner sekvensen etter annP:

- (76) a. *(...) en humoristisk måte å sette søkelyset på.
b. (...) en måte å tjene penger på.

Hovedverbet *tjene* i (76-b) får oppfylt sine krav om et objekt (*penger*), siden *Jacksons minne* i (75-b) kun er en del av en adjunkt. Hovedverbet i (76-a), derimot (*sette*) mangler et oblikt ledd for å få oppfylt kravet om kompletthet (som ble oppfylt i (75-a), ved *på et problem*). At (76-b) er grammatisk gir en mulig forklaring på hvorfor måteNP som inneholder annP ofte mangler måteP. Når parseren produserer setninger som (75-b) – og utelater måteP – er det i tråd med Kimballs 5. prinsipp: *A phrase is closed as soon as possible, i.e., unless the next node parsed is an immediate constituent of that phrase* (Kimball, 1973, s. 36). Selv om både *et problem* og *Jacksons minne* åpenbart er umiddelbare døtre av samme frase som *på* i (75-a) og (75-b), ser det ut som parseren har tolket annP som en avslutning på måteNP, og 'lukket' frasen. Kimballs syvende prinsipp sier *When a phrase is closed, it is pushed down into a syntactic (possible semantic) processing stage and cleared from short-term memory* (Kimball, 1973, s. 38). Når en frase først er lukket, er den slettet fra korttidsminnet og blir utilgjengelig.

Ved prosessering av en konstruksjon uten langdistanseavhengigheter, der elementene som hører sammen befinner seg ved siden av hverandre, ville alle elementene være 'om bord' i konstruksjonen, og lineært tilgrensende elementer ville være i umiddelbart naboskap også syntaktisk. Slik er det ikke med måteNP, der måteP befinner seg langt fra objektet sitt (*måte*). Dette krever av korttidsminnet at det husker relasjonene det må sørge for at er fullstendige (som at *måte* i måteNP får en måteP å være objekt til). Når et element med samme form som det som skal til for å lukke frasen forekommer underveis i ytringen, lukkes frasen (tilsynelatende automatisk). Med tanke på at performansen foregår lineært,

⁷⁰Heretter vi jeg kalle en *på* i måteF 'annP' ('en annen *på*').

er det forståelig at annP tolkes som måteP i ytringer som (75-b), der sekvensen frem til *på* er grammatisk. At *Jacksons minne* ikke kan få en syntaktisk rolle (siden *tjene* er transitivt og allerede har et objekt) 'oppdages' ikke før etter at måteNPen er lukket av den falske måtePen (annP). En viktig forskjell mellom eksemplene i (75) er at annP i (75-a), i motsetning til annP i (75-b), ikke kan tolkes som annet enn innledningen til *sette* sitt oblike ledd.

Tabell 9 viser andelen ugrammatiske og grammatiske (+/-G) måteNP med en annP som underveis ikke kan (-/+ T) tolkes som måteP (fra søkene i NA).

TABELL 9:

	-T	+T
-G	5	20
+G	1	6

Frekvensen er for lav til at tallene kan si noe om effekten muligheten for en lineær tolkning av annP som måteP har for om den egentlige måtePen utelates gir, men det ser ut som en slik tolkning er mulig langt oftere (26 forekomster) enn den ikke er det når annP forekommer (6 forekomster). Dette er ikke overraskende, siden det er svært få verb som tar en oblik agent med *på* som kjerne.⁷¹

10.3 RP og *på* som 'lukkere'

I 10.1.6 så vi hvordan preposisjoner, og i særlig grad *på* kunne 'avbryte' prosesseringen av måteNP, og føre til at parseren regner en annen *på* enn måteP for å være måteP, og slik produsere en ugrammatisk setning. Det synes som om parseren kan ha en forventning om hvilke elementer som skal forekomme i setningen for at frasen som er under produksjon skal kunne ferdigsstilles, jf. Kimballs femte prinsipp om at setningen lukkes så snart den *kan* lukkes. Dette gjelder trolig særlig når et element er obligatorisk, i tilfeller som måteNP, der *måte* signaliserer at det skal forekomme en *på* senere i setningen som skal ta *måte* som objekt. Når den første forekomsten av *på* viser seg, synes det ikke som parseren tar hensyn til om det er i den forventede posisjonen (som en *mulig* avslutning av måteNP) måteP forekommer i (se 10.2)⁷², men regner måteNP for å være lukket fordi det har forekommet

⁷¹Samtlige av disse måteNPene, der annP ikke kunne tolkes som måteP, hadde hovedverb som tok oblik agent (med ett unntak, som inneholdt direkte sitat).

⁷²Riktignok er dette svake tall – med for lav frekvens til å si noe sikkert – men funnene som er gjengitt i tabell 9 viser at det neppe spiller en stor rolle for sannsynligheten for om *på* skal avbryte parsingen av måteNP at en slik *på* har en syntaktisk posisjon som gjør at den kunne vært måteP.

et element som tilfredsstillende behovet til parseren: at dette elementet er *på*. Som vist i tabell 3 og 6 er det to ting som synes avgjørende for om parseren mislykkes i å lukke måteNP på riktig måte (med måteP): (1) forekomsten av *på* i måteF, og (2) lengden på måteF.

Uten å ha presentert statistikk for det, har jeg så langt vist at også andre preposisjoner har en tendens til å utelates fra bestemte posisjoner etter forholdsvis forutsigbare mønstre. I 7.2 viste jeg hvordan andre, strandede preposisjoner enn måteP kan utelates i tilfeller der en formlik preposisjon forekommer tidligere i setningen, noe som intuitivt synes å skyldes at den ikkestrandede preposisjonen kan se ut til å utgjøre samme funksjon som det den strandede preposisjonen vanligvis gjør. Eksempel (54-a) gjengis her som (77).

(77) *De bor ennå i den kommunen de er folkeregistrert __.

Både den forekommende og den manglende preposisjonen har som objekt en NP med *kommune* som kjerne (henholdsvis *den kommunen de er folkeregistrert i* og *den kommunen* (ved syntaktisk binding)). Når kjernen i en preposisjonsfrase har den samme kjernen som preposisjonsfrasens komplementator (som i (77)) har vi å gjøre med en konstruksjon ikke ulik langdistanseavhengigheter. Sammenlign (77) med f.eks. (79) på neste side. De har til felles at en 'kopi' av et manglende element befinner seg tidligere i setningen. I (79) er denne kopien filler, mens i (77) er kopien en formlik preposisjon, som ikke er en kopi etter reglene for syntaksen. Et annet fellestrekk mellom dem er at det manglende elementet ikke bare er kjerne i en frase av samme type som det forekommende elementet (NP og PP), men at også både **kjernen i denne frasen**, og kjernen i utfyllingen er den samme for de to elementene. For (79): [NP the [N **girl**] [CP [C that]]], for (77): [PP [P **i**] [NP den [N kommunen]]]. En slik utelatelse av strandede preposisjoner *kan* altså foregå etter analogi med utelatelsen av et fonologisk uttrykt nominalt element i gappens posisjon.

Jeg har også vist (i 7.1) at parseren kan utelate en av to formlike 'nabopreposisjoner', som *på* i (51-c), gjengitt her som (78-a), med det grammatiske motsvaret (78-b), hentet fra (52-c).

- (78) a. *Det er jo en helhetsvurdering, og det kan man jo veldig vanskelig svare *på* direkte.
b. Det er jo en helhetsvurdering, og det kan man jo veldig vanskelig svare på på direkte.

Når en preposisjon mangler i konstruksjoner som dette, synes det å være fordi parseren tolker forekomsten av *på* som både en strandet preposisjon (som kjerne i *på det*), og som en preposisjon i ikkestrandet posisjon, som kjerne i *på direkten*. I 7.1 vurderte jeg som en forklaring på hvorfor (78-a) foretrekkes fremfor (78-b) at det kjennes unaturlig å bruke sekvensen *på på*. En annen mulig forklaring kan være at den første preposisjonsfrasen (med en strandet *på*) krever mer å prosessere enn den 'nøytrale' preposisjonsfrasen, der objektet følger like etter kjernen. Når den første PPen er lukket og skyves ut av korttidsminnet (jf. Kimballs syvende prinsipp), er den utenfor parserens rekkevidde, og den egentlige årsaken til at *på* befinner seg i denne posisjonen (dvs. som kjerne i *på det*) kan være 'glemt' av parseren. Dermed er det fritt frem for parseren å bruke denne forekomsten av *på* som kjerne i *på direkten*.⁷³

På samme måte som lengden på måteF kan være avgjørende for hvorvidt parseren inkluderer måteP eller ikke, er lengden på en langdistanseavhengighet avgjørende for hvorvidt parseren inkluderer RP eller ikke (jf. 10.1.2). Det er ikke utenkelig at det er den samme årsaken til at parsingen av en måteNP avbrytes av en *på*, som at RP inkluderes: Parseren vil ha noe som kan ta rollen som objekt (som vanligvis opptas av gappen). Erteschik-Shirs eksempel (16-d) gjengis her som (79).

(79) This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to ?_/her.

I dette eksempelet står (etter det jeg har argumentert for å være grammatisk) en gap som objekt i leddsetningen *his mother had given some cakes to*, med en filler med *girl* som kjerne. Setninger der et nulluttrykk representerer et argument er forholdsvis sjelden; parseren håndterer vanligvis setninger der objektet er uttrykt, og setninger med nøytral SVO-stilling kan lukkes når objektet forekommer (Med *kan* lukkes mener jeg at kravene til kompletthet⁷⁴ for verbet er oppfylt). Hvis filleren befinner seg for langt unna gappen til at den er tilgjengelig for parseren når parseren skal lukke setningen, synes parseren å ha 'mistet' filler-gap-avhengigheten som følge av overbelastningen.⁷⁵ Parseren behøver et nominalt element for å lukke setningen, og bruker et pronomen som kan referere anaforisk til filleren.

Som vist i 1.1 (se bl.a. figur 1) er det ofte parseren som svikter ved produksjon av

⁷³Dette er kun en *mulig* forklaring, og jeg vil ikke gå nærmere inn på den i denne oppgaven.

⁷⁴Jf. prinsippet om kompletthet (Dalrymle, 1990, s. 37)

⁷⁵Erteschik-Shirs vurderinger av eksemplene i (16) i 3.2 taler for at større avstand mellom filler og gap gir større belastning for parseren.

akseptable setninger som bryter med de grammatiske reglene, som (79) med det resumptive pronomenet *her*. I tilfeller der RP foretrekkes fremfor en gap, kan det virke som RP ikke velges som en bevisst, kommunikativ løstning fra språkbrukerens side, men snarere som et element parseren bruker for å oppfylt (det tilsynelatende ellers uoppfylte) kravet om kompletthet. Siden dette kravet allerede er oppfylt, ved syntaktisk binding til filleren (som gappen til i eksempel (79) utgjør et preposisjonsobjekt), resulterer tilleggingen av enda et 'objekt' (RP) i at et annet prinsipp brytes (prinsippet om koherens).

11 Konklusjon

I denne oppgaven har jeg vist noen mulige årsaker til at dype, egeninnføyde setninger – på tross av at de er fullt ut grammatiske – er vanskelige å bruke og vanskelige å forstå. Jeg har referert til empiriske undersøkelser som viser hvordan to setninger, som begge er like grammatiske, akseptable og forståelige, kan kreve en ulik mengde arbeid av parseren. Jeg har også vist til studier av språkbrukeres respons på dype, egeninnføyde setninger, og slått fast at det ikke skal mer til enn tre innføyninger for å gi parseren store problemer med prosesseringen av en ytring.

Felles for setninger med en 'lange' langdistanseavhengigheter og dype, egeninnføyde setninger er inakseptabiliteten disse forårsaker. Det som skiller de to er muligheten man i langdistansekonstruksjonen har for å 'redde' akseptabiliteten og bedre forståeligheten ved hjelp av et resumptivt pronomen. En slik redningsaksjon – har jeg hevdet – vil i alle tilfeller resultere i en ugrammatisk setning (i alle fall for engelsk, norsk og svensk), enten den var grammatisk eller ugrammatisk (ved brudd på øyrestriksjonene) før det resumptive pronomenet ble lagt til. Til gjengjeld kan det resumptive pronomenet i noen tilfeller (i) øke akseptabiliteten (særlig ved brudd på øyrestriksjonene) og (ii) øke forståeligheten.

I kapittel 4.1 viste jeg ved å beskrive referansemuligheter og -begrensninger at 'parasitic gaps' (PG) ikke har like mange muligheter for referanse som et pronomen i PGs posisjon. På samme måte har jeg argumentert mot påstanden (som riktignok har vært omstridt) om at resumptive pronomen er fonologiske uttrykk for en gap: En gaps mulige referenter er begrenset til én, mens et pronomen kan ha flere mulige referenter. Intuitivt virker det (ironisk nok) som om RP brukes av de samme grunnene som det er til at PG foretrekkes brukt fremfor et pronomen. På samme måte som bruk av PG i PG-konstruksjoner kan gi mer presis referanse enn et pronomen (som kan være flertydig dersom det befinner seg flere nominale elementer med de samme attributtene for kjønn, tall og person i setningen), kan et resumptivt pronomen i slutten av en lang langdistansekonstruksjon minne lytteren (eller taleren, for den del) om hvilket av de nominale elementene gappen binder. Denne måten å referere på har likevel sine ulemper. En av dem har blitt nevnt i denne oppgaven: Setninger med resumptive pronomen er ugrammatiske. En annen ulempe ved bruk av resumptive pronomen er den manglende nøyaktigheten anaforisk referanse har i forhold til syntaktisk binding når det finnes flere nominale elementer med person- tall- og kjønnsattributter som samsvarer med pronomenet.

Del 2 av oppgaven har vist at en 'redningsaksjon' for akseptabiliteten tilsvarende den RP synes å brukes i også kan skje ubevisst, i lange måteNPer, der et element (måteP)

kan utelates snarere enn å tillegges (som RP). Utelatelsen av måteP ser ut til å ha den samme effekten for akseptabiliteten til lange måteNPer som det å inkludere RP har for langdistansekonstruksjoner.⁷⁶ Følgene dette får for grammatikaliteten er også de samme (inkoherens), ved at begge setningstypene har flere (eller andre typer) argumenter enn predikatene tillater. Det har blitt vist at bruken av RP kan gjøre setninger mer forståelige: Man plasserer et nominalt element (RP) med anaforisk referanse til filleren, i den posisjonen i setningen man i en nøytral setning vanligvis finner nominale elementer (foran et verb). Bruken av måteP i en lang måteNP synes (i motsetning til RP) ikke å verken gagne eller forringe forståeligheten til en ytring, siden lytteren allerede kjenner rollen *måte* må ha i en slik konstruksjon. På slutten av en lang måteNP kan måteP oppleves som malplassert, men siden konstruksjonen så langt allerede er behandlet av parseren og forstått av lytteren, kan måteP 'ignoreres' dersom bruken av den ikke forstås eller oppleves som akseptabel.

Av det denne oppgaven har vist kan det virke som et språk som en samling grammatiske setninger (som ikke alltid er akseptable eller forståelige) tilsynelatende ikke har blitt til først og fremst i kommunikasjonsøyemed, men snarere som et system for systemets skyld. Ugrammatiske setninger har vist seg å kunne være nyttigere i kommunikasjon enn grammatiske, og graden av grammatikalitet, akseptabilitet og forståelighet for visse konstruksjoner har vist seg å variere i stor grad (og på en uforutsigbar måte) selv mellom ellers så like språk som norsk, dansk og svensk. Chomsky (1991, s. 448) mener den utbredte bruken av ugrammatiske konstruksjoner i kommunikasjonsøyemed i seg selv ikke utgjør noe problem for lingvistikken:

There is no real paradox here; there is no reason to suppose, a priori, that the general design of language is conducive to efficient use. Rather, what we seem to discover are some intriguing and unexpected features of language design, not unlike those that have been discovered throughout the inquiry into the nature of language, though unusual among biological systems of the natural world.

Tendensen vi som språkbrukere har til å benytte oss av konstruksjoner som bryter med de grammatiske reglene (det finnes trolig flere enn dem jeg har vist eksempler på i denne oppgaven) behøver ikke å bety at vår språklige kompetanse ikke er et godt verktøy for kommunikasjon. Det er parserens utilstrekkelighet (*memory limitations*) som skaper skillet mellom kompetansen og performansen. Det er først og fremst parserens begrensninger som utgjør forskjellen mellom oss vanlige språkbrukere og Chomskys *ideal speaker-listener*. Det

⁷⁶Jf. den hyppige forekomsten av lange måteNPer uten måteP, som vist i tabell 3.

er parserens utilstrekkelighet som hindrer oss i å ta i bruk alle mulighetene grammatikken gir oss, og som forårsaker at vi gjør grammatiske feil. Likevel setter ikke dette oss ut av stand til å kommunisere ved bruk av språk, siden kommunikasjon foregår nettopp i en kontekst med tilgang på andre lingvistiske og ekstralingvistiske midler enn grammatikk (som diksjon og kontekst). I et slikt miljø er kanskje vanlige språkbrukere som oss selv bedre i stand til å formidle proposisjoner enn Chomskys ideelle taler-lytter.

Litteratur

- Allwood, J. (1982). The Complex NP Constraint in Swedish. I E. Engdahl & E. Ejerhed (red.), *Readings on unbounded dependencies in Scandinavian languages* (s. 15-32). Umeå: Almqvist & Wiksell.
- Behaghel, O. (1923). *Deutsche Syntax: eine geschichtliche Darstellung*. (Bind 4). Heidelberg: Winter.
- Bresnan, J. (2001). *Lexical-Functional Syntax*. Maiden, Massachusetts: Blackwell.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1963). Formal properties of grammars. I R. D. Luce, R. R. Bush & E. Galanter (red.), *Mathematical Psychology* (Vol. 2) (s. 323-418). New York: John Wiley and Sons, Inc.
- Chomsky, N. (1964). Current issues in linguistic theory. I J. A. Fodor & J. J. Katz (red.), *The structure of language: Readings in the philosophy of language* (s. 50-118). Englewood Cliffs, Prentice Hall.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1977). On *wh*-movement. I P.W. Culicover, T. Wasow & A. Akmajian (red.), *Formal syntax* (s. 71-132). New York: Academic Press.
- Chomsky, N. (1982). *Some concepts and consequences of government and binding*. Massachusetts: MIT Press.
- Chomsky, N. (1991). Some notes on economy of derivation and representation. I R. Freidin (red.), *Principles and parameters in comparative grammar*. Massachusetts: MIT Press.
- Chomsky, N. (2004). Degrees of Grammaticalness. I B. Aarts, D. Denison, E. Keizer & G. Popova (red.), *Fuzzy grammar: a reader* (s. 117-150). Oxford: Oxford University Press.
- Creider, C. A. (1987). Structural and pragmatic factors influencing the acceptability of sentences with extended dependencies in Norwegian. *Working Papers in Linguistics* 4. Universitetet i Trondheim.
- Dalrymple, M. (1990). *Lexical Functional Grammar*. San Diego: Academic Press.
- Engdahl, E. (1982). Restrictions on unbounded dependencies in Swedish. I E. Engdahl & E. Ejerhed (red.) *Readings on unbounded dependencies in Scandinavian languages* (s. 151-174). Umeå: Almqvist & Wiksell International.
- Engdahl, E. (1983). Parasitic gaps. *Linguistics and philosophy* 6 (s. 5-33).
- Engdahl, E. (1984a). Subject gaps: An asymmetry between questions and relatives in Norwegian, i C. Jones & P. Sells (red.) *Proceedings of NELS 14*. University of Massachusetts, Amherst.

- Engdahl, E. (1984b). Parasitic gaps, resumptive pronouns and subject extraction. *The syntax and semantics of constituent questions with special reference to Swedish*. Dordrecht: Reidel.
- Erteschik-Shir, N. (1992). Resumptive Pronouns in Islands. I H. Goodluck & M. Rochemont (red.), *Island constraints* (s. 89-108). Dordrecht: Kluwer.
- Falk, Y. N. (2002). *Resumptive pronouns in LFG*. Artikkel presentert ved Proceedings of the LFG02 Conference. Hentet fra <http://csli-publications.stanford.edu/LFG/7/lfg02falk.pdf>
- Fodor, J. D. (1981). Does performance shape competence?. *Philosophical transactions of the Royal Society of London. Series B, biological sciences* 295(1077) (s. 285-295).
- Faarlund, J., Lie, S. & Vannebo, K. (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Hawkins, R. (2001). *Second language syntax: A generative introduction*. Oxford: Blackwell.
- Hellan, L. & Christensen K. K. (1986). Introduction. I L. Hellan & K. K. Christensen (red.), *Topics in Scandinavian syntax* (s. 1-29). Dordrecht: D. Reichel Publishing Company.
- Hyltenstam, K. (1984). The use of typological markedness conditions as predictors of second language acquisition. I R. W. Andersen (red.), *Second languages: A cross-linguistic perspective* (s. 39-58). Rowley, Massachusetts: Newbury House.
- Just, M., Carpenter, P. A. & Keller, T. A. (1996). The Capacity Theory of Comprehension: New frontiers of evidence and arguments. *Psychology Review*, 103(4), 773-780.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition* 2(1), (s. 15-47).
- Korrekturavdelingen (2011, september). *Frekvensordliste – vanlige ord i norsk språk*. Hentet fra <http://www.korrekturavdelingen.no/K4Frekvensordliste500VanligsteNorsk.htm>
- Lie, S. (1982). Discontinuous questions and subjacency in Norwegian. I E. Engdahl & E. Ejerhed (red.) *Readings on unbounded dependencies in Scandinavian languages* (s. 151-174). Umeå: Almqvist & Wiksell International.
- Lyons, J. (1968). *Introduction to theoretical linguistics*. London: Cambridge University Press.
- Marcus, M. (1975). Diagnosis as a notion of grammar. In R. Schank & B. Nash-Weber (red.), *Theoretical issues in natural language processing* (s. 6-10). Cambridge, Massachusetts.
- Maling, J. (1977). *Preposition stranding in Germanic*. Presentert ved the Colloquium on bounded vs. unbounded transformation. University of Amsterdam, Amsterdam.
- Maling, J. (1978). An asymmetry with respect to wh-islands. *Linguistic Inquiry*, 9(1), 75-89.
- Milne, R. W. (1982). Predicting garden path sentences. *Cognitive Science*, 6, 349-373.

- Norsk Aviskorpus (2011, juli). *Søk i hele korpuset*. Hentet fra <http://avis.uib.no/sok/sok-i-hele-korpuset>
- Rizzi, L. (1990). *Relativized minimality*. Massachusetts: MIT Press.
- Ross, J. R. (1967). Constraints on variables in syntax. Upublisert doktorgradsavhandling. MIT, Cambridge, Massachusetts.
- Saltveit, L. (1994). *Bortskjemt av vår Herre. Trekk fra eget liv og fra en annerledes krig*. Haugesund: Ludvig Eriksen.
- Shlonsky, U. (1992). Resumptive pronouns as a last resort. *Linguistic Inquiry*, 23, 443-468.
- Stromswold, K., Caplan, D., Alpert, N. & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, 52, 452-473.
- Taraldsen, K. T. (1978). The scope of wh movement in Norwegian. *Linguistic Inquiry*, 9(4) (s. 623-640).
- Teleman, U. (1999). *Ser du stjärnan i det blå? Om grammatisk beskrivning och normering*. Hentet fra <http://www.ep.liu.se/ecp/006/025/ecp00625b.pdf>. Linköping: Linköping University Electronic Press.
- Teleman, U., Christensen, L., Hellberg, S. & Andersson, E. (1999). *Svenska Akademiens grammatik*. Stockholm: Svenska Akademien.
- Vangsnes, Ø. (2004). On *wh*-questions and V2 across Norwegian dialects: a survey and some speculations. *Working Papers in Scandinavian Syntax*, 73, (s. 1-59).

Abstract

This thesis presents a number of syntactic constructions that relatively often seem to cause the parser to produce an ungrammatical and/or unacceptable sentence. The first part of the thesis shows various types of deviant constructions, and discusses whether such constructions are ungrammatical and – if so – why they seem to be preferred to their well-formed equivalents. Filler-gap-dependencies involving resumptive pronouns are among such deviant constructions, and has been shown to vary greatly in acceptability even between closely related languages such as Norwegian and Swedish. The thesis also includes a discussion regarding to what extent island constraints (varying greatly in acceptability between e.g. English and Norwegian) constitute restrictions on grammaticality, or whether violations of them only cause unacceptability. In the second part of the thesis, empirical evidence is presented, of significantly frequent occurrences of ungrammatical attempts to form certain Norwegian noun phrases. These noun phrases (abbreviated 'måteNP') are equivalents to nominalization in sentences like 'Peter eats fish in an unusual way'. Whereas the preposition (*in*) is left out in such nominalizations in English ('Peter has an unusual way to eat fish $_/*in$ '), the Norwegian preposition *på* is (obligatorily) kept in the equivalent Norwegian NP: 'Peter spiser fisk på en uvanlig måte' / 'Peter har en uvanlig måte å spise fisk $*_/på$ '. Moreover, the thesis shows it to be probable that these and other constructions are found in language use – despite their ungrammaticality – because of errors made by the parser. In conclusion, the relations between the parser, competence and performance are briefly discussed, as are grammar's role in language, and language's role in human communication.