

Bestemt og ubestemt

En korpusbasert studie av morsmålets rolle for bruk av bestemte og ubestemte former hos norskinnlærere uten grammatisk bestemthet i morsmålet

Hilde Johansen

Avhandling for graden philosophiae doctor (ph.d.)
Universitetet i Bergen
2022

UNIVERSITETET I BERGEN



Bestemt og ubestemt

En korpusbasert studie av morsmålets rolle for bruk av bestemte og ubestemte former hos norskinnlærere uten grammatisk bestemthet i morsmålet

Hilde Johansen



Avhandling for graden philosophiae doctor (ph.d.)
ved Universitetet i Bergen

Disputasdato: 04.03.2022

© Copyright Hilde Johansen

Materialet i denne publikasjonen er omfattet av åndsverkslovens bestemmelser.

År: 2022

Tittel: Bestemt og ubestemt

Navn: Hilde Johansen

Trykk: Skipnes Kommunikasjon / Universitetet i Bergen

SUMMARY IN ENGLISH

Title: Definite and indefinite. A corpus-based study of first language transfer in the use of definite and indefinite forms by learners without grammatical definiteness in their first language.

This thesis presents a corpus-based study of first language influence (L1 transfer) on the use of definite and indefinite forms of the noun. It investigates texts written by learners of Norwegian whose first language is without grammatical definiteness (-GD-learners). The aim is to investigate two potential effects of L1 transfer: Investigation 1 examines a *quantitative* effect on the use of the definite form in obligatory contexts, and investigation 2 examines a *qualitative* effect on the choice between definite or indefinite forms in different linguistic contexts.

The study makes use of a set of methodological criteria for transfer research proposed by Jarvis (2000, 2010), which involves the investigation of four different *transfer effects* in learner language: *within-group similarity*, *between-group differences*, *cross-language congruity* and *intralingual contrasts*. When these four effects can be found together in learner language, they constitute evidence of transfer by proving that a specific linguistic behaviour is motivated by the first language and that it is also *specific* and *representative* for the group of learners that are studied.

The learner language data are drawn from the Norwegian learner corpus ASK, and the study is to a large extent corpus-based, but also includes additional manual annotation of parts of the data. Investigation 1 is based on a corpus of texts written by learners with Polish, Russian or Serbo-Croatian (-GD) as their first languages, and this group of learners is compared to a group of learners with English, Dutch and German (+GD) as their first languages. Investigation 2 studies a smaller corpus of learners whose first languages are Polish (-GD) and English (+GD), and also includes texts written in L1 Norwegian.

While investigation 1 simply looks at the percentages of correct suppliance of definite suffixes in obligatory contexts, investigation 2 investigates transfer through two different hypotheses on how L1 Polish may influence the choice between definite and indefinite forms in the Polish informants' learner language. The first hypothesis is about

transfer of pragmatic word order as a signal of a noun's identifiability, including an L1 based assumption that identifiability is already signalled when pragmatic word order is used. This hypothesis is based on earlier work on how communicative redundancy affects the use of articles in learner English, especially Young (1996). Like Young's study, Hypothesis 1 predicts that definite suffixes and indefinite articles will be considered redundant when their function (identifiability) is already signalled by word order, but suffixes and articles will be supplied when identifiability is *not* signalled, that is when pragmatic word order is not used.

The second hypothesis is about transfer of the functional distinctions underlying the choice of form within the grammatical category *case* in Polish. Hypothesis 2 predicts that definite and indefinite forms will be used to express the distinction between functionally marked and unmarked cases, that is between the nominative and the other cases (non-nominatives). This means that in the Polish learners' texts the indefinite forms will be used in linguistic contexts associated with the nominative in Polish (grammatical subjects), and that definite forms will be used in linguistic contexts associated with non-nominatives in Polish.

The results of the study are positive for investigation 1: There are significant differences between -GB-learners and +GB-learners when it comes to how often definite forms are used in obligatory contexts, and it is concluded that this is due to L1 transfer. Investigation 2 comes out negative for both hypotheses: The definite and indefinite forms in Norwegian are neither used in accordance with redundant and non-redundant contexts, nor to mark the distinction between nominatives and non-nominatives. Instead, it is revealed that the Polish learners to a great extent use forms in accordance with the rules of the target language: Indefinite forms are used when the nouns have indefinite reference, definite forms are used when the nouns have definite reference, and both forms are used for generic reference.

It should be emphasised that the Polish learners *did* have linguistic choices that were in accordance with the predictions of the hypotheses about L1 transfer in some of the contexts. But when these L1 like tendencies were compared to the linguistic choices of the Norwegian L1 group, it was revealed that the Polish learners' choices occurred in contexts where the implications of the hypotheses were the same as the typical choice

of form in the target language. In linguistic contexts especially elicited to highlight where the implications of the hypotheses were *not* the same as in the target language, the Polish learners' choice of form was *not* in accordance with the hypotheses about L1 transfer. This finding proved to be of special importance for the conclusion that the Polish learners' use of definite and indefinite forms probably was influenced by the target language, rather than their first language.

A secondary finding of the study was two cases of L1 transfer in the *English* learners' choice of form. The first case was an overrepresentation of the indefinite article when reference was indefinite, and the other case was an underrepresentation of definite forms when reference was generic. This transfer was explained by two differences between English and Norwegian: Indefinite articles are more often used in English than in Norwegian when reference is indefinite, and definite forms are less often used in English when reference is generic.

FORORD

Denne avhandlingen er skrevet som en del av forskningsprosjektet *ASKeladden – en korpusbasert tilnærming til morsmålstransfer i innlærerspråk*, et prosjekt som rettet seg mot morsmålpåvirkning og bruk av det elektroniske innlærerkorpuset Norsk andrespråskorpus (ASK). Prosjektet ble ledet av professor Kari Tenfjord, varte fra 2008 til 2013 og ble finansiert av Norges Forskningsråd.

Jeg har gjennom denne perioden hatt flere veiledere, først hovedveileder førsteamanuensis Jon Erik Hagen og biveileder professor Kari Tenfjord, og senere professor Ingvild Nistov og professor Lars Anders Kulbrandstad, som veiledet meg fram til innleveringen. Førsteamanuensis Bård Uri Jensen har også vært involvert for å gi råd om statistikk tidlig i prosjektet.

Gjennom ASKeladdens prosjektperiode hadde jeg glede og nytte av å diskutere deler av prosjektet med flere eksterne fagpersoner fra inn- og utland, og i tillegg har prosjektdeltakerne og medkontoristene Snorre Karkkonen Svensson, Silje Ragnhildstveit, Olivia Szymanska, Marte Nordanger og Cecilie Hammes Carlsen vært verdifulle og inspirerende samtalepartnere på mange forskjellige måter.

Paul Meurer, som står bak grensesnittet til ASK, har med jevne mellomrom hjulpet meg å løse små og store spørsmål og kriser knyttet til korpuset, og min tidligere kollega fra Norsk språkstest, Anna-Marie Kjøde Olsen, har hjulpet meg med korrektur på enkelte kapitler.

En stor takk til alle dere!

No book can ever be finished. While working on it we learn just enough to find it immature the moment we turn away from it.

Karl Popper

INNHold

SUMMARY IN ENGLISH	3
FORORD.....	6
INNHold.....	7
LISTE OVER TABELLER.....	13
LISTE OVER FIGURER.....	15
LISTE OVER FORKORTELSER	17
1 INNLEDNING	18
1.1 ASkeladden-prosjektet	19
1.2 Den teoretiske bakgrunnen for studien	19
1.2.1 Bestemthet som grammatisk kategori.....	20
1.2.2 Bestemthet innen andrespråksforskningen	21
1.2.3 Morsmålspåvirkning	22
1.2.4 Norsk versus engelsk som målspåk	25
1.3 Empiriske undersøkelser	26
1.3.1 Undersøkelse 1.....	27
1.3.2 Undersøkelse 2.....	27
1.4 Datamateriale og sentrale metodiske sider ved studien	30
1.4.1 Datamateriale	30
1.4.2 Påvisning av transfereffekter	30
1.5 Strukturen i avhandlingen	32
2 MORSMÅLSPÅVIRKNING.....	33
2.1 Definisjon og avgrensning.....	33
2.2 Fenomenet transfer: hva, hvordan og når?.....	35
2.2.1 Hva? Ulike perspektiv på hva transfer er.....	35
2.2.2 Hvordan? Utfall av og klassifisering av transfer	37
2.2.3 Når? Faktorer som påvirker forekomst av og type transfer	38
2.3 Tverrspråklige likheter og forskjellers rolle for transfer	39
2.3.1 Grad av samsvar mellom morsmål og målspåk.....	39
2.3.2 Subjektive og objektive likheter og forskjeller.....	40
2.3.3 Ulike typer kontrastive antakelser og transfertyper	42
2.3.4 Transfer som resultat av likhetsrelasjon	43
2.3.5 Transfer som resultat av ulikhetsrelasjon	44
2.3.6 Transfer som resultat av manglende tverrspråklig identifikasjon.....	44
2.3.7 Oppsummering av fenomenet transfer.....	48
2.4 Metodiske sider av studiet av transfer.....	48
2.4.1 Et enhetlig rammeverk for transferstudier og kriterier	48
2.4.2 Teorinøytral definisjon	49
2.4.3 Transfereffekter	50
2.4.4 Samspillet mellom transfereffektene	54
2.4.5 Kontroll på variabler.....	55
2.5 Oppsummering av kapittel 2	56
3 GRAMMATISK BESTEMTHET OG REFERANSE.....	57
3.1 Bestemthet som grammatisk kategori	57
3.1.1 Grammatisk bestemthet – en heterogen kategori.....	59

3.2	Referanse	60
3.2.1	Referensiell bruk	61
3.2.2	Ikke-referensiell bruk	63
3.3	Bestemthetsekvivalenter	65
3.4	Tall og tellelighet	68
3.5	Oppsummering av kapittel 3	69
4	BESTEMTHET I ANDRESPRÅKSFORSKNINGEN	70
4.1	Huebner (1983)	71
4.2	-GB-innlærere og bestemthet	74
4.2.1	Lingvistiske og diskursive universalier hos -GB-innlærere	75
4.2.2	Transfer hos -GB-innlærere	78
4.2.3	Standardforklaringen	87
4.3	Redundanseffekten	89
4.4	Nordiske studier	91
4.4.1	Nordanger 2017	91
4.4.2	Axelsson (1994)	92
4.4.3	Overførbarhet mellom studier med ulike målspråk	94
4.5	Oppsummering av kapittel 4	96
5	ELEKTRONISKE INNLÆRERKORPUS OG ASK	98
5.1	Elektroniske innlærerkorpus	98
5.1.1	Innholdssiden av elektroniske innlærerkorpus	100
5.2	Norsk Andrespråskorpus (ASK)	101
5.2.1	Design og utvalgsriterier	101
5.2.2	Innholdet i ASK	103
5.3	ASK i studiet av transfer	105
5.4	Oppsummering av kapittel 5	106
6	UNDERSØKELSE 1: KVANTITATIV EFFEKT AV MORSMÅLET	108
6.1	Metodiske sider av undersøkelse 1	109
6.1.1	Datamaterialet	109
6.1.2	Obligatoriske kontekster og markeringsprosent	109
6.1.3	Korpus-søk	109
6.1.4	Transfereffekter og statistisk testing	110
6.2	Hele GB-korpuset	113
6.2.1	Antall obligatoriske kontekster i GB-korpuset	113
6.2.2	-GB-korpuset: Forskjeller i markeringsprosent	116
6.2.3	GB-korpuset: antall tekster med 100 % markering	117
6.2.4	Oppsummering av funn i GB-korpuset	118
6.3	SP- og HN-korpuset	118
6.3.1	Størrelse og obligatoriske kontekster i SP- korpus og HN-korpus	119
6.3.2	SP-korpuset	120
6.3.3	HN-korpuset	123
6.3.4	Oppsummering av funn i SP-korpuset og HN-korpuset	125
6.3.5	Tidligere undersøkelse av B1- og B2-nivå	126
6.4	Polsk og engelsk delkorpus	127
6.4.1	Obligatoriske kontekster i polsk og engelsk delkorpus (Språkprøven) ..	128

6.4.2	Markeringsprosent i polsk og engelsk gruppe (Språkprøven)	129
6.4.3	Obligatoriske kontekster i polsk og engelsk delkorpus (Høyere nivå) ...	130
6.4.4	Markeringsprosent i polsk og engelsk gruppe (Høyere nivå)	131
6.5	Oppsummering og diskusjon av undersøkelse 1	133
6.5.1	Hovedfunn i undersøkelse 1	133
6.5.2	Transfereffekter: morsmålsamsvar og intralingval kontrast	135
6.5.3	Forbehold ved utvalget av substantivfraser	138
6.5.4	Manglende differensiering av referensielle typer	139
6.5.5	Konklusjon for undersøkelse 1	140
6.5.6	Neste delundersøkelse	141
7	SUBSTANTIV OG REFERANSE I NORSK, POLSK OG ENGELSK	143
7.1	Enkle substantivfraser	143
7.2	Norsk	144
7.2.1	Norske substantiv	144
7.2.2	Generelle forhold knyttet til substantivets form	145
7.2.3	Tall og tellelighet	146
7.2.4	Referanse og valg av bestemte og ubestemte former	147
7.3	Polsk	151
7.3.1	Polske substantiv og morfologiske kategorier	151
7.3.2	Bestemthetsekvivalenter i polsk	155
7.3.3	Referensielle typer i polsk	157
7.3.4	Oppsummering av beskrivelsen av polsk	158
7.4	Engelsk	159
7.4.1	Engelske substantiv og substantivfraser	159
7.4.2	Referanse og valg av artikkel i engelsk	161
7.4.3	Oppsummering engelsk	163
7.5	Oppsummering av språklige forskjeller og likheter	164
8	UNDERSØKELSE 2: HYPOTESER OG DATAMATERIALE	167
8.1	Bakgrunnen for de to hypotesene om polsk morsmålspåvirkning	168
8.1.1	Bakgrunn for hypotesene	168
8.1.2	Hypotese 1	171
8.1.3	Hypotese 2	178
8.1.4	Hypoteser og prediksjoner oppsummert	185
8.2	Datamateriale – utvalg av tekster fra ASK	186
8.2.1	Tekstene	187
8.2.2	Tekstskriverne	188
8.3	Datamateriale – utvalg av substantivfraser og søk i ASK	189
8.3.1	Søkemeter: grunnsøk og modifisering	189
8.3.2	Andel av enkle substantivfraser	191
8.4	Koding av kontekster	192
8.4.1	Syntaktisk funksjon	192
8.4.2	Setningsposisjon	193
8.4.3	Tekstuell aktivering	195
8.4.4	Referensiell type	198
8.4.5	Tellelighet	204
8.5	Generelt om gjennomføringen av undersøkelse 2	205

8.5.1	Prediksjoner om assosiasjon mellom form og kontekst.....	205
8.5.2	Undersøkelse på gruppenivå og individnivå.....	207
8.5.3	Undersøkelse av transfereffekter	208
8.5.4	Om inndeling i individuelle profiler	213
8.6	Oppsummering av kapittel 8	215
9	UNDERSØKELSE 2: HYPOTESE 1	217
9.1	Utvalg av materiale og prediksjoner.....	217
9.2	Transfereffekter.....	221
9.2.1	Morsmålssamsvar	221
9.2.2	Intragruppe-homogenitet	224
9.2.3	Intergruppe-heterogenitet og intralingval kontrast	224
9.3	Undersøkelse av hypotese 1 i hele materialet	226
9.3.1	Deskriptivt for alle gruppene	226
9.3.2	Fordeling av former i samlet materiale.....	239
9.3.3	Morsmålssamsvar i den polske gruppen	228
9.3.4	Sammenligning med engelsk og norsk gruppe	237
9.3.5	Individuelle profiler	245
9.3.6	Oppsummering av funn for hypotese 1 i det samlede materialet	249
9.4	Materiale differensiert etter ulike referensielle typer	252
9.5	Indefinitt referanse	253
9.5.1	Deskriptivt for alle gruppene	255
9.5.2	Sammenfalls- og markørkontekster for indefinitt referanse	257
9.5.3	Samlet oversikt for indefinitt referanse	258
9.5.4	Assosiasjon mellom UB-kontekster og ubestemt entall/flertall	260
9.5.5	Assosiasjon mellom BE-kontekster og bestemt form.....	263
9.5.6	Assosiasjon mellom ART-kontekster og artikkelfrase	266
9.5.7	Oppsummering av funn på gruppenivå.....	269
9.5.8	Individuelle profiler	272
9.5.9	Oppsummering av funn for hypotese 1 ved indefinitt referanse	277
9.6	Definitt referanse	283
9.6.1	Deskriptivt for alle gruppene	286
9.6.2	Sammenfalls- og markørkontekster for definitt referanse	286
9.6.3	Samlet oversikt for definitt referanse.....	287
9.6.4	Assosiasjon mellom UB-kontekster og ubestemt entall/flertall	289
9.6.5	Assosiasjon mellom BE-kontekster og bestemt form.....	291
9.6.6	Assosiasjon mellom ART-kontekster og artikkelfrase	294
9.6.7	Oppsummering av funn på gruppenivå.....	296
9.6.8	Individuelle profiler	297
9.6.9	Oppsummering av funn for hypotese 1 ved definitt referanse.....	299
9.7	Generisk referanse.....	303
9.7.1	Deskriptivt for alle gruppene	306
9.7.2	Sammenfalls- og markørkontekster ved generisk referanse	308
9.7.3	Samlet oversikt for generisk referanse.....	311
9.7.4	Assosiasjon mellom UB-kontekster og ubestemte former	311
9.7.5	Assosiasjon mellom BE-kontekster og bestemt form.....	313
9.7.6	Assosiasjon mellom ART-kontekster og artikkelfrase	316
9.7.7	Oppsummering av funn på gruppenivå.....	318
9.7.8	Individuelle profiler	319

9.7.9	Oppsummering av funn for hypotese 1 ved generisk referanse.....	321
9.8	Polsk morsmålssamsvar på tvers av referensiell type.....	324
9.9	Oppsummering hypotese 1	326
9.9.1	Morsmålssamsvar og intergruppe-heterogenitet i samlet materiale	327
9.9.2	Morsmålssamsvar og heterogenitet i det differensierte materialet	328
9.9.3	Sekundære funn i materialet	335
9.9.4	Hypotese 1: Avslutning	340
10	UNDERSØKELSE 2: HYPOTESE 2.....	341
10.1	Utvalg av materiale og prediksjoner.....	341
10.2	Transfereffekter	343
10.2.1	Intragruppe-homogenitet	344
10.2.2	Intergruppe-heterogenitet og intralingval kontrast	344
10.3	Undersøkelse av hypotese 2 i hele materialet	345
10.3.1	Deskriptivt for alle gruppene	345
10.3.2	Fordeling av former i samlet materiale.....	349
10.3.3	Morsmålssamsvar i den polske gruppen.....	349
10.3.4	Sammenligning med norsk og engelsk gruppe	350
10.3.5	Oppsummering av funn på gruppenivå.....	352
10.3.6	Individuelle profiler	352
10.3.7	Oppsummering av funn for hypotese 2 i det samlede materialet	355
10.4	Differensiert materiale – ulike referensielle typer	357
10.5	Indefinitt referanse	357
10.5.1	Deskriptivt for alle gruppene	358
10.5.2	Sammenfalls- og markørkontekster for indefinitt referanse	360
10.5.4	Samlet oversikt for indefinitt referanse	361
10.5.5	Form/kontekst-assosiasjoner på gruppenivå.....	364
10.5.6	Oppsummering av funn på gruppenivå.....	366
10.5.7	Individuelle profiler	366
10.5.8	Oppsummering av funn for hypotese 2 ved indefinitt referanse	369
10.6	Definitt referanse	370
10.6.1	Deskriptivt for alle gruppene	371
10.6.2	Sammenfalls- og markørkontekster for definitt referanse	373
10.6.3	Samlet oversikt for definitt referanse.....	374
10.6.4	Form/kontekst-assosiasjoner på gruppenivå.....	377
10.6.5	Oppsummering av funn på gruppenivå.....	378
10.6.6	Individuelle profiler	378
10.6.7	Oppsummering av funn for hypotese 2 ved definitt referanse.....	381
10.7	Generisk referanse.....	383
10.7.1	Deskriptivt for alle gruppene	384
10.7.2	Sammenfalls- og markørkontekster for generisk referanse	386
10.7.3	Samlet oversikt for generisk referanse.....	387
10.7.5	Form/kontekst-assosiasjoner på gruppenivå.....	389
10.7.6	Oppsummering av funn på gruppenivå.....	391
10.7.7	Individuelle profiler	391
10.7.8	Oppsummering av funn for hypotese 2 ved generisk referanse.....	393
10.8	Polsk morsmålssamsvar på tvers av referensielle typer	395
10.9	Oppsummering hypotese 2	398

10.9.1	Morsmålssamsvar og intergruppe-heterogenitet i samlet materialet	398
10.9.2	Morsmålssamsvar og intergruppeheterogenitet i oppdelt materiale	399
10.9.3	Målspråkslikhet.....	402
10.9.4	Transfer hos de engelske innlærerne	403
10.9.5	Hypotese 2: avslutning.....	404
11	OPPSUMMERING OG DISKUSJON AV STUDIEN	406
11.1	Forskningsspørsmål og studiens to delundersøkelser	406
11.2	Undersøkelse 1: Kvantitativt utslag av transfer	407
11.2.1	Utfordringer med gjennomføringen av undersøkelse 1	411
11.2.2	Konklusjon undersøkelse 1	412
11.3	Undersøkelse 2: kvalitativt utslag av transfer	412
11.3.1	Funn for hypotesene i undersøkelse 2.....	413
11.3.2	Hypotese 1 – oppsummering og diskusjon av resultater	414
11.3.3	Hypotese 2 – oppsummering og diskusjon av resultater	417
11.3.4	Konklusjon for forskningsspørsmål 2.....	421
11.3.5	Hypotesene og <i>inaktiv transfer</i>	421
11.3.6	Målspråksnære formvalg	428
11.3.7	Metodiske forbehold for undersøkelse 2	429
11.3.8	Oppsummering undersøkelse 2.....	437
11.4	Transfer i det engelske materialet.....	440
11.5	Hypoteser om transfer ved nullkontrast	442
11.5.1	Å finne transfer når det er nullkontrast.....	443
11.5.2	Nullkontrast og tverrspråklig identifikasjon – alternative hypoteser....	444
11.6	Transfereffekter og korpus.....	451
11.6.1	Transfereffekter	451
11.6.2	Korpuserfaringer.....	455
11.7	Utleddning	459
LITTERATURKISTE	461	
VEDLEGG	468	
Vedlegg 1: Personvariabler for delkorpuset i undersøkelse 2.....	468	
Vedlegg 2: Tema som er valgt for tekstene til de tre gruppene.....	470	
Vedlegg 3: Søkeuttrykk for grunnsøk fra ASK (undersøkelse 2)	471	

LISTE OVER TABELLER

Tabell 2-1 Oversikt over variabler som samspiller med transfer	55
Tabell 6-1: Deskriptive data for GB-korpuset (obligatoriske kontekster)	113
Tabell 6-2 Markeringsprosent (og rang) i de to innlærergruppene i GB-korpuset	116
Tabell 6-3 Andel tekster med 100 % markering i GB-korpuset	118
Tabell 6-4: Antall obligatoriske kontekster i HN-korpuset og SP-korpuset.	119
Tabell 6-5 Deskriptive data for SP-korpuset (obligatoriske kontekster)	120
Tabell 6-6 Markeringsprosent (og rang) i de to innlærergruppene i SP-korpuset	121
Tabell 6-7 Andel tekster med 100 % markering i SP-korpuset	122
Tabell 6-8: Deskriptive data for HN-korpuset (obligatoriske kontekster)	123
Tabell 6-9 Markeringsprosent (og rang) i de to innlærergruppene i HN-korpuset	124
Tabell 6-10 Andel tekster med 100 % markering i HN-korpuset	125
Tabell 6-11 Andel tekster med 100 % markering i B1-korpuset	127
Tabell 6-12 Andel tekster med 100 % markering i B2-korpuset	127
Tabell 6-13 Deskriptive data for polsk og engelsk delkorpus, SP (obligatoriske kontekster)	128
Tabell 6-14 Markeringsprosent (og rang) i polsk og engelsk gruppe, Språkrøven	129
Tabell 6-15 Andel tekster med 100 % markering i polsk og engelsk gruppe, Språkrøven	130
Tabell 6-16 Deskriptive data for polsk og engelsk delkorpus, HN (obl. kontekster)	130
Tabell 6-17 Markeringsprosent (og rang) i polsk og engelsk gruppe, HN	131
Tabell 6-18 Andel tekster med 100 % markering i polsk og engelsk gruppe, HN	132
Tabell 6-19 Oversikt for funn i de ulike delkorpusene i undersøkelse 1	133
Tabell 7-1 Det norske bøyningssystemet for substantiv	145
Tabell 7-2 Bøyningssystemet for polske substantiv i entall	154
Tabell 7-3 Bøyningssystemet for polske substantiv i flertall	154
Tabell 7-4 Det engelske systemet for substantiv	160
Tabell 8-1 Form/kontekst-assosiasjoner for de to hypotesene	185
Tabell 8-2 Oversikt over antall tekster i engelsk og polsk delkorpus	186
Tabell 8-3 Antall ord og gjennomsnittlig tekstlengde materialet for undersøkelse 2	187
Tabell 8-4 Selvrapporterte engelskferdigheter hos de polske informantene	189
Tabell 8-5 En forenklet framstilling av grunnsøk 1-5.	190
Tabell 8-6 Eksempler på de fem typene enkle substantivfraser	191
Tabell 8-7 Andelen av substantiv og enkle substantivfraser fordelt på de tre gruppene	192
Tabell 8-8 Eksempler på de ulike syntaktiske funksjonene	192
Tabell 8-9 Eksempler på de ulike setningsposisjonene	194
Tabell 8-10 Eksempler på tekstuelt aktivert og ikke-aktivert referent	198
Tabell 8-11 Eksempler på de ulike referensielle typene hentet fra ASK	203
Tabell 8-12 Eksempel på tellelig, ikke-tellelig og kollektivt substantiv hentet fra ASK.	204
Tabell 8-13 Oversikt over form/kontekst-assosiasjoner som er predikert av hypotesene	205
Tabell 9-1 Oversikt over forventede formvalg i de tre gruppene	225
Tabell 9-2 Andel av de ulike formtypene i de tre gruppene	227
Tabell 9-3 Andel av de ulike konteksttypene i de tre gruppene	227
Tabell 9-4 Polsk gruppe: Fordeling av former etter konteksttype for hypotese 1	229
Tabell 9-5 Polsk gruppe: Fordeling av former etter konteksttype for hypotese 1	238
Tabell 9-6 Engelsk gruppe: Fordeling av former etter konteksttype for hypotese 1	238
Tabell 9-7 Norsk gruppe: Fordeling av former etter konteksttype for hypotese 1	238
Tabell 9-8 Profiltyper i de tre gruppene	247
Tabell 9-9 Oversikt over hovedfunn for hypotese 1 i det samlede materialet	249
Tabell 9-10 Andel av de ulike formtypene i de tre gruppene (indefinit referanse)	256
Tabell 9-11 Andel av de ulike konteksttypene i de tre gruppene (indefinit referanse)	256
Tabell 9-12 Sammenfalls- og markørkontekster for indefinit referanse	258
Tabell 9-13 Polsk gruppe: Fordeling av former etter konteksttype (indefinit).	259
Tabell 9-14 Engelsk gruppe: Fordeling av former etter konteksttype (indefinit)	259
Tabell 9-15 Norsk gruppe: Fordeling av former etter konteksttype ved (indefinit)	259
Tabell 9-16 Profiltyper i de tre gruppene (indefinit referanse)	274
Tabell 9-17 Oversikt over hovedfunn for hypotese 1 ved indefinit referanse	277
Tabell 9-18 Andel av de ulike formtypene i de tre gruppene (definit referanse)	285

Tabell 9-19 Andel av de ulike konteksttypene i de tre gruppene (definitt referanse)	285
Tabell 9-20 Sammenfalls- og markørkontekster for definitt referanse	287
Tabell 9-21 Polsk gruppe: Fordeling av former etter konteksttype (definitt)	288
Tabell 9-22 Engelsk gruppe: Fordeling av former etter konteksttype (definitt)	288
Tabell 9-23 Norsk gruppe: Fordeling av former etter konteksttype (definitt)	288
Tabell 9-24 Profiltyper i de tre gruppene (definitt referanse)	298
Tabell 9-25 Oversikt over hovedfunn for hypotese 1 ved definitt referanse	300
Tabell 9-26 Andel av de ulike formtypene i de tre gruppene (generisk referanse)	307
Tabell 9-26 Andel av de ulike konteksttypene i de tre gruppene (generisk referanse)	307
Tabell 9-28 Sammenfalls- og markørkontekster for generisk referanse	309
Tabell 9-29 Polsk gruppe: Fordeling av former etter konteksttype (generisk)	310
Tabell 9-30 Engelsk gruppe: Fordelinger av former etter konteksttype (generisk)	310
Tabell 9-31 Norsk gruppe: Fordelinger av former etter konteksttype (generisk)	310
Tabell 9-32 Profiltyper i de tre gruppene (generisk referanse)	320
Tabell 9-33 Oversikt over hovedfunn for hypotese 1 ved generisk referanse	321
Tabell 9-34 Transfereffekter for assosiasjonen UB-kontekst og ubestemt entall/flertall	330
Tabell 9-35 Transfereffekter for assosiasjonen BE-kontekst og bestemte form	331
Tabell 9-36 Transfereffekter for assosiasjonen ART-kontekst og artikkelfrase	333
Tabell 9-37 Transfereffekter for individuelle profiler	334
Tabell 10-1 Oversikt over forventede formvalg i de tre gruppene	344
Tabell 10-2 Andel av subjekt og objekt i de tre gruppene	347
Tabell 10-3 Andel av ubestemt og bestemt form i de tre gruppene	347
Tabell 10-4 Polsk gruppe: Fordeling av former etter konteksttype	348
Tabell 10-5 Engelsk gruppe: Fordeling av former etter konteksttype	348
Tabell 10-6 Norsk gruppe: Fordeling av former etter konteksttype	348
Tabell 10-7 Profiltyper i de tre gruppene	354
Tabell 10-8: Oversikt over hovedfunn for hypotese 2 i det samlede materialet	356
Tabell 10-9 Andel av subjekt og objekt i de tre gruppene (indefinitt referanse)	359
Tabell 10-10 Andel av ubestemt og bestemt form i de tre gruppene (indefinitt referanse)	359
Tabell 10-11 Sammenfalls- og markørkontekster for indefinitt referanse	360
Tabell 10-12 Polsk gruppe: Fordeling av former etter konteksttype (indefinitt)	362
Tabell 10-13 Engelsk gruppe: Fordeling av former etter konteksttype (indefinitt)	362
Tabell 10-14 Norsk gruppe: Fordeling av former etter konteksttype (indefinitt)	362
Tabell 10-15 Profiltyper i de tre gruppene (indefinitt referanse)	367
Tabell 10-16 Oversikt over hovedfunn for hypotese 2 ved indefinitt referanse	369
Tabell 10-17 Andel av subjekt og objekt i de tre gruppene (definitt referanse)	372
Tabell 10-18 Andel av ubestemt og bestemt form i de tre gruppene (definitt referanse)	372
Tabell 10-19 Sammenfalls- og markørkontekster for definitt referanse	373
Tabell 10-20 Polsk gruppe: Fordeling av former etter konteksttype (definitt)	375
Tabell 10-21 Engelsk gruppe: Fordeling av former etter konteksttype (definitt)	375
Tabell 10-22 Norsk gruppe: Fordeling av former etter konteksttype (definitt)	375
Tabell 10-23 Profiltyper i de tre gruppene (definitt referanse)	380
Tabell 10-24 Oversikt over hovedfunn for hypotese 2 ved definitt referanse	382
Tabell 10-25 Andel av subjekt og objekt i de tre gruppene (generisk referanse)	385
Tabell 10-26 Andel av ubestemt og bestemt form i de tre gruppene (generisk referanse)	385
Tabell 10-27 Sammenfalls- og markørkontekster for generisk referanse	387
Tabell 10-28 Polsk gruppe: Fordeling av former etter konteksttype (generisk)	388
Tabell 10-29 Engelsk gruppe: Fordeling av former etter konteksttype (generisk)	388
Tabell 10-30 Norsk gruppe: Fordeling av former etter konteksttype (generisk)	388
Tabell 10-31 Profiltyper i de tre gruppene (generisk referanse)	392
Tabell 10-32 Oversikt over hovedfunn for hypotese 2 ved generisk referanse	394
Tabell 10-33 Transfereffekter for assosiasjonen mellom subjekt og ubestemt form	400
Tabell 10-34 Transfereffekter for assosiasjonen mellom objekt og bestemt form	400
Tabell 10-35 Transfereffekter for individuelle profiler	401

LISTE OVER FIGURER

Figur 6-1 Den individuelle spredningen av antall obligatoriske kontekster i GB-korpuset	114
Figur 6-2 Den individuelle spredningen for markeringsprosent i GB-korpuset	117
Figur 6-3 Den individuelle spredningen av antall obligatoriske kontekster i HN- og SP-korpuset	120
Figur 6-4 Den individuelle spredningen av antall obligatoriske kontekster i SP-korpuset	121
Figur 6-5 Den individuelle spredningen for markeringsprosent i SP-korpuset	122
Figur 6-6 Den individuelle spredningen av antall obligatoriske kontekster i HN-korpuset	123
Figur 6-7 Den individuelle spredningen for markeringsprosent i HN-korpuset	124
Figur 6-8 Individuell spredning av obligatoriske kontekster i polsk og engelsk gruppe, SP	128
Figur 6-9 Individuell spredning for markeringsprosent i polsk og engelsk gruppe, SP	129
Figur 6-10 Individuell spredning av obl. kontekster i polsk og engelsk gruppe, HN	131
Figur 6-11 Individuell spredning for markeringsprosent i polsk og engelsk gruppe, HN	132
Figur 8-1 Forenklet fremstilling av Hagens feltskjema for hovedsetninger (1998: 264)	175
Figur 8-2 Forenklet fremstilling av Hagens feltskjema for leddsetninger (1998: 269)	176
Figur 9-1 Andel av de ulike formtypene i de tre gruppene	227
Figur 9-2 Andel av de ulike konteksttypene i de tre gruppene	227
Figur 9-3 Andel form innen kontekst (polsk gruppe)	229
Figur 9-4 Andel kontekst innen form (polsk gruppe)	230
Figur 9-5 Andel form innen UB-kontekst	232
Figur 9-6 Andel kontekst innen ub. ent/flt.	232
Figur 9-7 Andel form innen BE-kontekst	234
Figur 9-8 Andel kontekst innen best. form	234
Figur 9-9 Andel form innen ART-kontekst	236
Figur 9-10 Andel kontekst innen art.frase	236
Figur 9-11 Andel av form innen UB-kontekster	239
Figur 9-12 Andel av kontekst innen ubestemt ent/flt	240
Figur 9-13 Andel av form innen BE-kontekster	241
Figur 9-14 Andel av kontekst innen bestemt form	242
Figur 9-15 Andel av form innen ART-kontekster	243
Figur 9-16 Andel av kontekst innen artikkelfraser	243
Figur 9-17 Antall tekster med tre, to eller bare én av konteksttypene	245
Figur 9-18 Profiltyper i de tre gruppene	247
Figur 9-19 Andel former i de tre gruppene (indefinit referanse)	256
Figur 9-20 Andel konteksttyper i de tre gruppene (indefinit referanse)	256
Figur 9-21 Andel av form innen UB-kontekst	261
Figur 9-22 Andel av kontekst innen ubest. ent/flt	261
Figur 9-23 Andel av form innen BE-kontekst	264
Figur 9-24 Andel av kontekst innen bestemt form	264
Figur 9-25 Andel av form innen ART-kontekst	267
Figur 9-26 Andel av kontekst innen artikkelfraser	267
Figur 9-27 Antall tekster med tre, to eller én konteksttype (indefinit referanse)	272
Figur 9-28 Profiltyper i de tre gruppene (indefinit referanse)	274
Figur 9-29 Andel av de ulike formtypene i de tre gruppene (definit referanse)	285
Figur 9-30 Andel av de ulike konteksttypene i de tre gruppene (definit referanse)	285
Figur 9-31 Andel av form innen UB-kontekst	290
Figur 9-32 Andel av kontekst innen ubest. ent/flt	290
Figur 9-33 Andel av form innen BE-kontekst	293
Figur 9-34 Andel av kontekst innen bestemt form	293
Figur 9-35 Andel av form innen ART-kontekst	295
Figur 9-36 Andel kontekst innen artikkelfraser	295
Figur 9-37 Antall tekster med tre, to eller én konteksttype (definit referanse)	297
Figur 9-38 Profiltyper i de tre gruppene (definit referanse)	298
Figur 9-39 Andel av de ulike formtypene i de tre gruppene (generisk referanse)	307
Figur 9-40 Andel av de ulike konteksttypene i de tre gruppene (generisk referanse)	307
Figur 9-41 Andel av form innen UB-kontekst	312
Figur 9-42 Andel av kontekst innen ubest. ent/flt	312

Figur 9-43 Andel av form innen BE-kontekst	315
Figur 9-44 Andel av kontekst innen bestemt form	315
Figur 9-45 Andel av form innen ART-kontekst	317
Figur 9-46 Andel av kontekst innen artikkelfraser	318
Figur 9-47 Antall tekster med tre, to eller én konteksttype (generisk referanse)	319
Figur 9-48 Profiltyper i de tre gruppene (generisk referanse)	320
Figur 9-49 Andel av form innen kontekst etter referensiell type i den polske gruppen	325
Figur 9-50 Andel av kontekst innen form etter referensiell type i den polske gruppen	326
Figur 10-1 Andel av subjekt og objekt i de tre gruppene	347
Figur 10-2 Andel ubestemt og bestemt form i de tre gruppene	347
Figur 10-3 Andel form innen kontekst	349
Figur 10-4 Andel kontekst innen form	349
Figur 10-5 Andel av form innen subjekter i gruppene	351
Figur 10-6 Andel av kontekst innen ubestemt form i gruppene	351
Figur 10-7 Andel av form innen objekter i gruppene	351
Figur 10-8 Andel av kontekst innen bestemt form i gruppene	351
Figur 10-9: Antall tekster som har subjekt og objekt	353
Figur 10-10 Profiltyper i de tre gruppene	354
Figur 10-11 Andel av subjekt og objekt i de tre gruppene (indefinit referanse)	359
Figur 10-12 Andel av ubestemt og bestemt form i de tre gruppene (indefinit referanse)	359
Figur 10-13 Andel av form innen subjekter i gruppene	363
Figur 10-14 Andel av kontekst innen ubest. former i gruppene	363
Figur 10-15 Andel av form innen objekter i gruppene	363
Figur 10-16 Andel av kontekst innen bestemt form i gruppene	363
Figur 10-17 Antall tekster som har subjekt og objekt (indefinit referanse)	367
Figur 10-18 Profiltyper i de tre gruppene (indefinit referanse)	367
Figur 10-19 Andel av subjekt og objekt i de tre gruppene (definit referanse)	372
Figur 10-20 Andel av ubestemt og bestemt form i de tre gruppene (definit referanse)	372
Figur 10-21 Andel av form innen subjekter i gruppene	376
Figur 10-22 Andel av kontekst innen ubestemt form i gruppene	376
Figur 10-23 Andel av form innen objekter i gruppene	376
Figur 10-24 Andel av kontekst innen bestemt form i gruppene	376
Figur 10-25 Antall tekster som har subjekt og objekt (definit referanse)	379
Figur 10-26 Profiltyper i de tre gruppene (definit referanse)	380
Figur 10-27 Andel av subjekt og objekt i de tre gruppene (generisk referanse)	385
Figur 10-28 Andel av ubestemt og bestemt form i de tre gruppene (generisk referanse)	385
Figur 10-29 Andel av form innen subjekter i gruppene	390
Figur 10-30 Andel av kontekst innen ubestemt form i gruppene	390
Figur 10-31 Andel av form innen objekter i gruppene	390
Figur 10-32 Andel av kontekst innen bestemt form i gruppene	390
Figur 10-33 Antall tekster som har subjekt og objekt (generisk referanse)	392
Figur 10-34 Profiltyper i de tre gruppene (generisk referanse)	392
Figur 10-35 Andel av form innen kontekst etter referensiell type i den polske gruppen	396
Figur 10-36 Andel av kontekst innen form etter referensiell type i den polske gruppen	396

LISTE OVER FORKORTELSER

AKK	Akkusativ
art.	Artikkelfrase (ubestemt artikkel + substantiv i ubestemt entall)
ART	Se ART-kontekst
ART-kontekst	Kontekst som er predikert å ha artikkelfrase (hypotese 1)
ASK	Norsk andrespråskorpus
be.	Bestemt form
BE	Se BE-kontekst
BE-kontekst	Kontekst som er predikert å ha bestemt entall/flertall (hypotese 1)
DAT	Dativ
def.	Definit referanse
Eng.	Engelsk
GB	Grammatisk bestemthet
+GB	Som har grammatisk bestemthet
-GB	Som ikke har grammatisk bestemthet
gen.	Generisk referanse
GEN	Genitiv
HK	Assumed Hearer Knowledge
HN el. HN-test	Test i norsk – høyere nivå
ind.	Indefinit referanse
INSTR	Instrumentalis
NOM	Nominativ
Nor.	Norsk
obj.	Objekt
OBL	Obligatorisk kontekst
Pol.	Polsk
S1	Morsmål/førstespråk
S2	Målspråk/andrespråk
SP el. SP-test	Språkprøven i norsk for voksne innvandrere
SR	Spesifikk referanse
St.avvik	Standardavvik
subj.	Subjekt
ub.	Ubestemt form
UB	Se UB-kontekst
UB-kontekst	Kontekst som er predikert å ha ubestemt entall/flertall (hypotese 1)

1 INNLEDNING

I denne avhandlingen tar jeg for meg bruk av bestemte og ubestemte former av substantivet i norskinnlæreres skriftlige tekster. Jeg undersøker de språklige valgene til andrespråksinnlærere som ikke kjenner den grammatiske kategorien bestemthet fra morsmålet, for å se om valg av form på substantivet er påvirket av språklige forhold i morsmålet. Valget av innlærergruppe er basert på at disse innlærerne har et morsmål som mangler en grammatisk kategori som er obligatorisk i norsk, altså det som kan kalles en *nullkontrast*.¹

Studien i avhandlingen tar for seg både et potensielt *kvantitativt* utslag av en slik morsmålspåvirkning, altså *hvor ofte* bestemthet markeres i innlærerspråket i språklige kontekster der det er obligatorisk i norsk, og et potensielt *kvalitativt* utslag, nemlig om morsmålet påvirker valget mellom ulike former av substantivet i ulike språklige kontekster i innlærerspråket. Det kvantitative utslaget blir undersøkt i et større datamateriale av tekster skrevet av innlærere med tre ulike morsmål uten grammatisk bestemthet, mens det kvalitative utslaget blir undersøkt i et mindre materiale med tekster skrevet av polske innlærere. Det er dermed to overordnede forskningsspørsmål i studien:

Forskningsspørsmål 1:

Påvirker morsmålet hvor ofte bestemte former brukes i obligatoriske kontekster i innlærerspråket til innlærere uten grammatisk bestemthet i morsmålet?

Forskningsspørsmål 2:

Påvirker morsmålet distribusjonen av bestemte og ubestemte former i innlærerspråket til polske innlærere?

Disse to forskningsspørsmålene blir forsøkt besvart gjennom undersøkelser av substantivformer i tekster hentet fra Norsk andrespråskorpus (ASK), et elektronisk innlærerkorpus utviklet ved Universitetet i Bergen. I undersøkelsene tar jeg i bruk et rammeverk for studier av morsmålspåvirkning foreslått av Jarvis (2000, 2010), som blant annet innebærer en sammenligning av innlærere *med* og *uten* grammatisk bestemthet i morsmålet for å avdekke om språklige mønstre i innlærerspråket kan

¹ Dette omtales også av Stockwell, Bowen og Martin (1965) som *nyhet*, men jeg har valgt å bruke termen *nullkontrast*, som er i tråd med Berggreen og Tenfjord (1999: 203).

forklares med påvirkning fra morsmålet. Datamaterialet mitt består derfor ikke bare av tekster skrevet av innlærere uten bestemthet i morsmålet, men også av tilsvarende mengde tekster skrevet av innlærere som kjenner denne kategorien fra morsmålet.

Jeg vil videre i dette innledende kapittelet kort beskrive bakgrunnen for studien, datamaterialet og de ulike delundersøkelsene, samt sentrale sider av den metodiske tilnærmingen.

1.1 ASKeladden-prosjektet

Studien som presenteres i denne avhandlingen, har vært en del av et større forskningsprosjekt, *ASKeladden*, som ble gjennomført ved Institutt for lingvistiske, litterære og estetiske studier ved Universitetet i Bergen. Alle delprosjektene i ASKeladden-prosjektet har hatt til felles at de har 1) brukt ASK som datamateriale, og 2) undersøkt fenomenet morsmålpåvirkning hos innlærere av norsk.

ASKeladden-prosjektet var regnet som en oppfølging av ASK-prosjektet, som utviklet selve korpuset. Fordi jeg selv har vært med på arbeidet med å utvikle ASK siden tidlig på 2000-tallet, og senere også har drevet med opplæring i bruk av korpuset, har denne typen arbeid med ASK også vært en viktig del av doktorgradsprosjektet mitt, i tillegg til at korpuset utgjør datamaterialet for avhandlingens undersøkelser.

1.2 Den teoretiske bakgrunnen for studien

Denne avhandlingen tar for seg to relativt velstuderte fenomener innen internasjonal andrespråksforskning: *morsmålpåvirkning* (transfer) og *grammatisk bestemthet*. Den teoretiske bakgrunnen for undersøkelsene mine er derfor knyttet til forhold som har med disse to fenomenene å gjøre. Det dreier seg for det første om allmennlingvistisk kunnskap om grammatisk bestemthet, for det andre tidligere forskning på bestemthet i innlærerspråk, for det tredje kontrastive sammenligninger mellom de ulike språkene som er involvert i studien, og for det fjerde kunnskap om fenomenet morsmålpåvirkning og hvordan innlærerspråk kan påvirkes av likheter og forskjeller mellom morsmålet og språket som skal læres (målspråket). Jeg vil her kort presentere de disse fire områdene, som alle også blir nærmere beskrevet i ulike kapitler i avhandlingen.

1.2.1 Bestemthet som grammatisk kategori

Bestemthet som grammatisk kategori fins ikke i alle verdens språk, og for de språkene som har denne kategorien, kan innholdet variere noe. Felles er likevel at kategorien er knyttet til begrepet *referanse*, som fins i alle språk og innebærer bruk av referensielle uttrykk, ofte substantivfraser eller pronomen, for å vise til referenter (J. Lyons 1977: 177). Grammatisk bestemthet kan ses på som en grammatikalisering av referanse, der man ved hjelp av obligatoriske språklige markører uttrykker en referents *identifiserbarhet*, altså om man antar at referenten man viser til, er identifiserbar eller ikke identifiserbar for den man snakker eller skriver til (C. Lyons 1999: 6; Hagen 1998: 54). Når referenten er identifiserbar, foreligger *definit referanse*, som typisk markeres med bestemte former, mens når referenten ikke er identifiserbar, er det snakk om *indefinit referanse*, som typisk uttrykkes med ubestemte former.

Bruken av bestemte og ubestemte former i norsk er i stor grad styrt av om det foreligger definit eller indefinit referanse, men formene brukes også ved referanse til *typer* av referenter, altså *generisk referanse*, og de kan brukes *ikke-referensielt*, som innebærer at substantivet først og fremst *beskriver* en referent, heller enn å referere til den (Faarlund, Lie & Vannebo 1997: 293-294; Hagen 1998: 51). I tillegg til disse fire referensielle forholdene er det dessuten syntaktiske faktorer involvert i valg av form på substantiv, for eksempel tilstedeværelse av determinativer i substantivfrasen (Faarlund et al. 1997: 296; Hagen 1998: 54). Grammatisk bestemthet er på denne måten en heterogen kategori i norsk, ettersom det er flere forhold som styrer om et substantiv skal ha bestemt eller ubestemt form i en gitt språklig kontekst. Dette vil jeg se nærmere på i kapittel 7.

I språk som *ikke* har grammatisk bestemthet, er de nevnte referensielle forholdene ofte underforstått heller enn eksplisitt uttrykt med obligatoriske språklige markører. Substantivfrasen har i disse språkene ikke noen spesiell markør som sier om en referent er identifiserbar eller ikke, men det fins likevel andre språklige midler som man mener legger til rette for at substantivfrasene kan bli forstått som definite eller indefinite (Szwedek 1976: 203). Et eksempel på dette er *pragmatisk ordstilling*, altså en inndeling av setningen i tematiske og rematiske ledd, der tematiske ledd typisk viser til *kjent* informasjon og er plassert tidlig i setningen, mens rematiske ledd viser til *ny* informasjon

og plasseres mot slutten av setningen (Szwedek 1976: 203; Firbas 1992: 118). En substantivfrases plassering i setningen vil dermed signalisere om referenten skal forstås som kjent og identifiserbar, eller ny og ikke-identifiserbar.

Jeg vil i avhandlingen omtale språk med grammatisk bestemthet (GB) som +GB-språk, mens språk *uten* grammatisk bestemthet vil bli omtalt som -GB-språk. Innlærere med disse språkene som morsmål vil på samme måte bli omtalt som +GB-innlærere og -GB-innlærere.² Kategorien grammatisk bestemthet, referanse og bestemthetsekvivalenter vil bli nærmere omtalt i kapittel 3.

1.2.2 Bestemthet innen andrespråksforskningen

Innen andrespråksforskning er grammatisk bestemthet et utbredt studieobjekt nettopp fordi bruk av nominale bestemthetsmarkører har vist seg å være spesielt vanskelig å lære for andrespråksinnlærere, framfor alt gjelder dette innlærere med morsmål uten bestemthet, for eksempel polsk, tsjekkisk, finsk, russisk, thai og koreansk. Det er veletablert kunnskap at læring av bestemthet utgjør en utfordring for akkurat -GB-innlærere, men det er mer uavklart *hvorfor* det er vanskelig og hvilke forhold som styrer utelatelse og bruk av bestemthetsmarkører i innlærerspråket til disse innlærerne.

To hovedtyper av forklaringer kan trekkes fram fra ulike studier av bestemthetsmarkører innen engelsk som andrespråk: På den ene siden har man tilnærminger som har som utgangspunkt at -GB-innlærere baserer de språklige valgene sine på *språklige universalier* (Chaudron & Parker 1990; Ionin & Wexler 2003; Ionin, Ko & Wexler 2004; Ionin, Zubizarreta og Maldonado 2008). Mangel på erfaring med grammatisk bestemthet fra morsmålet antas nemlig å føre til at innlærere heller forstår denne kategorien gjennom språklige universalier som er nært knyttet til bestemthet, for eksempel identifiserbarhet, topikalitet og spesifisitet, og det er slike trekk som antas å styre bruk og utelatelse av bestemthetsmarkører i innlærerspråk. På den andre siden har man tilnærminger som antar at innlæreres språklige valg kan være påvirket av forhold i

² Denne gruppen av innlærere omtales gjerne som '-ART-learners' i den engelskspråklige litteraturen fordi det gjerne er fokus på om morsmålet har eller ikke har bestemte og ubestemte *artikler*. I denne avhandlingen har jeg valgt å kalle dem '-GB-innlærere', som tar utgangspunkt i den grammatiske kategorien heller enn i den språklige markøren, noe som virker mer naturlig siden man i norsk uttrykker grammatisk bestemthet med både artikler og bøyning.

morsmålet til tross for at morsmålet ikke har grammatisk bestemthet. Det kan for eksempel være morsmålspåvirkning gjennom bruk av pragmatisk ordstilling (Young 1996), bruk av demonstrativer (Young 1996; Ekiert 2007, 2010) eller en preferanse for naken form av substantivet (Trenkic 2009; Trenkic & Pongpaibroj 2013). Noen studier kombinerer dessuten universalier og morsmålspåvirkning som forklaringsfaktor (Young 1996; Jarvis 2002), og Young mener at disse to faktorene er virksomme på ulike tidspunkt i andrespråkutviklingen (Young 1996: 172). Disse ulike tilnærmingene til grammatisk bestemthet i innlærerspråk vil jeg komme tilbake til i kapittel 4, hvor jeg presenterer et utvalg av tidligere studier som undersøker bruk av bestemthetsmarkører hos andrespråksinnlærere.

Den foreliggende avhandlingen retter seg først og fremst mot morsmålets rolle for bruk av bestemthetsmarkører i innlærerspråk. Forskningsspørsmål 1 har som mål å avdekke om -GB-innlærerne som er representert i ASK-korpuset, har større utfordringer med bruk av bestemt form enn +GB-innlærerne, mens forskningsspørsmål 2 undersøker om det er forhold i morsmålet som påvirker hvilke språklige kontekster de ulike bestemte og ubestemte formene brukes i. Når det gjelder tidligere forskning, er det derfor først og fremst studiene med morsmålstilnærming som er relevante. Spesielt Young (1996), som framhever hvordan forholdet mellom et substantivs identifiserbarhet og dets plassering i setningen kan få markering av bestemthet til å virke overflødig (redundant) i visse språklige kontekster, har vært relevant for min egen studie.

1.2.3 Morsmålspåvirkning

Fenomenet *morsmålspåvirkning*, som jeg her også vil vise til som *transfer* eller *overføring*, innebærer at morsmålet til en andrespråksinnlærer på en eller annen måte påvirker læring og bruk av et andrespråk. Jarvis og Pavlenko (2008: 1) definerer transfer som den påvirkningen kunnskap om ett språk har på en persons kunnskap om eller bruk av et annet språk: «[...] the influence of a person's knowledge of one language on that person's knowledge or use of another language [...]».

Transfer er nært knyttet til, og ofte et resultat av, de kontrastive relasjonene som foreligger mellom morsmålet og målspråket, altså språklige forskjeller og likheter. Det er av flere framhevet at det er innlærernes *subjektive* oppfatninger av likheter og forskjeller mellom språkene som er utløsende for transfer, og en kan dele inn i tre ulike

kontrastive antakelser som innlærere kan ha om forholdet mellom morsmål og målspråk: *likhetsrelasjoner*, *ulikhetsrelasjoner* og *nullrelasjoner* (Ringbom 2007: 5; Ringbom & Jarvis 2009: 109-110).

Alle disse tre kontrastive relasjonene antas å kunne føre til transfer, men kanskje først og fremst *likhetsrelasjoner*, som innebærer at innlærere oppfatter strukturer i de to språkene som like og dermed overførbare fra morsmålet til målspråket. Dette kalles for en *positiv tverrspråklig identifikasjon*, men transfer kan også være resultat av en *negativ tverrspråklig identifikasjon*: Innlærere oppfatter noe som *ulikt* i de to språkene og unngår dermed å overføre trekk fra morsmålet til målspråket, noe som omtales som en *ulikhetsrelasjon*. Både likhetsrelasjoner og ulikhetsrelasjoner innebærer altså det som kalles *tverrspråklig identifikasjon* (Weinreich 1963).³ I tilfeller der forskjellene mellom morsmål og målspråk er veldig store eller likhetene er for abstrakte, vil det ikke være grunnlag for en slik tverrspråklig identifikasjon, og man snakker da om *nullrelasjoner*. Konsekvensen av en nullrelasjon kan være at innlærerne sitter med en ubevisst og umotivert antakelse om at målspråket fungerer på samme måte som morsmålet, en antakelse som ikke er basert på observasjoner i målspråksinnputt, men heller på mangelen av slike observasjoner (Kellerman 1995: 137; Ringbom & Jarvis 2009: 108).

Når det gjelder innlæreres antakelser om kontrastive relasjoner, kan en dermed skille mellom *oppfattede* og *antatte* likheter og forskjeller (Jarvis & Pavlenko 2008: 179).⁴ Mens likhetsrelasjoner og ulikhetsrelasjoner handler om oppfattede likheter, motivert av bevisste eller ubevisste observasjoner av målspråksinnputt, er nullrelasjoner kun *antatte* likheter som kun er motivert av en generell forestilling om at målspråket fungerer på samme måte som morsmålet.

Som nevnt over, er det innlæreres subjektive antakelser om de kontrastive forholdene mellom morsmål og målspråk som antas å føre til transfer, noe Kellerman (1983) kalte *psykotypologi*. Det er derfor nødvendig å skille de subjektive, psykotypologiske likhetene og forskjellene fra de mer *objektive* forskjellene og likhetene som kan beskrives med utgangspunkt i lingvistisk fagkunnskap om språkene (Jarvis & Pavlenko 2008: 177). Det er psykotypologi som er drivkraften bak transfer,

³ Eng. 'cross linguistic identification'

⁴ Eng. 'Perceived similarities and differences' og 'assumed similarities and differences'

mens objektive forskjeller og likheter først og fremst er egnet til å *beskrive* ulike tilfeller av transfer (Jarvis & Pavlenko 2008: 179). I denne avhandlingen har jeg i tråd med dette valgt å skille mellom begrepene *nullkontrast* og *nullrelasjon*: «Nullkontrast» viser til en lingvistisk beskrivelse av forskjellen mellom morsmål og målspråk, i dette tilfellet at morsmålet mangler en grammatisk kategori som fins i målspråket, mens «nullrelasjon» derimot viser til innlæreres subjektive antakelser om forskjeller og likheter mellom språkene. Skillet er altså motivert av forskjellen mellom objektive beskrivelser og psykotypologi, og i tråd med Ringbom og Jarvis (2009: 109) er nullrelasjoner da definert som subjektive antakelser.

Om vi vender tilbake til avhandlingens spørsmål om bruk og ikke bruk av bestemthetsmarkører, er de tre typene kontrastive antakelser om forholdet mellom morsmål og målspråk relevante. Mellom morsmål (-GB) og målspråket norsk (+GB) foreligger det en objektiv nullkontrast, der målspråket har en kategori som morsmålet ikke har. Som nevnt i omtalene av andrespråksstudier over, har enkelte studier av bestemthet i innlærerspråk hatt som utgangspunkt at transfer *ikke* kan finne sted når det foreligger en slik nullkontrast, siden det ikke antas å være noe å overføre fra morsmålet. Som nevnt fins det også en rekke studier som viser hvordan transfer likevel kan foreligge ved nullkontrast, og dette vil også være mitt utgangspunkt. Forskjellene mellom et -GB-morsmål og et +GB-målspråk kan imidlertid ses på som for store til å fremme tverrspråklig identifikasjon, noe som gjør det sannsynlig at innlærere oppfatter forholdet mellom morsmål og målspråk som en nullrelasjon, som igjen kan ha ulike utfall, hvorav ett er en umotivert antakelse om at målspråket fungerer på samme måte som morsmålet (Ringbom & Jarvis 2009: 108).

I denne avhandlingens studie er det først og fremst nullrelasjonen som er sentral, da begge hypotesene som stilles for forskningsspørsmål 2, tar utgangspunkt i transfer uten opphav i tverrspråklig identifikasjon, noe jeg kommer nærmere inn på under. Dette er imidlertid et vanskelig terreng: Kan man skille entydig mellom transfer som er resultat av en umotivert antakelse om at morsmål og målspråk er like, og transfer som er resultat av tverrspråklig identifikasjon? Dette er et spørsmål som vil bli diskutert i avhandlingens siste kapittel (11.5.2).

1.2.4 Norsk versus engelsk som målspråk

I norsk sammenheng har det vært relativt få større studier som ser på bruk av bestemte og ubestemte former i innlærerspråk, men for noen år siden kom en omfattende doktoravhandling om bestemthet (Nordanger 2017), og i tillegg er Nistovs (2001) doktoravhandling relevant, selv om denne ikke bare ser på substantiv, men på valg av referensielle uttrykk generelt. Det fins også noen masteroppgaver og artikler (Jin, Åfarli & van Drommelen 2009; Nordanger 2009; Paulsen 2009; Soboleva 2016; Moraczewska 2017; Premovic 2018). Det fins dessuten et par doktorgradsavhandlinger innen *svensk* som andrespråk, som kan ha overføringsverdi til norsk kontekst (Axelsson 1994; Nyqvist 2013).

De aller fleste studier på dette området er likevel innen engelsk som andrespråk, og det er naturlig å stille seg spørsmålet om hvorvidt hypoteser og funn fra studier med *engelsk* som målspråk er overførbare til studier med *norsk* som målspråk, ettersom det er formelle forskjeller i hvordan bestemthetskategorien kommer til uttrykk i disse to språkene: Engelsk uttrykker bestemthet med bestemt og ubestemt artikkel (*the* og *a*), mens norsk bruker en kombinasjon av ubestemt artikkel og bøyninger. Dette er en problemstilling jeg kommer tilbake til i kapittel 4, hvor jeg vurderer ulike andrespråksstudier med engelsk som målspråk med tanke på om hypoteser og funn er anvendelig i en studie hvor norsk er målspråk.

Studier som har skandinaviske språk heller enn engelsk som målspråk, kan for øvrig være et interessant bidrag til forståelsen av hvordan bestemthet læres ettersom de skandinaviske språkene skiller seg fra engelsk på flere måter, selv om de er like med tanke på å ha kategorien grammatisk bestemthet. Studier av bestemthet der norsk er målspråk, kan altså ikke bare bidra generelt med ny kunnskap om hvordan morsmålet eller andre faktorer påvirker læring og bruk av bestemthet, men slike studier kan også fungere som en innsiktsgivende kontrast til studier med engelsk som målspråk, der en for eksempel kan finne ut mer om hvilken rolle *måten* bestemthet uttrykkes på i målspråket, spiller. Både Axelsson (1994) og Nordanger (2017) har berørt denne problematikken og begge studiene har funn som tyder på at målspråkets uttrykksmåte kan være relevant for hvilke utfordringer innlærere møter når de lærer svensk og norsk,

og dessuten at dette avhenger av trekk ved innlærernes morsmål, noe jeg vil komme inn på i 4.4.

1.3 Empiriske undersøkelser

Formålet med studien i denne avhandlingen har vært å undersøke morsmålpåvirkning i bruken av bestemte og ubestemte former hos norskkinnlærere som ikke kjenner bestemthet fra morsmålet, og dette undersøkes både med tanke på hvor ofte substantiv som krever bestemt form, faktisk har det, og med tanke på hvilke språklige forhold i morsmålet som kan tenkes å ligge bak distribusjonen av bestemte og ubestemte former i innlærerspråket. Den empiriske undersøkelsen er derfor delt i to med utgangspunkt i disse to problemstillingene: Undersøkelse 1, som tar for seg det kvantitative utslaget av transfer, og undersøkelse 2, som ser på det mer kvalitative utslaget transfer kan ha på formvalg i ulike språklige kontekster.

Undersøkelse 1 og 2 har ulike formål, men er også ulike med hensyn til datamaterialet som er brukt. Undersøkelse 1 tar for seg et utvalg på 600 tekster skrevet av -GB-innlærere med tre ulike morsmålsbakgrunner (polsk, russisk og bosnisk-kroatisk-serbisk), mens undersøkelse 2 bare ser på 41 tekster skrevet av innlærere med polsk som morsmål (41 tekster). I begge undersøkelsene sammenligner jeg -GB-innlærerne med tilsvarende antall tekster skrevet av +GB-innlærere; engelske, nederlandske og tyske i undersøkelse 1, og engelske innlærere i undersøkelse 2. Undersøkelse 2 har i tillegg et kontrollkorpus med tekster skrevet av personer med norsk som morsmål. Korpuset i undersøkelse 1 har totalt 1200 tekster, mens det for undersøkelse 2 er snakk om 123 tekster. Bruk av kontrollgrupper med +GB-innlærere og morsmålsbrukere er motivert av metode for studier av transfer, nærmere bestemt påvisning av *transfereffekter* (Jarvis 2000), noe jeg kommer tilbake til under (1.4.2).

Forskjellen i størrelsen til de to delkorpuserne for undersøkelse 1 og 2, gjør at det er ulike muligheter for å bruke statistikk. I undersøkelse 1 har jeg brukt statistiske tester for å teste forskjeller mellom ulike innlærergrupper, mens undersøkelse 2 har for få substantivfraser per tekst til at jeg i særlig grad har kunnet bruke statistiske tester for å si om forskjeller mellom gruppene er statistisk signifikante eller ikke. At materialet er mindre i undersøkelse 2, har derimot gjort det mulig å legge til manuell annotasjon av ytterligere lingvistisk informasjon enn det som allerede er tilgjengelig i ASK, noe som

utvider muligheten for å undersøke forholdet mellom former og språklige kontekster, ettersom annotasjonen kan tilpasses de hypotesene som stilles.

Jeg vil videre kort beskrive hovedinnholdet i de to undersøkelsene.

1.3.1 Undersøkelse 1

Undersøkelse 1 knytter seg til forskningsspørsmål 1, og har som mål å se på *kvantitative utslag* av morsmålpåvirkning ved å undersøke andelen av substantivfraser som har bestemt form i obligatoriske kontekster for bestemt form i norsk. Utgangspunktet er den utbredte oppfatningen om at mangel på en grammatisk kategori i morsmålet vil føre til læringsutfordringer og dermed også feil i bruken av bestemthetsmarkører.

I denne delundersøkelsen ser jeg på et utvalg av 1200 tekster skrevet av innlærere med seks ulike morsmål, og disse seks er som nevnt over, fordelt på to grupper etter om de har grammatisk bestemthet i morsmålet eller ikke: -GB-gruppen består av innlærere med polsk, russisk og bosnisk-kroatisk-serbisk som morsmål, mens +GB-gruppen består av innlærere med engelsk, tysk og nederlandsk som morsmål. Prediksjonen er at -GB-gruppen sjeldnere vil markere substantivene med bestemt form enn +GB-gruppen, og denne forskjellen er ventet å gi seg utslag både gjennom at -GB-innlærere generelt sett har lavere markeringsprosent i obligatoriske kontekster enn +GB-innlærerne, og gjennom at -GB-gruppen har færre innlærere med 100 % markering enn +GB-gruppen. Metode, datamateriale og resultater fra undersøkelse 1 presenteres samlet i kapittel 6.

1.3.2 Undersøkelse 2

Undersøkelse 2 knytter seg til forskningsspørsmål 2, og tar for seg et mer begrenset utvalg av tekster da jeg kun ser på tekster skrevet av polske innlærere (-GB) sammenlignet med engelske innlærere (+GB), i tillegg til tekster skrevet av personer med norsk som morsmål. Det er 41 tekster fra hver gruppe, til sammen 123 tekster. Til tross for at det er snakk om engelskspråklige og polskspråklige innlærere med ulike nasjonaliteter, vil jeg for enkelthets skyld omtale disse som «polske» og «engelske», da i betydningen *innlærere med polsk eller engelsk som morsmål*.

Målet med undersøkelse 2 er å komme nærmere det jeg over omtalte som *kvalitative utslag av morsmålpåvirkning*, det vil si der noe i morsmålet påvirker

distribusjonen av bestemte og ubestemte former i innlærerspråket, altså hvilke kontekster hvilke former er brukt i, heller enn hvor ofte formene brukes. I motsetning til undersøkelse 1, ser jeg her på bruk av både bestemte og ubestemte former.

En annen viktig forskjell fra undersøkelse 1 er at jeg bare ser på *enkle substantivfraser*, altså substantivfraser som ikke har modifierende adledd, men kun selve substantivet med eller uten en bestemthetsmarkør (C. Lyons 1999: 2). Jeg har dermed kun hentet ut fem typer substantivfraser fra korpuset, fraser som består av enten: 1) ubestemt entallssubstantiv 2) artikkel + ubestemt entallssubstantiv, 3) ubestemt flertallssubstantiv 4) bestemt entallssubstantiv, og 5) bestemt flertallssubstantiv. Denne avgrensningen er først og fremst motivert av at forholdet mellom form og innhold er mer direkte i enkle substantivfraser enn i substantivfraser som inneholder modifierende adledd.

Forskningsspørsmål 2 undersøkes gjennom to ulike hypoteser om hvordan polsk som morsmål kan tenkes å påvirke distribusjonen av bestemte og ubestemte former i innlærerspråket. Hypotesene er formulert ut fra kunnskap om forholdet mellom transfer og tverrspråklige kontraster mellom morsmål og målspråk, men er også delvis basert på tidligere undersøkelser av bestemthet hos innlærere. Den ene hypotesen har som utgangspunkt at polske innlærere overfører pragmatisk ordstilling som signal om referenters identifiserbarhet og dermed også en oppfatning om at markering av identifiserbarhet er unødvendig så lenge pragmatisk ordstilling brukes. Den andre hypotesen handler om overføring av innholdsdistinksjoner knyttet til grammatisk kasus.

Begge disse forholdene ventes å ha konsekvenser for hvordan bestemte og ubestemte former brukes i innlærerspråket, og det er ikke pragmatisk ordstilling og kasusmorfologi i seg selv som undersøkes, men hvordan bestemthetsmarkørene i norsk blir tatt i bruk for å uttrykke forhold som er relatert til dette. Hypotesene bygger nemlig på det som blant annet går under navnet *transfer to nowhere* (Kellerman 1995), som kort sagt viser til at når det ikke skjer noen tverrspråklig identifikasjon mellom morsmål og målspråk, så kan språklige innholdselementer som er sentrale i morsmålet likevel bli overført til innlærerspråket og komme til uttrykk gjennom de språklige midlene som er tilgjengelig i målspråket (Kellerman 1995: 141). Dette vil jeg komme tilbake til i kapittel 8. Hypotesene vil bli omtalt som *hypotese 1* og *hypotese 2*, og er gjengitt under.

Hypotese 1:

De polske innlærerne vil overføre morsmålets bruk av pragmatisk ordstilling som signal om tekstuell identifiserbarhet og en oppfatning om at markering av identifiserbarhet er overflødig når pragmatisk ordstilling brukes. De norske bestemthetsmarkørene vil bli tatt i bruk når identifiserbarhet ikke går fram av ordstillingen.

Hypotese 2:

De polske innlærerne vil overføre kasus fra morsmålet, og norsk substantivmorfologi vil bli tatt i bruk for å uttrykke kasusdistinksjoner.

Hypotese 1 handler altså om at polske innlærere overfører pragmatisk ordstilling som signal om identifiserbarhet til innlærerspråket. Når pragmatisk ordstilling får denne rollen i innlærerspråket, vil bruk av bestemthetsmarkører oppfattes som overflødige (*redundante*) når identifiserbarheten allerede går fram av ordstillingen, og de vil dermed ikke brukes. Dette gjør at bestemthetsbøyning eller ubestemt artikkel ikke vil brukes når substantivfrasen er plassert i tråd med tema/rema-prinsippet, men at de brukes når ordstillingen *bryter* med dette prinsippet, altså når tematiske fraser står i slutten av setningen eller når rematiske fraser står i begynnelsen av setningen.

Bestemthetsmarkørers redundans er et tema som har vært brakt fram i flere studier av engelsk som andrespråk, men med et noe ulikt innhold (Young 1996; Jarvis 2002; Ekiert 2010; Trenkic 2009; Trenkic & Pongpairoj 2013). Jeg har i min egen studie valgt å legge meg nært opp til Young (1996), som ser på hvordan samspillet mellom identifiserbarhet og setningsplassering bidrar til at bestemthetsmarkører brukes eller ikke brukes i innlærerspråket. Dette står i motsetning til Trenkic (2009) og Trenkic og Pongpairoj (2013) som mener at redundans kun kan forklare at bestemthetsmarkører utelates, og ikke vekslingen mellom bruk og ikke bruk.

Hypotese 2 er knyttet til at man i polsk uttrykker grammatisk kasus gjennom substantivmorfologi, og at dette overføres til innlærerspråket. Hypotesen bygger på at grammatikalisererte kategorier er obligatoriske og automatiserte, og dermed spesielt vanskelige å endre eller legge fra seg når man formulerer seg på et annet språk enn morsmålet. På grunn av dette vil et behov for å uttrykke kasus overføres til innlærerspråket, og norsk substantivmorfologi vil bli brukt for å uttrykke de relevante kasusdistinksjonene. I undersøkelsen av hypotese 2 vil jeg rette fokus mot skillet mellom

nominativ og de øvrige kasusene (ikke-nominativ), og antakelsen er at ubestemte former vil bli tatt i bruk for å uttrykke nominativ, mens bestemte former vil brukes for ikke-nominativ.

Bakgrunnen for de to hypotesene, i tillegg til metode og datamateriale for undersøkelse 2, presenteres i kapittel 8. Deretter presenteres resultatene for hypotese 1 i kapittel 9 og resultatene for hypotese 2 i kapittel 10.

1.4 Datamateriale og sentrale metodiske sider ved studien

1.4.1 Datamateriale

Datamaterialet i studien er hentet fra Norsk andrespråskorpus (ASK), et elektronisk innlærerkorpus med tekster skrevet av norskinnlærere. Korpuset består av nærmere 2000 tekster skrevet av innlærere med 10 ulike morsmålsbakgrunner, i tillegg til et kontrollkorpus skrevet av personer med norsk som morsmål. Tekstene er langvarsoppgaver fra to offisielle tekster i norsk: *Språkprøven i norsk for voksne innvandrere* og *Test i norsk – høyere nivå*. Korpuset tilbyr gode muligheter for å søke fram akkurat de språklige strukturene jeg er ute etter, i tillegg til relevant grammatisk informasjon og mulighet til å se substantivfrasene i hele sin tekstlige sammenheng. I tillegg har jeg kunnet legge inn egne annoteringer for å klassifisere substantivfrasene etter de forholdene som hypotesene i undersøkelse 2 krever, der dette ikke allerede er tilgjengelig gjennom ASK. De to ulike delkorpusene som er brukt i undersøkelse 1 og 2, vil bli beskrevet med tanke på utvalg og bearbeiding i henholdsvis kapittel 6 og 8.

1.4.2 Påvisning av transfereffekter

Studien har som mål å ta i bruk en metode for å avdekke transfer som først ble foreslått av Jarvis (2000, 2010). I tillegg til å understreke viktigheten av en teorinøytral definisjon av transfer og utelukkning av andre påvirkningsfaktorer enn morsmålet, beskriver Jarvis fire såkalte *transfereffekter* som han mener bør påvises for å kunne konkludere om at noe faktisk er transfer (Jarvis 2000: 253-354). Disse fire effektene vil jeg omtale som *intragruppe-homogenitet*, *intergruppe-heterogenitet*, *morsmålssamsvar* og *intralingval kontrast*, som er oversettelser av Jarvis' termer *intragroup homogeneity*, *intergroup*

heterogeneity, cross-language congruity og intralingual contrast (Jarvis 2010).⁵ For å undersøke disse effektene trenger man ikke bare språklig materiale fra innlærergruppen man vil studere, men også en kontrollgruppe med innlærere som har et annet morsmål, noe jeg har både i undersøkelse 1 og 2. I tillegg til dette har jeg i undersøkelse 2 et kontrollkorpus med tekster skrevet av morsmålsbrukere av norsk.

Å påvise transfereffekter handler for det første om å vise at språklige valg i innlærerspråk både er *representative* og *spesifikke* for en gruppe innlærere med samme morsmål sammenlignet med en innlærergruppe med et annet morsmål, altså at det er likheter *innad* i gruppene, men forskjeller *mellom* dem (intragruppe-homogenitet og intergruppe-heterogenitet). I tillegg handler det om å vise at det er et samsvar mellom de språklige valgene i innlærerspråket og språklige forhold i morsmålet (morsmålssamsvar), og at morsmålssamsvaret varierer i språklige kontekster der morsmål og målspråk er like eller ulike (intralingval kontrast). Disse effektene vil bli nærmere presentert i kapittel 2.4.3.

Slike transfereffekter blir undersøkt både i undersøkelse 1 og undersøkelse 2, men på ulike måter, som vil bli redegjort for underveis i avhandlingen. Transfereffektene *intragruppe-homogenitet* og *intergruppe-heterogenitet* blir undersøkt ved å sammenligne -GB-innlærere med +GB-innlærere, men det har vært nødvendig å finne ulike måter å gjøre det på for de to undersøkelsene. Også transfereffekten *morsmålssamsvar* er ulikt definert i de to undersøkelsene: Morsmålssamsvar i undersøkelse 1 er definert ut fra nullkontrasten som foreligger mellom -GB-morsmål og norsk, altså at fravær av kategorien i morsmålet fører til mindre korrekt bruk av bestemte former i innlærerspråket. I undersøkelse 2 er derimot morsmålssamsvar mer konkret definert gjennom de to hypotesene, som begge sier noe spesifikt om hvilke trekk fra morsmålet som ventes å bli overført til innlærerspråket. Transfereffekten *Intralingval kontrast* er ikke studert i undersøkelse 1, men i undersøkelse 2 har jeg valgt å dele opp materialet i ulike språklige kontekster etter om prediksjonene om formvalg for den polske gruppen sammenfaller eller ikke sammenfaller med det typiske formvalget i

⁵ Disse går under noe ulike navn i Jarvis 2000 og 2010, men jeg har valgt å framheve disse engelske termene da de ligger nærmest opp til mine norske oversettelser.

målspråket. På denne måten har jeg kunnet vise om språklige valg som ser ut til å være motivert av morsmålet, alternativt kan forklares som målspråkstilegnelse.

Utgangspunktet for oppdeling av materialet for intralingval kontrast er ulike typer av referanse, altså indefinit, definit og generisk referanse. Denne inndelingen av materialet gjør det også mulig å skille mellom språklige kontekster der den polske og den engelske gruppen er ventet å skille seg fra hverandre i språklige valg. Selv om polsk og engelsk som morsmål i utgangspunktet kan påvirke innlærerne til ulike typer språklige valg i innlærerspråket, vil formvalgene også kunne sammenfalle i noen tilfeller. Oppdelingen av materialet vil dermed bidra til at transfereffekten *intergruppeheterogenitet* undersøkes på en måte som sørger for at slike tilfeldige sammenfall ikke tilslører faktiske forskjeller mellom gruppene.

Operasjonaliseringen av transfereffekter for undersøkelse 2 er beskrevet mer detaljert i kapittel 8.5.

1.5 Strukturen i avhandlingen

Avhandlingen består av 11 kapittel, inkludert dette innledningskapittelet som har gitt en kort oversikt over de teoretiske og metodiske hovedtrekkene i studien. Alt som er omtalt så langt, blir nøyere gjennomgått i avhandlingens ulike kapittel.

I kapittel 2 redegjør jeg for teoretiske og metodiske sider av andrespråksfenomenet morsmålspåvirkning, mens kapittel 3 tar for seg bestemthet og referanse i språk med og uten grammatisk bestemthet, og kapittel 4 tar for seg andrespråksforskning knyttet til bestemthet. Kapittel 5 gir en kort redegjørelse for elektroniske innlærerkorpus som datakilde, i tillegg til å presentere ASK-korpuset. I kapittel 6 presenterer jeg metode og resultat fra undersøkelse 1, mens de neste fire kapitlene (7-10) er viet til undersøkelse 2: Kapittel 7 gir en kontrastiv beskrivelse av bestemthet, substantivmorfologi og referanse i norsk, polsk og engelsk, kapittel 8 presenterer de to hypotesene om polsk morsmålspåvirkning, i tillegg til datamaterialet og metodiske sider av undersøkelsen. Resultatene for hypotese 1 og 2 presenteres så i henholdsvis kapittel 9 og 10. Til sist, i kapittel 11, oppsummerer jeg funnene fra undersøkelse 1 og 2 og drøfter disse ut fra ulike forhold som er kommet fram i arbeidet med studien enten teoretisk eller metodisk.

2 MORSMÅLSPÅVIRKNING

I dette kapittelet vil jeg redegjøre for begrepet morsmålspåvirkning (transfer) med tanke på hva det er, når det forekommer og hvordan en kan gå fram metodisk for å finne morsmålspåvirkning i innlærerspråk. I 2.1 avgrensner jeg morsmålspåvirkning opp mot andre beslektede fenomener, i 2.2 redegjør jeg for hva man mener morsmålspåvirkning er, og hvilke forhold som er utløsende for morsmålspåvirkning, og i 2.3 ser jeg nærmere på forholdet mellom morsmålspåvirkning og tverrspråklige likheter og forskjeller. Til slutt, i 2.4, presenterer jeg metodiske sider av studiet av morsmålspåvirkning med utgangspunkt i Jarvis (2000, 2010), som har foreslått et samlet metodisk rammeverk for å bedre kunne avdekke når noe er morsmålspåvirkning og når det ikke er det.

2.1 Definisjon og avgrensning

Morsmålets rolle for andrespråklæring kan forstås på grunnlag av to videre og delvis overlappende fenomen: *tverrspråklig innflytelse* og *læringstransfer*. Begrepet tverrspråklig innflytelse (Kellerman & Sharwood Smith 1986) viser til den gjensidige påvirkningen mellom ulike språk, enten på samfunnsnivå eller individnivå, mens læringstransfer (Ellis 1965) er et generelt psykologisk fenomen som innebærer bruk av tidligere ervervet kunnskap i nye situasjoner, eller at én læringsoppgave påvirker en annen. Ringbom (2007: 5) skriver at det å lete etter likheter er essensielt ved all læring, i form av en kobling mellom nye oppgaver og eksisterende kunnskap:

The search for similarities is an essential process in learning. The natural procedure in learning something new is to establish a relation between a new proposition or task and what already exists in the mind.

I denne sammenhengen er det snakk om å etablere koblinger mellom tidligere lærte språk og et nytt språk som skal læres.

Selv om tverrspråklig innflytelse alltid innebærer påvirkning fra ett språkssystem til et annet, er det ikke alltid relatert til språklæring, og selv om læringstransfer alltid innebærer at man benytter seg av eller påvirkes av tidligere kunnskap, er ikke all tidligere kunnskap lingvistisk, og en innlærer har tilgjengelig mange ulike typer tidligere kunnskap som kan brukes med eller uten intensjon for å lette læringsprosessen. Det spesifikke fenomenet som blir undersøkt i denne avhandlingen, *morsmålspåvirkning*, er altså en del av begge disse to fenomenene: Morsmålspåvirkning innebærer den

påvirkningen allerede etablert kunnskap fra morsmålet (S1) har på læring eller bruk av et andrespråk (S2). Et vanlig og kortere ord som ofte blir brukt om morsmålspåvirkning, er *transfer*, og jeg vil i denne avhandlingen bruke begge disse termene om hverandre, i tillegg til *overføring*.

Når tverrspråklig innflytelse her avgrenses til morsmålspåvirkning, innebærer det å undersøke hvilken påvirkning morsmålskunnskap har på innlærerspråket til individuelle språkbrukere (innlærere). Det dreier seg altså ikke om tverrspråklig innflytelse i form av konvensjonaliserte språkkontaktsvarieteter på samfunnsnivå. Det begrenser seg også til en studie av transfer *fra morsmål til andrespråk*, og tar ikke for seg transfer fra andrespråket til morsmålet, eller transfer mellom to andrespråk.⁶ Kapittel 2 vil først og fremst dreie seg om denne typen tverrspråklig innflytelse fra morsmålet, og trekke fram de aspektene av fenomenet som er mest relevante for min spesifikke studie, selv om mye av det som blir omtalt, også gjelder generelt for tverrspråklig innflytelse.

For denne studien definerer jeg morsmålspåvirkning som *den påvirkningen en persons kunnskap fra morsmålet har på denne personens kunnskap om og/eller bruk av andrespråket*, altså basert på Jarvis og Pavlenko (2008: 1): «[...] the influence of a person's knowledge of one language on that person's knowledge or use of another language [...].»

En annen definisjon som brukes mye, er Odlin (1989: 27), som framhever at transfer er en type påvirkning som oppstår på grunn av likheter og forskjeller mellom målspråk og morsmål, eller andre tidligere lærte språk:

Transfer is the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired.

Selv om denne definisjonen har vært mye brukt og framhever et viktig aspekt av transferfenomenet, nemlig tverrspråklige likheter og forskjeller, kan den ses på som problematisk ettersom den ikke skiller eksplisitt mellom hva transfer *er* og det som *utløser* transfer, eller faktorer som samspiller med transfer (Alonso 2002: 86). Grunnen til at det er viktig å skille mellom disse forholdene, er at samspillet mellom fenomenet

⁶ I engelsk faglitteratur omtales disse som 'forward', 'reverse' og 'lateral transfer' (Jarvis & Pavlenko 2008: 22), men 'borrowing' og 'substratum' transfer er også brukt (Odlin 1989: 12).

og det som utløser det, er et empirisk spørsmål i seg selv. Odlin (1989: 27) påpeker selv svakheter ved denne definisjonen, og kaller det en arbeidsdefinisjon.

En annen definisjon som er sentral i undersøkelsen min, er en teorinøytral arbeidsdefinisjon formulert av Jarvis (2000: 252), hvor transfer først og fremst er definert ut fra hvordan vi kan observere og identifisere fenomenet, nærmere bestemt som en korrelasjon mellom et trekk i innlærerspråket og morsmålet. Denne definisjonen vil jeg komme nærmere tilbake til i 2.4.2.

2.2 Fenomenet transfer: hva, hvordan og når?

I dette underkapittelet vil jeg gi en oversikt over ulike perspektiv på hva transfer er, ulike klassifiseringer av transfer og ulike forhold som styrer hvorvidt transfer forekommer eller ikke.

2.2.1 Hva? Ulike perspektiv på hva transfer er

Det er ulike syn på hva transfer egentlig er, og det er vanlig å skille mellom fire forskjellige perspektiv: transfer som *prosess*, *begrensning*⁷, *strategi* eller *inaktivt utfall* (Jarvis 2000: 250; Alonso 2002: 92).

Transfer-som-prosess innebærer å se på transfer som en kognitiv prosess hvor språkinnlærere danner seg mentale assosiasjoner mellom språklige elementer i morsmål og i målspråk, såkalt *tverrspråklig identifikasjon*, der språklige elementer vurderes som like og derfor overførbare til andrespråket (Jarvis 2000: 250; Alonso 2002: 91). *Transfer-som-begrensning* er en avvisning av ideen om at transfer er en prosess, og man hevder at morsmålet heller fungerer som en begrensning på utviklingen av innlærerspråket og hvilke hypoteser innlærerne danner seg om hvordan målspråket fungerer. *Transfer-som-strategi* er et perspektiv på transfer som ser ut til å være mer kontroversielt siden det ofte er forbundet med et syn på transfer som uviktig i andrespråklæringen (Jarvis & Pavlenko 2008: 9). Det innebærer å se på transfer som en kommunikativ strategi hvor trekk eller ord fra morsmålet brukes for å fylle igjen hull i innlærerens andrespråkskompetanse. Dette perspektivet impliserer nødvendigvis en avvisning av at det er mulig med transfer som går motsatt vei, altså *fra* andrespråket til morsmålet, noe man vet at forekommer (Jarvis & Pavlenko 2008: 9). Selv om mange

⁷ Eng. 'constraint'. Jeg har her valgt oversettelsen 'begrensning' i tråd med Berggreen og Tenfjord (1999: 218)

har avvist ideen om at transfer utelukkende er en kommunikativ strategi, er det likevel allment akseptert at transfer *også* kan være en vanlig strategi brukt av innlærere, men det er altså bare én av mange typer transfer. *Transfer-som-inaktivt utfall* omtales av Jarvis (2007: 299) som ‘inert outcome’, og er en form for *udynamisk og ikke-reaktiv* tilstand, hvor transfer kan forekomme uten noen form for tverrspråklig identifikasjon mellom morsmål og målspråk:

Inert means nonreactive and nondynamic. According to this view, L1-based conceptual influence can take place even when the learner has not made any overt comparisons or interlingual identifications between L1 and L2 forms and features.

Jeg har i denne avhandlingen valgt å kalle dette for *inaktivt utfall*, for å best dekke det som handler om det ikke-dynamiske og ikke-reaktive i forklaringen over. Denne typen transfer er også koblet til lingvistisk relativitet og begrepet *konseptuell transfer*, som kommer til å bli omtalt nærmere lenger ned.

Disse fire perspektivene på hva transfer er, kan virke som motsetninger: Begrensnings-perspektivet avviser prosess-perspektivet, strategi-perspektivet avviser alle de andre perspektivene, og prosess- og inaktivt utfall-perspektivene står i motsetning til hverandre ettersom det ene forutsetter en tverrspråklig identifikasjonsprosess, mens den andre ikke gjør det. Alonso (2002: 99) argumenterer likevel for at dette ikke er perspektiv som ekskluderer hverandre, og mener at selv om transfer ofte baseres på tverrspråklig identifikasjon, kan det også være en strategi, eller en begrensning: «[...] these views are not mutually exclusive and though transfer is a process based on interlingual identifications, learners can resort to it as a strategy, while constraints acts as filters.» Også Jarvis (2000: 253) nevner at alle disse fire perspektivene på transfer bør integreres i en framtidig definisjon av transfer.

I de neste underkapitlene kommer jeg tilbake til noen av disse perspektivene på transfer, spesielt transfer som prosess og som inaktivt utfall, ettersom jeg i denne avhandlingen vil fokusere spesielt på transfer som resultat av mangel på tverrspråklig identifikasjon.

2.2.2 Hvordan? Utfall av og klassifisering av transfer

Selv om det er mange måter å klassifisere transfer på, er en av de mest vanlige inndelingene ut fra *utfallet* transfer har i forhold til målspråksnormen, der transfer kan være *positiv* eller *negativ*, eller ha en *kumulativ effekt* på tilegnelsen. Odlin (1989: 36-38) skiller mellom *positivt/negativt* transferutfall og *tilegnelsestid*. Negativ transfer er når utfallet av transfer er språkbruk som avviker fra målspråksnormen. Dette er gjerne ulike typer språklige feil, men det kan også være avvik med tanke på hvor ofte et språklig element brukes, altså *overrepresentasjon* eller *underrepresentasjon* i forhold til det som er vanlig hos en morsmålsbruker. I andre tilfeller kan innlærerne ha preferanser for visse trekk som brukes ofte i morsmålet, eller de kan fullstendig unngå strukturer som er unike for målspråket (Odlin 1989: 36-37). I motsetning til rene feil vil disse utfallene ikke være så lett å legge merke til, og heller ikke nødvendigvis ha noen negativ effekt på innlærerspråket annet enn at formuleringen skiller seg fra det en morsmålsbruker typisk ville sagt. Positiv transfer har man derimot når utfallet er konvensjonell eller målspråkslik språkbruk, og knyttes gjerne til hvordan morsmålet kan *fasilitere* andrespråklæringen, altså gjøre det lettere å lære noe.

Mens negativ og positiv transfer er utfall som er identifiserbare som spesifikke og direkte observerbare trekk ved innlærernes språkbruk, fins det også noen mer generelle utfall av transfer, noe Odlin (1989: 38) viser til som *kumulative effekter*. Dette kan være positive eller negative effekter som morsmålet har på tempoet eller rekkefølgen tilegnelsen av andrespråket følger, altså at det fører til at utviklingen blir forsinket, går raskere eller at det påvirker rekkefølgen trekk læres i, eventuelt det endelige resultatet (Jarvis & Pavlenko 2008: 11). Disse effektene er kumulative i den forstand at de er *sekundær-effekter* av en mer spesifikk transfer eller av generell fasilitering eller utfordringer som oppstår på grunn av likhet og forskjell mellom morsmål og målspråk.

Hva som overføres, eller om det er positivt eller negativt, er bare to av flere måter å klassifisere tilfeller av transfer på, og målspråket er bare ett av flere referansepunkter. Jarvis og Pavlenko (2008: 20) gjør rede for 10 dimensjoner transfer kan klassifiseres innen. Noen av disse har jeg allerede vært inne på tidligere i dette kapittelet: *retning* (forward/reverse/lateral), *utfall* (positivt/negativt) og *intensjonalitet* (strategi/mental

assosiasjon). I tillegg nevner de *språklig nivå* (fonologi, syntaks, pragmatikk osv.), *kognitivt nivå* (lingvistisk/konseptuelt), *kunnskapstype* (implisitt/eksplisitt), *modus* (produksjon/resepsjon), *kanal* (auditiv/visuell), *form* (verbal/ikke-verbal) og *manifestasjon* (overt/covert). Alle disse dimensjonene er relevante for klassifiseringen av transfer, men framhever ulike sider av fenomenet.

En annen viktig dimensjon er forskjellen på transfer på *eksemplarnivå* og transfer på *systemnivå* (Ringbom 2007: 54; Jarvis & Pavlenko 2008: 58; Ringbom og Jarvis 2009: 110).⁸ Om transfer på systemnivå skriver Jarvis og Pavlenko (2008: 58) at dette er tilfeller av transfer hvor hele språklige delsystem er påvirket av et annet språk:

System-wide effects include cases where an entire lexical, morphological, phonological, or syntactic class in a person's use of one language is affected by the corresponding system from another language that she knows.

Dette står i motsetning til tilfeller av transfer som begrenser seg til spesifikke språklige enkeltelementer eller avhenger av en egenskap ved et språklig element. Forskjellen kan illustreres med en polsk innlærer som generelt støtter seg på det morfologiske systemet i morsmålet i bruken av bestemthetsbøyning i innlærerspråket, versus en innlærer som viser feil bruk av bestemt form på grunn av en feilkategorisering av et spesifikt substantiv som tellelig eller ikke-tellelig (Jarvis & Pavlenko 2008: 58).

2.2.3 Når? Faktorer som påvirker forekomst av og type transfer

Jarvis og Pavlenko (2008: 174) skriver at en av de viktigste vendingene i transferforskningen var da man endret fokus fra *transfer* til '*transferability*' (overførbarhet), altså hvor sannsynlig det er at noe vil bli overført fra morsmålet til innlærerspråket:

This was a shift from the documentation of cases of transfer to the more fundamental investigation of what makes something likely to be transferred in the first place.

Overførbarhet er ofte sett på som innlærernes egen vurdering av om et element fra morsmålet kan brukes i målspråket, og er en av mange faktorer som kan påvirke om

⁸ Eng. '*item-specific*' og '*system-wide*'

transfer forekommer. Jarvis og Pavlenko (2008: 175) deler inn slike faktorer i fem ulike hovedområder:

- Lingvistiske og psykolingvistiske faktorer
- Kognitive og oppmerksomhets- eller utviklingsmessige faktorer
- Kumulativ språkerfaring og kunnskap
- Læringsmiljørelaterte faktorer
- Språkbruksrelaterte faktorer

Mens sentrale lingvistiske og psykolingvistiske faktorer, som språklig nivå eller en strukturs frekvens og saliens i morsmål eller målspråk, påvirker overførbarheten til strukturen direkte, kan de øvrige, ikke-lingvistiske, faktorene begrense om transfer vil forekomme i en spesifikk kontekst eller om den individuelle innlæreren er tilbøyelig til å ha transfer. Noen av effektene av begrensningene er *læringsrelaterte*, mens andre er *performanserelaterte*. Læringsrelaterte effekter er om det skjer eller ikke skjer en tverrspråklig identifikasjon i selve læringsprosessen, noe som påvirker hvordan et språklig trekk læres, mens performanserelaterte effekter handler om hvorvidt transfer forekommer i faktisk språkbruk, og styres av forhold ved språkbrukssituasjonen (Jarvis & Pavlenko 2008: 175).

I det følgende vil det først og fremst være de lingvistiske og psykolingvistiske begrensningene som blir omtalt, da dette er det viktigste utgangspunktet for hypotesene og diskusjonene senere i denne avhandlingen.

2.3 Tverrspråklige likheter og forskjellers rolle for transfer

At likheter og forskjeller mellom morsmål og målspråk er viktig for transfer, er noe man har ment siden de aller tidligste teoriene om transfer, og selv om det er oppnådd ny og endret forståelse av hvilken rolle forskjeller og likheter spiller, er slike kontrastive forhold fremdeles regnet som grunnleggende for transfer. I dette underkapittelet vil jeg gjøre rede for ulike nivå av tverrspråklige likheter og forskjeller, og hvordan disse antas å føre til transfer.

2.3.1 Grad av samsvar mellom morsmål og målspråk

Tverrspråklige likheter og forskjeller kan identifiseres på flere nivå. På et overordnet nivå er det nødvendig å skille mellom *den generelle graden av samsvar* mellom to språk

og *de spesifikke likhetene og forskjellene* (Jarvis & Pavlenko 2008: 177).⁹ Det første går på hvor like de to språkene er generelt, mens det siste viser til spesifikke (ikke-)korrespondanser på ulike språklige nivå. Det meste av det som blir presentert under, vil handle om spesifikke likheter og forskjeller, men det generelle samsvaret mellom språk er en viktig faktor for transfer og andrespråkslæring, ettersom det fasiliterer eller begrenser innlærerens muligheter til å dra nytte av innputt, eller muligheten til å oppdage viktige likheter mellom morsmål og målspråk på et mer spesifikt nivå.

2.3.2 Subjektive og objektive likheter og forskjeller

Det er videre nødvendig å skille mellom *objektive* og *subjektive* likheter og forskjeller, altså forskjeller og likheter mellom to språk med utgangspunkt i en lingvistisk tilnærming, i motsetning til forskjeller og likheter sett fra innlærerens subjektive ståsted. Det siste kalles gjerne *psykotypologi*, et begrep som ble lansert av Kellerman (1983). En av de viktigste innsiktene en har kommet til innen transferforskningen, er nettopp betydningen av innlærernes egne antakelser om hva som er likt og forskjellig, og dermed også overførbart fra morsmålet til innlærerspråket.

Objektive likheter og forskjeller viser ideelt sett til faktiske forhold mellom språk, men som Ringbom og Jarvis (2009: 107) påpeker, eksisterer det ingen egentlig enighet om hvordan en skal måle eller stadfeste hva som er «faktisk», og utsagn om likheter og forskjeller vil ofte være teoriavhengige. Et annet problem med en objektiv målestokk er at lingvistenes blikk og kategorisering av språk ikke nødvendigvis sammenfaller med innlærernes, og de kan derfor være et mindre egnet redskap for å identifisere transfer (Jarvis & Pavlenko 2008: 179). Det er også en annen viktig forskjell mellom disse to kontrastive nivåene: Objektive likheter og forskjeller (hvordan de nå enn fastsettes) hører til språkssystemene og er mer eller mindre *konstante*, mens subjektive likheter og forskjeller opererer på individnivå, og de kan derfor i langt større grad både forandre seg over tid og være ulike fra innlærer til innlærer (Jarvis & Pavlenko 2008: 178; Ringbom & Jarvis 2009: 107). De to ulike nivåene relaterer seg også ulikt til transfer: De subjektive *trigger* transfer, mens de objektive først og fremst fungerer som et redskap

⁹ Eng. 'overall similarities/differences' og 'specific similarities/differences' (Jarvis & Pavlenko 2008: 177)

for å klassifisere tilfeller av transfer som positive eller negative, eller for å danne hypoteser om transfer (Jarvis & Pavlenko 2008: 179).

2.3.2.1 *Oppfattede og antatte likheter og forskjeller*

Når det gjelder subjektive kontrastive antakelser som innlæreren kan ha, er det vanlig å skille mellom *oppfattede* likheter og forskjeller, altså antakelser som er basert på noe innlærerne har oppfattet i målspråksinnputt, og *antatte* likheter og forskjeller, som ikke er basert på slike observasjoner.¹⁰ Oppfattede likheter og forskjeller er ifølge Jarvis og Pavlenko (2008: 179) når et språklig trekk i innputt bevisst eller ubevisst vurderes av innlæreren som likt noe i morsmålet:

A perceived similarity is a conscious or unconscious judgment that a form, structure, meaning, function, or pattern that an L2 user has encountered in the input of the recipient language is similar to a corresponding feature of the source language.

Antatte likheter forekommer derimot uavhengig av observasjoner av innputt (Jarvis & Pavlenko 2008: 179). Som jeg vil komme tilbake til under, påvirker disse to typene subjektive kontrastive antakelser morsmålspåvirkningen på ulike måter.

Selv om innlærernes oppfattede likhets- og ulikhetsrelasjoner ikke nødvendigvis reflekterer objektive likheter og forskjeller, er det åpenbart at de objektive likevel kan påvirke hva som oppfattes av innlæreren: Jo flere likheter mellom to språk, jo flere likheter og forskjeller kan bli identifisert. Dette betyr at objektive likheter og forskjeller, både på generelt og på spesifikt nivå, kan fungere som en begrensning på psykotypologien i mange tilfeller, og at jo likere språkene er typologisk sett, jo mer kan innlæreren dra nytte av (Kellerman 1995: 130).

Psykotypologi antas å samhandle et trekks overførbarhet. Å vurdere noe som overførbart eller ikke innebærer både en vurdering av mulige korrespondanser mellom morsmål og målspråk, og en vurdering av hvor sannsynlig det er at en antatt likhet faktisk er reell og overførbart. Disse to prinsippene, *psykotypologi* og *overførbarhet*, som begge ble formulert først av Kellerman (1983), er altså de to viktigste begrensningene på transfer og styrer dermed sannsynligheten for at transfer skal forekomme.

¹⁰ Eng. 'Perceived similarities and differences' og 'assumed similarities and differences'

Sannsynligheten øker om språkene oppfattes som like, men avtar om innlæreren oppfatter den aktuelle strukturen som for markert til å også forekomme i målspråket:

The essence of the psychotypology constraint is that transfer is more likely to occur when the L2 user perceives the L1 and L2 as being similar, whereas the essence of the transferability constraint is that structures perceived by the L2 user as marked (or language-specific) are less likely to transfer (Jarvis & Pavlenko 2008: 174).

Jeg vil nå rette fokus mot hvordan kontrastive antakelser hos innlærere kan føre til ulike typer transfer.

2.3.3 Ulike typer kontrastive antakelser og transfertyper

Ifølge Ringbom og Jarvis (2009: 109-110) kan psykotypologi ha tre ulike utfall:

- 1) Innlærere har oppfattede antakelser om spesifikke *likheter* mellom språkene.
- 2) Innlærere har oppfattede antakelser om spesifikke *forskjeller* mellom språkene.
- 3) Forskjellene mellom språkene er for store til at det oppfattes forskjeller og likheter.

Disse tre utfallene vil refereres til som henholdsvis *likhetsrelasjoner*, *ulikhetsrelasjoner* og *nullrelasjoner*. De tre kontrastive relasjonene ser ikke ut til å beskrives som objektive forskjeller og likheter av Ringbom og Jarvis, men som *subjektive*, ettersom de tar utgangspunkt i innlærerens oppfatning av hvordan relasjonene mellom morsmål og målspråk er. For eksempel skriver de at likhetsrelasjoner innebærer at noe i målspråket *oppfattes* som likt noe i morsmålet:

The similarity relation means that an item or pattern in the TL is **perceived** as formally and/or functionally similar to a form in the L1 (or other previously learned language) (Ringbom & Jarvis 2009: 109, min utheving).

Begrepet *nullrelasjon*, som er beskrevet her, skal dermed i denne avhandlingen skilles fra begrepet *nullkontrast*, som er termen jeg bruker for å beskrive de objektive forskjellene mellom språk med og uten grammatisk bestemthet, altså at målspråket har en kategori som morsmålet ikke har.

Mens det foreligger mye støtte for antakelsen om at likhetsrelasjoner er den sterkeste triggeren for transfer, kan transfer også forekomme som resultat av de andre typene kontrastive relasjoner (Jarvis & Pavlenko 2008: 179). Dette betyr at transfer kan være resultat av at innlærerne antar at noe er *likt* eller av at de antar at noe er *ulikt*, og i

tillegg kan transfer forekomme i tilfeller hvor det ikke er noe særskilt grunnlag for å danne seg slike kontrastive antakelser. Det siste kan være tilfelle når de generelle forskjellene mellom morsmål og målspråk er store, eller når likhetene er på et høyt abstraksjonsnivå (Ringbom & Jarvis 2009: 110). Denne kompleksiteten i forholdet mellom tverrspråklige likheter og forskjeller og transfer, ligger til grunn for de videre underkapitlene, hvor jeg går nærmere inn i hvordan de ulike typene kontrastive antakelser alle kan føre til transfer, men til ulike typer transfer.

2.3.4 Transfer som resultat av likhetsrelasjon

Den første typen transfer jeg vil se på, er den som kan forekomme som resultat av at innlærerne antar at noe i morsmål og målspråk er likt, altså en *likhetsrelasjon*. Denne typen antakelse blir hevdet å være en av de vanligste triggerne av transfer, ettersom det å lete etter likheter er grunnleggende ved all læring (Ringbom 2007: 5).

Ringbom og Jarvis (2009: 108) skriver at i reseptive språkbrukssituasjoner er likhetene ofte *oppfattede*, mens de i produksjonssituasjoner heller er *antatte*, siden innlærere da ikke observerer innputt, men formulerer seg på grunnlag av det de allerede har lært. Likevel kan transfer i produksjonssituasjoner være resultat av oppfattede likheter fra tidligere reseptive situasjoner eller andre møter med målspråksinnputt.

Noen lingvistiske nivå er mer tilgjengelige enn andre med hensyn til å oppfatte likheter, siden noen språklige trekk er mer saliente i innputt enn andre. Dette betyr at formelle trekk (ortografiske, fonologiske, morfologiske) lettere blir oppfattet enn for eksempel semantiske trekk eller trekk som innebærer et høyt abstraksjonsnivå. På grunn av dette mener man at formelle likheter vanligvis er *oppfattede*, mens semantiske og pragmatiske oftere bare er *antatte* uavhengig av målspråksinnputt (Ringbom & Jarvis 2009: 108; Jarvis & Pavlenko 2008: 179).

Et annet viktig aspekt av antatt/oppfattet-distinksjonen er hvordan disse to typene kontrastive antakelser noen ganger virker sammen. Når en viss terskel av likhet er foreliggende, kan både generelle og spesifikke oppfattede likheter føre til *overantakelser* om ytterligere likheter som aldri har vært oppfattet i innputt og som heller ikke behøver å være faktiske likheter mellom språkene (Jarvis & Pavlenko 2008: 180; Ringbom & Jarvis 2009: 109).

Transfer som resultat av likhetsrelasjoner, i kombinasjon med naturlige tilegnelsesprinsipper, kommer inn under det som kalles *transfer to somewhere-prinsippet* (Andersen 1983), som blant annet sier at en struktur kun blir overført om den oppfattes å ha et passende sted å overføres til i målspråket:

[...] a language structure will be susceptible to transfer only if it is compatible with natural acquisitional principles or is perceived to have a similar counterpart (a *somewhere* to transfer to) in the recipient language (Jarvis & Pavlenko 2008: 174).

Hovedpoenget med prinsippet er at transfer ikke forekommer med mindre innlæreren antar at det fins visse likheter i målspråket som gjør det mulig å ta i bruk et trekk fra morsmålet, altså er vurderingen av overførbarhet sentral for transfer som er resultat av likhetsrelasjoner.

2.3.5 Transfer som resultat av ulikhetsrelasjon

På samme måte som innlærere kan identifisere trekk som de antar er like, kan de også identifisere trekk som de antar er *ulike* i de to språkene, noe som naturligvis vil føre til at et språklig trekk ikke blir overført fra morsmålet, siden ulikheter impliserer lav grad av overførbarhet. Ringbom og Jarvis (2009: 109) skriver dessuten at antatte forskjeller gjerne vokser ut av underliggende likheter, og kan resultere i at innlærerne *ikke* overfører noe fra morsmålet, men også i språkbruk som likevel kan regnes som transfer, for eksempel *unngåelse*, hvor innlærerne unngår et trekk som er ansett som forskjellig (Schachter 1974: 210; Odlin 1989: 37). Dette er altså ikke en type morsmålspåvirkning som innebærer en ren overføring av elementer fra morsmålet, men likevel en språklig adferd som har sitt opphav i de kontrastive relasjonene mellom morsmål og målspråk.

I tillegg nevner Jarvis og Pavlenko at ulikhetsrelasjoner ikke nødvendigvis fører til vansker eller unngåelse, men at det faktisk kan gjøre et trekk lettere å tilegne seg. I tilfeller med lett oppfattbare forskjeller kan kontrasten i seg selv være fasiliterende, ikke i form av transfer, men heller gjennom at man blir mer oppmerksom på innputt på de områdene man ikke kan støtte seg på morsmålet (Jarvis & Pavlenko 2008: 11).

2.3.6 Transfer som resultat av manglende tverrspråklig identifikasjon

Det som er felles for transfer som resultat av likhetsrelasjoner og ulikhetsrelasjoner, er at de begge innebærer en prosess der innlæreren etablerer mentale assosiasjoner mellom

trekk i morsmålet og målspråket, altså tverrspråklig identifikasjon. I tilfeller der det er store forskjeller mellom språkene, eller likhetene er veldig abstrakte, kan denne identifikasjonsprosessen imidlertid være fraværende, og det er dette som kalles for en nullrelasjon.

Til tross for enigheten om at tverrspråklig identifikasjon er grunnleggende for transfer, har mange forskere hevdet at transfer også kan forekomme uten en tverrspråklig identifikasjon. Kellerman (1995: 131) påpeker at om det skulle være slik at likheter er den eneste årsaken til transfer, så ville det aldri forekomme transfer når det ikke er grunnlag for å oppfatte likheter mellom språkene. Dette er ifølge Kellerman ikke i tråd med det man har funnet i forskning. Som eksempel bruker han de velkjente vanskene innlærere som ikke har bestemthet i morsmålet, har i tilegnelsen av bestemthet (Kellerman 1995: 131). For å inkludere denne typen morsmålpåvirkning, foreslår Kellerman et komplementært prinsipp til Andersens (1983) *transfer to somewhere*, som han heller *transfer to nowhere* (Kellerman 1995: 137). Mens *transfer to somewhere* vektlegger viktigheten av tverrspråklig likhet og innlæreres positive tverrspråklige identifikasjon, handler *transfer to nowhere* om den typen transfer som forekommer når ingen slike likheter oppfattes, eller kanskje ikke fins i det hele tatt:

This principle states that there can be transfer which is not licensed by similarity to the L2, and where the way the L2 works may very largely go unheeded; hence *transfer to nowhere* [...] (Kellerman 1995: 137).

Dette beskrives altså som en situasjon hvor et trekk er unikt for enten morsmål eller målspråk, eller hvor de eventuelle likhetene er for abstrakte til å kunne oppfattes. I slike tilfeller vil enten morsmålet begrense innlærernes hypoteser om hvordan målspråket fungerer, eller fraværet av grunnlag for å finne likheter fører til en generell, umotivert antakelse om at målspråket fungerer på samme måte som morsmålet (Ringbom & Jarvis 2009: 108). I begge tilfeller vil det være snakk om ikke-oppfattede og antatte likheter.

Kellerman (1995: 137) påpeker at *transfer to nowhere* kan ha sitt opphav i morsmålets begreppssystem, altså at det ikke først og fremst handler om *formelle* forskjeller mellom språkene, men heller at man i ulike språk konseptualiserer verden på ulike måter.

2.3.6.1 *Lingvistisk versus konseptuell transfer*

Kellermans refleksjoner om forskjeller i hva transfer dypest sett dreier seg om, kan kobles til et mer differensiert syn på tverrspråklige likheter og forskjeller, nemlig skillet mellom *lingvistiske* og *begrepsmessige* tverrspråklige likheter og forskjeller (Kellerman 1995: 140).¹¹ Forskjeller mellom språk kan enten handle om språkstrukturene og hva de uttrykker, eller det kan handle om begrepsinventaret og organiseringen av begreper. Disse ulike typene forskjeller og likheter antas å føre til to ulike typer transfer, nemlig *lingvistisk* og *konseptuell* transfer (Jarvis 2007; Jarvis & Pavlenko 2008: 61). Jarvis og Pavlenko (2008: 61) skriver at lingvistisk transfer har opphav i språklig strukturer, mens konseptuell transfer har opphav i begrepene som ligger bak strukturene:

[...] the distinction between types of transfer that are examined primarily in relation to linguistic forms and structures versus types of transfer that are analyzed in relation to the mental concepts that underlie these forms and structures.

Konseptuell transfer blir koblet til teorier om *språklig relativitet*, altså ideen om at språket påvirker tanken og at språkbrukere med ulike morsmål derfor oppfatter verden på ulike måter (Jarvis 2007; Jarvis & Pavlenko 2008: 115). Dette er en ide som tidligere har kommet til uttrykk for eksempel hos Whorf (1956), som antydte at ulike språk legger ulike føringer på hvordan språkbrukerne observerer verden utenfor språket, noe som gjør at mennesker ikke kan se verden på samme måte:

Users of markedly different grammars are pointed by their grammar towards different types of observations and different observations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world (Whorf 1956: 221).

Denne måten å tenke på møtte mye motstand på 1960-tallet, og ble lenge avvist av forskere, men etter hvert har språklig relativitet gjenoppstått innen kognitiv lingvistikk, for eksempel hos Lakoff (1987). Mens den tidlige relativismen kanskje først og fremst var opptatt av hvordan leksikon påvirket menneskelig persepsjon, er de nyere retningene mer orientert mot språkets mer abstrakte og ikke-perseptuelle kategorier. Lindsay (2003: 275) skriver for eksempel at språkets påvirkning på hvordan vi tenker, er sterkere

¹¹ 'Cross-linguistic similarities and differences' og 'cross-linguistic conceptual similarities and differences'

på områder som ligger utenfor det vi kan sanse, altså på abstrakte områder. Dette gjelder for eksempel kategorier som genus, temporalitet og romlige kategorier, og det går et sentralt skille mellom såkalt *språkmedierte* begreper og *språkuavhengige* begreper (Jarvis & Pavlenko 2008: 113). De språkmedierte begrepene vil utvikles parallelt med språket slik at det er en gjensidig påvirkning mellom språk og begrep, noe som ikke er tilfelle for de språkuavhengige begrepene (Jarvis & Pavlenko 2008: 114).

Slobin (1996) har også argumentert for at relativistiske effekter kan være relevant for læring av et nytt språk, men heller enn å argumentere for språkets påvirkning på selve tanken og generell kognisjon, foreslår han fenomenet *thinking for speaking*, som er en type påvirkning som kun forekommer når man bruker språket og som ikke er så sterk at den påvirker måten vi ser verden på generelt (Slobin 1996: 71). Påstanden er at vi er opplært gjennom morsmålet til å fokusere på visse aspekter av virkeligheten *når vi omtaler den språklig*, og dette hevder Slobin (1996: 89) er noe som er vanskelig å endre når man skal lære et nytt språk.

Slobin (1996: 91) nevner bestemthet som en av flere kategorier knyttet til fenomenet *thinking for speaking*, altså kategorier som kun eksisterer i og for språket, og dermed kun kan læres gjennom språket:

Distinctions of aspect, definiteness, voice, and the like, are, *par excellence*, distinctions that can only be learned through language, and have no other use except to be expressed in language. They are not categories of thought in general, but categories of thinking for speaking.

Jarvis (2007) samler de ulike relativisme-inspirerte tilnærmingene i andrespråksforskningen under paraplybegrepet *konseptuell transfer*, og disse har som felles antakelse at det er snakk om transfer som har opphav i kognitive eller begrepsmessige forskjeller mellom språkbrukere av ulike språk (Jarvis 2007: 44). Noen av disse forskjellene på begrepsnivå kan være direkte knyttet til *språklige* forskjeller mellom språk, mens andre begrepsforskjeller kan ha sitt opphav i kultur, altså er de en effekt av *kulturell* relativitet, heller enn språklig (Jarvis 2007: 44).

Forskjellen mellom *lingvistisk* og *konseptuell* transfer blir av Jarvis og Pavlenko (2008: 112) knyttet til den tidligere nevnte forskjellen mellom transfer som resultat av tverrspråklig identifikasjon og transfer som resultat av nullrelasjon uten tverrspråklig identifikasjon. Til tross for de åpenbare parallellene, påpeker Jarvis og Pavlenko

(2008: 112) at selv om konseptuell transfer ofte blir undersøkt med utgangspunkt i tverrspråklige *forskjeller*, for eksempel nullrelasjoner, kan likheter også være relevante.

2.3.7 Oppsummering av fenomenet transfer

I denne første delen av kapittelet om transfer har jeg forsøkt å gjøre rede for de mest grunnleggende innsiktene om transfer som virker relevante for min egen studie. Jeg har vektlagt forskjellen mellom *transfer i seg selv* (den faktiske språklige kunnskapen som blir overført fra morsmålet) og *utfallet av transfer*. Jeg har også presentert en rekke ulike faktorer som styrer om transfer forekommer, med hovedfokus på den rollen tverrspråklige likheter og forskjeller og psykotypologi spiller, og hvordan transfer kan være resultat av både likheter og forskjeller mellom morsmål og målspråk, i tillegg til at det både kan være et produkt av en tverrspråklig identifikasjonsprosess og mangel på en slik.

I forbindelse med undersøkelse 2, som ser på morsmålspåvirkning hos polske innlærere, er det spesielt transfer som oppstår som resultat av nullrelasjon og *transfer to nowhere* som er relevante ettersom polsk mangler en språklig kategori som fins i norsk.

2.4 Metodiske sider av studiet av transfer

I resten av dette kapittelet vil jeg presentere et enhetlig rammeverk for transferstudier foreslått i Jarvis (2000), og delvis videreutviklet i Jarvis (2010). Jarvis (2010: 170) skiller mellom to typer metodiske tilnæringer til transfer: *comparison-based* og *detection-based*, og det er den første typen metode som vil bli presentert nedenfor, altså hvordan man identifiserer transfer ved hjelp av ulike sett av sammenligninger mellom språk og grupper av språkbrukere, som sammen danner argument for at det foreligger transfer (Jarvis 2010: 170).

2.4.1 Et enhetlig rammeverk for transferstudier og kriterier

Jarvis (2000: 247) argumenterer for en felles prosedyre for undersøkelser av tverrspråklig innflytelse og transfer for å sikre at ulike forskere studerer det samme fenomenet og oppnår valide konklusjoner. Jarvis problematiserer gapet mellom ulike transferstudier både på teoretisk og metodisk nivå, i tillegg til å påpeke at det er en inkonsistens i konklusjonene om transfer så langt i forskningen. Teoretisk sett er det ingen konsensus om hva fenomenet transfer egentlig er, og som jeg var inne på i 2.2.1,

kan transfer ses på som en begrensning, en prosess, en strategi eller som inaktivt utfall (Jarvis 2000: 250). Metodisk sett mener Jarvis at man har tilnærmet seg transfer på ulike måter, med ulike krav til hva som regnes som evidens for transfer og med ulike typer sammenligninger, noe som har ført til ulike konklusjoner om transfer. Jarvis (2000: 246) påpeker at transfer ofte blir behandlet som et «*you-know-it-when-you-see-it*»-fenomen, og at det har manglet enighet om hva transfer er og hvordan det skal undersøkes, noe som gjør at ulike forskere kanskje ikke har sett eller lett etter de samme effektene:

Until now, L1 influence has been treated largely as a you-know-it-when-you-see-it phenomenon, and although most researchers may indeed recognize L1 effects when they see them, the lack of consensus concerning what L1 influence is and how it should be investigated may mean that different researchers have not seen (or even looked for) the same effects.

Ett eksempel på dette er hvordan noen studier undersøker transfer ved å sammenligne innlærerspråket til innlærere med ulike morsmålsbakgrunner, mens andre sammenligner innlærerspråk med morsmålsperformanse. Som jeg vil komme tilbake til under, hevder Jarvis (2000: 252) at disse typene sammenligninger isolert ikke utelukker andre påvirkningsfaktorer enn morsmålet som kilde til språkbruken, og han understreker nødvendigheten av begge typer sammenligninger.

Mangelen på enighet og felles prosedyrer er bakgrunnen for Jarvis' enhetlige rammeverk for transferstudier. Det er ment som felles retningslinjer som kan brukes på tvers av teoretiske tilnærminger til transfer, for å sikre at alle transferstudier forholder seg til de samme kravene for å påvise transfer. Rammeverket består av tre hovedkriterier (Jarvis 2000: 249):

1. en teorinøytral definisjon av transfer
2. en avklaring av hvilke typer evidens som utgjør et argument for transfer
3. en liste over andre variabler som må kontrolleres for

I de følgende underkapitlene vil jeg presentere disse tre kriteriene nærmere.

2.4.2 Teorinøytral definisjon

Et av problemene som ofte påpekes ved transferstudier, er at det ikke er noen felles enighet om hva transfer er, og dermed ikke noen felles definisjon av transfer. I tillegg er det slik at når man ikke har en full forståelse av fenomenet, er det umulig å oppnå eller

formulere en fullstendig definisjon basert på fenomenets natur (Odlin 1989: 28; Jarvis 2000: 249). Når det er ulike definisjoner, blir studiene mindre sammenlignbare, og som løsning på dette foreslår Jarvis (2000: 250) en arbeidsdefinisjon som i tillegg til å være *teorinøytral*, også er *metode-basert* heller enn *fenomenbasert*. En slik definisjon har som mål å uttrykke eksplisitt hvordan transfer kan undersøkes og hva som teller som evidens for transfer, ikke å definere hva transfer *er*. I dette tilfellet handler det om signifikante korrelasjoner mellom trekk i innlærerspråket og i morsmålet:

[...] I propose the following working definition: L1 transfer refers to any instance of learner data where a statistically significant correlation (or probability-based relation) is shown to exist between some feature of learners' IL performance and their L1 background (Jarvis 2000: 252).

En metodisk definisjon som denne har som utgangspunkt den mest fundamentale ideen om tverrspråklig innflytelse: Et språk påvirker et annet, og dette må logisk sett manifestere seg som en korrelasjon mellom morsmålsbakgrunn og språklige mønster i innlærerspråket.

2.4.3 Transfereffekter

Den evidensen for transfer som definisjonen til Jarvis nevner, finner man altså som en korrelasjon mellom innlærernes morsmålsbakgrunn og innlærerspråket deres, som er en kompleks type evidens å finne, siden det både innebærer å koble innlærerspråket til morsmålet og å utelukke andre faktorer enn morsmålet som kilde til språklige valg. Det er her Jarvis' andre kriterium kommer inn i bildet, nemlig en spesifisert liste over hvilke typer evidens som er nødvendige. Tanken er at man kan undersøke fenomenet transfer *indirekte* ved å se etter logiske, observerbare effekter transfer vil ha på et datamateriale (Jarvis 2000: 253-254). Slike effekter kaller Jarvis for *transfereffekter*, og de fire transfereffektene han foreslår, er i denne avhandlingen omtalt som *intragruppe-homogenitet*, *intergruppe-heterogenitet*, *morsmålssamsvar* og *intralingval kontrast*.¹² *Intragruppe-homogenitet* betyr at innlærere med samme morsmål har lik språklig adferd for et gitt trekk, mens *intergruppe-heterogenitet* betyr at innlærere med *ulike* morsmål

¹² Jarvis (2010) kaller dem også hhv. 'within-group homogeneity', 'between-group heterogeneity', 'cross-language congruity', 'intralingual contrast'.

har ulik språklig adferd. *Morsmålssamsvar* innebærer at det er likheter mellom morsmålet og innlærerspråket, mens *intralingval kontrast* innebærer at innlærerspråket vil variere etter om morsmål og målspåk er like eller ulike på et bestemt område. De fire transfereffektene er altså:

1. **Intragruppe-homogenitet:** Likhet innad i en innlærergruppe.
2. **Intergruppe-heterogenitet:** Forskjeller mellom to ulike innlærergrupper.
3. **Morsmålssamsvar:** Samsvar mellom innlærernes morsmål og innlærerspråk.
4. **Intralingval kontrast:** Variasjon etter om morsmål og målspåk er like/ulike.

Ifølge Jarvis (2010: 181) opererer de fire transfereffektene i par og gir evidens på ulike aspekter av korrelasjonen mellom morsmålsbakgrunn og innlærerspråk. Det første paret (1 og 2) refererer han til som *gruppebaserte* argumenter, mens det andre paret (3 og 4) er *kildebaserte* argumenter. Det første paret sikrer at den gitte adferden er *spesifikk* og *representativ* for en viss morsmålsgruppe, mens det andre sikrer at det faktisk er morsmålet som motiverer adferden, og ikke andre *ikke-lingvistiske* faktorer som også er felles for en gruppe innlærere med samme morsmål.

2.4.3.1 *Gruppebaserte argumenter: homogenitet og heterogenitet*

Kjernen i de gruppebaserte argumentene for transfer er at hvis transfer foreligger, så skal det også foreligge en korrelasjon mellom gruppetilhørighet og lingvistisk adferd (Jarvis 2010: 174). Å undersøke homogenitet og heterogenitet i og mellom ulike morsmålsgrupper innebærer å sammenligne innlærerspråket til innlærere med *samme* morsmål, og innlærere med morsmål som er *ulike* med hensyn til det språklige trekket som studeres. Formålet er å vise at en viss språklig adferd er både spesifikk og representativ for én av innlærergruppene. Ved høy *homogenitet* innad i en gruppe viser innlærerne lik adferd og man kan se på adferden som representativ for morsmålsgruppen. Slik representativitet kan forventes siden innlærerne sitter på mer eller mindre den sammen språklige kunnskapen fra morsmålet. Ved høy *heterogenitet* mellom gruppene har innlærere med ulike morsmål ulik adferd, og en kan da se på adferden som spesifikk for én av gruppene. Dette sikrer at en gitt adferd ikke er noe som fins hos alle innlærere, og utelukker dermed utviklingsmessige eller universelle påvirkningsfaktorer (Jarvis 2000: 254).

Generelt er det slik at sterke transfereffekter indikerer transfer. Det kan likevel være visse tilfeller hvor en effekt kan være svak, men hvor transfer likevel foreligger, eller motsatt, altså tilfeller hvor det tilsynelatende foreligger en transfereffekt, men hvor transfer *ikke* er årsaken. Jarvis (2000: 255) skriver at homogeniteten innad i en gruppe kan være svak, altså utvise stor variasjon, i tilfeller der det er mye variasjon eller mange varianter å velge mellom i morsmålet, noe som gjør at vi kan forvente tilsvarende variasjon i innlærerspråket selv om transfer foreligger. I tillegg kan homogenitet innad i en gruppe være en effekt av målspråkspåvirkning eller tilegnelsesuniversalier, og heterogenitet mellom grupper kan dermed være lav til tross for at det er transfer, om det er slik at de to morsmålene tilfeldigvis fører til den samme språklige adferden (Jarvis 2000: 256-257).

Det er også et poeng at disse to transfereffektene må ses relativt til hverandre: Homogeniteten *mellom* gruppene skal være svakere enn homogeniteten *innad* i hver av gruppene (Jarvis & Pavlenko 2008: 46). Hvordan dette kan måles, kommer jeg tilbake til i kapittel 6.1.4 i forbindelse med undersøkelse 1.

2.4.3.2 *Kildebaserte argumenter: morsmålssamsvar og intralingval kontrast*

Begge de kildebaserte argumentene for transfer innebærer en sammenligning av trekk i morsmålet og innlærerspråket, og har som hensikt å finne evidens for at et trekk som korrelerer med en spesifikk morsmålsbakgrunn, faktisk har sitt opphav i morsmålet språkssystem og ikke i andre faktorer som er felles for språkbrukere med samme morsmål, for eksempel kulturelle eller utdanningsmessige forhold (Jarvis 2000: 258). Transfereffekten *morsmålssamsvar* beskrives som en undersøkelse av hva i morsmålet som har ført til de språklige valgene i innlærerspråket:

[...] involves showing more explicitly what it is in the language users' source-language knowledge (and performance) that has brought about the observed patterns in their recipient-language performance (Jarvis & Pavlenko 2008: 47).

Den andre kildebaserte transfereffekten, *intralingval kontrast*, utfyller denne sammenligningen av morsmål og innlærerspråk ved at man deler språklige trekk inn etter om de er like eller ulike i morsmål og målspråk, og målet er å undersøke om innlærerspråket varierer i tråd med dette. Jarvis (2010: 175-176) skriver at for språklige

trekk der morsmål og målspråk er ulike, forventer man også betydelige forskjeller i innlærerspråket i sammenlignet med trekk som er felles for språkene:

Perhaps the simplest way to find evidence for this type would be to divide features of the target language into those that are congruent and those that are incongruent with the source language, and then look at whether learners' performance on the two types of features in the target language differs significantly.

Evidens for intralingval kontrast kan en dermed finne innen strukturer i de to språkene hvor noen strukturer er felles og andre ikke, men det er også mulig å dele inn i kontraster på kontekstnivå, der språkene har lik bruk av et trekk i noen språklige kontekster, men ikke i andre (Jarvis 2010: 178).

Som ved de gruppebaserte transfereffektene, trenger ikke høy grad av morsmålssamsvar å bety transfer. Det kan også være et resultat av språkuniversalier eller målspråkspåvirkning. Morsmålssamsvar kan i mange tilfeller være vanskelig å finne, både fordi det kan være diffust, eller rett og slett fordi en ikke har det rette lingvistiske redskapet (Jarvis 2010: 169). Som nevnt tidligere i diskusjonen av tverrspråklige likheter og forskjeller, kan effekten av morsmålet være av en mer indirekte type og vanskeligere å sette fingeren på. Derfor trenger ikke tverrspråklig samsvar nødvendigvis å være en direkte språklig korrespondanse mellom morsmål og innlærerspråk, men kanskje heller en forklaring på hva det er i morsmålet som motiverer et spesifikt mønster i innlærerspråket (Jarvis 2000: 255).

Mens det virker åpenbart hvordan sammenligning mellom morsmål og innlærerspråk kan bidra til identifisering av transfer, er effekten *intralingval kontrast* vanskeligere å gripe, og det er ikke lett å se nøyaktig hvordan denne effekten styrker et argument for transfer, slik Jarvis (2010) mener at den gjør. Den foreslåtte effekten ser ut til å være basert på en antakelse om at ulike typer kontrastive relasjoner mellom morsmål og målspråk (likheter, ulikheter og nullrelasjoner) påvirker forekomst av og type transfer på ulike måter, men i motsetning til de andre tre effektene, som handler om rene observasjoner av språkbruken, ser intralingval kontrast ut til å innebære en undersøkelse av hvilke *utfall* transfer har, sett i forhold til målspråket, og nært knyttet til antakelsen om at ulike typer kontrastive relasjoner fører til ulike typer transfer, eller

ulik grad av mestring. Jeg vil komme tilbake til en diskusjon av dette i forbindelse med undersøkelse 2 (se 8.5.3.4).

2.4.4 Samspillet mellom transfereffektene

De to evidensparene, gruppebaserte og kildebaserte, danner sammen et valid argument for transfer og det er viktig at transfereffektene ikke undersøkes i isolasjon (Jarvis 2000: 255). Hver for seg kan de fire effektene foreligge uten at det er snakk om transfer, og de utgjør kun et argument om de undersøkes sammen.

Intragruppe-homogenitet sikrer representativiteten til adferden innad i morsmålsgruppen, men kan ikke utelukke universalier eller utviklingstrekk uten å bli sett i sammenheng med intergruppe-heterogenitet for å utelukke at dette er en adferd som er felles for alle innlærere uavhengig av morsmål. Videre evaluerer de gruppebaserte argumentene kun i hvilken grad en spesifikk adferd korrelerer med gruppetilhørighet, og kan ikke sikre argumentet om transfer uten at de kildebaserte argumentene viser mer eksplisitt hva i morsmålet som motiverer adferden. Dette utelukker andre faktorer knyttet til en morsmålsgruppe, men som ikke har med språkssystemet i morsmålet å gjøre, for eksempel målspråkstilegnelse.

Jo flere effekter man kan finne, og jo sterkere hver effekt viser seg å være, jo mer solid er argumentet for transfer. Ideelt sett skal alle fire effektene undersøkes, men Jarvis (2010: 181) understreker at *minst én fra hvert par* bør bekreftes for å ha et gyldig argument for transfer, altså at man har evidens for at språklig adferd er både gruppebasert og morsmålsbasert:

It is certainly not necessary to present every possible pairing, but one should be skeptical of any argument for the presence or absence of CLI that does not use these types of evidence in combination with one another in a way that, at the minimum, establishes (either quantitatively or qualitatively, or both) whether the target-language behaviour in question is a group-based phenomenon and whether it is also a source-language-based phenomenon.

Med dette som et minimum kan en påvise transfer ved å minst etablere 1) et forhold mellom språklig adferd og morsmålsgruppetilhørighet, og 2) et forhold mellom den språklige adferden og morsmålet språkssystem. Men siden formålet med transferstudier

ikke bare er å identifisere transfer, men også å lære mer om fenomenet, er forslaget til Jarvis å undersøke alle så langt det er mulig (Jarvis 2000: 255).

Ideelt sett bør også alle fire effektene undersøkes ved å se på *performansedata* både fra morsmålet og innlærerspråket, og morsmålsdataene bør komme fra de samme individene som står for innlærerspråksdataene, siden transfer er noe som skjer på individnivå (Jarvis & Pavlenko 2008: 49). Slike data er ikke alltid mulig å samle inn, og Jarvis og Pavlenko (2008: 48) skriver at bruk av eksterne data, som tidligere studier eller andre informasjonskilder, også er en mulighet.

Jeg vil komme tilbake til hvordan jeg har gått fram for å undersøke transfereffekter i kapittel 6 og 8.

2.4.5 Kontroll på variabler

Det siste av kriteriene i Jarvis' rammeverk er det mer generelle metodiske kravet om at andre variabler skal kontrolleres for, både fordi transfer da kan isoleres som den faktiske påvirkningsfaktoren, og fordi man kan undersøke samspillet mellom transfer og andre faktorer. Dette kan være enten hvor sterkt transfer påvirker i forhold til de andre faktorene, eller hvordan andre variabler forsterker eller modererer transfer (Jarvis & Pavlenko 2008: 52). Jarvis presenterer en liste over variabler som man empirisk har vist at samspiller med transfer, som jeg har oppsummert i tabell 2-1 under. Jarvis deler dem inn i *innlærervariabler*, *lingvistiske variabler* og *kombinerte variabler*, hvor den siste typen berører både innlæreren og lingvistiske forhold (Jarvis 2000: 260-261).

Tabell 2-1 Oversikt over variabler som samspiller med transfer

Innlærervariabler	Kombinerte variabler	Lingvistiske variabler
Alder	Språkbakgrunn	Språklig distanse mellom morsmål og målspråk
Personlighet, motivasjon, anlegg for språk	Type og mengde språklig inntutt	Oppgavetype og lingvistisk nivå
Sosial bakgrunn Utdanning	Målspråksferdigheter	Hvor prototypisk eller markert det språklige trekket er

Som nevnt i 2.2.3, kan slike variabler enten påvirke hvilke mentale assosiasjoner innlærerne danner seg mellom morsmål og målspråk, såkalte læringsrelaterte effekter, eller variablene kan påvirke mengden og typen transfer i faktisk språkbruk, såkalt performanse-relaterte effekter (Jarvis & Pavlenko 2008: 175). Dette betyr at de nevnte

variablene kan både tildekke og interagere med transfer, og de kan enten påvirke den tverrspråklige identifikasjonen direkte, eller om transfer vil forekomme i en gitt språkbrukssituasjon.

Jarvis påpeker at alle disse faktorene burde kontrolleres for i en transferstudie, men at ingen studier så langt (per år 2000) faktisk har klart å kontrollere for samtlige (Jarvis 2000: 261). Jarvis og Pavlenko (2008: 210) påpeker dessuten at dette er en liste som vil utvide seg med årene ettersom vi oppnår mer kunnskap om hvordan transfer fungerer, noe som gjør det til en vanskelig oppgave å kontrollere for alle.

Det er ulike måter å kontrollere for en variabel på, og Jarvis og Pavlenko (2008: 53) nevner fire: 1) eliminere variabelen, 2) holde variabelen konstant, 3) danne en tilfeldig eller lik distribusjon av variabelen, eller 4) undersøke variabelen aktivt. Ved hjelp av én eller flere av disse metodene kan en gjøre innlærergruppene mer sammenlignbare og utelukke andre faktorer enn morsmålet som kilde til språklige mønster i innlærerspråket.

2.5 Oppsummering av kapittel 2

Jeg har i dette kapittelet gitt en oversikt over fenomenet morsmålspåvirkning (transfer) både fra et teoretisk og et metodisk perspektiv. I avhandlingsstudien vil spesielt spørsmålet om transfer som resultat av nullkontrast og antakelser om nullrelasjon mellom morsmål og målspråk være viktig, ettersom jeg har fokus på innlærere som ikke har grammatisk bestemthet i morsmålet. Jeg vil komme tilbake til diskusjonene av de fire transfereffektene i forbindelse med både undersøkelse 1 (se 6.1.4 og 6.5.2) og undersøkelse 2 (8.5.3).

3 GRAMMATISK BESTEMTHET OG REFERANSE

Grammatisk bestemthet er en kategori som på langt nær fins i alle språk. Likevel regner man deler av det innholdet som bestemthet uttrykker, som universelt, og man snakker gjerne om *definitthet*, som fins i alle språk, og enten kan være et underforstått og pragmatisk fenomen eller komme til uttrykk gjennom elementer som er iboende definite, for eksempel pronomer, demonstrativer eller tema-markører (C. Lyons 1999: xv). Det er ikke noe en-til-en-forhold mellom den språkspesifikke, grammatiske bestemtheten og den universelle semantiske/pragmatiske definittheten. De relaterer seg begge til *referanse* og *identifiserbarhet*, men utover dette har grammatisk bestemthet og bruk av bestemte og ubestemte former diakront spredt seg til å gjelde mer enn bare det referensielle, og spredningsmønsteret er noe forskjellig i ulike språk. Dette betyr at grammatisk bestemthet er en grammatisk kategori som bare delvis knytter seg til universelle semantiske forhold, og som bare delvis er enhetlig på tvers av språk som har denne kategorien grammatikalisert.

I dette kapitlet vil jeg hovedsakelig redegjøre for tre forhold: for det første den grammatiske kategorien bestemthet (3.1), for det andre det pragmatiske begrepet referanse (3.2), og for det tredje det som kan kalles *bestemthetsekvivalenter*, altså elementer i språk uten grammatisk bestemthet, som hevdes å uttrykke noe av det samme som den grammatiske kategorien (3.3).

3.1 Bestemthet som grammatisk kategori

Selv om det fins ulike lingvistiske tilnæringer til bestemthet, ser de fleste ut til å enes om at bestemthet er nært knyttet til begrepet *referanse* og referenters *identifiserbarhet*, altså om taleren (senderen) anser en substantivfrases referent som identifiserbar eller ikke for mottageren av ytringen. Om identifiserbarhet skriver C. Lyons (1999: 5-6) at bestemt form signaliserer at tilhøreren er i stand til å identifisere referansen til substantivet:

The idea is that the use of the definite article directs the hearer to the referent of the noun phrase by signalling that he is in a position to identify it.

Forskjellen mellom en identifiserbar og en ikke identifiserbar referent er illustrert i de to setningene under, hvor mottageren av setning 3.1-a forventes å kjenne til den omtalte avhandlingen, som dermed er identifiserbar, mens dette ikke gjelder for setning 3.1-b:

3.1-a Hilde skriver på avhandlingen.

3.1-b Hilde skriver på en avhandling.

I 3.1-a har substantivet «avhandling» *definit referanse*, som typisk uttrykkes med bestemt form, mens substantivet i 3.1-b har *indefinit referanse*, noe som typisk uttrykkes med ubestemt form.

Definit referanse handler altså om at en referent er identifiserbar, eller *kjent*, for tilhøreren (mottageren av ytringen), og et viktig aspekt av dette er på hvilket grunnlag en referent kan regnes som identifiserbar. De aller vanligste domenene for etablering av definit referanse er *anafori*, *situasjon*, *kunnskap om verden* og *assosiasjon* (C. Lyons 1999: 4, 158). Anaforisk bestemthet innebærer at referenten er kjent gjennom å ha vært nevnt tidligere i samtalen eller teksten, som «boken» i setning 3.1-c (under). Situasjonell bestemthet innebærer at referenten er identifiserbar gjennom de umiddelbare fysiske omgivelsene, som «boken» i 3.1-d, eller det kan være snakk om en videre situasjon hvor referenten ikke er direkte synlig, men hvor den kan identifiseres innen et visst område, for eksempel i et hus, et lokalsamfunn eller et land (C. Lyons 1999: 4). Dette er vist med de ulike alternativene i 3.1-e. Andre referenter kan være identifiserbare på grunnlag av generell kunnskap om verden, ofte fordi de er den eneste av sitt slag, altså at de er absolutt *unike*, som «sola» i 3.1-f (C. Lyons 1999: 4). Referenter kan også være unike utelukkende gjennom sitt beskrivende innhold (Hagen 1998: 52). Dette foreligger når substantivfrasen er så informativ at den utelukker alle andre referenter og frasens referent dermed må være unik (3.1-g, Hagens eksempel). Den siste hovedtypen, assosiativ bestemthet, strekker seg over de tre andre hovedtypene, og blir også omtalt som *slutningsbasert*. Dette er referenter som introduseres indirekte via andre referenter, og de er identifiserbare fordi de følger med på lasset som en assosiasjon til en annen referent. I eksempel 3.1-h er det mulig å identifisere «døren» gjennom referenten «huset», fordi vi vet at alle hus normalt har dører. «Døren» kan dermed introduseres i bestemt form uten å være nevnt direkte tidligere.

- 3.1-c Jeg fikk en bok hos broren min. Boken var veldig bra.
3.1-d Gi meg boken!
3.1-e Kanskje du kan ta en tur på kjøkkenet / til tannlegen / til slottet.
3.1-f Se, der skinner sola igjen!
3.1-g Dyra i Hakkebakkeskogen lever i fred og fordragelighet.
3.1-h Inne i skogen kom hun til et hus. Hun banket på døra.

Om en referent ikke er identifiserbar gjennom et av disse domeneene, er det som regel snakk om indefinit referanse, og man vil da typisk *introdusere* referenten ved hjelp av ubestemt form, slik vi ser med «en bok» i den første setningen og «et hus» i den siste.

3.1.1 Grammatisk bestemthet – en heterogen kategori

Selv om identifiserbarhet av de fleste regnes som kjernen av bestemthet, er bildet mer komplekst enn som så. Grammatisk bestemthet handler ofte om mer enn identifiserbarhet, og det er interne forskjeller mellom språk som har denne kategorien. Chesterman (1991: 39) uttrykker at enten må bestemthet være *mer* enn bare referanse, eller så har bestemthetsmarkører med mer enn bare definitthet og identifiserbarhet å gjøre. Det er derfor vanlig å regne grammatisk bestemthet som en relativt heterogen kategori som dekker flere områder, som alle er nært beslektet med referanse.

Grammatisk bestemthet i verdens språk har gjerne utviklet seg langs to dimensjoner: 1) ved at omfanget for kjerneinnholdet har utvidet seg, altså fra anaforisk og situasjonell bruk til andre typer identifiseringsgrunnlag, og 2) ved at kategorien har utvidet seg til å handle om *mer* enn spesifikke referenters identifiserbarhet, til for eksempel referanse til *typer* av referenter heller enn individuelle referenter (generisk referanse) og kontekster *uten* referanse (ikke-referensiell bruk) (C. Lyons 1999: 278).

Når det gjelder utvidelsen til flere typer identifiseringsgrunnlag, er det uenighet om hva som er det mest grunnleggende og opphavelige identifiseringsgrunnlaget for definit referanse. Noen mener det er anafori, mens andre mener at det er den situasjonelle bruken, som er overlappende med demonstrasjon (C. Lyons 1999: 160). C. Lyons (1999: 161) foreslår at anaforisk bruk har utviklet seg fra demonstrasjon, og at anafori senere har kommet til som en form for temporal utpekning, der en peker tilbake i tid (innen teksten/samtalen), heller enn ut i rommet slik en gjør ved demonstrasjon og situasjonell bestemthet.

Utvidelsen av bestemthetskategorien utover definitt/indefinit-distinksjonen, finner vi framfor alt ved bruk av (u)bestemte former i substantivfraser som enten er *ikke-referensielle*, eller som har *generisk referanse*. I begge tilfeller er identifisering av spesifikke referenter ikke relevant. Ikke-referensielle uttrykk er først og fremst en beskrivelse av en referent og viser dermed ikke til en spesifikk referent som skal identifiseres (3.1-i, under), mens generisk referanse viser til klasser eller hele mengder av en type referenter, og det er derfor ikke noen spesifikk referent som skal identifiseres (3.1-j). En annen type substantivfraser har det Hagen (1998: 61) kaller *idiomatisk bruk* av bestemt form, hvor referenten ikke er entydig identifiserbar (3.1-k).

3.1-i Magne er rektor.

3.1-j Pangolinen er utrydningstruet.

3.1-k Sjefen fikk flis i fingeren.

Det siste eksemplet viser hvordan «fingeren» her kan være identifiserbar gjennom kunnskap om at de fleste mennesker har fingre, selv om vi ikke kan identifisere nøyaktig hvilken finger det er snakk om. Dette er en type identifiserbarhet som kommer som konsekvens av at personer har visse ting som hører til deres personlige sfære, for eksempel kroppsdeler, slektninger eller andre ting folk vanligvis har, også omtalt som «inalienable possession» (Halmøy 2010: 55). Dette er altså fraser som innebærer possessive forhold («fingeren min»), men hvor det possessive determinativet ikke trenger å være til stede i norsk (Halmøy 2010: 55).

Jeg vil videre gå nærmere inn i begrepet referanse.

3.2 Referanse

Begrepet referanse er sentralt for forståelsen av bestemthet, og det er også et sentralt begrep i den metodiske gjennomføringen av undersøkelse 2 i denne avhandlingen, derfor vil jeg nå gå dypere inn i dette begrepet.

Ulike faglige tilnærminger opererer med ulike definisjoner og syn på hva referanse er. Referanse i den logiske tradisjonen er en relasjon mellom et referensielt uttrykk og en ekte eller imaginær referent i verden, mens referanse i den pragmatiske tradisjonen ikke er en egenskap ved uttrykket i seg selv, men noe en *gjør* med uttrykket (J. Lyons 1977: 177). J. Lyons (1977: 177) uttrykker at referanse er en *handling* som taleren/senderen utfører ved hjelp av et passende uttrykk:

It should be noted that, according to this conception of the relation of reference, it is the speaker who refers (by using some appropriate expression): he invests the expression with reference by the act of referring.

Diskusjonen om hva referanse er, kommer ikke til å være et vesentlig tema her, men det er viktig å understreke at referanse, i likhet med identifiserbarhet, er universelt. Man refererer i alle språk, men språk med grammatisk bestemthet kan sies å ha grammatikalisert deler av det som går inn under referanse (C. Lyons 1999: 226). I språk uten grammatisk bestemthet går identifiserbarhet fram på andre måter, noe jeg vil komme inn på i slutten av kapitlet, hvor jeg omtaler bestemthetsekvivalenter i språk uten grammatisk bestemthet.

Jeg vil under gjøre rede for fire referensielle undertyper som er avgjørende for valg av bestemthetsmarkører i norsk og andre språk med grammatisk bestemthet: *indefinit*, *definit*, *generisk* og *ikke-referensiell*. Jeg vil se på både hva disse referensielle typene innebærer og hvor grensene går mellom dem.

3.2.1 Referensiell bruk

Referensiell bruk av substantivfraser innbefatter *generisk referanse* og *individuell referanse* (Hagen 1998: 51). Ved individuell referanse viser man til én eller flere spesifikke og individuelle enheter eller objekter, som pinnsvinet i setning 3.2-a under, mens man ved generisk referanse viser mer generelt til en hel klasse av objekter, som pinnsvinene i setning 3.2-b. I den første setningen vises det til et individuelt pinnsvin, mens den andre setningen er en uttalelse om alle pinnsvin.

3.2-a Vilde så et pinnsvin i hagen.

3.2-b Pinnsvin er sjenerte dyr.

Generisk referanse er av C. Lyons (1999: 179) definert som når man refererer til en hel klasse eller generaliserer over en klasse i sin helhet: «Generic noun phrases are those in which reference is made to an entire class, or, perhaps more accurately, which are used to express generalizations about a class as a whole [...]»

Individuell referanse er i denne avhandlingen brukt som en fellesbenevnelse for definit og indefinit referanse, som er referanse til én eller flere individuelle entiteter som av taleren/senderen regnes som enten identifiserbare for mottageren (definit) eller

ikke identifiserbare for mottageren (indefinit) (Hagen 1998: 52).¹³ Siden forskjellen mellom definit og indefinit referanse allerede er omtalt over i 3.1, vil videre fokusere på generisk referanse og ikke-referensiell bruk.

3.2.1.1 *Generisk referanse*

Ved generisk referanse uttrykker man seg om klasser av objekter heller enn individuelle objekter, og det kan virke enkelt å holde generisk og individuell referanse fra hverandre, men skillelinjen mellom dem varierer med hvordan en definerer generisitet. Carlson og Pelletier (1995: 2-3) påpeker en viktig distinksjon mellom to ganske ulike fenomener som begge tradisjonelt anses som generiske. Det første fenomenet er *referanse til klasse* (eng. 'kinds'), hvor referansen er til klassen i sin helhet, som i 3.2-b over, hvor påstanden om sjenerteth gjelder *alle* pinnsvin. Det andre fenomenet kalles *karakteriserende setninger*, og innebærer generaliseringer om hendelser og fakta, som i følgende to setninger:

3.2-c Pinnsvin legger ikke egg.

3.2-d Pinnsvin kommer til hagen vår hver natt.

Både 3.2-c og 3.2-d (over) er generiske utsagn, men er forskjellige med tanke på hvilket nivå de er generiske: 3.2-c er generisk på nominalt nivå og generaliserer dermed over objekter, mens 3.2-d er karakteriserende på setningsnivå og generaliserer dermed kun over hendelser (Carlson and Pelletier 1995: 4). Pinnsvinene som kommer til hagen er spesifikke, individuelle pinnsvin, selv om hendelsen har noe generelt og habituet ved seg, mens manglende egglegging gjelder for alle pinnsvin.

Når jeg senere (undersøkelse 2) deler inn datamaterialet etter referensiell type, vil jeg ikke behandle individuelt refererende fraser i karakteriserende setninger som generiske. Substantivfraser som den i 3.2-c vil derfor regnes som generiske, mens fraser som den i 3.2-d vil regnes som indefinitte.

Det vanlig å regne med flere undertyper av generisk referanse. Dette er typer som representerer ulike måter å snakke om og konseptualisere en klasse på, og valg av form på substantivet ved generisk referanse styres til en viss grad av dette. Jeg vil her kort si

¹³ Termen «individuell referanse» er hentet fra Hagen (1998: 51), og brukes om det som noen omtaler som «ikke-generisk» eller «spesifikk referanse».

litt om *taksonomisk generisk referanse, referanse til klasse og generaliserende referanse*.

Taksonomisk refererende generika fungerer som en konstituerende definisjon av en klasse (Faarlund et al. 1997: 292; C. Lyons 1999: 183; Hagen 1998: 60). Det refereres da til egenskaper som er iboende eller definerende for en klasse, som i eksempel 3.2-e (under). *Generaliserende generika* uttrykker en generell tendens for en klasse, og utsagnet er sant for alle individ i klassen, som i 3.2-f. Ved referanse til klasse har man å gjøre med utsagn om klassen som helhet, men som *ikke* gjelder for alle de individuelle medlemmene i klassen, som i 3.2-g.

3.2-e En potet er en grønnsak. (taksonomisk)

3.2-f Poteter inneholder proteiner. (generaliserende)

3.2-g Poteten kom til Europa med båt. (klasse-referanse)

Den første setningen viser en slags definerende egenskap ved poteter, og mens den andre setningen er en generalisering om alle individuelle poteter *og* klassen som helhet (alle inneholder proteiner), er den tredje setningen bare et utsagn om poteten som klasse, men ikke om alle individuelle poteter (Halmøy 2010: 62-63, eksemplene over er basert på Halmøy).

3.2.2 Ikke-referensiell bruk

Et annet aspekt av referanse er det komplementære begrepet *ikke-referensialitet*. Substantivfraser kan som vi har sett, brukes til å referere til objekter, substanser og abstrakte begrep, men de kan også brukes *uten* å referere, og heller til å *beskrive* en referent. Et eksempel er setningen «Magne er rektor», hvor substantivet «rektor» ikke viser til en rektor, men til en egenskap ved Magne, på samme måte som adjektivet «flink» i «Magne er flink».

Declerck (1986) nevner tre tilfeller som ofte regnes som ikke-referensiell bruk av substantivfraser: *predikativ, ikke-spesifikk og attributiv* bruk. De tre følgende setningene illustrerer forskjellen mellom de tre typene slik det kommer til uttrykk i norsk:

3.2-h Magne er en god rektor. (predikativ)

3.2-i Vi trenger en rektor. Han må være flink. (ikke-spesifikk)

3.2-j Drapsmannen må være gal. Han gjorde grusomme ting. (attributiv)

Declerck (1986: 27) skriver om predikative fraser at de ikke viser til en referent, men tildeler referenten en karakteristikk eller kvalitet, noe som gjør dem tydelig ikke-referensielle. En predikativ substantivfrase som «en god rektor» i 3.2-h (over) er altså ikke-referensiell siden den ikke viser til noen referent, men heller beskriver subjektet «Magne». *Ikke-spesifikk og attributiv bruk* er derimot ikke like tydelig ikke-referensielle, men mer i grenselandet ettersom de faktisk viser til referenter, men likevel bare er *svakt* referensielle siden referentene ikke er identifiserbare for den som ytrer seg. I eksempel 3.2-i (over) refereres det til en hvilken som helst referent som tilfredsstillende beskriver «rektor», men senderen har ikke en spesifikk rektor i tankene, derfor kan ikke rektoren identifiseres. I 3.2-j vises det til en spesifikk person, en drapsmann, men siden drapsmannen ikke er kjent for noen, er han heller ikke identifiserbar.

Ikke-spesifikke og attributive substantivfraser er også ulike predikative ved at de kan introdusere nye referenter. Dette kan man for eksempel se på at det er mulig å bruke pronomen (han) i de etterfølgende setningene for å referere til både rektoren og drapsmannen i 3.2-i og 3.2-j, noe som ikke er mulig ved predikativ (Declerck 1986: 27).

Hvorvidt en substantivfrase introduserer en ny diskursreferent eller ikke, er også viktig for å skille predikativ bruk fra en annen substantivtype som opptrer etter kopulaverb, nemlig de *spesifiserende*, som spesifiserer en referent (Declerck 1986: 27). Både i det tidligere eksemplet «Magne er en god rektor» og i 3.2-k (under) er substantivfrasen etter kopulaverbet «er» en beskrivelse av subjektet, men i 3.2-k er det i tillegg en identitetsrelasjon mellom de to frasene. Både «Magne» og «rektoren jeg anbefalte» har begge referanse til et spesifikt individ, men referansen er sammenfallende:

3.2-k Magne er rektoren jeg anbefalte.

Som vi har sett, er hovedforskjellene mellom predikativ bruk og de andre bruksområdene at man ved predikativ bruk *beskriver* og ikke introduserer nye referenter, noe ikke-spesifikk og attributiv bruk gjør. På grunn av denne forskjellen er det også vanlig å si at ikke-spesifikk og attributiv bruk ikke ses på som ikke-referensielle, men

at de heller har *svak referanse* (Declerck 1986: 26). Mellom den tydelig ikke-referensielle bruken en finner ved predikativ, og den helt tydelig referensielle bruken en finner ved definnitt og indefinnitt referanse, fins det altså et grenseområde av svak referanse.

3.3 Bestemthetsekvivalenter

Til tross for at grammatisk bestemthet ikke eksisterer som grammatisk kategori i de fleste språk, er det, som jeg var inne på innledningsvis, visse aspekter av bestemthet som er universelle, og spørsmålet om hvorvidt det fins bestemthetsekvivalenter i språk uten grammatisk bestemthet, er relevant i denne sammenheng. Bestemthetsekvivalenter er språklige trekk i -GB-språk som påstås å korrelere med bestemte og ubestemte former i +GB-språk, selv om det kan diskuteres om disse faktisk *uttrykker* bestemthet, eller om de heller bare *legger til rette for* en bestemthetsfortolkning.

Ideen om bestemthetsekvivalenter er for eksempel uttrykt av Szwedek (1976) i en kontrastiv sammenligning av engelsk og polsk: Selv om språk som polsk ikke har bestemthetsmarkører, mener Szwedek (1976: 203) at man likevel kan slutte seg til om et substantiv har definnitt eller indefinnitt referanse, og at det derfor må være *noe* som legger opp til en slik slutning:

Although there is no article in Polish we seldom have doubts whether a noun in a text is definite or indefinite. It is obvious that there must exist linguistic mechanisms in Polish which perform the function parallel to the function of the article in English.

Den vanligst nevnte ekvivalenten til bestemthet er *pragmatisk ordstilling*, altså organisering av setningen etter *tema/rema-prinsippet*. Kort sagt er ideen at setninger er delt inn i to hovedkomponenter: *tema*, som er *kjent* informasjon, og *rema*, som er *ny* informasjon. Tema/rema samspiller med leddstillingen i en setning på den måten at tematiske ledd har en tendens til å komme tidlig i setningen og danne et fundament for setningens hovedinnhold, mens rematiske ledd kommer mot slutten av setningen og sier noe om temaet (Firbas 1992: 118; Hagen 1998: 260). Dette gjelder ikke bare nominale ledd i setningen, men jeg vil her konsentrere meg om det som angår substantivfraser som tema og rema.

Et tematisk ledd er alltid *kjent* eller *identifiserbart*. Derfor er det stor grad av sammenfall mellom tematiske substantivfraser i -GB-språk, og definte substantivfraser med bestemt form i +GB-språk. Rematiske ledd vil tilsvarende være *ukjente* og *uidentifiserbare*, og sammenfaller ofte med indefinitte substantivfraser med ubestemt form. På grunn av dette sammenfallet mener man at substantivets plassering i setningen i -GB-språk kan indikere referentens status som defintitt eller indefintitt, slik bestemte og ubestemte former gjør i +GB-språk. Setningsinitial plassering av en substantivfrase signaliserer defintitt referanse, mens setningsfinal plassering signaliserer indefintitt referanse (Szwedek 1976: 208)

Trenkic (2002: 114) er kritisk til denne sammenkoblingen, og hevder at begrepene *kjent* (tema) og *ny* (rema) ikke egentlig er ekvivalente i språk med og uten grammatisk bestemthet. Hun mener *kjent/ny* i -GB-språk er en *bevissthetskategori*, mens *kjent/ny* i +GB-språk er en *kunnskapskategori*. Forskjellen ligger i at pragmatisk ordstilling innebærer en vurdering av hvilke referenter mottageren har i bevisstheten i ytringsøyeblikket, mens grammatisk bestemthet er en vurdering av hvilke referenter mottageren har *kunnskap* om:

The main error in identifying given and new information with definiteness/indefiniteness, lies in the fact that definiteness is related to 'given' and 'new' as a category of knowledge, while word order and information structuring are related to 'given' and 'new' as a category of consciousness. Word order reflects the speaker's assumptions about the hearer's state of consciousness at the time of utterance, rather than his assumptions of the hearer's state of knowledge.

Det referensielle identifiseringsgrunnlaget for tematiske ledd ved pragmatisk ordstilling er med andre ord først og fremst basert på om en referent er *tekstuel*t identifiserbar, altså anafori, og i mindre grad basert på situasjonell eller kunnskapsbasert identifiserbarhet, slik grammatisk bestemthet kan være. Det vil si at selv om både pragmatisk ordstilling og grammatisk bestemthet forholder seg til identifiserbarhet, så framhever de ulike aspekter: Ordstillingen følger et *aktiveringsprinsipp* som tilsvarer anaforisk bestemthet, mens grammatisk bestemthet følger et bredere *identifiseringsprinsipp* hvor referenter identifiseres som definte også på andre grunnlag enn anafori. Det vil si at referenter som kan regnes som kjente ved grammatisk bestemthet, ikke nødvendigvis kan regnes

som kjente ved pragmatisk ordstilling. Ved pragmatisk ordstilling blir dermed *kjent* et smalere begrep enn ved grammatisk bestemthet, mens *rema* blir et videre begrep.

Det er også stilt spørsmål ved om informasjonsstruktur er det faktiske ordstillingsprinsippet i -GB-språk. Siewierska (1993) viser at forhold som prosesseringstyngde og leddtyngde også kan forklare hvilket ledd som kommer først og sist i polske setninger. Hun mener at kanskje informasjonsstruktur og prosesseringstyngde *jobber sammen*, heller enn at det bare er informasjonsstruktur som styrer ordstillingen. Dette er plausibelt ettersom tematiske ledd har en tendens til å være lette, mens rematiske ledd er tyngre.

Andre antatte bestemthetsekvivalenter er *demonstrativer*, som er en naturlig kandidat i og med at bestemthetsmarkører diakront mest sannsynlig har utviklet seg fra demonstrativer, som er iboende definitte (C. Lyons 1999: 107). Semantisk sett er demonstrasjon det området der grammatisk bestemthet og demonstrativer overlapper, altså gjennom identifisering i den situasjonelle konteksten. Til tross for at demonstrativer fins i alle språk, er det likevel grunn til å hevde at demonstrativer ikke i seg selv *uttrykker* bestemthet, men at de understreker en allerede underforstått bestemthet (C. Lyons 1999: 236).

Også et samspill mellom demonstrativer og pragmatisk ordstilling har blitt regnet som bestemthetsekvivalent. Demonstrativer kan nemlig brukes for å markere at setningsplasseringen til en substantivfrase *bryter* med tema/rema-prinsippet, altså om en rematisk substantivfrase plasseres initialt eller en tematisk substantivfrase plasseres finalt i setningen. Venneri (2011: 207) skriver at om frasene skal få sin rette fortolkning som tema (definit) eller rema (indefinit) når de ikke står i sin typiske setningsposisjon, må dette signaliseres eksplisitt gjennom bruk av tallord eller demonstrativer:

[...] in order to be interpreted as indefinite a nominal phrase in topic position has to be explicitly marked for indefiniteness, for example with *jeden* 'one'. In contrast, the bare noun in the comment position is interpreted as indefinite [...] For definite interpretation of a noun phrase in the comment a definite determiner is required, for example the demonstrative pronoun *ten* [...].

Spørsmålet om ekvivalenter er interessant, men kan kritiseres for å være *translatørgrammatisk*, altså det at en beskriver et språks oppbygging innenfor rammene til et annet språk, og ikke på egne premisser (Berggreen & Tenfjord 1999: 204).

Chesterman (1991: 4) skriver for eksempel at det egentlig er på bakgrunn av språk *med* grammatisk bestemthet at man i det hele tatt snakker om at det skal finnes bestemthet i språk uten. Bestemthetsekvivalenter har likevel vært trukket fram som et mulig område for morsmålspåvirkning hos innlærere med morsmål uten grammatisk bestemthet (Young 1996; Ekiert 2004, 2007, 2010), noe jeg vil komme tilbake til i kapittel 4. Dette er også ett av utgangspunktene for hypotesene i undersøkelse 2 i denne avhandlingen.

3.4 Tall og tellelighet

Avslutningsvis i dette kapittelet er det to andre relevante sider ved substantivfraser som er relevante å si noe om, nemlig *tall* og *tellelighet*.

Den grammatiske kategorien tall (numerus) er i utgangspunktet et gjennomiktig begrep, hvor *entall* viser til enkeltreferenter og *flertall* viser til to eller flere referenter. Det er likevel et viktig poeng at dette gjelder bare for individuell og ikke-referensiell bruk. Ved generisk referanse avspeiler nemlig ikke valget mellom entall- eller flertallsform antall i seg selv, ettersom generisk referanse *alltid* involverer flere referenter. Ifølge Halmøy (2010: 71) handler valget mellom entall og flertall ved generisk referanse om ulike måter å konseptualisere flertallsmengder på; enten som *ett sett* av individer (entall) eller som mange individer (flertall).

En annen kategori, som er nært knyttet til tall, er *tellelighet*, altså skillet mellom tellelige og ikke-tellelige substantiv. Tellelige substantiv viser til enheter som kan telles, mens ikke-tellelige substantiv viser til abstrakte begreper eller substanser som typisk ikke kan telles. Dette blir ofte behandlet som en fast *leksikalsk* egenskap ved et substantiv, som er nært knyttet til referenten i seg selv. Som flere forskere har påpekt, er dette problematisk, og tellelighet kan alternativt regnes som en *kontekstuell* egenskap, hvor tellelig/ikke-tellelig handler om fortolkning og kan endre seg fra kontekst til kontekst, heller enn å være en fast egenskap (Chesterman 1991: 7; Borthen 2003: 12). Dette kan ses på at substantiv som typisk anses som tellelige, likevel kan få en ikke-tellelig fortolkning i visse kontekster, som i 3.4-a under, og tilsvarende kan typisk ikke-tellelige substantiv bli individualiserte og tolket som tellelige, som i 3.4-b (Faarlund et al. 1997: 145-148).

3.4-a Frode har hund i fryseren.

3.4-b Det fins mange forskjellige intelligenser

Hvilke substantiv som regnes som tellelige og ikke-tellelige, kan også til en viss grad være språkspesifikt. For eksempel hevder Chesterman (1991: 7-8) at mens hunder i vår del av verden typisk ses på som objekter og tellelige, er dette også *en matrett* i andre deler av verden, noe som gjør det mulig å oppfatte hund som noe ikke-tellelig: «Thus in our culture *dog* is normally thought of as a discrete object and hence count, but in another culture *some dog* might denote desirable food and thus become mass noun.»

3.5 Oppsummering av kapittel 3

Dette kapittelet har beskrevet den grammatiske kategorien bestemthet og knyttet den til begrepene referanse og identifiserbarhet. Jeg har dessuten beskrevet fire ulike referensielle typer: *indefinit referanse*, *definit referanse*, *generisk referanse* og *ikke-referensiell bruk*. I kapittel 7 vil jeg komme nærmere inn på hvordan disse kommer til uttrykk i norsk, engelsk og polsk, og en inndeling etter referensiell type er sentral for undersøkelse 2 med hensyn til å avdekke transfereffektene intergruppe-heterogenitet og intralingval kontrast (se 8.5.3.5).

Bestemthetsekvivalenter i språk uten grammatisk bestemthet er noe jeg også vil komme tilbake til både i omtalen av andrespråksforskning på bestemthet i kapittel 4, og som utgangspunkt for en av hypotesene i undersøkelse 2.

4 BESTEMTHET I ANDRESPRÅKSFORSKNINGEN

Jeg vil i dette kapittelet se nærmere på tidligere forskning som har sett på læring og bruk av bestemthetsmarkører hos andrespråksinnlærere. Et sentralt tema i andrespråksstudier av bestemthet har vært innlærere med morsmål uten grammatisk bestemthet, og hvordan disse tilegner seg og tar i bruk bestemthetsmarkører når målspråket har bestemthet. Som nevnt i innledningskapittelet, er det en relativt utbredt oppfatning at -GB-innlærere har større vansker med bruk av bestemthetsmarkører enn +GB-innlærere, men det er likevel ulike oppfatninger om hva som er årsaken til dette eller hva som styrer bruken av markørene i innlærerspråk. I dette kapittelet vil jeg se nærmere på tidligere forskning på dette området. Jeg vil ha hovedfokus på -GB-innlærere og hvordan morsmålet deres kan antas å påvirke bruk av bestemthetsmarkører, men jeg vil også komme innom tilnæringer med et mer universelt læringsperspektiv på bestemthet.

En stor del av forskningen på bestemthet i innlærerspråk har rettet seg mot engelsk som målspråk. Innen disse studiene har det vært vanlig å undersøke *tilegnelse av artikler* ('article acquisition') eller *valg av artikkel* ('article choice'), heller enn kategorien grammatisk bestemthet som sådan. Ofte ser man på bruken av *ubestemt artikkel* ('a') og *bestemt artikkel* ('the'), i tillegg til *nullartikkel*, en benevnelse som ofte brukes om substantiv uten eksplisitt artikkel.¹⁴ Man har vært opptatt av når de ulike artiklene tilegnes i forhold til hverandre, altså læringsløyper, eller hvilke forhold som styrer valg av den ene artikkelen framfor den andre. Spesielt har variasjonsmønster med veksling mellom bruk og utelatelse av artikkel vært mye undersøkt i innlærerspråk, i tillegg til ombygging av artikler, ettersom dette kan si oss noe om hvordan innlærerne distribuerer formene sine og om forholdet mellom form og kontekst i innlærerspråket. I en studie av norsk som andrespråk er det ikke i like stor grad naturlig å snakke om «artikler», så jeg vil først og fremst bruke den mer generelle termen «bestemthetsmarkører», eventuelt «bestemt form» og «ubestemt artikkel», med mindre det er spesifikt snakk om engelsk språk eller språket til innlærere av engelsk.

¹⁴ Dette vises gjerne til som 'zero article' på engelsk, og kan dermed oversettes til 'nullartikkel' på norsk. Også uttrykket 'null article' fins i noen tilnæringer til engelsk språk, for eksempel i Chesterman (1991: 16) hvor det viser til fravær av eksplisitt artikkel ved *egennavn*, men ikke ved substantiv generelt.

Selv om engelsk som andrespråk dominerer innen studier av bestemthet, fins noen større nordiske doktorgradsavhandlinger, for eksempel Nordanger (2017) innen norsk som andrespråk, og Axelsson (1994) og Nyqvist (2013) innen svensk som andrespråk. Siden norsk formelt sett er delvis ulikt engelsk, for eksempel i å bruke bestemt suffiks heller enn bestemt artikkel, vil spørsmålet om hvorvidt funn og tilnærming fra studier med engelsk som målspråk er overførbare til norsk som målspråk, også bli diskutert på slutten av kapittelet.

I framstillingen av studier er jeg først og fremst opptatt av de studiene som har forsøkt å si noe om hvordan morsmålet påvirker bruk og tilegnelse av bestemthet hos -GB-innlærere. I den sammenhengen vil jeg imidlertid også komme inn på noen studier som *avviser* at transfer kan finne sted når det foreligger en nullkontrast mellom morsmål og målspråk, og som derfor heller ser på språklige universalier som styrende for valg av form. Jeg vil dessuten rette fokus mot hva tidligere studier sier om morsmålet påvirkning på valg av form og veksling mellom former, og ikke i særlig grad komme inn på korrekthet og læringsløyper.

I 4.1 presenteres en av de første undersøkelsene av engelske artikler som tar i betraktning både artiklenes *form* og *funksjon* i innlærerspråk (Huebner 1983). I 4.2 går jeg gjennom en rekke senere studier av bestemthet både med universell tilnærming, transfertilnærming og med mer sammensatte tilnærminger. I kapittel 4.3 går jeg nærmere inn i en mye omtalt forklaring på mangel på bestemthetsmarkører, kalt *redundansforklaringen*. Til slutt presenterer jeg i 4.4 noen nordiske studier av bestemthet i innlærerspråk og diskuterer overførbarheten mellom studier med norsk og engelsk som målspråk.

4.1 Huebner (1983)

Huebner (1983) var en av de første studiene av bestemthetsmarkører som tok i betraktning både form og funksjon ved å knytte bruk av de ulike artiklene i engelsk til fire ulike *semantiske kontekster*. Definisjonene av semantiske kontekster var hentet fra Bickerton (1975) og basert på fravær og/eller tilstedeværelse av trekkene *±spesifikk*

referanse ([±SR]) og *Assumed Hearer Knowledge* ([±HK]) (Huebner 1983: 133).¹⁵ I tillegg tok Huebner i betraktning om substantivfrasene var tematiske eller rematiske ([±Topic]).¹⁶

I tidligere studier av bestemthet hadde man det man kan kalle en *kun form-*tilnærming, hvor man først og fremst var interessert i hvorvidt bestemthet var tilegnet eller ikke, noe som i praksis var et spørsmål om hvor stor andel av de obligatoriske kontekstene for hver av artiklene som hadde korrekt artikkel (Huebner 1983: 27). Om en for eksempel hadde mer enn 90 % korrekte markeringer, ble formen regnet som tilegnet. Huebner var derimot i større grad opptatt av hva som skjer *før* bestemthet regnes som tilegnet, altså hvilke funksjoner artiklene ser ut til å uttrykke i innlærerspråket uavhengig av korrekthet, og hva dette kan si oss om innlærernes egne hypoteser om målspråket og hvordan disse endres over tid:

The variation in the use of a given form before it reaches criterion may tell us more about the nature of language acquisition than the order in which a given set of morphemes reach the level of acquisition can (Huebner 1983: 27).

Undersøkelsen til Huebner hadde bare én informant, *Ge*, som ble fulgt over ett år. Informanten hadde hmong som morsmål og lærte engelsk uten opplæring (Huebner 1983: 50). Hubners studie skiller seg fra de fleste senere studier, både i å være en kasusstudie og i å undersøke naturlig læring.

Hubners metodiske framgangsmåte innebar også en forskjell fra tidligere studier både i utvalget av kontekster som ble undersøkt, og i hvordan disse ble klassifisert. Hubner tok *alle* substantivkontekster i betraktning, for eksempel tok han med *all bruk* av artikkelen 'the', ikke bare bruk i kontekster hvor bestemt artikkel er påkrevd i målspråket (obligatoriske kontekster).

Kombinasjonen av trekkene innen de to kategoriene ±SR og ±HK gir fire varianter som tilsvarer de fire referansetyperne jeg omtalte i kapittel 3, og som utgjør bruksområdene for de tre engelske artiklene (bestemt, ubestemt og nullartikkel (Ø)):

+ SR -HK Indefinit referanse = Bruk av 'a' og Ø

¹⁵ Kategorien 'Hearer Knowledge' (HK) tilsvarer det som i kapittel 3 ble omtalt som *identifiserbarhet*, altså om taleren (senderen) vurderer at mottageren av ytringen kan identifisere referenten, derav vises det til *tilhørerkunnskap*.

¹⁶ Huebner (1983) bruker termen 'topic' (norsk «emne»), som tilsvarer det som jeg tidligere har kalt «tema».

+ SR	+HK	Definit referanse	= Bruk av 'the'
- SR	+HK	Generisk referanse	= Bruk av 'a', 'the' og Ø
- SR	- HK	Ikke-referensiell bruk	= Bruk av a' og 'the'

Huebner fant en utvikling i bruken av bestemte former styrt av disse trekkene i kombinasjon med tema/rema-status. I begynnelsen ble bestemt artikkel *kun* brukt i definte kontekster, men bare om de i tillegg var tematiske ([+HK +SR +topic]). I neste omgang ble artikkelen brukt i *samtlige* kontekster, noe Huebner omtaler som en *nøytralisering*. Etter nøytraliseringen ble bruk av 'the' systematisk eliminert i de kontekstene som har minst til felles med definit ([+HK +SR]), altså først i [-HK -SR], før bruken til slutt var målspråkslik (Huebner 1983: 204).

Det er dermed hos Huebner snakk om en kontekstbasert utvikling av 'the', som begrenses av kategoriene HK, SR og tema/rema-distinksjonen, og Huebner (1983: 142) mente at denne utviklingen er knyttet til informasjonsorganisering i innlærerspråket. Det at artiklene i starten brukes for å skille mellom tematiske og ikke-tematiske referenter, tolker Huebner som en *topik-prominent* organisering av informasjonen. Han knytter tilstedeværelsen av en slik organisering av innlærerspråket til at informantens morsmål er topik-prominent,¹⁷ og mener at stadiet med nøytralisering er nødvendig for å nullstille det topik-baserte systemet slik at innlærerspråket kan bli mer målspråkslikt. Huebner (1983: 146) understreker at en slik nøytralisering kanskje ikke er nødvendig hos innlærere med andre morsmål.

Et annet viktig poeng for Huebner er at informanten allerede tidlig nådde kriteriet for tilegnelse av bestemt form på tidligere studiers premisser, siden det var utstrakt bruk av 'the' i obligatoriske kontekster, men det tok lenger tid før bruken av 'the' var forsvunnet fra kontekster hvor 'the' er ugrammatisk (Huebner 1983: 137). Dette viser hvordan Huebners studie er *funksjonsorientert*, da det er et hovedmål å se om *både* form og funksjon er lært, noe som innebærer både å bruke artiklene i de obligatoriske kontekstene, og å unngå å bruke dem i kontekster de ikke skal brukes i.

Selv om Huebner tar utgangspunkt i universalier, er han ikke entydig på at utviklingen han finner hos informanten (Ge) er universell for innlærere, men antyder at det like gjerne *ikke* kan være det, og at dette må utforskes videre:

¹⁷ Språk som organiserer setningene sine etter strukturen *topic-comment* (Li & Thompson 1976), på norsk *emne-kommentar*.

Whether the paths Ge followed in arriving at this discovery are universal paths inherent in all language learning or not remains to be determined. It may very well be that they are not (Huebner 1983: 146).

Med dette åpner Huebners studie både for universelle og morsmålsspesifikke forklaringsfaktorer i tilegnelsen av bestemthet hos -GB-innlærere, og den åpner også for at det er en forskjell i læringsgangen hos innlærere med og uten bestemthet i morsmålet.

4.2 -GB-innlærere og bestemthet

Det virker å være en rådende antakelse at grammatisk bestemthet er spesielt vanskelig å tilegne seg for innlærere som ikke kjenner kategorien fra morsmålet. Dette er nok grunnen til at en stor del av studiene på dette området først og fremst er opptatt av nettopp innlærere med -GB-språk som morsmål, for eksempel russisk, polsk, bosnisk-kroatisk-serbisk, tsjekkisk, thai og koreansk, for å nevne noen av morsmålene som forekommer i studiene jeg omtaler under. Det er imidlertid en viktig forskjell mellom noen av disse studiene, nemlig hvorvidt man antar at fraværet av bestemthet i morsmålet kan føre til transfer eller ikke. På den ene siden har man nemlig tilnærminger som avviser at nullkontrasten kan føre til transfer, ettersom man mener at det ikke er noe å overføre fra morsmålet til innlærerspråket så lenge morsmålet mangler bestemthet (for eksempel Chaudron & Parker 1990; Ionin & Wexler 2003; Ionin, Ko & Wexler 2004; Ionin, Zubizarreta og Maldonado 2008). Disse studiene fokuserer på at det er tilgang på kunnskap gjennom språklige universalier som styrer læring og bruk av bestemthetsmarkører. På den andre siden har man tilnærminger som mener at transfer kan forekomme også ved nullkontrast, ettersom det fins andre språklige forhold i morsmålet som kan overføres og påvirke valg av form i innlærerspråket (for eksempel Ekiert 2004, 2007, 2010; Young 1996; Jarvis 2002; Trenkic 2007, 2008, 2009; Trenkic & Pongpaioj 2013; Nordanger 2017). Noen av disse har også sett på hvordan universalier og morsmål virker sammen (Young 1996; Jarvis 2002). Jeg vil under først presentere noen av de studiene som antar at universalier styrer bruk av artikler hos -GB-innlærere av engelsk, før jeg går over til ulike forslag til hvordan morsmålet antas å påvirke når det mangler grammatisk bestemthet.

4.2.1 Lingvistiske og diskursive universalier hos -GB-innlærere

Flere studier avviser muligheten for transfer når morsmålet ikke har grammatisk bestemthet. Ionin og Wexler (2003: 249) skriver for eksempel at de ikke forventer transfer hos innlærere med russisk og koreansk som morsmål, på grunn av fraværet av artikler i morsmålet:

Both Russian and Korean lack articles, and neither language has obligatory marking of definiteness or specificity on DPs. Thus, we did not expect any direct transfer effects.

Chaudron og Parker (1990: 50) viser en lignende tankegang, og skriver at de strukturelle forskjellene mellom engelsk (målspråk) og japansk (morsmål) gjør transfer lite sannsynlig, og de forventer derfor at de japanske innlærerne heller vil ha universelle tendenser:

Because of the structural differences between English and Japanese as to how NPs encode topics [...] we expect that Japanese learners of English will reveal these universal tendencies in development, as a direct transfer of the Japanese system is not likely [...].

Begge de to studiene som er nevnt over, fokuserer altså på om innlæreres bruk av bestemthet er styrt av språklige universalier, men de ser på to ulike typer universalier: Ionin og Wexler (2003) ser på de lingvistiske universalierne *spesifikk referanse* og *identifiserbarhet*, mens Chaudron og Parker (1990) ser på diskursive universalier i form av språklig markering av om referenter er *nye*, *kjente* eller *nylig nevnt*.

4.2.1.1 *Fluktueringshypotesen*

Fluktueringshypotesen er en hypotese som undersøkes i flere ulike studier (Ionin & Wexler 2003; Ionin et al. 2004; Ionin et al. 2008). Denne hypotesen baserer seg på tilsvarende kategorier som Huebner (\pm HK og \pm SR), men da som en antatt del av *den universelle grammatikken* (UG). Utgangspunktet er at når innlærerne ikke har bestemthet i morsmålet, vil det ikke være noe grunnlag for transfer, og de vil da basere seg på de trekkene som er tilgjengelig gjennom UG, altså HK (definit) og SR (spesifikk) (Ionin & Wexler 2003: 249; Ionin et al. 2004: 18; Ionin et al. 2008: 559).

Fluktueringshypotesen har som utgangspunkt at UG har en parameter for artikkelvalg (*Article Choice Parameter*): Verdens språk har sin grammatiske bestemthet styrt enten av hvorvidt referenten er *spesifikk* eller av hvorvidt den er *definit*, derfor vil

UG inneholde begge disse innstillingene og disse vil begge være tilgjengelig kunnskap for alle språkbrukere (Ionin & Wexler 2003: 248; Ionin et al. 2004: 12). Fluktueringshypotesen går ut på at -GB-innlærere vil ha tilgang til disse innstillingene, men vil være usikker på hvilken som gjelder for målspråket, noe som fører til en *fluktuering*, en veksling mellom formene, i visse kontekster. I kontekster hvor de to trekkene har likt fortegn ([+definitiv/+spesifikk] og [-definitiv/-spesifikk]), vil det være en kategorisk assosiasjon mellom én av artiklene og én av kontekstene, som vil si at innlærerne bruker 'the' i den ene konteksten og 'a' i den andre, men det vil variere fra innlærer til innlærer hvilken artikkel som assosieres med hvilken kontekst (Ionin et al. 2004: 17; Ionin et al. 2008: 559). I kontekster hvor de to trekkene har *motsatt* fortegn ([+definitiv/-spesifikk] og [-definitiv/+spesifikk]), vil innlærerne derimot ha større problemer med å velge artikkel, noe som fører til at de noen ganger vil basere valget på trekket *definitiv* og andre ganger på trekket *spesifikk*; en fluktuering som vil vedvare inntil innputt har gitt innlæreren tilstrekkelig informasjon om målspråket og artiklene blir brukt korrekt (Ionin & Wexler 2003: 249; Ionin et al. 2004: 16; Ionin et al. 2008: 574).

Ionin et al. (2008) har også undersøkt om fluktueringshypotesen gjelder for +GB-innlærere. Studien sammenligner russiske (-GB), spanske (+GB) innlærere, og finner at mens de russiske innlærerne har slik fluktuering, så forekommer ikke dette hos de spanske innlærerne (Ionin et al. 2008: 565). De spanske innlærerne har derimot høy grad av korrekthet, og valg av artikkel ser ut til å være basert på kategorien *definitiv*, som i morsmålet og målspråket. Hos de spanske innlærerne vil altså morsmålet eliminere vekk alternativet *spesifikk* (Ionin et al. 2008: 565).

Med dette viser disse ulike studiene om fluktueringshypotesen at mens +GB-innlærere har transfer, så vil ikke dette forekomme hos -GB-innlærere, som da må basere seg på universalier som hjelper dem å «å forstå» innputt. I og med at antall mulige hypoteser om målspråket vil være stort, mener forfatterne at det er lite sannsynlig at større grupper av -GB-innlærere ville falt ned på et så entydig mønster som det som har blitt funnet i disse studiene, dersom universalier ikke gjorde seg gjeldende (Ionin et al. 2008: 570). For -GB-innlærere avvises både transfer og undervisning som forklaringer ettersom det ikke fins språklige forhold i morsmålet som kan forklare mønstrene de

finner, og lærebøker presenterer flere ulike forklaringer på hvordan artikler brukes (Ionin et al. 2008: 570). Forklaringen mener de derimot ligger i samspillet mellom UG og målspråksinnputt: UG bidrar med kunnskap om mulige parameterinnstillinger, og innputt sørger for kunnskap om hvilke som gjelder for det spesifikke målspråket (Ionin et al. 2008: 574).

4.2.1.2 *Diskursive universalier*

Chaudron og Parker (1990) har en annen type universell vinkling på valg av artikler hos japanske (-GB) innlærere av engelsk. De ser på *diskursive* universalier knyttet til referenters informasjonsstatus i løpet av en samtale eller en tekst: Referenter kan være *nye*, *kjente* eller *nylig nevnt* (eng. 'current'). Mens *nye* referenter er referenter som er foreløpig ukjente, og derfor må introduseres, er *kjente* referenter allerede introdusert. Referenter som er *nylig nevnt*, er kjente, men i tillegg er de framme i bevisstheten på et gitt tidspunkt, for eksempel fordi det er den sist nevnte referenten (Chaudron & Parker 1990: 45).

Det antas at det er universelt at brukere av ulike språk vil veksle mellom ulike referensielle uttrykk for å skille mellom referenters varierende informasjonsstatus i løpet av en tekst, for eksempel veksle mellom pronomen og substantiv, eller mellom ubestemt og bestemt form i språk som har dette. Matchingen mellom kontekst og form varierer derimot fra språk til språk, da ulike språk foretrekker ulike former for å uttrykke disse distinksjonene (Chaudron & Parker 1990: 45). I +GB-språk brukes typisk ubestemt form ved nye referenter, bestemt form ved kjente referenter og pronomen ved referenter som er nylig nevnt. I -GB-språk har man derimot ikke muligheten til å skille mellom kjente og nye referenter på denne måten.

Chaudron og Parker kobler både valg av form og de ulike informasjonsstatusene til *markerthet* (strukturell markerthet og diskursiv markerthet), i form av to skalaer som er relatert til hverandre: Strukturelt sett er nullanafori den minst markerte formen, mens eksistensielle setninger og dislokeringer er de mest markerte formene. På den diskursive skalaen er referenter som er *nylig nevnt*, minst markerte, mens de som er *kjente* er mest markerte (Chaudron & Parker 1990: 47). Antakelsen er at siden japanske innlærere ikke har noen mulighet for direkte overføring fra morsmålet, vil de vise universelle tendenser

i sine formvalg gjennom å bruke de minst markerte formene i de minst markerte kontekstene, og motsatt (Chaudron & Parker 1990: 50).

Chaudron og Parker konkluderer med at innlærerne tidlig skiller mellom diskurskontekster ved å variere mellom pronomen, bestemte substantiv og ubestemte substantiv (Chaudron & Parker 1990: 59). Over tid og med økende ferdighetsnivå sprer korrekt bruk av artikler seg fra de minst markerte formene og kontekstene, til de mest markerte (Chaudron & Parker 1990: 43, 60).

Også Jarvis (2002) har tatt for seg de samme diskursive universalierne og undersøkt om svenske (+GB) og finske (-GB) engelskinnlærere er sensitive overfor disse diskurskontekstene, og om engelske artikler blir brukt til å skille mellom dem i innlærerspråket. Mens Chaudron og Parker kun ser på tematiske referensielle uttrykk, undersøker Jarvis både tema og rema, og dermed også om denne tema/rema blir formelt markert av innlærerne, slik Huebner (1983) mente hans informant gjorde (Jarvis 2002: 391). Jarvis (2002: 387) konkluderer med at tema/rema-distinksjonen ikke ser ut til være viktig for valg av form i innlærerspråket, mens forskjellen mellom *ny*, *kjent* og *nylig* ser ut til å være påvirket av hvordan disse distinksjonene kommer til uttrykk i morsmålet:

The quantitative results of the study cast some doubt on learners' sensitivity to the topic-comment distinction and also suggest that learners' tendency to mark distinctions between new, continuous, and reintroduced NP referents is influenced by the prominence of such distinctions in the L1.

Når det gjelder selve valget av form, er Jarvis (2002) mer opptatt av morsmåletts effekt, derfor vil denne studien bli nærmere presentert sammen med transferstudiene under (4.2.2.2).

4.2.2 Transfer hos -GB-innlærere

I studier som ser på transfer hos -GB-innlærere, er det foreslått flere ulike typer morsmålpåvirkning, og i de transferstudiene som skal omtales under, er det noen som knytter transfer til det som tidligere ble omtalt som bestemthetsekvivalenter, altså pragmatisk ordstilling og visse demonstrativer, mens andre studier viser til en overføring av preferanse for nakne substantiv fra morsmålet.

4.2.2.1 Young (1996): Overføring av pragmatisk ordstilling

Young (1996) har undersøkt om det foreligger transfer fra morsmålets pragmatiske ordstilling i innlærerspråket til tsjekkiske og slovakiske innlærere (-GB) av engelsk (+GB). Pragmatisk ordstilling er som tidligere nevnt, foreslått som en bestemthetsekvivalent, da organiseringen av setningen etter setningsinitial tematisk og setningsfinal rematisk informasjon delvis sammenfaller med bruk av bestemt og ubestemt form av substantivet i +GB-språk (se 3.3). Young ser for seg en overføring av pragmatisk ordstilling til innlærerspråket og at valg av artikkel hos -GB-innlærere av engelsk vil kunne være påvirket av om det referensielle uttrykket står setnings-initialt eller setnings-finalt (Young 1996: 150-151). Dette knyttes til at artikler kan oppleves som innholdsmessig redundante og overflødige når substantivfrasen står i visse setningsposisjoner: Når setningen har tema/rema-rekkefølge vil kjent/ny-distinksjonen gå fram av ordstillingen, og det vil oppleves som overflødig med artikler, som dermed utelates. Når plasseringen av substantivfrasen derimot *bryter* med tema/rema-prinsippet, vil artikler brukes, siden det da ikke går fram av ordstillingen om referenten er kjent eller ny (Young 1996: 156).

Youngs undersøkelse viser at de tsjekkiske og slovakiske innlærerne bruker de to engelske artiklene 'the' og 'a' for å markere at ordstillingen bryter med tema/rema-prinsippet, altså ved setningsfinale tema eller setningsinitiale rema (Young 1996: 172), men bare til en viss grad: Informasjonsflyt signaliseres ifølge Young (1996: 169) ved å veksle mellom *ingen* artikkel ved finale rema og *bestemt* artikkel ved tematiske ledd uavhengig av plasseringen i setningen. Ubestemte artikler ser derimot ikke ut til å spille noen særlig rolle:

Thematic constituents (old/given information) tend to be marked with definite articles; final rhemes (new information presented toward the end of the clause) tend to have no surface articles; indefinite articles appear to play no part in the signalling of information flow.

Young (1996: 171-172) viser imidlertid at valg av artikkel kan ses som skalært, fra finale tema, som får mest markering, og til finale rema, som får minst, og han knytter dette til transfer.

Young antyder dermed at det ved en grammatisk nullkontrast mellom morsmål og målspråk kan skje en overføring av kunnskap på tvers av lingvistiske nivå: Morsmålet

har ikke grammatisk bestemthet, men innlærerne har likevel en forståelse av identifiserbarhet siden dette ligger til grunn for ordstillingen i morsmålet. Han skriver at innlærerne allerede har kunnskap om funksjonen til engelske artikler gjennom ekvivalenter til grammatisk bestemthet, men at måten det markeres på ikke vil være målspråklig (Young 1996: 141-142).

Young mener også at han finner annen påvirkning fra morsmålet knyttet til *tellelighet* og *demonstrativer*. Når det gjelder tellelighet er dette noe Young finner ut at er sentralt for valg av ubestemt artikkel versus ingen artikkel. I engelsk er mangel på artikkel sterkt knyttet til ikke-tellelige substantiv, men inndelingen i tellelig og ikke-tellelig kan være språkspesifikk, og Young (1996: 137) mener også at noen av feilene til innlærerne kan knyttes til en feilaktig og morsmålsbasert klassifisering av telleligheten til substantivet. Når det gjelder demonstrativer, finner han at bruk av demonstrativer, først og fremst 'this', favoriseres framfor bruk av bestemte artikler i anaforiske kontekster når det er definit referanse, noe han knytter til en assosiasjon mellom funksjonene til demonstrativer i morsmålet og målspråket (Young 1996: 171).

Youngs studie ser imidlertid ikke bare på morsmålsfaktoren, men også trekkene \pm HK og \pm SR. Tendenser til å markere i tråd med disse universelle distinksjonene finner han først og fremst hos innlærere som er på et høyt ferdighetsnivå (Young 1996: 172). Dette får Young (1996: 172) til å konkludere med at innlærere kun på et tidlig stadium i utviklingen er påvirket av morsmålet, mens innlærere på et høyere nivå derimot ser ut til å bruke artiklene til å markere de universelle distinksjonene. Young (1996: 171) mener altså at det foreligger en forandring av forståelsen av artikkelsystemet, som går fra å være *morsmålsbasert* til å bli mer *målspråklig* via universalier, men at det bare delvis overlapper med målspråket.

Youngs hypoteser om transfer av pragmatisk ordstilling, og hvordan dette fører til en oppfatning av redundans ved bruk av bestemthetsmarkører, er sentral for den ene av mine hypoteser i undersøkelse 2, og jeg vil derfor komme tilbake til denne studien i kapittel 8 og 9. Jeg vil dessuten ta opp igjen diskusjonen om fenomenet redundans under i 4.3, da dette er et fenomen som flere studier har berørt, men på delvis ulike måter.

4.2.2.2 *Jarvis (2002): Diskursuniversalier og redundans*

Jarvis (2002) ser på valg av artikkel hos svenske (+GB) og finske (-GB) innlærere av engelsk, og undersøker om -GB-innlærere er sensitive for diskursuniversaliene *tema/rema*, og *ny/kjent/nylig nevnt*, på den måten at de bruker artikler eller andre referensielle uttrykk for å markere disse distinksjonene (Jarvis 2002: 388). Siden man markerer disse universelle distinksjonene på ulike måter i ulike språk, undersøker Jarvis også på hvilken måte morsmålet påvirker formvalget. Han ser altså både på om distinksjonene *er* uttrykt, og dermed psykologisk reelle, og på *hvordan* de er uttrykt.

Materialet i studien er skrevne gjenfortellinger av en stumfilm, gjort både av finske og svenske innlærere av engelsk, og av finske, svenske og amerikanske morsmålsbrukere. Materialet viser referanse til den samme kvinnelige aktøren i filmen, og er kategorisert etter hvilke referensielle uttrykk som brukes for å referere til henne, og hvilke konteksttyper det er snakk om (Jarvis 2002: 294).

Prediksjonene om de referensielle valgene til de svenske og finske innlærerne av engelsk, er basert på en delundersøkelse av hvilke former morsmålsbrukere av svensk og finsk bruker (Jarvis 2002: 392). Funnene for morsmålsbrukerne mener Jarvis er en tydelig avspeiling av at bestemthet er obligatorisk i svensk og ikke i finsk: Siden det er obligatorisk i svensk, skiller svenske morsmålsbrukere strukturelt mellom *kjente* og *nye* referenter, der nye typisk markeres med ubestemt form, kjente med bestemt form og referenter som er nylig nevnt, med pronomen (Jarvis 2002: 401). Siden grammatisk bestemthet ikke fins i finsk, bruker de finske morsmålsbrukerne substantiv uten artikkel både i kontekster med nye og kjente referenter (Jarvis 2002: 404)

I innlærermaterialet finner Jarvis at begge de to innlærergruppene er tydelig påvirket av morsmålet i måten de bruker artiklene på. Om svenske innlærere skriver han at resultatene tydelig viser morsmålpåvirkning som er fasiliterende (Jarvis 2002: 401). De svenske innlærerne gjør nemlig det samme kategoriske skillet mellom ny og kjent som svenske morsmålsbrukere, med bruk av ubestemt artikkel ved nye referenter, i tillegg til sterke tendenser til å bruke pronomen når referenten er nylig nevnt, og bestemt artikkel når referenten er kjent. De finske innlærerne viser derimot et annet system enn de svenske, og har betraktelig mer bruk av substantiv uten artikkel, i likhet med de finske morsmålsbrukerne (Jarvis 2002: 406). Svensk som morsmål har altså en fasiliterende

effekt på læring av bestemthet i engelsk, mens finsk som morsmål derimot bidrar til utelatelse av artikler.

Også Jarvis, i likhet med Young (1996), viser til kommunikativ redundans som forklaring på den utstrakte bruken av substantiv uten artikkel hos de finske innlærerne. Den høye frekvensen av nakne substantiv tolkes som en overføring fra morsmålet, og et resultat av at man vil unngå markering av bestemthet siden det allerede framgår av ordstillingen. Markering oppleves altså som redundant av de finske språkbrukerne, i morsmålet så vel som i innlærerspråket:

[...] it represents the L1 Finnish convention of avoiding (what the Finns perceives to be) redundant markers of definiteness and indefiniteness when those properties of an NP are already salient in a given discourse context. (Jarvis 2002: 416)

Jarvis viser altså også til redundans, men mener det kanskje handler om andre forhold enn Youngs (1996) kobling mellom setningsplassering og tema/rema. Forklaringen til Jarvis peker likevel på det samme fenomenet: Finske innlærere overfører fra morsmålet en oppfatning av at det er redundant med bestemthetsmarkører når dette går fram av konteksten. Jarvis (2002: 416) sikter også til at det er flere forhold, ikke bare diskursrelaterte, som fører til at artikkel blir brukt, men skriver at de alle kan avspeile en slik redundanseffekt: «[...] these factors may correspondingly reflect the conditions that determine whether Finns perceive *the* to be redundant.»

I studien til Jarvis (2002: 409-410) viser det seg at omtrent halvparten av tekstene mangler artikkel ved tema og bruker 'the' ved rema, mens den andre halvparten har motsatt mønster. Det samme gjelder imidlertid for distinksjonen ny/kjent, og Jarvis ser derfor på alternative forklaringer på disse mønstrene, for eksempel at bruk av artikkel uttrykker skillet mellom subjekt og ikke-subjekt, eller initial og ikke-initial plassering i setningen. Nesten halvparten av tekstene viser seg å ha kategorisk fravær av 'the' setningsinitialt, og Jarvis (2002: 411) mener at dette kanskje kan belyse hvorfor Huebner (1983) og Young (1996) finner tendenser til at innlærere bruker artikler til å skille mellom tema og rema, ettersom tema og rema ofte sammenfaller med disse setningsposisjonene.

4.2.2.3 Ekiert (2007; 2010): Morsmålpåvirket selektiv bruk av artikler

Ekiert ser i flere studier på polske innlæreres engelsk (Ekiert 2007; 2010), og mener at mangelen på grammatisk bestemthet og artikler i morsmålet fører til en *selektiv og meningsbasert* bruk av artikler i innlærerspråket, og at dette er direkte påvirket av hvordan morsmålet uttrykker identifiserbarhet gjennom blant annet pragmatisk ordstilling (Ekiert 2007: 27; 2010: 147).

Ekiert (2007) tar for seg én polsk informant over tid med fokus på bruk av ubestemt artikkel i spesifikke og ikke-spesifikke kontekster. Utgangspunktet er at artikler blir brukt på en meningsfull måte, selv om det ikke er i tråd med målspråkets norm (Ekiert 2007: 26-27), og hun antyder at forskjellene mellom morsmålet og målspråket for indefinitte kontekster gir seg systematiske utslag i innlærerens språk (Ekiert 2007: 1). Ekiert (2007: 31) finner at bruk av ubestemt artikkel har en tendens til å koble seg til ikke-spesifikke referenter, og mener at distribusjonen av artikkelen blir et uttrykk for en morsmålpåvirket tilnærming til identifiserbarhet.

I Ekiert (2010) følges tre polske innlærere av engelsk over tid, og Ekiert finner en mulig morsmålpåvirket forskjell i bruk av den bestemte artikkelen i henholdsvis *tekstuelle* og *situasjonelle* kontekster. Tekstuell bruk av artikkel innebærer at referenten er identifiserbar i selve teksten, mens situasjonell bruk innebærer at en må søke ut i de fysiske omgivelsene eller i generell kunnskap om verden for å identifisere referenten (Ekiert 2010: 131). Informantene ser ut til å vise størst utvikling i korrekthet for tekstuell bruk av både 'a' og 'the', altså ved introduksjon av referent og opprettholdelse av samme referent, mens bruken av situasjonell 'the' forblir variabel gjennom hele innsamlingsperioden (Ekiert 2010: 140). Forklaringen på hvorfor tekstuelle kontekster i større grad enn situasjonelle får markering, mener Ekiert (2004: 145) er at pragmatisk ordstilling gjør innlærerne sensitive for de aspektene av bestemthet som handler om tekstuell identifiserbarhet. Som vi så i 3.3 er det en forskjell i identifiserbarhet slik den kommer til uttrykk gjennom pragmatisk ordstilling og grammatisk bestemthet, der pragmatisk ordstilling først og fremst retter seg mot hvorvidt en referent er tidligere nevnt eller ikke, mens grammatisk bestemthet har et videre område for etablering av identifiserbarhet som både omfatter tekstuell (anaforisk) og situasjonell bruk.

Ekiert (2010) viser også at en av informantene utelater 'the' når substantivet står setningsinitialt, og 'a' når substantivet er setningsfinalt, noe som altså er i tråd med Youngs (1996) hypotese som ble omtalt tidligere. Ekiert (2010: 146) forklarer dette, i likhet med Young (1996), med at innlærerne opplever artikler som redundante når identifiserbarheten går fram av ordstillingen.

Ekiert (2010) forklarer altså funnene med morsmålpåvirkning. Hun viser til at mangelen på artikler i morsmålet fører til at de polske innlærerne ikke ser nødvendigheten av å markere identifiserbarhet (Ekiert 2010: 146). Konsekvensen er likevel ikke at artikler ikke brukes i det hele tatt, men derimot en *selektiv bruk av artikler* som gir seg utslag i hyppigere bruk av artikler ved tekstuell identifiserbarhet og mer variabel bruk ved situasjonell identifiserbarhet. Ekiert antyder at morsmålpåvirkningen hun mener å ha funnet, har sitt opphav i morsmålets begrepssystem, og knytter dette til Slobins (1996) og *thinking for speaking* (se 2.3.6.1). Ekiert (2010: 147) mener at begrepsdistinksjoner som er relevante i morsmålet kommer til uttrykk gjennom de språklige midlene som er tilgjengelig i målspråket.

Ekiert (2010) berører dermed noen av de samme fenomenene som er omtalt for Young (1996) og Jarvis (2002). Det er fokus på pragmatisk ordstilling som et potensielt opphav til transfer, men da først og fremst som en kilde til forståelse av fenomenet identifiserbarhet, hvor tekstuell identifiserbarhet er tilgjengelig for innlærerne fra morsmålet og denne kunnskapen overføres til innlærerspråket, noe som fører til en selektiv bruk av artikler.

Det som er felles for den transferforskningen som er nevnt så langt, er at man ser for seg at morsmålet påvirker *når* engelske artikler brukes og ikke brukes i innlærerspråket. Det vises til at det er forhold i morsmålet som styrer denne vekslingen mellom former og at formvalget kan være systematisk og uttrykke et språklig innhold selv om det ikke er snakk om samme systematikk og innhold som i målspråket. Både Young (1996), Jarvis (2002) og Ekiert (2010) viser på denne måten hvordan morsmålet kan bidra til å forklare både bruk av og mangel på artikler i språket til innlærere av engelsk. De neste studiene som vil bli omtalt, har en mer sammensatt forklaring på bruk og ikke bruk av artikler, der morsmålet først og fremst forklarer at artikler ikke blir brukt

i visse språklige kontekster, mens det er andre forhold som bidrar til bruk av artikler i andre kontekster.

4.2.2.4 *Trenkic: Syntactic Misanalysis Hypothesis*

Mens flere studier har vært opptatt av forhold knyttet til spesifikk referanse, tema/rema og identifiserbarhet, enten som en direkte påvirkning fra morsmålet eller som universelt tilgjengelige kategorier, har *Syntactic Misanalysis Hypothesis* (Trenkic 2007; 2008; 2009) en mer sammensatt natur. Denne hypotesen berører både spørsmålet om hvilket språklig innhold -GB-innlærere legger i de engelske artiklene, og hvordan kapasitetsbetingelser fører til større sannsynlighet for utelatelse av artikler i visse språklige kontekster. Hypotesen har et sentralt element av morsmålpåvirkning, men innebærer også innputt- og prosesseringsforhold.

Syntactic Misanalysis Hypothesis tar utgangspunkt i en antakelse om at bestemte og ubestemte artikler i engelsk ikke har noe semantisk innhold, men kun utfyller en syntaktisk funksjon i substantivfrasen. Artikler er som regel innholdstomme, pleonatiske elementer i engelsk, da referentens bestemthetsstatus uansett er tilgjengelig gjennom konteksten (Trenkic 2007: 294; 2008: 9; 2009: 8).¹⁸ Trenkic mener at -GB-innlærere oppfatter en likhet mellom artikler i engelsk og tallord eller demonstrativer i morsmålet, noe som fører til at engelske artikler *feiltolkes* som en type adjektiver og ikke som syntaktiske elementer. For disse innlærerne har engelske artikler dermed et leksikalsk meningsinnhold, og blir behandlet deretter (Trenkic 2007: 312; 2008: 9). Så lenge artiklene oppfattes som leksikalske, vil bruken av dem være *pragmatisk motivert* i innlærerspråket, de vil kun bli brukt i tilfeller hvor de faktisk bidrar til kommunikasjonen, og de kan dermed i noen tilfeller bli nedprioritert på grunn av kapasitets-begrensninger (Trenkic 2007: 314; 2008: 10; 2009: 7). Utelatelse av bestemthetsmarkører er dermed ifølge Trenkic et resultat av både en morsmålsbasert feiltolkning og prosesseringsforhold.

Meningsinnholdet som antas å bli lagt i artiklene, er relatert til identifiserbarhet, men en mer konkret type identifiserbarhet enn det en finner i målspråket engelsk, da det

¹⁸ *Pleonastisk* innebærer her en form for overflødighet/redundans, men merk at det her er snakk om å være innholdsmessig redundant i *målspråket*, ikke at de oppleves som redundante i innlærerspråket, som er den typen redundans Young (1996) og Jarvis (2002) har omtalt. Disse fenomenene er relaterte, men som jeg vil komme tilbake til, mener jeg det er noen forskjeller.

heller representerer en objektiv egenskap ved referenten, for eksempel at den er fysisk til stede eller har blitt tidligere introdusert, altså tekstuell eller situasjonell identifiserbarhet (Trenkic 2007: 312; 2008: 11).

I Trenkic (2009) og Trenkic og Pongpairøj (2013) er forklaringen noe revidert i forhold til tidligere artikler. I tillegg til innlærernes feiltolkning av engelske artikler, beskrives en konkurranse mellom påvirkning fra morsmålet og påvirkning fra målspråksinnputt, en konkurranse som avgjøres av kapasitetsbetingelser: Innlærerne vil på den ene siden analysere innputt gjennom strukturelle forventninger basert på hvordan morsmålet fungerer, men på den andre siden er de oppmerksomme på statistiske regelmessigheter i innputt og observerer at artikler er noe som skal brukes i målspråket (Trenkic 2009: 8). I møtet mellom disse to påvirkningsfaktorene oppstår det en konkurranse mellom morsmålsstrukturen *nakent substantiv* (uten artikkel) og målspråksstrukturen *'the' + substantiv*. Utfallet av konkurransen er variasjon mellom de to strukturene, i praksis variasjon mellom bruk og utelatelse av artikkel. Hvilken form som vinner fram til enhver tid, styres av hvor sterkt de to ulike typene uttrykk er knyttet til begrepet som uttrykkes, og kapasitetsbetingelser. Morsmålsstrukturen, som ikke har artikkel, vil vinne når kapasitetsbelastningen er stor, både fordi den er mindre kompleks og fordi den står sterkere, men styrken til målspråksstrukturen og prosesseringskapasiteten vil begge øke over tid. Etter hvert som innlærerne blir mer og mer eksponert for innputt, vil målspråksstrukturen styrkes og morsmålsstrukturen svekkes, noe som fører til mer og mer bruk av fraser med artikkel:

Thus, early on, while L2 experience is still limited, the new L2 patterns should be highly vulnerable to competition from the established, and hence active, L1 alternatives; the probability of selecting bare *NPs* will, therefore, be high. With more L2 experience, though, the *Det +NP* structure is likely to become stronger and more active, and so less vulnerable to competition (Trenkic & Pongpairøj 2013: 161- 162).

Trenkic (2009) viser spesielt til to typer tendenser som har vært funnet i ulike andrespråksstudier: At -GB-innlærere oftere utelater artikler i *saliente* kontekster eller når substantivet er modifisert av et adjektiv, mens artiklene ofte brukes i ikke-saliente eller umodifiserte kontekster (Trenkic 2008: 9-10; Trenkic & Pongpairøj 2013: 152). Saliente kontekster har man for eksempel i temaposisjon, ved andregangs nevning av en referent, eller når referenten er nevnt direkte i motsetning til å komme fram via

assosiasjoner til en annen referent. Disse har til felles at referentene i slike kontekster er saliente i minnet (Trenkic & Pongpairoj 2013: 152). Dette omtales som «salienseffekten», og påstanden er at saliente og adjektivmodifiserte kontekster utgjør en større belastning på kapasiteten til innlærerne enn øvrige kontekster, og at dette forklarer hvorfor artikler blir utelatt i slike kontekster (Trenkic & Pongpairoj 2013: 163).

Trenkic (2009) og Trenkic og Pongpairojs (2013) forklaringer på hvorfor innlærere veksler mellom bruk og ikke bruk av artikler, er altså ikke bare knyttet til morsmålspåvirkning, men også til målspåksinnputt og språklig prosessering. Morsmålet blir for det første knyttet til en feiltolkning av engelske artikler som leksikalske størrelser, noe som igjen knyttes til en tverrspråklig identifikasjon mellom artikler i målspåket og demonstrativer i morsmålet.¹⁹ Morsmålet bidrar for det andre med det man kan kalle en preferanse for nakent substantiv, som overføres til innlærerspråket. Morsmålspåvirkningen konkurrerer imidlertid med det innlæreren observerer i målspåket, altså bruk av ‘the’ + substantiv.

4.2.3 Standardforklaringen

Trenkic (2009) tar også kritisk for seg den utbredte redundansforklaringen som jeg allerede har vært inne på tidligere i kapitlet, noe som også omtales som «standardforklaringen» på utelatelse av artikler hos -GB-innlærere (Trenkic & Pongpairoj 2013: 152). Standardforklaringen går altså ut på at artikler utelates i innlærerspråket fordi de oppleves som kommunikativt redundante av innlærerne når identifiserbarheten går fram av konteksten, først og fremst gjennom ordstillingen.

Trenkic (2009) viser til at redundans er en utbredt forklaring på salienseffekten i tidligere studier, altså at artikler ofte utelates i saliente kontekster, men mener at denne standardforklaringen om redundans ikke er tilstrekkelig. Kommunikativ redundans tilsier nemlig at artikler *alltid* utelates, siden de alltid er redundante, og dermed kan ikke redundans forklare den vekslingen mellom bruk og ikke bruk som salienseffekten innebærer (Trenkic 2009: 10). Trenkic (2009: 12) mener i stedet at redundans inngår som en del av den tidligere omtalte konkurransen mellom morsmålsstruktur og

¹⁹ Termen ‘tverrspråklig identifikasjon’ brukes ikke av verken Trenkic (2009) eller Trenkic og Pongpairoj (2013), men slik jeg tolker dem, er det en slik prosess de viser til.

målpråksstruktur. Mens morsmålet og redundans begge favoriserer utelatelse av artikler (naken form) favoriserer målpråksinnputt tilstedeværelse av artikkel, og valget mellom disse to strukturene avgjøres av kapasitetsbetingelsene i en gitt språkbrukssituasjon. Redundans blir dermed bare én av to faktorer som fører til en preferanse for nakent substantiv i innlærerspråket, mens det er målpråksinnputt som fører til at artikler faktisk brukes.

I Trenkic og Pongpairoj (2013) undersøkes forholdet mellom salienseffekten og standardforklaringen empirisk hos innlærere med fransk (+GB) og thai (-GB) som morsmål. Studien ser for det første på om salienseffekten faktisk fins, uavhengig av ferdighetsnivå eller om morsmålet har GB eller ikke, og for det andre spør de seg om redundans er hovedårsaken til utelatelse av artikler (Trenkic & Pongpairoj 2013: 153). I studien manipuleres en referents saliens systematisk ved hjelp av en film om en fisk som spiser en annen, og hvor fiskene veksler på å være den visuelt mest saliente. Forfatterne ser på om visuell saliens påvirker bruk av artikkel i innlærernes beskrivelse av fiskene i filmen (Trenkic & Pongpairoj 2013: 156). Studien viser at salienseffekten foreligger og at artikler først og fremst utelates i saliente kontekster, til tross for at redundansen er holdt konstant. Effekten ble funnet kun hos innlærerne med thai som morsmål, uavhengig av ferdighetsnivå, mens saliens hadde ikke noen innvirkning på bruk av artikkel hos de franske innlærerne eller hos de engelske morsmålsbrukerne (Trenkic & Pongpairoj 2013: 160-161). Dette mener forfatterne skyldes at thai tillater naken substantivfrase i disse kontekstene, mens engelsk og fransk ikke gjør det, noe som gjør at det kun er innlærerne med thai som morsmål som opplever en konkurranse mellom naken form og artikler, mens både morsmål og målpråk favoriserer artikkel hos de franske innlærerne (Trenkic & Pongpairoj 2013: 162).

Trenkic og Pongpairoj (2013) mener altså at studien viser at redundans ikke kan forklare salienseffekten, men de anerkjenner det likevel som en del av et større bilde, der redundans sammen med morsmålpåvirkning bidrar til en preferanse for substantiv uten artikkel. Forfatterne påpeker også to sentrale problemer ved redundans som forklaring, nemlig at det er vanskelig å si hvor grensen mellom redundant og ikke-redundant går, i tillegg til at redundans ikke kan forklare utvikling over tid med mer og mer bruk av artikler i innlærerspråket (Trenkic 2009: 11; Trenkic & Pongpairoj

2013: 155). Jeg vil under se nærmere på de ulike tilnærmingene til redundans som har vært presentert så langt.

4.3 Redundanseffekten

Flere av studiene som er omtalt så langt, berører fenomenet *kommunikativ redundans* i forbindelse med bruk av artikler i engelsk som andrespråk, altså at artikler utelates i innlærerspråket fordi innholdet går fram av konteksten, og dermed vil være overflødig å uttrykke. Når man ser nærmere på hva de ulike studiene sier om redundans, er det likevel noen forskjeller mellom dem, og redundans ser også ut til å ha ulik rolle i forhold til andre faktorer som påvirker innlærerspråket. En sentral forskjell virker å være hvorvidt man mener at redundans kan fungere som forklaring på *både* bruk og utelatelse av artikler i innlærerspråk, slik det framstår hos Young (1996) og Jarvis (2002), eller om det bare kan forklare *utelatelse* av artikler, som hos Trenkic (2009) og Trenkic og Pongpairoj (2013). Forskjellen ser ut til å kunne beskrives som at Young (1996) og Jarvis (2002) ser på redundans som en egenskap ved *innlærerspråket* og som nært knyttet til transfer, mens Trenkic (2009) og Trenkic og Pongpairoj (2013) ser på det som en egenskap ved *målspråket*, altså at artikler alltid er redundante i engelsk, i tillegg til at det også er et generelt kommunikativt fenomen knyttet til kvantitetsmaksimen til Grice (1975).

Hos Young (1996) er redundanseffekten noe som virker sammen med transfer av pragmatisk ordstilling fra morsmålet til innlærerspråket. Slik jeg tolker Young, er det slik at når innlærere overfører pragmatisk ordstilling som signal om identifiserbarhet, overfører de også en oppfatning av når identifiserbarheten går fram av konteksten, og når den *ikke* gjør det, og det er denne morsmålspåvirkete oppfatningen som ligger til grunn for at markering oppleves som redundant i noen kontekster, men ikke i andre. På denne måten mener jeg redundans er et *innlærerspråksfenomen*, der en søker å unngå redundans når identifiserbarhet går fram av ordstillingen, men det er morsmålet som legger føringer på hva som oppfattes som redundant, på grunn av transfer av pragmatisk ordstilling som signal om identifiserbarhet.

Jarvis' (2002) beskrivelse av redundansfenomenet er relativt lik Youngs, da han hevder at stort fravær av artikler hos finske innlærere handler om at man i finsk unngår markering av bestemthet i kontekster der det likevel går fram av konteksten (Jarvis 2002:

416). Jarvis viser imidlertid til en del andre forhold enn Young med tanke på hva som kan forklare vekslingen mellom bruk og ikke bruk av artikler, og mener at det kan være styrt kun av hvor i setningen substantivfrasen er plassert, heller enn kombinasjonen av plassering og tema/rema-status, og han viser også til andre forhold som han likevel mener kan være knyttet til en oppfatning av markering som redundant (Jarvis 2002: 411).

Young (1996) og Jarvis (2002) er altså like i at de viser til redundans som noe som oppstår i samspillet mellom pragmatisk ordstilling i morsmålet, transfer og ønsket om å unngå redundans i innlærerspråket, og begge ser ut til å hevde at redundanseffekten kan forklare både utelatelse og bruk av artikler i innlærerspråket. Spørsmålet er om det er naturlig å skille mellom transfer og redundans som separate fenomen i disse tilnærmingene. Ringbom (2011: 19) problematiserer forholdet mellom *unngåelse av redundans*²⁰ og transfer, og viser til at det kanskje ikke er hensiktsmessig eller mulig å skille de to fenomenene fra hverandre.

Hos Trenkic (2009) og Trenkic og Pongpairoj (2013) hevdes det tydelig at transfer og redundans er to ulike fenomener, og at redundans kun forklarer utelatelse av artikkel, eller *bruk av naken form*, som de også omtaler det (Trenkic & Pongpairoj 2013: 163). Det er en preferanse for naken form som utgjør selve morsmålspåvirkningen, mens redundans holdes adskilt som et generelt kommunikativt fenomen som sier at når noe ikke er innholdsmessig nødvendig, så trenger en ikke uttrykke det, noe som knyttes til Grices kvantitetsmaksime, altså at man ikke skal bruke en mer informativ form enn nødvendig (Grice 1975). Både preferanse for naken form i morsmålet og redundans bidrar til at naken form brukes i innlærerspråket og at artikler dermed utelates, men det er målspårsinnputt som forklarer at artikler faktisk brukes noen steder, i tillegg til at kapasitetsbetingelser avgjør konkurransen mellom naken form og substantivfrase med artikkel.

Det kan dermed se ut til at det er to noe ulike redundansfenomener her. For det første et generelt kommunikativt prinsipp som kun vil kunne forklare utelatelse av bestemthetsmarkører, og dessuten et mer spesifikt innlærerspråksprinsipp som knyttes til transfer, og som har potensial til å forklare både bruk og utelatelse av

²⁰ Eng. 'opposition to redundancy'

bestemthetsmarkører. Dette vi jeg komme tilbake til i kapittel 8.1 i forbindelse med den ene hypotesen i undersøkelse 2.

4.4 Nordiske studier

Større studier av bestemthet innen norsk eller andre skandinaviske språk er det ikke så mange av, men jeg vil her omtale to slike studier som jeg har funnet relevant for min egen studie, nemlig doktorgradsavhandlingene til Nordanger (2017) innen norsk som andrespråk, og Axelsson (1994) innen svensk som andrespråk. Jeg har valgt å omtale akkurat disse to studiene siden de begge ser blant annet på morsmålets rolle, og begge sammenligner -GB-innlærere med +GB-innlærere, samt berører koblingen mellom form og funksjon eller kontekst i innlærerspråk. I tillegg undersøker begge studiene språket til innlærere med slaviske språk som morsmål, altså russisk og polsk.

4.4.1 Nordanger 2017

Nordanger (2017) undersøker samspillet mellom universalier og morsmåleffekter hos russiske (-GB) og engelske (+GB) innlærere av norsk. Studien undersøker bruk av bestemte og ubestemte former både i lys av Bickertons (1975) universalier \pm SR og \pm HK (se 4.1), men også i lys av Princes (1981) *kjenhetstaksonomi*, som deler inn diskursreferenter etter om det er snakk om nye referenter, tidligere nevnte referenter ('evoked'), eller referenter som er indirekte introdusert, altså det som i 3.1 ble omtalt som *slutningsbaserte referenter* ('inferred') (Nordanger 2017: 276).

En av delundersøkelsene til Nordanger (2017) ser på korrekthet hos de engelske og russiske innlærerne i de fire ulike konteksttypene basert på kategoriene \pm SR og \pm HK (indefinit, generisk, ikke-referensiell og indefinit). Prediksjonene er at det å mangle GB i morsmålet fører til større utfordringer hos russiske innlærere, mens det å ha GB i morsmålet er fasiliterende for engelske innlærere, og studien finner relevante forskjeller mellom de to innlærergruppene. Begge gruppene har en god del feil, spesielt i defintitte og indefintitte kontekster, men innlærergruppene skiller seg fra hverandre med tanke på hvilke kontekster feilene er mest framtrede i: Den russiske gruppen utelater nemlig mest *ubestemte artikler i indefintitte kontekster*, mens den engelske har mest utelatelse av *bestemt form i defintitte kontekster* (Nordanger 2017: 259). Dette er imidlertid en tendens som går raskt over for de engelske innlærerne, mens utfordringene vedvarer

over flere innsamlingspunkt for de russiske. At de engelske innlærerne tidlig i utviklingen har utfordringer ved definit referanse, knytter Nordanger (2017: 375) til de strukturelle forskjellene som foreligger mellom engelsk og norsk i måten bestemthet uttrykkes på, mens mindre utfordringer med indefinit referanse knyttes til at strukturelle likheter (ubestemt artikkel i begge språkene) har en fasiliterende effekt hos de engelske innlærerne.

Nordanger ser også på redundans, men da i form av kjenthetsgrader i Princes (1975) kjenthetsstaksonomi, der det antas en sammenheng mellom redundans og markering av identifiserbarhet. Hvor kjent en referent er, altså om en referent er *ny*, *slutningsbasert* eller *kjent*, innebærer ulik grad av redundans: Markering er mest redundant ved tidligere nevnte referenter, og minst redundant ved nye (Nordanger 2017: 305-306). På grunn av dette ventes det at markering oftest er utelatt ved tidligere nevnte referenter, og sjeldnest ved nye referenter. Nordanger finner at en del av de russiske innlærerne har formvalg som er i tråd med redundansforklaringen, men at det faktisk er enda flere engelske innlærere som har dette, noe som ikke er ventet siden disse har GB i morsmålet (Nordanger 2017: 299). Begge grupper viser seg å ha minst markering ved slutningsbaserte referenter, mens tidligere nevnte referenter *oftest* er markert, noe som er motsatt av det som var forventet. Nordanger (2017: 316) mener derfor hun ikke har funnet støtte for at oppfatning av redundans påvirker markeringen av bestemthet hos de russiske innlærerne, men mener at hun heller ikke helt kan avvise redundans som forklaring.

4.4.2 Axelsson (1994)

Axelssons studie (1994) ser på svenskinnlærere med morsmålene finsk (-GB), polsk (-GB) og spansk (+GB), og undersøker ni ulike typer substantivfraser delt inn etter strukturell kompleksitet og om de er bestemte eller ubestemte (Axelsson 1994: 30). Utgangspunktet er at noen typer fraser er vanskeligere å tilegne seg enn andre, og at dette vil reflekteres i formvalgene til innlærerne. Studien undersøker dessuten hvilke semantiske og pragmatiske kontekster de ulike typene fraser brukes i, for eksempel *kjent*, *ny*, *generisk*, og *ikke-spesifikk*, i tillegg til *syntaktisk funksjon* (Axelsson 1994: 29).

Et av de viktigste funnene til Axelsson er en tilegnelsesrekkefølge med tre ulike stadier for de ulike typene fraser (Axelsson 1994: 99), men hun finner også tendenser til transfer hos alle de tre innlærergruppene, og det er disse jeg først og fremst vil gjengi her.

Axelsson finner spesielt forskjeller mellom finske og spanske innlærere med tanke på substantivfraser som 'en bok' og 'boken', altså enkle ubestemte og bestemte substantivfraser. De spanske innlærerne ser ut til å mestre de ubestemte bedre og tidligere enn de bestemte frasene, mens det motsatte gjelder de finske, som mestrer bestemte fraser først (Axelsson 1994: 101). Dette stemmer også med Nordangers funn for russiske og engelske innlærere, der de russiske innlærerne hadde mest utelatelse av ubestemte artikler, mens de engelske innlærerne hadde mest utelatelse av bestemt form (Nordanger 2017: 259). Axelsson (1994: 148) mener at denne forskjellen mellom spanske og finske innlærere kan ha å gjøre med at de finske innlærerne er mer vant til å være oppmerksomme mot *slutten av ord*, mens de spanske er mer vant til foranstilte artikler, og at de finske innlærerne derfor har større mulighet til å oppdage bestemthetsbøyning enn de spanske. Axelsson sier ikke direkte hvorfor finske innlærere skal ha lettere for å oppdage bøyning, men jeg antar dette har å gjøre med at finske substantiv ofte har bøyningssuffikser på substantivene, noe spansk har mindre av. Axelsson nevner heller ikke dette som en forklaring for de språklige mønstrene hun finner hos *polske* innlærere, som har et morsmål som i likhet med finsk har mye substantivmorfologi. At dette ikke nevnes, kan ha å gjøre med at de polske innlærerne skiller seg fra de andre to innlærergruppene med å ha utfordringer *både* med bestemte og ubestemte fraser. Axelsson (1994: 102) viser til at dette kan handle om en forenklingstrategi hvor de polske innlærerne kun bruker selve substantivet og utelater artikler eller bøyninger, noe som resulterer i at de ofte har naken form av substantivet. Mens polske innlærere har en slik forenkling ved både bestemte og ubestemte fraser, skjer altså dette bare ved bestemte fraser hos de spanske innlærerne og ubestemte hos de finske (Axelsson 1994: 102).

Et annet transferrelatert forhold Axelsson (1994) finner, er at de spanske innlærerne ikke bare gjør sjeldnere feil, men også oftere bruker fraser med ubestemt artikkel enn de finske og polske, som i motsetning til de spanske innlærerne ofte bruker

det svenske pronomenet *någon* (*noen*), noe Axelsson (1994: 149) antyder kan være transfer ettersom både polsk og finsk kan bruke ubestemte pronomen i slike kontekster: ‘joku’ i finsk og ‘jakis’ i polsk (Se også omtale av ‘jakis’ i beskrivelsen av polsk i 7.3.2).

4.4.3 Overførbarhet mellom studier med ulike målpråk

Store deler av den forskningen som har vært omtalt over, ser på språket til innlærere av engelsk, og det er derfor naturlig å gjøre en vurdering av om slike studier av engelsk som målpråk har overføringsverdi til en studie med norsk som målpråk. Selv om norsk og engelsk er like i å ha grammatisk bestemthet, er det både strukturelle og innholdsmessige forskjeller mellom de to språkene, noe jeg vil komme nærmere inn på i kapittel 7. Det er de strukturelle forskjellene som er mest markante: Mens engelsk har både bestemt og ubestemt artikkel, har norsk bare ubestemt artikkel, mens trekket bestemt uttrykkes morfologisk.

Siden vi vet at de kontrastive relasjonene mellom morsmål og målpråk legger føringer på transfer i innlærerspråk (se 2.3), vil enhver morsmålsbasert forklaring på bruken av bestemthetsmarkører i innlærerspråk potensielt være spesifikk for et gitt målpråk. Når en skal vurdere overførbarheten mellom ulike målpråk, virker det derfor nødvendig å skille mellom morsmålsbaserte forklaringer som kan betraktes som målpråksuavhengige, og forklaringer som må ses på som målpråksavhengige.

Siden både engelsk og norsk har den samme grammatiske kategorien, grammatisk bestemthet, vil morsmålsbaserte forklaringer som tar utgangspunkt i de mer semantiske og begrepsmessige sidene av denne kategorien, i prinsippet kunne være brukbare for både norsk og engelsk som målpråk.

Forklaringer på transfer som har sitt opphav i tverrspråklig identifikasjon mellom strukturer i morsmål og målpråk, vil derimot ikke være anvendelig på tvers av målpråk om målpråkene uttrykker bestemthet ulikt, slik norsk og engelsk gjør. Et eksempel på dette er Trenkics (2007; 2008; 2009) hypotese om at engelske artikler feiltolkes som leksikalske elementer etter å ha blitt oppfattet som paralleller til demonstrativer i morsmålet (Trenkic 2007: 312; 2008: 9). Hypotesen er altså basert på at det skjer en tverrspråklig identifikasjon mellom morsmål og målpråk, og dermed er måten bestemthet uttrykkes på vesentlig for denne hypotesen. For norsk, som bare har ubestemte artikler, ikke bestemte, er tverrspråklig identifikasjon av denne typen mindre

sannsynlig og kanskje begrenset til ubestemte artikler. Om både morsmålet til innlærerne og målspråket derimot har substantivmorfologi, kan en se for seg tverrspråklig identifikasjon mellom *bøyningsendelser* i språkene som et alternativ til den tverrspråklige identifikasjon mellom artikler og demonstrativer når engelsk er målspråk.

Som nevnt i 2.3.4 skal formelle likheter mellom språkene lettere være gjenstand for tverrspråklig identifikasjon enn semantiske, noe som gjør at en potensielt kan se for seg en kobling mellom *substantivmorfologi* i morsmål og målspråk, der begge har dette. Et relevant moment her er likheten mellom norsk og de -GB-morsmålene som er med i studien min, når det kommer til å ha substantivmorfologi i større grad enn det engelsk har. Her kan man altså se for seg en mulig tverrspråklig identifikasjon mellom *bøyningsendelser* i språkene som alternativ til den tverrspråklige identifikasjon mellom artikler og demonstrativer som er omtalt over.

Ideen om morsmålpåvirkning i form av overføring av naken form til innlærerspråket som vi så hos Trenkic (2009) og Trenkic og Pongpaioj (2013), er derimot en hypotese som er direkte overførbart til norsk fordi den ikke er avhengig av målspråkets strukturer. En slik transfer skjer på bakgrunn av hva som uttrykkes i morsmålet og kan dermed skje uavhengig av hvilken måte målspråket uttrykker bestemthet på. Også ideen om at pragmatisk ordstilling som signal om identifiserbarhet overføres til innlærerspråket, er overførbart til norsk som målspråk, ettersom dette også handler om en overføring av trekk fra morsmålet uavhengig av noen tverrspråklig identifikasjon.

Også andre nordiske studier har vært inne på dette med overførbareheten mellom engelsk som målspråk og norsk eller svensk som målspråk. Axelsson (1994: 151) mener sammenligninger på tvers av målspråk er vanskelig, men Axelssons studie er også en studie med sterkt fokus på strukturelle forhold, så påstanden må ses i lys av dette. Nordanger (2017: 129) er enig med Axelsson om at det er viktig å ta hensyn til at strukturelle forskjeller kan påvirke sammenlignbarheten, men Nordanger mener at mange studier har fokusert på semantiske og pragmatiske sider av grammatisk bestemthet, og at det for slike studier ikke er noe i veien for å sammenligne hypoteser og funn mellom målspråk. Nordanger (2017: 129) viser også til at dette virker mindre problematisk enn at en del studier har en tendens til å behandle -GB-innlærere med ulike

morsmål som én type innlærergruppe til tross for de store forskjellene som kan foreligge også mellom -GB-språk.

Hvorvidt tverrspråklig identifikasjon er opphavet til transfer eller hypoteser om transfer, kan med andre ord være avgjørende for om funn og hypoteser fra en studie med engelsk som målspråk kan overføres til norsk eller ikke, men en slik vurdering vil uansett være noe usikker av andre grunner: I og med at vi vet at det ofte er subjektive kontrastive antakelser som ligger til grunn for transfer, er det vanskelig å vurdere reell overførbarhet uten å vite om tverrspråklig identifikasjon som ligger bak innlærernes språklige valg eller ikke, og dette er i seg selv empiriske spørsmål. I så måte er sammenligninger mellom ulike målspråk i seg selv veldig relevant for å belyse om det er tverrspråklig identifikasjon eller nullrelasjonsantakelser som ligger bak et gitt transfermønster i innlærerspråk. Dette er noe jeg vil komme tilbake til i slutten av avhandlingen (11.5.2).

4.5 Oppsummering av kapittel 4

Dette kapittelet har jeg presentert ulike studier av grammatisk bestemthet eller bruk av bestemthetsmarkører i innlærerspråk med fokus på -GB-innlærere, og jeg har blant annet framhevet to hovedscenarier for den grammatiske nullkontrasten mellom -GB-språk og +GB-språk: På den ene siden at nullkontrasten *ikke* fører til transfer, men at innlærere baserer seg på språklige universalier, og på den andre siden at det *kan* føre til transfer til tross for nullkontrasten. Inndelingen i ulike typer kontrastive relasjoner (likhetsrelasjoner, ulikhetsrelasjoner, nullrelasjoner) kan belyse disse to scenarioene, da det som vi så i kapittel 2, ikke må være faktiske objektive likheter mellom morsmål og målspråk for at transfer skal forekomme: Selv om det objektivt sett er en nullkontrast mellom språkene, er ikke dette et hinder for at transfer kan skje som resultat av en umotivert antatt likhet mellom språkene, nemlig at språkene fungerer på samme måte, eller at innlærer overfører formpreferanser fra morsmålet til innlærerspråket.

Jeg har også sett nærmere på redundansforklaringen, som viser til at bestemthetsmarkører utelates når identifiserbarheten går fram på andre måter, og satt opp et skille mellom studier som ser ut til å se på redundans som et generelt kommunikativt fenomen adskilt fra transfer, og studier som ser på redundans mer som et innlærerspråksfenomen som ikke nødvendigvis kan skilles fra transfer.

Et tredje moment som ble diskutert mot slutten av kapittelet, er overførbarheten studier med engelsk som målspråk har til studier med norsk som målspråk. Her mente jeg det var vesentlig for overførbarheten til morsmålsbaserte forklaringer om de så på transfer som et resultat av tverrspråklig identifikasjon eller ikke. Hovedpåstanden var at de studiene som tar utgangspunkt i den semantiske eller pragmatiske innholdssiden av grammatisk bestemthet og ikke antok transfer utløst av tverrspråklig identifikasjon, har større overførbarhet enn studier hvor måten bestemthet uttrykkes på og tverrspråklig identifikasjon mellom morsmål og målspråk var en del av forklaringen. Det er dermed ikke bare de strukturelle forskjellene mellom ulike målspråk som er relevante for overførbarheten, men også selve opphavet til transferen, som kan være uavhengig av hvordan målspråket uttrykker bestemthet: Om opphavet til transfer i innlærerspråket er en nullrelasjon der morsmål og målspråk antas å fungere på samme måte uten at det er motivert av målspråksinnputt, er det kanskje større grunn til å anta like funn på tvers av målspråk enn om opphavet til transferen antas å være tverrspråklig identifikasjon.

Studien i denne avhandlingen vil med sine to delundersøkelser belyse ulike deler av bestemthet i innlærerspråk, og mens undersøkelse 1 søker å belyse spørsmålet om hvorvidt bestemthet er en større utfordring for -GB-innlærere enn +GB-innlærere, vil undersøkelse 2 nærme seg spørsmålet om morsmålets rolle for valg av form i ulike språklige kontekster i innlærerspråket. I forbindelse med undersøkelse 2 (kapittel 8) vil jeg derfor komme tilbake til noen av de transferrelaterte problemstillingene som har blitt nevnt i dette kapittelet.

5 ELEKTRONISKE INNLÆRERKORPUS OG ASK

I dette kapittelet vil jeg beskrive elektroniske innlærerkorpus generelt, og mer spesifikt det innlærerkorpuset som er brukt i denne studien, *Norsk andrespråkskorpus* (ASK). I 5.1 redegjør jeg kort for elektroniske innlærerkorpus, i 5.2 presenterer jeg Norsk andrespråkskorpus både med tanke på design og innhold, og i 5.3 ser jeg kort på bruk av ASK til å undersøke transfer. Det mer spesifikke utvalget av datamateriale fra ASK som er brukt i de to delundersøkelsene i denne avhandlingen, blir ikke beskrevet i dette kapittelet, men hver for seg i kapittel 6 og kapittel 8. Deler av informasjonen i dette kapittelet er basert på dokumentasjon jeg selv har utviklet eller vært med på å utvikle som del av ASK-prosjektet, for eksempel brukerveiledning (Johansen 2011).

5.1 Elektroniske innlærerkorpus

Granger (2002: 48) beskriver elektroniske innlærerkorpus som elektroniske samlinger av autentiske andrespråksdata som er samlet inn etter eksplisitte kriterier og deretter beriket med koding etter standardiserte metoder. Granger (2002: 46) framhever videre hvordan data fra elektroniske innlærerkorpus har noen fordeler sammenlignet med andre typer data som typisk har blitt brukt i andrespråksforskning, for eksempel at andrespråksforskning har en tendens til å være basert på eksperimentelle og introspektive data, og at man har lett for å avvise naturlige språkbruksdata, eller at andrespråksforskning ofte mangler viktig informasjon om informantene og dermed også muligheten til å kontrollere for ulike variabler som påvirker språklæring og språkbruk. Granger mener også at det er en tendens til å basere seg på små mengder data. Disse påstandene om metodologiske svakheter er sannsynligvis ikke gyldige for enhver andrespråksstudie som ikke bruker et elektronisk innlærerkorpus, men Granger (2004: 124) peker likevel på noen metodiske områder som elektroniske innlærerkorpus kan bidra til å styrke i kraft av å inneholde autentisk språkbruk, store mengder data, og direkte tilgang til informasjon som gjør det mulig å kontrollere for ulike variabler.

Et viktig kriterium for et innlærerkorpus er at det inneholder autentisk, naturlig språkbruk, men autentisitet og naturlighet kan ses på som problematiske begreper når det gjelder innlærerspråk. Granger (2002: 49; 2008: 2) skriver derfor at korpusdata vil

innebære *ulike grader* av naturlighet; fra autentisk dagligspråk til autentisk språkbruk i klasserommet.

En annen stor fordel med elektroniske innlærerkorpus er størrelsen og mengden språkdata, i tillegg til teknologi til å annotere store mengder data automatisk eller halvautomatisk, og til å gjøre effektive søk. Dette er kanskje en av de sterkeste sidene av elektroniske innlærerkorpus, og det har stor betydning for representativiteten og generaliserbarheten til andrespråksstudier, i tillegg til at blir det lettere å utføre kvantitativ forskning med bruk av statistiske tester (Granger 2004: 125). Granger (2008: 3) påpeker likevel at størrelse ikke er tilstrekkelig, om ikke korpuset har blitt utformet etter klare kriterier. Et eksempel er at språkbruksdata må kobles til biografisk informasjon om innlærerne og forholdene rundt språkbrukssituasjonen, noe som bidrar til at man kan kontrollere for de mange påvirkningsfaktorene i andrespråklæring.

Selv om elektroniske innlærerkorpus deler egenskaper med mer generelle korpus, er det likevel noen viktige forskjeller, for eksempel at det er noen variabler som er spesifikke for andrespråklæring som må tas med i korpusdesignet, altså variabler som er av spesiell interesse for andrespråksforskningen. Granger (2004: 125; 2008: 5) skiller her mellom *generelle variabler*, som alder, kjønn, morsmål, medium, sjanger, tema og lengde; og *andrespråksspesifikke variabler*, som ferdighetsnivå, læringskontekst, andre andrespråk, målspråkseksposering og dessuten oppgavetype og forholdene rundt språkbrukssituasjonen. Granger (2008: 6) framhever spesielt variabelen ferdighetsnivå, og hvordan denne er vanskelig å kartlegge. Ofte brukes for eksempel oppholdstid, antall år på skole eller resultat fra standardiserte språktester som mål for språkferdighet, men disse er ifølge Granger (2008: 6) kun *indirekte* mål for ferdighet, og kan i virkeligheten inkludere flere ulike nivå. Dette påvirker reliabiliteten til forskningen, og det er behov for mer direkte mål for ferdighetsnivå om en vil unngå at språklige trekk som egentlig har opphav i ferdighetsnivå, blir tolket til være et resultat av morsmålsbakgrunn eller andre faktorer (Granger 2008: 6). Dette problemet er også diskutert i Carlsen (2012), som gir en detaljert beskrivelse av hvordan påliteligheten til ferdighetsnivå kan sikres ved hjelp av nivåbeskrivelsene i Det felles europeiske rammeverket for språk (Europarådet 2001), som kan fungere som et eksternt og holistisk grunnlag for ferdighetsnivå i innlærerkorpus. Dette vil jeg komme tilbake til i omtalen av ASK under.

5.1.1 Innholdssiden av elektroniske innlærerkorpus

De fleste elektroniske innlærerkorpus inneholder i tillegg til språkbruksdata og ekstralingvistiske variabler, annen språkrelatert informasjon, for eksempel leksikalsk og grammatisk tagging av språkbruken. Dette gjør det lettere å søke på et høyere abstraksjonsnivå enn bare selve råteksten. Noen innlærerkorpus har også tagget tekstene med avvik fra målspråket (feilkoding), en praksis som har blitt kritisert for å representere et slags mangelsyn på innlærerspråk fordi det ikke følger prinsippet om å studere innlærerspråk på egne premisser uten målspråksnormen som referansepunkt. Dette vises gjerne til som *the comparative fallacy*, eller *sammenligningsfella*, og er et sentralt metodisk forbehold i studiet av innlærerspråk (Bley-Vroman 1983). Det kan likevel argumenteres for at feilkoding ikke nødvendigvis alltid bryter med verdien av å studere innlærerspråk på egne premisser. For eksempel mener Granger (2002: 53) det er fullt forenelig å både kunne undersøke innlærerspråk på egne premisser samtidig som man sammenligner med målspråket for å vurdere avvikene i forhold til dette:

It is important to stress that the two positions are not irreconcilable. One can engage in close investigation of interlanguage in order to understand the system underlying it and concurrently or subsequently compare the interlanguage with one or more native speaker norms in order to assess the extent of the deviation.

Også i Tenfjord, Hagen og Johansen (2006; 2009) er bruken av feilkoder diskutert med bakgrunn i ASK-korpuset, og konklusjonen er at bruken av feilkoder *ikke* bryter med prinsippet om at innlærerspråk skal studeres på egne premisser, ettersom feilkodene kun er et verktøy for beskrivelse av innlærerspråket, og ikke en teoretisk analyse:

Vårt eneste mål har vært å tilby en objektiv og preanalytisk beskrivelse av informantenes språkbruk, og for dette har vi funnet målspråksgrammatikken som velegnet, både for kategorisering av språklige kategorier og for beskrivelse av avvik fra målspråksnormen (Tenfjord et al. 2009: 68).

Et annet viktig moment er at feilkoding bare representerer en ganske liten del av et innlærerkorpus (Tenfjord et al. 2009: 53). Som beskrevet over inkluderer et korpus både språkbruksdata, ulike former for annotasjon og personvariabler, og innlærerspråk kan fint studeres gjennom korpuset også uten bruk av feilkoder. Feilkoding kan dessuten brukes som et verktøy for å gjøre mer spesifikke søk i korpuset, eller som en måte å

identifisere fenomener som kan være interessante studieobjekt, og alt dette kan gjøres uten å gå i sammenligningsfella eller å være for fokusert på innlæreres feil i seg selv.

5.2 Norsk Andrespråskorpus (ASK)

Norsk andrespråskorpus (ASK) er et korpus med tekster skrevet av innlærere av norsk som andrespråk, og er utviklet ved Universitetet i Bergen. Korpuset inneholder rundt 1700 tekster skrevet av innlærere med ti ulike morsmålsbakgrunner, samt et kontrollkorpus med 200 tekster skrevet av personer med norsk som morsmål (Tenfjord et al. 2009: 53). I tillegg til selve tekstene har korpuset informasjon om innlærerne, tekstvariabler, feilkoder og tagging av lemma, ordklasse og morfologiske og syntaktiske trekk.

Dataene i ASK kan søkes i ved å bruke en søkemeny eller ved å bruke standardiserte søkeuttrykk, og gir muligheten til å vise treffene i form av lister av ord og fraser man har søkt etter, vist i sin setningskontekst (konkordans), eller i form av hele teksten treffene står i. Det er også mulighet for ulike typer opptellinger og statistisk informasjon av selvvalgte distribusjoner og kollokasjoner. En nyttig funksjon er dessuten brukeres mulighet til å legge til selvvalgt annotasjon i konkordansene, som også blir søkbar etter innlegging. Dette gjør at man ikke bare kan bruke den informasjonen som ASK har gjort tilgjengelig i utgangspunktet, men også legge til egen informasjon tilpasset den forskningen man bruker korpuset til.

5.2.1 Design og utvalgskriterier

Tekstene i ASK er opprinnelig skrevet som svar på langsvarsoppgaver som er delprøver i to ulike standardiserte norske språktester: *Språkprøven i norsk for voksne innvandrere*, og *Test i Norsk – Høyere Nivå* (Tenfjord et al. 2009: 53). Disse vil fra nå av bli omtalt som «Språkprøven» og «Høyere nivå-testen».²¹ De to ulike testene er laget for å teste språkferdighet på to ulike nivå i Det felles europeiske rammeverket for språk (Europarådet 2001): Språkprøven er regnet som en B1-test, mens Høyere nivå-testen er regnet som en B2-test, og testtakerne må derfor ligge *på* eller *over* disse nivåene for å bestå.

²¹ Språkprøven er i dag erstattet av Språkprøve 2 og 3 (Carlsen & Moe 2019: 49).

Forskjellen mellom de to testene er i utgangspunktet knyttet til at de er ulike *testtyper*, og selv om de i prinsippet er knyttet til ulike ferdighetsnivå, kan man ikke automatisk anta at tekstene fra hver av testene faktisk er på B1- eller B2-nivå, da mange av tekstene ligger både over og under det nivået testen setter som grense for bestått. Dette skyldes at angivelsen av nivå gjelder *hele testen* og ikke bare den ene delprøven som er lagt inn i ASK (skriftlig langsvarsoppgave). For å få en mer reliabel vurdering av ferdighetsnivået i selve tekstene, ble 1200 av tekstene i korpuset vurdert på nytt som en del av ASKeladden-prosjektet (Carlsen 2010), noe jeg vil komme tilbake til lenger ned.

Det har vært tre utvalgs-kriterier for tekster til ASK-korpuset: *morsmålskriteriet*, *beståttkriteriet* og *infokriteriet* (Tenfjord et al. 2009: 55-56). *Morsmålskriteriet* gikk ut på at det skulle være et utvalg av tekster som var skrevet av innlærere med 10 ulike morsmål, med noe typologisk variasjon. Språkene som er representert er: albansk, bosnisk/kroatisk/serbisk, engelsk, nederlandsk, polsk, russisk, somali, spansk, tysk og vietnamesisk, og omfanget skulle i utgangspunktet være på 200 tekster fra hver gruppe, med 100 fra hver av testtypene, men dette antallet er ikke nådd for albansk, somali og vietnamesisk på høyere nivå (Tenfjord et al. 2009: 56).

Beståttkriteriet for utvalg av tekster til ASK innebar at kandidaten som skrev teksten måtte ha fått bestått på testen som helhet, både av personvern-hensyn og for å forsikre seg om at teksten hadde et minimum av kommunikativ funksjonalitet, noe som var viktig for å kunne feilkode tekstene (Tenfjord et al. 2009: 56-57). *Infokriteriet* er knyttet til det sentrale kriteriet om dokumentasjon av variabler. For at en tekst skulle være med i korpuset, måtte det være tilgjengelig biografisk informasjon om innlæreren som hadde skrevet den (Tenfjord et al. 2009: 57).

I tillegg til innlærertekstene fra de 10 morsmålsgruppene har korpuset et *kontrollkorpus* med 2 x 100 tekster skrevet av personer med norsk som morsmål. Informantene ble hentet fra kor, studenter eller andre grupper med en viss heterogenitet av mennesker (Tenfjord et al. 2009: 96). Disse informantene gjennomførte ikke hele språktesten, bare langsvarsoppgaven, men under så like forhold som mulig som ved ordinær testavvikling.

5.2.2 Innholdet i ASK

Som andre elektroniske innlærerkorpus, inneholder ASK både språkbruksdata, personvariabler, tekstvariabler, feilkoding og ulike former for lingvistisk informasjon. Disse vil bli presentert under.

5.2.2.1 Språkbruksdata

En stor fordel med tekstene i ASK er at språkbrukssituasjonen er relativt lik for alle tekstene, ettersom alle har blitt skrevet som del av en av to standardiserte tester. Tekstene handler likevel om en rekke ulike tema, og oppgaveformuleringene de er basert på, er noe ulike for de to testtypene. Mens tekstene fra Språktesten er ment å være åpne og handle om mer personlige tema, legger Høyere nivå-testen mer opp til argumenterende tekster med tema av mer allmenn interesse (Carlsen 2012: 169). I tillegg til dette er tekstene gjerne av ulik lengde, da anbefalt lengde på Språkprøven er rundt 200 ord, mens det er rundt 400 ord på Høyere nivå-testen, men det er stor variasjon i hvor lange enkelttekstene faktisk er.

5.2.2.2 Personvariabler

De ulike personlige variablene som er dokumentert kan deles i tre grupper: *generelle personvariabler*, som alder, kjønn og info knyttet til utdanning og arbeid; *språkvariabler*, som sier noe om hvilke andre språk innlærerne kan; og *målspråksvariabler*, som handler om opplæring i og bruk av norsk. Ikke alle variablene er oppgitt av alle kandidatene, faktisk er *morsmål* den eneste variabelen som er dokumentert for samtlige kandidater, men også hjemland, alder og kjønn er oppgitt av så å si alle.

5.2.2.3 Tekstvariabler og ferdighetsnivå

De tekstrelaterte variablene som er dokumentert og søkbare i ASK, er *teksttema* og *ferdighetsnivå*. Som nevnt over, var ikke ferdighetsnivå opprinnelig en tilgjengelig variabel i ASK, bortsett fra inndelingen i to ulike testtyper med B1 og B2 som beståttgrense. For å få en mer nyansert og valid inndeling i ferdighetsnivå ble 1200 av tekstene i ASK vurdert på nytt. Her ble det brukt opptil 10 sensorer per tekst, i motsetning til kun 2 i den opprinnelige vurderingen. For en mer omfattende beskrivelse av denne prosessen, se Carlsen (2012).

Ferdighetsnivå ses ofte på som en personlig variabel som avspeiler ferdighetsnivået til et individ, men vurderingsmåten som er brukt i ASK, er det Carlsen (2012: 165) kaller *tekst-sentrert*, altså at vurderingen av nivået kun baserer seg på selve teksten, i utgangspunktet uavhengig av innlærervariabler eller andre språkferdigheter. Denne typen vurdering skiller seg fra en *innlærer-orientert* vurdering som tildeler ferdighetsnivå med utgangspunkt i trekk ved innlæreren.

Nivåene som er brukt for å vise ferdighetsnivå i ASK er hentet fra Det felles europeiske rammeverket for språk, og går fra A1 til C2. Stort sett ligger tekstene fra Språkprøven på et lavere ferdighetsnivå enn på Høyere nivå-testen, slik en kan forvente med tanke på at testene måler forskjellige nivå, men det er også en del overlapping mellom testene i form av at noen tekster fra Språkprøven ligger høyere enn B1, og at noen tekster fra Høyere nivå-testen ligger lavere enn B2 (Carlsen 2010: 142; 2012: 178). Carlsen (2010: 142) viser også at nivåene ikke er jevnt fordelt i de ulike morsmålsgruppene: Mens engelske, polske og russiske testkandidater i gjennomsnitt ligger på B1-nivå, ligger somaliske og vietnamesiske kandidater noe under (A2/B1-nivå), mens de tyske ligger noe over (B1/B2). Carlsen (2010: 141; 2012: 177) argumenterer for at dette viser nødvendigheten av en re-evaluering av ferdighetsnivå i forhold til inndelingen i «lavere» og «høyere» nivå som var eneste indikator på ferdighetsnivå i ASK i utgangspunktet.

5.2.2.4 Feilkoder

Feilkodingen av tekstene i ASK innebærer identifisering og koding av alle avvik fra normert norsk bokmål. Kodingen bidro til søkbare *feilkoder*, som også er koblet til korreksjonsforslag. Det er fem hovedtyper av feilkoder i ASK: 1) leksikalske feil, 2) morfosyntaktiske feil, 3) syntaktiske feil, 4) tegnsetningsfeil og 5) uidentifiserbare feil (Tenfjord et al. 2009: 63). For en nærmere beskrivelse av disse feilkodene med autentiske eksempler, se Johansen 2007. En viktig forutsetning for implementeringen av feilkoder er det tidligere nevnte bestått-kriteriet:

Grunnen til at bestått-kriteriet har vært så viktig, er at det å identifisere språkbruksfeil er meningsløst og logisk umulig hvis man ikke i utgangspunktet har en rimelig klar oppfatning av hva som er ytringens intenderte innhold (Tenfjord et al. 2009: 59).

I tillegg til dette kriteriet ble det operert med to prinsipper i gjennomføringen av selve feilkodingen for å unngå for mye subjektivitet i kodingen. *Det pragmatiske probabilitetsprinsippet* gikk ut på å velge den mest sannsynlige fortolkningen med utgangspunkt i tekst og situasjon, mens *det minimale modifikasjonsprinsippet* gikk ut på å velge det korreksjonsforslaget som innebar minst mulig inngrep i teksten (Tenfjord et al. 2009: 60).

5.2.2.5 POS og grammatisk tagging

I tillegg til denne spesifikke annoteringen med feilkoder har ASK også mer generell språklig annotering. For det første er det automatisk ordklassetagging som gir hvert ord en ordklasse og morfologisk form, for det andre er det lemmatisering, som tildeler hvert ord et leksem, og for det tredje er det parsing som angir de morfosyntaktiske kategoriene og viser hvordan ordene er forbundet syntaktisk (Tenfjord, Meurer & Hofland 2006 1822-1823). Denne informasjonen har blitt lagt til språkdataene automatisk ved hjelp av Oslo-Bergen-taggeren, som er en grammatisk tagger for bokmål og nynorsk (Johannessen, Hagen, Lynum & Nøklestad 2012).²²

Å bruke en slik automatisk tagger på språkbruk som er mer eller mindre avvikende i forhold til den standarden som taggeren er laget for (korrekt norsk bokmål), kan være problematisk, og for å bøte på noen av problemene har man i ASK latt taggeren basert seg på korreksjonsforslagene til ord som har ortografiske feil heller enn på de faktiske ordforekomstene i innlærertekstene (Tenfjord et al. 2009: 62).

5.3 ASK i studiet av transfer

Siden ASK består av nesten 2000 tekster skrevet av informanter med flere forskjellige morsmålsbakgrunner, og inneholder dokumentasjon av en rekke variabler, er det altså et velegnet utgangspunkt for studier av transfer etter Jarvis' (2000) rammeverk, som ble omtalt i 2.4. Korpuset legger til rette for både kvantitative og kvalitative undersøkelser av de to gruppebaserte transfereffektene, altså sammenligning av innlærere med ulike morsmål (*intragruppe-homogenitet* og *intergruppe-heterogenitet*). Transfereffekten *morsmålssamsvar*, som viser til paralleller mellom morsmål og innlærerspråk, ligger

²² Oslo-Bergen-taggeren (Johannessen et al. 2012) er en grammatisk tagger for bokmål og nynorsk utviklet av Universitetet i Oslo og UniComputing i Bergen. Nettside: <http://tekstlab.uio.no/obt-ny/index.html>

derimot utenfor korpusets område siden ASK bare har innlærertekster, og ikke morsmålstekster for de ulike innlærergruppene. Informasjon om morsmålet må da enten oppnås gjennom en formell beskrivelse av språket, for eksempel fra eksterne kilder, eller data fra andre korpus som har morsmålsdata (Jarvis & Pavlenko 2008: 49).

Den tredje komponenten i Jarvis' rammeverk, kontroll på ulike variabler, er relativt godt tilrettelagt for gjennom ASK, selv om ikke alle variabler er dekket, for eksempel sosial bakgrunn, personlighet og anlegg for språklæring, men dette er uansett variabler som verken er lette å innhente eller å måle. Granger (2004: 126) mener uansett at det er urealistisk å skulle kontrollere for alle variabler siden det er så mange faktorer som påvirker innlærerspråk:

[...] there are so many variables that influence learner output that one cannot realistically expect ready-made learner corpora to contain all the variables for which one may want to control.

I praksis kan dette bety at når en bruker et korpus, vil det gjerne være utvalget i korpuset som styrer populasjonen i undersøkelsen. ASK har nok ikke et utvalg som er representativt for alle norskinnlærere, for eksempel fordi det er begrenset til de innlærerne som faktisk har *tatt* en av to språktester og dessuten *bestått*. Disse to forholdene kan ha medført enkelte skjevheter når det gjelder ulike innlærervariabler, for eksempel ser man at det er flere kvinner enn menn (hhv. 1347 og 585 stk.) som er representert, og at det er flest kandidater med høyskole-/universitetsutdannelse (1243 stk.), noe som kan være en avspeiling av hvem som tar testene og består dem.

5.4 Oppsummering av kapittel 5

Kapittel 5 har beskrevet noen generelle egenskaper og problemstillinger knyttet til elektroniske innlærerkorpus, i tillegg til å beskrive ASK, som er datagrunnlaget for min egen studie.

I foreliggende studie har jeg brukt ASK til å lage to tilpassede delkorpus. For det første et større delkorpus på 1200 tekster til undersøkelse 1, og for det andre et mindre delkorpus på 123 tekster til undersøkelse 2. Det første delkorpuset bidrar med et relativt stort datagrunnlag som tillater bruk av statistiske tester, mens det andre delkorpuset er betraktelig mindre, noe har åpnet for manuell annotering av egenskaper som er relevante

for hypotesene om morsmålspåvirkning. Jeg vil komme tilbake til utvalg og bruk av korpuset hver for seg for undersøkelse 1 (6.1.1) og undersøkelse 2 (8.2).

6 UNDERSØKELSE 1: KVANTITATIV EFFEKT AV MORSMÅLET

Følgende kapittel presenterer første del av avhandlingens empiriske undersøkelser, som er en kvantitativ undersøkelse av markering av bestemthet. Undersøkelse 1 er utformet for å besvare forskningsspørsmål 1:

Påvirker morsmålet hvor ofte bestemte former brukes i obligatoriske kontekster i innlærerspråket til innlærere uten grammatisk bestemthet i morsmålet?

Dette gjøres gjennom å undersøke om det er lavere andel av markering i tekster skrevet av -GB-innlærere, enn det er i tekster skrevet av +GB-innlærere. Dette er en ren kvantitativ undersøkelse som teller opp antall substantiv som har bestemt form i obligatoriske kontekster, altså kontekster hvor det ifølge målspråksnormen skal være bestemt form. Deler av denne undersøkelsen er tidligere publisert i en artikkel (Johansen 2010), men er her utvidet og noe endret i forhold til den publiserte utgaven.

Undersøkelse 1 ser på 1200 tekster fra ASK, skrevet av innlærere med seks ulike morsmål, tre -GB-språk og tre +GB-språk. Dette delkorpuset fra ASK vil bli omtalt som «GB-korpuset», og vil i løpet av undersøkelsen deles opp i mindre deler, hvor faktorer som testtype/-nivå og morsmålsbakgrunn holdes konstant.

Kvantitative forskjeller mellom de to innlærergruppene antas å kunne gi seg utslag på to måter:

- 1) -GB-gruppen vil ha lavere markeringsprosent enn +GB-gruppen.
- 2) -GB-gruppen vil ha færre tekster med 100 % markering enn +GB-gruppen.

Innledningsvis i dette kapittelet vil jeg presentere de metodiske sidene av undersøkelse 1: I delkapittel 6.1 beskriver jeg datamaterialet, størrelsene *obligatorisk kontekst* og *markeringsprosent*, hvordan dataene er søkt fram i ASK, og hvordan jeg har gått fram for å måle de ulike transfereffektene. Delkapittel 6.2 viser resultatene fra GB-korpuset, mens **Feil! Fant ikke referansekilden.** bryter ned materialet og ser på de to testtypene hver for seg (Høyere nivå og Språkprøven). I 6.4 ser jeg kun på delmengden av tekster skrevet av polske og engelske innlærere. I kapittel 6.5 diskuterer jeg funnene, spesielt med tanke på transfereffekten *morsmålssamsvar*, men også med et kritisk blikk

på utvalget av substantivfraser og egenskaper ved disse som burde ha vært tatt hensyn til i måten undersøkelsen ble gjort på.

6.1 Metodiske sider av undersøkelse 1

6.1.1 Datamaterialet

GB-korpuset inneholder 1200 tekster skrevet av innlærere med seks ulike morsmål - tre med grammatisk bestemthet i morsmålet (engelsk, tysk og nederlandsk), og tre uten grammatisk bestemthet (polsk, russisk og bosnisk-kroatisk-serbisk). Hver morsmålsgruppe er representert med 200 tekster hver, 100 fra Språkprøven (SP), og 100 fra Test i norsk - høyere nivå (HN). Selv om ASK inneholder tekster skrevet av innlærere fra ytterligere fire morsmålsgrupper, tre som har grammatisk bestemthet (albansk, somali og spansk), og ett som ikke har det (vietnamesisk), er de seks valgt fordi de representerer språklig nærhet innad i hver av gruppene: tre slaviske språk uten bestemthet, og tre vest-germanske, alle med bestemthet uttrykt ved hjelp av foranstilte artikler.

6.1.2 Obligatoriske kontekster og markeringsprosent

Undersøkelse 1 retter seg mot hvor ofte substantiv har bestemt form i obligatoriske kontekster. Størrelsen *obligatorisk kontekst* viser til språklige kontekster hvor et gitt grammatisk trekk er obligatorisk i målspråket, og dette er et mye brukt referansepunkt for å kvantifisere mestring og bruk av grammatiske trekk i innlærerspråk. *Markeringsprosenten* måler antall kontekster hvor formen er brukt, relativt til det totale antall kontekster hvor formen er obligatorisk. I denne sammenhengen vil det altså være antall substantiv i bestemt form, delt på det totale antall kontekster hvor bestemt form skal brukes, ganget med 100 (se under). Dette målet ekskluderer kontekster hvor bestemt form er overgeneralisert, altså brukt der det ikke skal brukes.

6.1.3 Korpus-søk

For å hente ut informasjon om markeringsprosenten i tekstene fra ASK, har jeg kombinert ulike søk hvor jeg henter ut kvantitative data for hver og en av tekstene i

GB-korpuset. Søkene er basert på en kombinasjon av ordklasse, grammatiske trekk, feilkoden F (gal form), og foreslått korreksjon.²³ Søkene jeg har gjort, er følgende:

SØK 1:

Substantiv i bestemt form som ikke er kodet med feilkoden F

Treff: Alle substantiv hvor det er brukt bestemt form, og hvor det er korrekt.

SØK 2:

Substantiv i ubestemt form som er kodet med feilkoden F og korrigert til bestemt form

Treff: Alle substantiv hvor det er brukt ubestemt form, men hvor det korrekte er bestemt form.

Summen av disse to søkene gir antall *obligatoriske kontekster*, mens *markeringsprosenten* er den relative frekvensen av alle forekomstene som faktisk har bestemt form (søk 1), noe som kan gjengis med formelen:

$$\text{Markeringsprosent} = \frac{\text{Søk 1 (antall substantiv med korrekt bestemt form)}}{\text{Søk 1} + \text{Søk 2 (antall obligatoriske kontekster)}} \times 100$$

Informasjonen er hentet ut gjennom distribusjonsfunksjonen i ASK, som kan gi antall for hver og en av tekstene i GB-korpuset, og dette er overført til et regneark hvor jeg har foretatt utregningene og fått en markeringsprosent for hver enkelttekst. Tilgang til antall per tekst har gjort det mulig å utføre en statistisk test som tar individuelle tall i betraktning.²⁴

Det er viktig å understreke at disse søkene har gitt meg alle substantiv som skal være i bestemt form, uavhengig av *hvilken type substantivfrase de er kjerne i*. I norsk kan det være flere grunner til at bestemt form er påkrevd på substantivet, og både referensielle forhold og den syntaktiske sammenhengen styrer om det skal, eventuelt kan, være bestemt eller ubestemt form av substantivet (Hagen 1998: 54). Dette er noe jeg vil komme tilbake til i 6.5.3.

6.1.4 Transfereffekter og statistisk testing

Siden det er et overordnet spørsmål i undersøkelse 1 hvorvidt markeringsprosent kan kobles til fravær eller tilstedeværelse av grammatisk bestemthet i innlærernes morsmål,

²³ Feilkoden F står for «gal form» og brukes i alle tilfeller der innlæreren har valgt feil bøyingsform av et ord, for eksempel bestemt form der det skal være ubestemt (Johansen 2007).

²⁴ Å hente ut antall på individnivå er mulig i ASK ved å vise søkene med attributtet 'tekst-id' i distribusjonsvisningen (Johansen 2011).

er det et poeng å kunne teste statistisk om det er forskjeller mellom de to innlærergruppene som har og ikke har GB i morsmålet. I tråd med Jarvis' kriterium om påvisning av transfereffekter, har jeg derfor ønsket å sammenligne gruppene med hensyn til *intragruppe-homogenitet* og *intergruppe-heterogenitet* i språkbruken, altså om de to innlærergruppene er ulike hverandre i språklige valg, men homogene innad i sine respektive grupper (se 2.4.3.1). Heterogenitet og homogenitet kan ses på som en skala fra total forskjell til total likhet, og i de fleste tilfeller vil forholdet mellom to grupper ligge et sted imellom (Jarvis & Pavlenko 2008: 46). Heterogenitet og homogenitet må derfor ses *i forhold til* hverandre, på den måten at homogeniteten *mellom* gruppene skal være mindre enn homogeniteten *innad* i gruppene (Jarvis & Pavlenko 2008: 46). Variansanalyse er en statistisk prosedyre som måler nettopp dette. Varians er et mål for spredningen i en gruppe, altså de individuelle avvikene fra gjennomsnittet, og med en variansanalyse sammenligner man variansen *mellom* grupper med variansen *innenfor* hver av gruppene. Når man bare har to grupper som skal sammenlignes, slik jeg har, sammenfaller variansanalyse med det som kalles en *t-test* (Rowntree 2000: 143).

En t-test er altså aktuelt å bruke for å avgjøre spørsmålet om både heterogenitet og homogenitet samtidig, men kun om dataene tilfredsstillt kravene til å bruke en slik test. Kravene til å bruke en t-test er blant annet at dataene er normalfordelte og at gruppene har lik varians, men dataene mine tilfredsstillt ikke disse kravene da de ikke er normalfordelte.²⁵ I stedet for har jeg brukt en ikke-parametrisk variant av t-testen, *Mann-Whitney U*, for å teste om forskjellene mellom de to gruppene er signifikante. Dette er en test som ikke baserer seg direkte på gjennomsnittet i gruppene, men som i stedet rangerer hver individuelle innlærertekst etter markeringsprosent, tildeler et rangeringsnummer, og deretter sammenligner den gjennomsnittlige rangeringsverdien («rang») i de to gruppene (Larsen-Hall 2010: 377). Siden markeringsprosenten ved en slik test blir omgjort til et rangeringstall, innebærer den en forenkling av dataene mine, men selv om testene jeg bruker ikke direkte baserer seg på gjennomsnitt og de faktiske markeringsprosentene i hver tekst, vil jeg oppgi disse tallene i beskrivelsen av

²⁵ Dette har jeg funnet ut ved å bruke en Kolmogorov-Smirnov-test som sjekker normalfordeling (Hinton et al. 2004: 97).

resultatene. I 6.2.1 kommer jeg inn på nok en årsak til å velge en ikke-parametrisk test, nemlig at markeringsprosentene i mitt materiale ikke er *intervalldata*, noe en parametrisk test også krever for å bli brukt.

Den statistiske testingen vil altså vise om det foreligger et *gruppebasert* argument for transfer, der en viser om en bestemt språklig tendens (høyere eller lavere markeringsprosent) er spesifikk og representativ for én innlærergruppe. Den tredje transfereffekten, *morsmålssamsvar*, er i denne delen av studien knyttet til at fravær av grammatisk bestemthet i morsmålet fører til *mindre* bruk av bestemt form i obligatoriske kontekster, mens det å ha grammatisk bestemthet i morsmålet fører til *mer* bruk. Denne måten å se morsmålsamsvar på skal jeg imidlertid komme tilbake til og diskutere etter selve undersøkelsen (6.5.2).

I tillegg til å undersøke forskjellene mellom gruppene når det gjelder markeringsprosent, har jeg ønsket å undersøke om det er forskjeller mellom gruppene med tanke på hvor mange tekster som har 100 % markering i obligatoriske kontekster. Dette innebærer en opptelling av antall tekster i hver gruppe som oppfyller kriteriet, og for å teste forskjellene mellom gruppene har jeg brukt en *kjikkvadrat-test*. Dette er også en ikke-parametrisk test, og den tester forskjeller mellom en eller to kategoriske variabler (Larsen-Hall 2010: 206).

Ved hjelp av metodene som er beskrevet over, får jeg vist gruppetendenser med fokus på individuell variasjon. Jeg får testet om det er slik at den ene gruppen generelt sett har lavere markeringsprosent enn den andre, og jeg får testet forskjeller i mestringsnivå, i tråd med følgende prediksjoner:

- 1) Tekstene skrevet av -GB-innlærere vil generelt sett ha lavere markeringsprosent enn tekstene skrevet av +GB-innlærerne, på den måten at gruppegjennomsnittet er signifikant lavere, og at variasjonen innad i gruppen er homogen opp mot dette gjennomsnittet.
- 2) -GB-gruppen vil ha signifikant færre innlærertekster med 100 % markering enn +GB-gruppen.

Grenseverdien for den statistiske testingen min (alfa) er 0,05. Det innebærer at eventuelle forskjeller mellom gruppene regnes som statistisk signifikante dersom p-verdien som kommer ut av den statistiske testen, er lavere enn 0,05, noe som betyr at

det er mindre enn 5 % sjanse for at forskjellene mellom gruppene er tilfeldige. Siden begge prediksjonene er det en kaller *ensidige*, altså at man ikke bare predikerer at det er en forskjell mellom gruppene, men også hvilken retning forskjellen vil slå ut (-GB-gruppen vil ligge *lavere* enn +GB-gruppen), skal p-verdiene halveres (Hinton, Brownlow, McMurray & Cozens 2004: 112). P-verdiene som oppgis i gjennomgangen av resultatene, er derfor den halverte verdien.

6.2 Hele GB-korpuset

Jeg vil nå presentere resultatene for undersøkelse 1, først for GB-korpuset i sin helhet, og deretter brutt ned til de to ulike testtypene Språkprøven (SP) og Høyere nivå (HN). Til slutt vil jeg også undersøke delkorpuset av tekster skrevet av polske og engelske innlærere. For hver av delene vil jeg først se deskriptivt på antall obligatoriske kontekster i de ulike gruppene, og deretter presenteres resultatet for henholdsvis markeringsprosent og tekster med 100 % markering.

6.2.1 Antall obligatoriske kontekster i GB-korpuset

Tabell 6-1 (under) viser deskriptive data for de to innlærergruppene i GB-korpuset i form av antall substantiv og antall obligatoriske kontekster på gruppenivå. Obligatoriske kontekster (OBL) er oppgitt både i absolutte tall og i prosentandel i forhold til antall substantiv. Den nederste raden viser tall på individnivå, både som gjennomsnittlig antall per tekst, standardavvik og maksimum- og minimumsverdi for antall obligatoriske kontekster. Tallene i parentes angir maksimumsantallet når tekster som avviker spesielt mye fra gjennomsnittet («utliggere») ses bort fra. Jeg vil komme tilbake til hva dette innebærer under.

Tabell 6-1: Deskriptive data for GB-korpuset (obligatoriske kontekster)

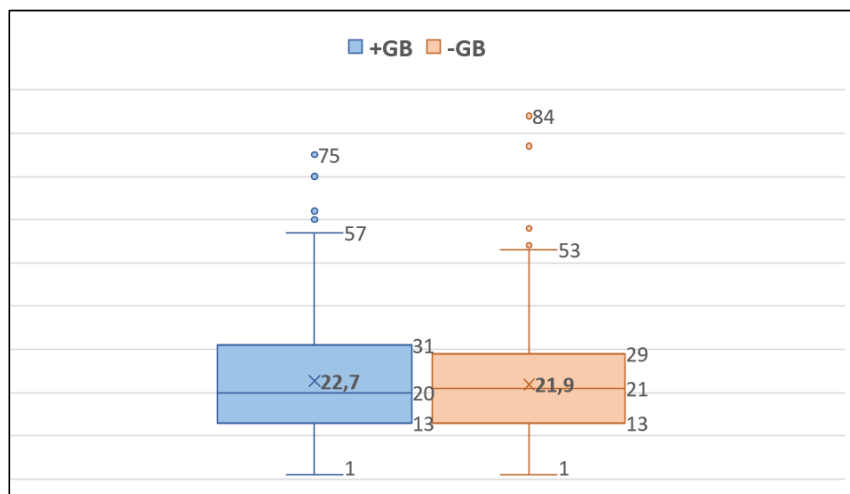
GB-korpus	+GB-innlærere		-GB-innlærere			
Tekster	600		600			
Substantiv	46 812		45 886			
OBL	13620	29,1 %	13142	28,6 %		
Gj.snitt	23 St. avvik: 12,7	Min. 1	Maks. 75 (57)	22 St. avvik: 11,6	Min. 1	Maks. 84 (53)

De to gruppene er veldig jevne med tanke på antall substantiv, da begge ligger på rundt 46000. Om en sammenligner andelen av obligatoriske kontekster, er denne også veldig jevn, da det i begge grupper er rundt 29 % av alle substantiv som krever bestemt form.

Gjennomsnittlig antall obligatoriske kontekster per tekst ligger på henholdsvis 22 og 23 forekomster i de to gruppene, med et relativt likt standardavvik (-GB 12,7; +GB 11.6), men om en går nærmere inn i den individuelle variasjonen, er det en god del variasjon fra tekst til tekst, der noen tekster bare har én obligatorisk kontekst, mens andre er oppe i 80 stykker. Denne relativt store spredningen gjelder imidlertid i begge gruppene.

Figur 6-1 (under) er et boksdiagram som viser hvordan antall obligatoriske kontekster er fordelt i tekstene i de to gruppene. Et boksdiagram skiller seg fra vanlige søylediagram ved å oppgi mer informasjon om individuell variasjon (Hinton et al. 2004: 44-46): Selve boksen angir hvilket antall av obligatoriske kontekster 50 % av tekstene ligger innen, mens streken gjennom boksen viser til medianverdien, og krysset viser til gjennomsnittet. Prikkene viser utliggere i materialet, som er verdier som ligger mer enn tilsvarende 1,5 bokslengder fra boksen, og dette er enkelttilfeller som ikke helt passer inn i resten av distribusjonen. Strekene som går loddrett ut over og under boksen kalles *værhår*, og kobler sammen de minimums- og maksimumsverdiene som ikke er utliggere.

Figur 6-1 Den individuelle spredningen av antall obligatoriske kontekster i GB-korpuset



Boksdigrammet på forrige side viser altså at 50 % av tekstene i begge gruppene har mellom 13 og rundt 30 obligatoriske kontekster. Av de resterende tekstene, har de aller fleste mellom 30 og cirka 55, og relativt få tekster har færre enn 13 obligatoriske kontekster. I hver av gruppene er det 4-5 utliggerne, men i et korpus på 1200 tekster er disse et klart mindretall. Om en fjerner utliggerne, ligger minimum- og maksimumsverdien mellom 1 og 53 i +GB-gruppen og 1 og 57 i -GB-gruppen. Gruppene er med andre ord relativt like i sin spredning, selv om det er en god del variasjon fra tekst til tekst i begge gruppene.

Tallene over sier noe om hvor like gruppene er med tanke på det å produsere obligatoriske kontekster, som videre kan ha konsekvenser for sammenlignbarheten mellom gruppene. Gruppene viser seg imidlertid å være veldig like her, både i det totale antallet obligatoriske kontekster og i måten de individuelle tallene er spredt på. Begge gruppene har en god del spredning, men de sprer seg på relativt like måter, og det er dermed ikke snakk om noen gruppespesifikk spredning som kan påvirke sammenlignbarheten.

Det store spennet i antall obligatoriske kontekster som fins i begge gruppene, har likevel noen konsekvenser, nemlig at markeringsprosentene for hver og en av tekstene risikerer å være mindre sammenlignbare og mindre egnet for visse statistiske tester. Dette kan illustreres med at i en tekst med to obligatoriske kontekster vil 50 % markering tilsvare én kontekst, mens i en tekst med 26 obligatoriske kontekster, vil 50 % tilsvare 13. Hvor mange kontekster 1 % tilsvare vil på grunn av dette variere fra tekst til tekst, og man kan dermed ikke anse prosenttallet i de ulike tekstene som en tallrekke med *like lang distanse* mellom hvert punkt, altså det som kalles *intervalldata*, og som er et av kravene som stilles til bruk av parametriske tester (Hinton 2004: 210).²⁶ Det ideelle for undersøkelsen hadde selvsagt vært om alle tekstene hadde like mange obligatoriske kontekster, men dette er vanskelig å få til ved fri språklig produksjon, hvor det vil være individuelt hvor mange kontekster språkbrukeren produserer. Problemet kan løses med å velge ut visse tekster med tilnærmet likt antall obligatoriske kontekster, men det kan også løses med å unngå den type tester som krever intervalldata, slik t-tester gjør.

²⁶ Takk til Bård Uri Jensen, førsteamanuensis ved Høgskolen i Innlandet, for å ha gjort oppmerksom på dette problemet med markeringsprosentene.

Hinton et al. (2004: 124) skriver at ikke-parametriske tester er det naturlige valget når man ikke har intervalldata, og at om man er i tvil om man har intervalldata, så skal man anta at man ikke har det, og heller bruke ikke-parametriske tester (Hinton 2004: 210). Som jeg avklarte tidligere, vil jeg uansett bruke ikke-parametriske tester på grunn av mangel på normalfordeling i datasettet mitt.

For å komme tilbake til dataene vi har å gjøre med her, ser vi altså at gruppene har et godt sammenligningsgrunnlag, da de har relativt lik mengde obligatoriske kontekster og spredningen er ganske lik i begge gruppene. Men vi ser også at den spredningen en finner i antall obligatoriske kontekster i begge gruppene, skaper et problem for hvordan markeringsprosentene kan brukes ved statistiske tester, og at tallene må observeres med det i bakhodet at én og samme markeringsprosent hos ulike innlærere, kan representere to ulike situasjoner hva angår antall av obligatoriske kontekster.

6.2.2 -GB-korpuset: Forskjeller i markeringsprosent

Jeg vil nå presentere *markeringsprosenten* i de to gruppene i hele GB-korpuset, og undersøke om det er en forskjell mellom gruppene som kan tilskrives noe annet enn tilfeldigheter, og som dermed gir grunnlag for en konklusjon om heterogenitet mellom gruppene. I tabell 6-2 (under) vises tall på gruppenivå og individnivå, i tillegg til teststatistikken og p-verdien fra den statistiske testen.

Tabell 6-2 Markeringsprosent (og rang) i de to innlærergruppene i GB-korpuset

GB-korpus	+GB-innlærere	-GB-innlærere	Mann-Whitney <i>U</i>
Antall tekster	600	600	Z = - 6.342 p = 0,001
OBL	13620	13143	
Markerte	12 237	11 179	
Gjennomsnitt (markeringsprosent)	87,7 % St. avvik: 13,9	83,7 % St. avvik:13,8	
Gjennomsnitt (rang)	663,74	537,26	

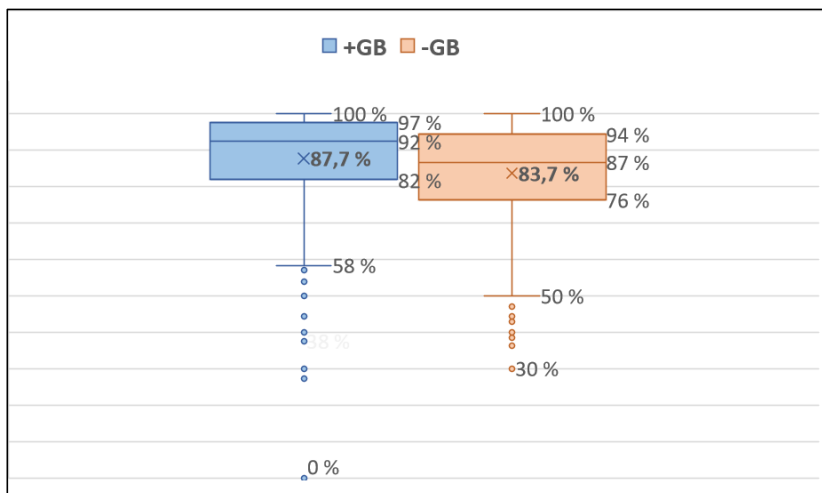
Når det gjelder antall obligatoriske kontekster som er markert med bestemt form, er det en forskjell mellom de to gruppene: Det viser seg at -GB-gruppen markerer i gjennomsnitt 83,7 % av de obligatoriske kontekstene, noe som er fire prosentpoeng mindre enn +GB-gruppen, som har markert 87,7 %. Denne forskjellen ble testet med Mann Whitney U, og forskjellen er signifikant ($p < 0,01$). Disse dataene tyder altså på at

det er forskjeller mellom de to gruppene, og at det er -GB-gruppen som har lavest grad av markering, som forventet.

Også her kan funnene illustreres med boksdiagrammet under (figur 6-2). At boksene som representerer de to gruppene, ikke er helt jevnt plassert i diagrammet, viser oss at det er en viss forskjell mellom dem: Halvparten av +GB-innlærerne ligger mellom 82 og 97 % markering, mens halvparten av -GB-innlærerne ligger mellom 76 og 94 %, altså et hakk lavere. Boksene overlapper i stor grad, men som den statistiske testen har vist, er forskjellen stor nok til at den kan regnes som signifikant.

Begge gruppene har en del utliggerer (-GB 8 stk.; +GB 10 stk.), men ser vi bort ifra disse, som til sammen utgjør bare 1,5 % av materialet, ligger begge gruppene samlet sett på over 50 % markering, men -GB-gruppen ligger altså generelt lavere enn +GB-gruppen på de fleste områdene som boksdiagrammet illustrerer.

Figur 6-2 Den individuelle spredningen for markeringsprosent i GB-korpuset



6.2.3 GB-korpuset: antall tekster med 100 % markering

I tillegg til å sammenligne markeringsprosjenter, har jeg sammenlignet antall tekster i hver gruppe som har 100 % bruk av bestemt form i obligatoriske kontekster, altså innlærere som ikke gjør noen feil. Tabell 6-3 (neste side) viser fordelingen i de to gruppene. Som forventet er antallet høyere hos +GB-innlærerne (23 %) enn hos -GB-innlærerne (18 %), og denne forskjellen er også statistisk signifikant ($\chi^2=14,2277$, $p<0.01$).

Tabell 6-3 Andel tekster med 100 % markering i GB-korpuset

GB-korpus	100 % markering		< 100 % markering		Kjikkvadrat-test
+GB-innlærere	138	23 %	462	77 %	$\chi^2=14,2277$ p=0.0002
-GB- innlærere	87	15 %	513	85 %	

6.2.4 Oppsummering av funn i GB-korpuset

Det viser seg altså å være forskjeller mellom innlærere med og uten bestemthet i morsmålet for begge de to forholdene jeg har undersøkt. Det var en statistisk signifikant forskjell mellom gruppene i markeringsprosent, og det var også en statistisk signifikant forskjell med hensyn til antall tekster som har 100 % markering. Så langt har jeg altså funnet støtte for prediksjonen om at -GB-innlærere i mindre grad markerer obligatoriske kontekster for bestemt form enn +GB-innlærere. Det er med andre ord funn av både transfereffekten *intergruppe-heterogenitet* i form av forskjeller mellom gruppene, og effekten *intragruppe-homogenitet* i form av likheter innad i hver av gruppene (se 6.1.4 for sammenhengen mellom disse transfereffektene og statistisk testing). Jeg vil videre se på om funnene er de samme når materialet er delt opp i de to ulike språktesttypene som er representert i ASK-korpuset, altså Språkprøven (SP) og Test i norsk - høyere nivå (HN).

6.3 SP- og HN-korpuset

Som vi så i 6.2.1, var det store forskjeller i antall obligatoriske kontekster i de ulike tekstene i GB-korpuset, fra så lite som én kontekst til over 80. Dette skyldes delvis at tekstene fra HN-testen generelt sett er mye lengre enn de som er skrevet til SP-testen, og dermed vil de gjerne også ha et høyere antall obligatoriske kontekster. I tillegg til tekstlengde er det en generell forskjell i ferdighetsnivå mellom testene, siden HN-testen måler B2 og over, mens SP-testen måler B1 og over. Forskjellen i testtype vil derfor også til en viss grad avspeile variabelen ferdighetsnivå, selv om ferdighetsnivåene varierer noe mellom kandidatene innad i hver av testene (se 5.2.2.3). For å redusere effekten av ulik tekstlengde og ferdighetsnivå, har jeg undersøkt tekstene fra de to språktestene hver for seg, i det jeg har kalt henholdsvis «SP-korpuset» og «HN-korpuset».

6.3.1 Størrelse og obligatoriske kontekster i SP- korpus og HN-korpus

At antall obligatoriske kontekster er forskjellig for de to ulike testene, kommer fram av tabell 6-4 (under), som sammenligner de to testtypene uten å skille mellom de to innlærergruppene.

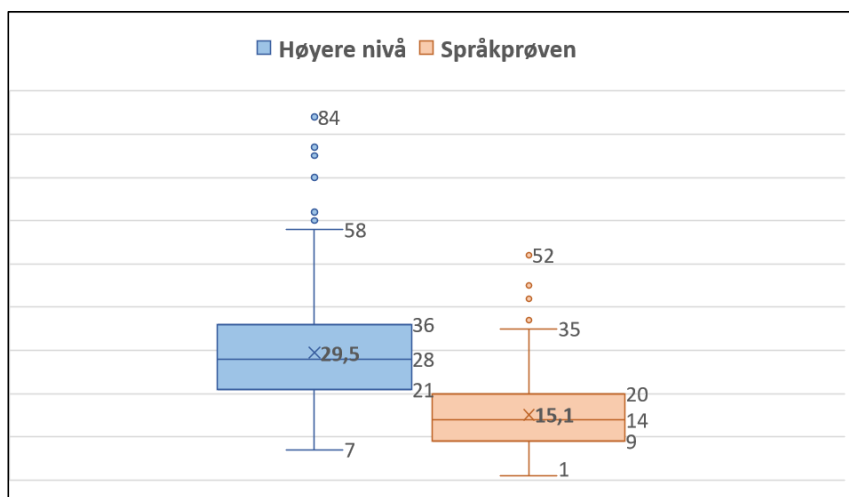
Tabell 6-4: Antall obligatoriske kontekster i HN-korpuset og SP-korpuset.

	HN-korpus			SP-korpus		
Tekster	600			600		
Substantiv	59824			32874		
OBL totalt	26774					
OBL	17691	66 % av total OBL 29,6 % av ant. subst		9083	34 % av total OBL 27,6 % av ant. subst	
Gj.snitt	29	Min. 7	Maks. 84 (58)	15	Min. 1	Maks. 52 (35)

Av alle obligatoriske kontekster i hele GB-korpuset samlet, er hele 66 % fra HN-korpuset, som dermed har nesten dobbelt så mange obligatoriske kontekster som SP-korpuset med sine 34 %. I tillegg ser vi at gjennomsnittlig antall per tekst er veldig ulikt: Gjennomsnittet for HN-korpuset er 29 kontekster per tekst, mens det for SP-korpuset er bare 15, så også her utgjør HN-korpuset det dobbelte av SP-korpuset. Til tross for forskjellene i størrelse, ser vi at den relative andelen av obligatoriske kontekster i forhold til antall substantiv, er ganske lik i de to delkorpusene (HN 26,9 %; SP 27,6 %).

Av figur 6-3 (neste side) kan vi også se at boksene som representerer de sentrale 50 % av innlærerne, ikke overlapper i det hele tatt i de to delkorpusene. 50 % av HN-tekstene har mellom 21 og 36 obligatoriske kontekster, mens SP-tekstene har mellom 9 og 20. Om vi ser bort fra utliggere, har HN-tekstene mellom 7 og 58 obligatoriske kontekster, mens SP-tekstene har mellom 1 og 35. Forskjellene i antall obligatoriske kontekster mellom de to korpusene er altså veldig tydelig, men på ingen måte uventet ettersom de to språktestene etterspør tekster av ulik lengde. I seg selv er dette en god grunn til at de to delkorpusene bør undersøkes hver for seg, både med tanke på de ulike ferdighetsnivåene, men også med tanke på det tidligere omtalte problemet med hvordan stor variasjon i antall obligatoriske kontekster per tekst påvirker kvalitetene til markeringsprosenten (se 6.2.1). Når spredningen i antall obligatoriske kontekster innad i gruppen reduseres, vil også markeringsprosentene bli mer sammenlignbare, selv om det fremdeles ikke er snakk om et likt antall kontekster i hver enkelt tekst.

Figur 6-3 Den individuelle spredningen av antall obligatoriske kontekster i HN- og SP-korpuset



Jeg vil nå undersøke om det er forskjeller i markeringsprosent mellom de to GB-gruppene også innen hver av de to språktestene, og dette gjøres på samme måte som for GB-korpuset: Først ser jeg på forskjellen i markeringsprosent, deretter på forskjellen i antall tekster med 100 % markering.

6.3.2 SP-korpuset

6.3.2.1 SP-korpuset: antall obligatoriske kontekster for bestemt form

Tabell 6-5 (under) viser antall og andel av obligatoriske kontekster i de to innlærergruppene i SP-korpuset, og vi ser at det gjennomsnittlige antallet obligatoriske kontekster er relativt likt i de to gruppene (-GB 16; +GB 14). Minimum- og maksimumsantall strekker seg fra 1 til 31 (+GB) og 37 (-GB) om utligger tas vekk.

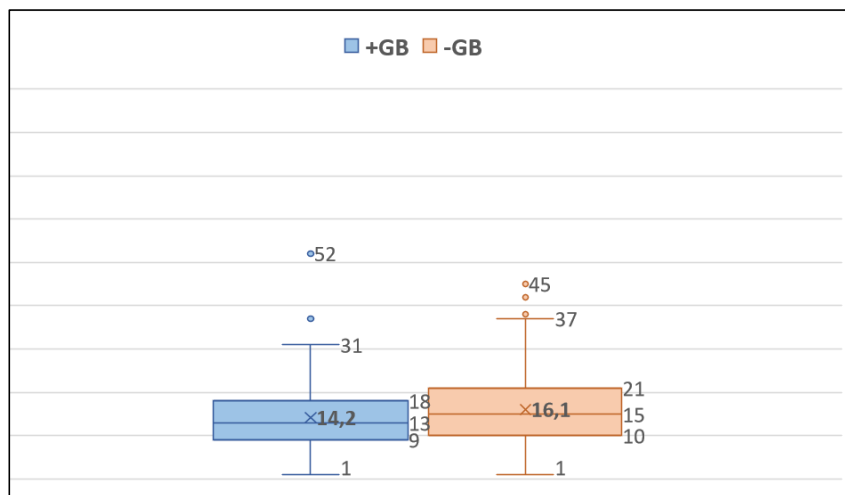
Tabell 6-5 Deskriptive data for SP-korpuset (obligatoriske kontekster)

SP-korpus	+GB-innlærere			-GB-innlærere		
Tekster	300			300		
Substantiv	16071			16803		
OBL	4265	26,5 %		4818	28,7 %	
Gj.snitt	14	Min.	Maks.	16	Min.	Maks.
		1	52 (31)		1	45 (37)

I figur 6-4 (neste side) ser vi også at det er stor grad av overlapping mellom boksene som representerer de to gruppene, men -GB-gruppen ligger litt høyere opp, både når det

gjelder gjennomsnitt og hvor de sentrale 50 % av tekstene plasserer seg. +GB-boksen er imidlertid litt smalere, og strekene strekker seg kortere, noe som tyder på at denne gruppen har noe mindre variasjon enn -GB-gruppen, som da har et noe større spenn i variasjonen. Forskjellene virker likevel ikke store nok til at de har noen konsekvenser for sammenlignbarheten.

Figur 6-4 Den individuelle spredningen av antall obligatoriske kontekster i SP-korpuset



6.3.2.2 SP-korpuset: forskjeller i markeringsprosent

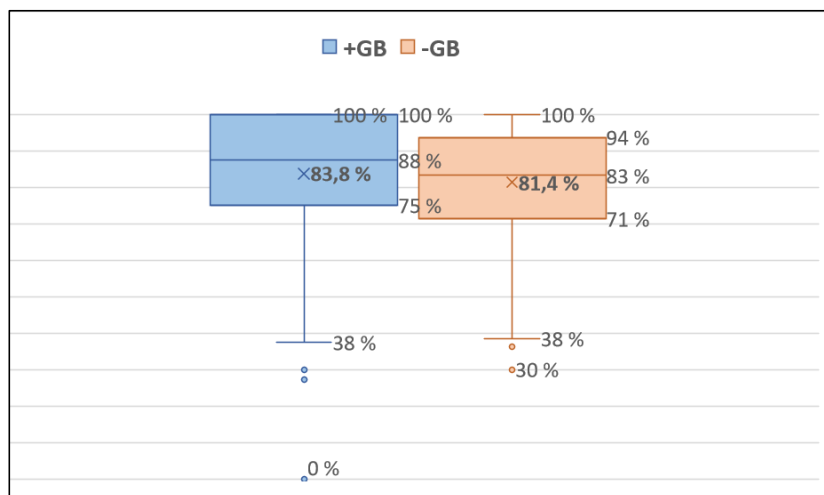
Tabell 6-6 (under) viser statistikk og testinformasjon for markeringsprosent i SP-korpuset, og vi ser at det er en liten forskjell i den gjennomsnittlige markeringsprosenten til de to gruppene: +GB-gruppen markerer ca. 84 % av kontekstene (st.avvik: 17,1), mot ca. 81,5 % i -GB-gruppen (st.avvik: 15,4). Til tross for at forskjellen kan virke ganske liten, er den statistisk signifikant ($Z=-2.601$, $p=<0.01$).

Tabell 6-6 Markeringsprosent (og rang) i de to innlærergruppene i SP-korpuset

SP-korpus	+GB-innlærere	-GB-innlærere	Mann-Whitney U
Tekster	300	300	Z = - 2.601 p = 0.0045
OBL	4265	4818	
Markerte	3622	3979	
Gjennomsnitt (markeringsprosent)	83,8 %	81,4 %	
	St.avvik: 17,1	St.avvik:15,4	
Gjennomsnitt (rang)	318,80	282,20	

Figur 6-5 (under) illustrerer denne forskjellen mellom gruppene, som gir seg utslag i at den blå boksen som representerer +GB-gruppen, ligger noe høyere enn den oransje boksen som representerer -GB-gruppen, i tillegg til at den sentrale halvparten av +GB-tekstene ligger mellom 75 og 100 % markering, mens den sentrale halvparten av -GB-tekstene ligger mellom 71 og 94 %. Om utliggerne fjernes, går begge grupper så lavt som 38 % markering, men +GB-gruppen strekker seg nærmere 100 % markering i flere individuelle tekster enn -GB-gruppen gjør.

Figur 6-5 Den individuelle spredningen for markeringsprosent i SP-korpuset



6.3.2.3 SP-korpuset: antall tekster med 100 % markering

Om vi ser på antall innlærertekster med 100 % markering i de to gruppene, gjelder dette 83 av innlærerne i +GB-gruppen (28 %), men bare 52 i -GB-gruppen (17 %), en forskjell som ifølge kjikvadrat-testen er signifikant ($\chi^2=9.1852$, $p<0.01$).

Tabell 6-7 Andel tekster med 100 % markering i SP-korpuset

SP-korpus	100 % markering		<100 % markering		Kjikvadrat-test
+GB-innlærere (N=300)	83	28 %	217	72 %	
-GB-innlærere (N=300)	52	17 %	248	83 %	$\chi^2=9.1852$ $p=0,00122$

6.3.3 HN-korpuset

6.3.3.1 HN-korpuset: antall obligatoriske kontekster

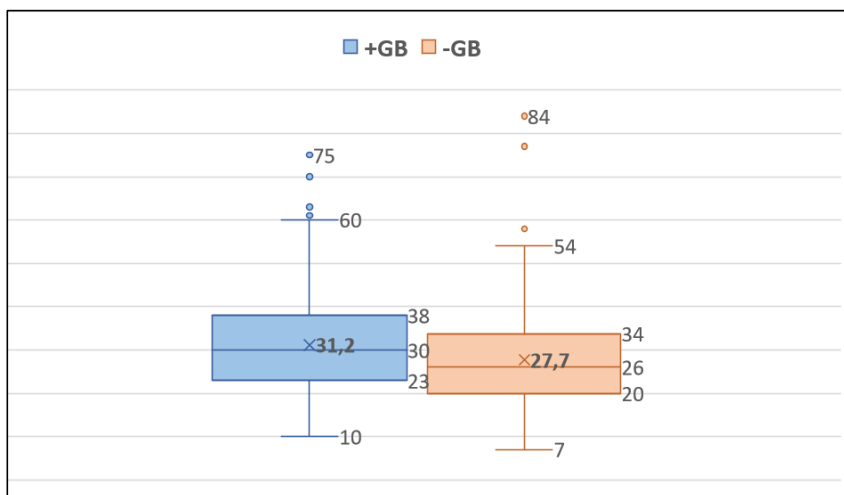
Om vi går videre og sammenligner antall obligatoriske kontekster i de to innlærergruppene i HN-korpuset, er det som tabell 6-8 (under) viser, også her et jevnt gjennomsnitt i de to gruppene (-GB 28; +GB 31), og en relativt lik spredning i antall kontekster (-GB 7-54; +GB 10-60).

Tabell 6-8: Deskriptive data for HN-korpuset (obligatoriske kontekster)

HN-korpus	+GB-gruppen			-GB-gruppen		
Tekster	300			300		
Substantiv	30741			29083		
OBL	9355	30,4 %		8324	28,6 %	
Gj.snitt	31	Min.	Maks.	28	Min.	Maks.
		10	75 (60)		7	84 (54)

Figur 6-6 (under) viser at det er stort sett overlapping mellom boksene som representerer de to gruppene, men i motsetning til i SP-korpuset er det her +GB-gruppen som ligger litt høyere enn -GB-gruppen. Som i SP-korpuset er det altså ganske jevne tall for gruppene, og den gruppeinterne spredningen er også ganske lik, faktisk noe mer lik enn for SP-gruppene, noe vi kan se på at bokser og streker har omtrent like stor utstrekning i de to gruppene her, mens +GB-boksene var noe smalere i SP-korpuset. Sammenlignbarheten mellom gruppene er dermed ikke et problem i HN-korpuset heller.

Figur 6-6 Den individuelle spredningen av antall obligatoriske kontekster i HN-korpuset



6.3.3.2 HN-korpuset: forskjeller i markeringsprosent

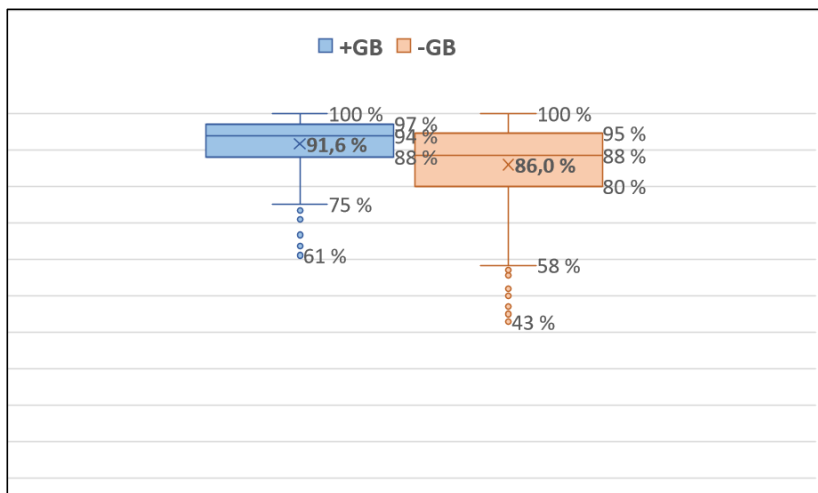
Tabell 6-9 (under) viser forskjellen i markeringsprosent mellom innlærergruppene i HN-korpuset, og også her finner vi en signifikant forskjell mellom 91,6 % markering i +GB-gruppen og ca. 85,9 % markering i -GB-gruppen ($Z = -6.686$, $p = <0.01$).

Tabell 6-9 Markeringsprosent (og rang) i de to innlærergruppene i HN-korpuset

HN-korpus	+GB-innlærere	-GB-innlærere	Mann-Whitney U
Tekster:	300	300	$Z = -6.686$ $p = 0,001$
OBL	9355	8324	
Markerte	8615	7199	
Gjennomsnitt (markeringsprosent)	91,6 %	85,9 %	
	St.avvik: 8,2	St.avvik: 11,6	
Gjennomsnitt (rang)	347,74	253,27	

I figur 6-7 (under) ser vi hvordan +GB-gruppen ligger noe høyere i markeringsprosent enn -GB-gruppen, i tillegg til å ha en smalere spredning: Den sentrale halvparten av +GB-tekstene ligger mellom 88 og 97 % markering, mens den sentrale halvparten av -GB-tekstene ligger mellom 80 og 95 %. I tillegg strekker markeringsprosenten seg generelt lavere i -GB-gruppen, da den ligger mellom 58 og 100 % om utligger fjernes, mens +GB-gruppen strekker seg fra 75 til 100 %.

Figur 6-7 Den individuelle spredningen for markeringsprosent i HN-korpuset



6.3.3.3 HN-korpuset: Antall tekster med 100 % markering

Tabell 6-10 (under) viser antall og andel av tekster med 100 % markering i de to gruppene. I -GB-gruppen er det snakk om 12 % av tekstene (35 stk.), altså lavere enn +GB-gruppen, som har 18 % (55 stk.). Denne forskjellen mellom gruppene signifikant ($\chi^2=5.2288$, $p<0.05$).

Antallet innlærertekster med 100 % markering er faktisk noe lavere i HN-korpuset enn i SP-korpuset (-GB 17 %; +GB 28 %). Dette kan selvsagt virke noe overraskende med tanke på at HN representerer et *høyere* ferdighetsnivå, men høyere ferdighetsnivå betyr jo også høyere kompleksitet og vanskegrad, så på den måten er det ikke nødvendigvis uventet.

Tabell 6-10 Andel tekster med 100 % markering i HN-korpuset

HN-korpus	100 % markering		< 100 % markering		Kjikkvadrat-test
+GB-innlærere (N=300)	55	18 %	245	82 %	$\chi^2=5.2288$ $p= 0,0111$
-GB-innlærere (N=300)	35	12 %	265	88 %	

6.3.4 Oppsummering av funn i SP-korpuset og HN-korpuset

Vi har over sett at forskjellene mellom gruppene også er til stede når materialet brytes ned på testnivå, altså at det er en generell tendens innen begge språktestene til at -GB-gruppen sjeldnere enn +GB-gruppen bruker bestemte former i obligatoriske kontekster, selv om markeringsprosenten er ganske høy i begge gruppene.

Undersøkelsen har først og fremst fokusert på forskjellene i markeringsgrad mellom -GB-gruppen og +GB-gruppen *innen* hver av språktesttypene, og ikke så mye på forskjellene *mellom* de to testtypene. De to delkorpusene fra HN og SP representerer som tidligere nevnt, til en viss grad ulike ferdighetsnivå, og de kunne derfor vært brukt til å gjøre en sammenligning av ulike ferdighetsnivå, men dette er en tråd jeg ikke har valgt å følge opp her, både fordi det tross alt er en viss spredning i ferdighetsnivå innen tekstene fra hver testtype, men også fordi det er flere andre forskjeller mellom dem, for eksempel lengde, teksttype og kompleksitetsnivå, som jeg allerede har vært inne på. Forskjeller mellom +GB- og -GB-innlærere innen ferdighetsnivå så jeg imidlertid på i Johansen (2010), og denne undersøkelsen vil jeg kort gjengi under.

6.3.5 Tidligere undersøkelse av B1- og B2-nivå

Studien i Johansen (2010) baserte seg på samme GB-korpus som her, og det ble der gjort en videre undersøkelse av et mindre delutvalg av tekster som var blitt plassert på henholdsvis B1- og B2-nivå, og hvor også noen andre faktorer var holdt konstant for å kunne sammenligne mer homogene innlærergrupper. Undersøkelsen besto av to delkorpus (B1-korpuset og B2-korpuset), og innen hver av disse to delkorpusene ble to faktorer holdt konstant: *oppholdstid i Norge på 1-2 år* og *daglig bruk av målspråket*, mens to faktorer var ulike: *ferdighetsnivå* og *testtype*. B1-korpuset besto altså av tekster fra B1/Språkprøven, og B2-korpuset av B2/Høyere nivå. I praksis betyr dette at B1- og B2-korpuset ikke bare skilte seg fra hverandre med tanke på ferdighetsnivå, men også når det gjaldt testforholdene, teksttype og tekstlengde. Siden flere faktorer ble holdt konstant var disse delkorpusene ganske små: B1-korpuset hadde til sammen 59 tekster (-GB 32; +GB 27), mens B2-korpuset hadde til sammen 49 tekster (-GB 24; +GB 25).²⁷

Undersøkelsen i Johansen (2010) viste at det i begge delkorpusene var en forskjell i markeringsprosent mellom innlærere med og uten GB i morsmålet, en forskjell på ca. *seks* prosentpoeng i begge delkorpusene (B1-korpus 5,7; B2-korpus 6,6). Forskjellene ble testet på tilsvarende måte som jeg har gjort her i avhandlingen, altså med en ikke-parametrisk test, og de var statistisk signifikante i begge delkorpusene (Johansen 2010: 169). Undersøkelsen viste altså at -GB-innlærerne markerte obligatoriske kontekster i mindre grad enn +GB-innlærerne på begge ferdighetsnivåene, og funnene i denne undersøkelsen var styrket av at gruppene var homogene med hensyn til både oppholdstid i Norge og at norsk ble brukt daglig.

Antall tekster med 100 % markering ble også undersøkt, men forskjellene ble ikke testet statistisk i Johansen (2010). En test gjort i forbindelse med denne avhandlingen viser imidlertid at forskjellene mellom innlærergruppene *ikke* er signifikante (B1: $\chi^2=3.5411$, $p> 0.05$; B2: $\chi^2=1.8111$, $p> 0.05$). Tall og statistisk informasjon er gjengitt i tabell 6-11 og tabell 6-12 på neste side.

²⁷ Disse delkorpusene inneholdt bare tekster skrevet av engelske, tyske, polske og russiske innlærere, ettersom tekstene fra innlærere med bosnisk-kroatisk-serbisk og nederlandsk ikke har blitt vurdert etter rammeverksnivåene.

Tabell 6-11 Andel tekster med 100 % markering i B1-korpuset

B1-korpus	100 % markering		<100 % markering		Kjikkvadrat-test
+GB-innlærere (N= 27)	10	37 %	17	63 %	$\chi^2=3.5411$ p= 0.0598
-GB-innlærere (N=32)	5	16 %	27	84 %	

Tabell 6-12 Andel tekster med 100 % markering i B2-korpuset

B2-korpus	100 % markering		<100 % markering		Kjikkvadrat-test
+GB-innlærere (N=25)	7	28 %	18	72 %	$\chi^2=1.8111$ p= 0.1783
-GB-innlærere (N= 24)	3	13 %	21	87 %	

Undersøkelsen av B1- og B2-korpuset viste også at markeringsprosenten gikk *opp* i begge innlærergruppene, slik en kan forvente med økende ferdigheter, men disse forskjellene var ikke signifikante (Johansen 2010: 169). Når det gjelder antall tekster som hadde 100 % markering, var det derimot en *nedgang* i begge gruppene, slik vi også så at det var mellom SP- og HN-korpuset. Nedgangen var størst i +GB-gruppen (9 prosentpoeng), og ganske liten i -GB-gruppen (3 prosentpoeng), men nedgangen er ikke signifikant i noen av gruppene ifølge en kjikkvadrat-test tatt i ettertid (+GB: $\chi^2=0.4818$, p=0.49; -GB= $\chi^2=0.746857$, p= 0.1094).

I likhet med HN-korpuset og SP-korpuset som er undersøkt i denne avhandlingen, viste altså B1- og B2-korpuset i Johansen (2010) også signifikante forskjeller mellom innlærergruppene, selv om det ikke var signifikante forskjeller i antall innlærertekster med 100 % markering. Dette kan forsterke antakelsen om at det også er forskjeller mellom gruppene når ferdighetsnivå holdes konstant.

6.4 Polsk og engelsk delkorpus

Siste del av undersøkelse 1 bryter ned materialet fra GB-korpuset etter morsmålsbakgrunn. GB-korpuset inneholder seks morsmålsgrupper som alle kunne ha vært undersøkt nærmere for å se om de spesifikke morsmålsbakgrunnene gir ulike utslag på markeringsprosent, men jeg vil her nøye meg med å fokusere på de to gruppene som er valgt ut til undersøkelse 2, altså polske og engelske innlærere. Framgangsmåten er den samme som over, og jeg ser på SP-tekstene og HN-tekstene hver for seg.

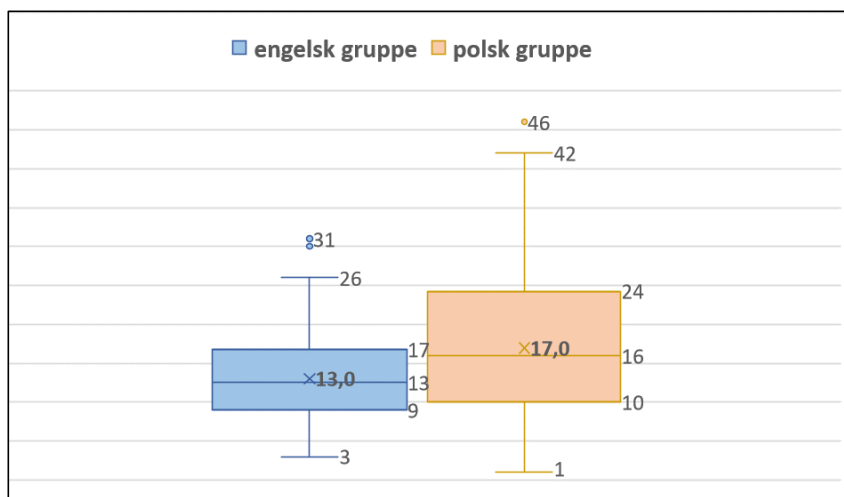
6.4.1 Obligatoriske kontekster i polsk og engelsk delkorpus (Språkprøven)

Andelen av obligatoriske kontekster er mer forskjellig mellom den polske og den engelske gruppen fra Språkprøven enn det vi har sett mellom +GB og -GB-gruppene i de andre delkorpusene, da det polske delkorpuset har en andel på 30 % obligatoriske kontekster av alle substantiv, mens det engelske bare har 25 % (tabell 6-13, under). Forskjellen viser seg også gjennom at det polske delkorpuset har gjennomsnittlig 17 obligatoriske kontekster per tekst, mot 13 i det engelske, og dessuten gjennom at de polske innlærertekstene strekker seg fra 1 til 42 kontekster, mens de engelske strekker seg fra 3 til 30, når utliggerer er fjernet.

Tabell 6-13 Deskriptive data for polsk og engelsk delkorpus, SP (obligatoriske kontekster)

Språkprøven	Engelsk gruppe			Polsk gruppe		
Substantiv	5215			5642		
OBL	1304	25 %		1696	30 %	
Gj.snitt	13	Min. 3	Maks. 30 (26)	17	Min. 1	Maks. 46 (42)

Figur 6-8 Individuell spredning av obligatoriske kontekster i polsk og engelsk gruppe, SP



Figur 6-8 (over) viser at boksen som representerer den polske gruppen er videre, og at værhårene strekker seg lenger vekk fra boksen, mens boksen for de engelske innlærerne er smalere og dermed viser til en variasjon som samler seg mer rundt gjennomsnittet enn det vi finner i den polske gruppen. Det er uklart hva denne forskjellen har opphav i,

men siden markeringsprosenten er et relativt tall basert på antall obligatoriske kontekster, vil ikke denne forskjellen slå ut på markeringsprosentene på noe vis.

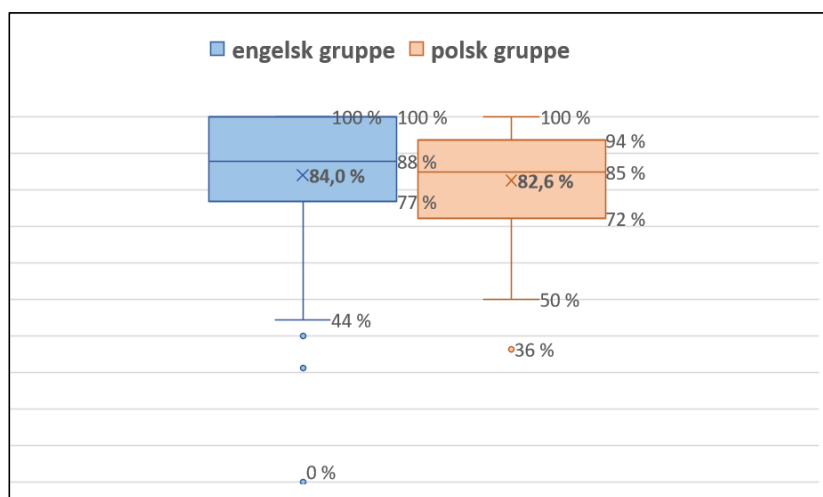
6.4.2 Markeringsprosent i polsk og engelsk gruppe (Språkprøven)

Når en sammenligner markeringsprosenten i den polske og den engelske gruppen, viser det seg at det kun er en liten forskjell, der engelske innlærere i gjennomsnitt markerer 84 % av kontekstene (st.avvik 17,8), og de polske markerer 82,6 % (st.avvik 14,2). Denne forskjellen er *ikke* signifikant ($Z = -1.410$, $p > 0.05$).

Tabell 6-14 Markeringsprosent (og rang) i polsk og engelsk gruppe, Språkprøven

Språkprøven	Engelsk gruppe	Polsk gruppe	Mann-Whitney U
Tekster	100	100	Z = - 1.410 p= 0,079
OBL	1304	1696	
Markerte	1096	1413	
Gjennomsnitt (markeringsprosent)	84,0 % St.avvik 17,8	82,6 % St.avvik 14,2	
Gjennomsnitt (rang)	106,23	94,77	

Figur 6-9 Individuell spredning for markeringsprosent i polsk og engelsk gruppe, SP



Figur 6-9 (over) visualiserer forskjellen mellom gruppene, og vi ser at selv om boksen som representerer den engelske gruppen, ligger noe høyere enn den polske, er både gjennomsnitt og median ganske jevne i de to gruppene. Boksene overlapper bare delvis: De sentrale 50 % av tekstene til den engelske gruppen ligger mellom 77 og 100 %, mens de ligger mellom 72 og 94 % i den polske gruppen. Inspeksjonen av boksdiagrammet

viser altså noen forskjeller mellom gruppene som kanskje ikke er tilfeldige, men det er mulig den statistiske testen ikke fanget forskjellen godt nok opp, og derfor ikke slo ut som signifikant.

Som boksdiagrammet over også viser, er det en overvekt av engelske tekster som strekker seg helt opp til 100 % markering, og tabell 6-15 (under) viser hva det nøyaktige antallet er i hver av gruppene: Den engelske gruppen har hele 31 tekster med 100 % markering, mens den polske gruppen bare har 17. Denne forskjellen *er* signifikant ($\chi^2=5.3728$, $p<0.05$).

Tabell 6-15 Andel tekster med 100 % markering i polsk og engelsk gruppe, Språkprøven

Språkprøven	100 % markering		<100 % markering		Kjikkvadrat-test
Engelske innlærere (N=100)	31	31 %	69	69 %	
Polske innlærere (N=100)	17	17 %	83	83 %	

I tekstene skrevet av engelske og polske innlærere til Språkprøven, er det altså ikke noen signifikant forskjell i markeringsgrad, men det er en signifikant forskjell i antall tekster med 100 % markering.

6.4.3 Obligatoriske kontekster i polsk og engelsk delkorpus (Høyere nivå)

Tabell 6-16 (under) viser deskriptive tall for obligatoriske kontekster i den polske og engelske gruppen i tekster fra HN-testen. Antall obligatoriske kontekster er relativt likt de to gruppene (pol. 27,6 %; eng. 28 %), og det samme gjelder gjennomsnittlig antall per tekst (pol. 27 stk.; eng. 30 stk.). Distansen fra minimum til maksimumsverdier er også veldig lik når utliggerer er utelatt (pol. 7-53; eng. 11-53). Det er altså to veldig like grupper hva angår andel obligatoriske kontekster både i gjennomsnitt og i spredning.

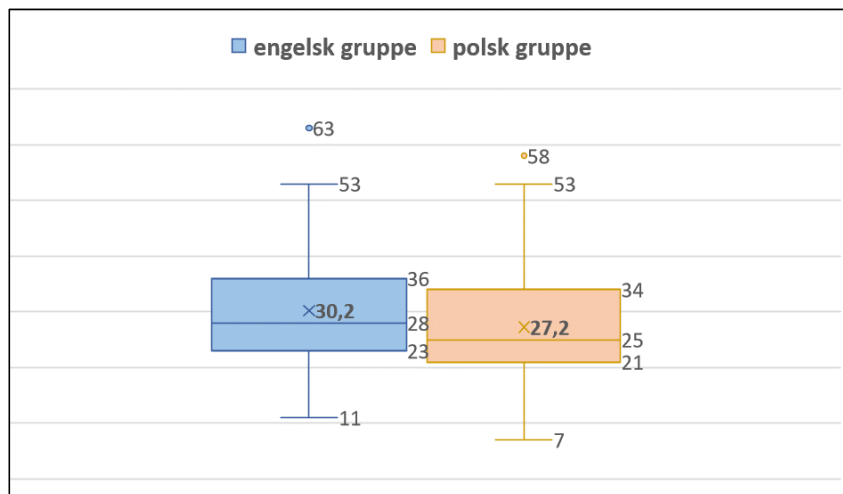
Tabell 6-16 Deskriptive data for polsk og engelsk delkorpus, HN (obl. kontekster)

HN-korpus	Engelsk gruppe			Polsk gruppe		
Substantiv	10682			9863		
OBL	3020	28 %		2724	27,6 %	
Gj.snitt	30	Min. 11	Maks. 63 (53)	27,2	Min. 7	Maks. 58 (53)

Figur 6-10 (neste side) illustrerer denne likheten i form av et jevnt antall og relativt lik spredning i de to gruppene. Det er jevn størrelse på boksene, cirka like lange værhaar, og

boksen for de engelske innlærertekstene ligger bare litt høyere enn for de polske. Sammenlignbarheten for disse to gruppene på HN-testen bør dermed være mindre problematisk enn det var på SP-testen.

Figur 6-10 Individuell spredning av obl. kontekster i polsk og engelsk gruppe, HN



6.4.4 Markeringsprosent i polsk og engelsk gruppe (Høyere nivå)

I HN-korpuset er forskjellen mellom engelske og polske innlærere større enn i SP-korpuset når det gjelder markeringsprosent (tabell 6-17). Den engelske gruppen har en gjennomsnittlig markeringsprosent på 92,3 %, mot 85,5 % i den polske gruppen, noe som er signifikant ($Z = -4.438$, $p < 0.01$).

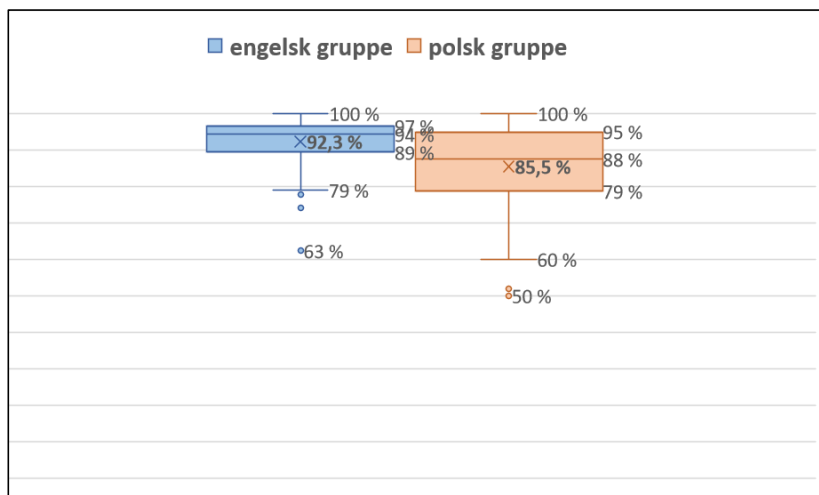
Tabell 6-17 Markeringsprosent (og rang) i polsk og engelsk gruppe, HN

HN-korpus	Engelsk gruppe	Polsk gruppe	Mann-Whitney U
Tekster	100	100	$Z = -4.438$ $p = 0,001$
OBL	3020	2724	
Markerte	2778	2328	
Gjennomsnitt (markeringsprosent)	92,3 % St.avvik: 6,6	85,5 % St.avvik: 11,6	
Gjennomsnitt (rang)	118,65	82,736	

Standardavvikene for de to gruppene er ganske ulike, med 6,6 for den engelske gruppen og 11,6 for den polske. Dette, og størrelsene på boksene i figur 6-11 (neste side), viser at spredningen er langt større i den polske gruppen enn i den engelske. 50 % av de engelske innlærertekstene ligger mellom 89 og 97 % markering, mens 50 % av de polske

ligger mellom 79 og 95 %, som er et større spenn. Også distansen mellom minimum og maksimum er betraktelig høyere i den polske gruppen: fra 60 til 100 %, mot 79 til 100 % i den engelske gruppen. Det er altså i tillegg til å være en signifikant forskjell mellom gruppene, også en merkbar forskjell i spredning, der den polske gruppen har langt mer gruppeintern variasjon.

Figur 6-11 Individuell spredning for markeringsprosent i polsk og engelsk gruppe, HN



I tabell 6-18 (under) ser vi at det derimot ikke er noen utslagsgivende forskjeller mellom gruppene i antall tekster med 100 % markering: Den engelske gruppen har 13, mens den polske har 12, en forskjell som er så liten at det ikke er nødvendig med noen statistisk test for å konstatere at det ikke er snakk om en signifikant forskjell.

Tabell 6-18 Andel tekster med 100 % markering i polsk og engelsk gruppe, HN

Høyere nivå	100 % markering		<100 % markering		Kjikkvadrat-test
Engelske innlærere (N=100)	13	13 %	87	87 %	Ikke utført
Polske innlærere (N=100)	12	12 %	88	88 %	

Boksdigrammet over (figur 6-11) avslører imidlertid at det er spesielt når markeringsprosenten kommer under 90 % at forskjellene mellom gruppene er størst: Den polske gruppen har langt flere innlærere som ligger mellom 79 og 90 % enn den engelske, som i all hovedsak ligger over 90 %. En statistisk testing av en inndeling som

ikke bare skiller mellom «100 %» og «under 100 %», ville kanskje i større grad fanget opp disse forskjellene.

6.5 Oppsummering og diskusjon av undersøkelse 1

6.5.1 Hovedfunn i undersøkelse 1

Målet med undersøkelse 1 har vært å finne ut om det er en sammenheng mellom det å *ikke* ha grammatisk bestemthet i morsmålet, og det å ikke bruke substantivets bestemte form, der det er obligatorisk ifølge målspråket. Undersøkelsen har tatt utgangspunkt i to prediksjoner: At -GB-gruppen vil ha lavere markeringsprosent enn +GB-gruppen, og at det i tillegg vil være færre tekster med 100 % markering i -GB-gruppen. GB-korpuset ble først undersøkt samlet i alle 1200 tekstene, men etterpå delt opp etter de to testnivåene HN og SP, og til slutt så jeg bare på tekster skrevet av polske og engelske innlærere.

Tabell 6-19 (under) gir en oversikt over de ulike delkorpusene som ble undersøkt, og om forskjellene mellom de to gruppene ble funnet signifikant eller ikke. Ikke-signifikante resultater er markert med grått. Resultatene fra Johansen (2010) er også tatt med.

Tabell 6-19 Oversikt for funn i de ulike delkorpusene i undersøkelse 1

Prediksjon: Delkorpus:	Markeringsprosent Signifikant forskjell?	100 %-tekster Signifikant forskjell?
GB-korpus	Ja	Ja
HN-korpus	Ja	Ja
SP-korpus	Ja	Ja
Polsk/engelsk SP	Nei	ja
Polsk/engelsk HN	Ja	Nei <i>(ikke nødvendig å teste)</i>
Johansen 2010		
B1-korpus	Ja	Nei
B2-korpus	Ja	Nei

Undersøkelse 1 har i stor grad vist at innlærergruppene *er* forskjellige når det gjelder begge de to forholdene som ble undersøkt, altså både markeringsprosent og antall tekster med 100 % markering. Spesielt tydelig er resultatene fra sammenligningen mellom de større gruppene av +GB/-GB-innlærere, hvor forskjellen var signifikant både i GB-korpuset, SP-korpuset og HN-korpuset. Resultatene for korpuset av bare engelske og polske innlærertekster var derimot mer blandede: I tekster fra Språkprøven var selve markeringsprosenten i de to gruppene ganske jevn, men den engelske gruppen hadde signifikant flere tekster med 100 % markering enn den polske. Det motsatte fant vi for HN-testen, der markeringsprosenten var signifikant lavere for den polske gruppen, men antall tekster med 100 % markering var så og si likt i de to gruppene.

De ulike delkorpusene som er undersøkt her, overlapper i stor grad, siden alle de mindre delkorpusene var en del av det større GB-korpuset, så det er kanskje ikke overraskende at funnene peker i samme retning. Samtidig gjorde oppdelingen i HN-korpuset og SP-korpuset at spennet i antall obligatoriske kontekster per tekst gikk ned i forhold til i GB-korpuset, noe som til en viss grad har økt sammenlignbarheten mellom markeringsprosenten på tvers av innlærertekster. Det viste oss også at det er forskjeller mellom gruppene på begge testnivå, noe som er interessant med tanke på at det kan indikere at det også er forskjell mellom gruppene på ulike ferdighetsnivå, i tillegg til å være til stede i forskjellige *testtyper* og *teksttyper*. Den tidligere undersøkelsen av B1- og B2-tekster som ble referert til (Johansen 2010), som også var en delmengde tekster fra det samme GB-korpuset, støttet også antakelsen om at forskjellene i markeringsprosent er til stede på ulike ferdighetsnivå, men det ble der ikke funnet noen forskjell i antall tekster med 100 % markering. I Johansen (2010) var gruppene delt inn etter faktiske ferdighetsnivå, ikke bare antatte ferdighetsnivå basert på hvilken test teksten var skrevet til.

Delkorpuset med SP-tekster skrevet av engelske og polske innlærere er interessant fordi det er disse to gruppene jeg går videre med i undersøkelse 2, og selv om resultatene var noe blandede og forskjellene mindre tydelige, så har vi sett at det er interessante forskjeller her også. Om vi følger en anbefaling fra Larsen-Hall (2010: 102) om å øke grenseverdien for hva som regnes som signifikant (alfa) innen andrespråksstudier til 0.10, kan også den polske og engelske gruppen regnes som

signifikant forskjellige for Språkprøve-tekstene, ettersom p-verdien for forskjellen mellom markeringsprosent var 0.079 ($p < 0.10$).²⁸

Nordanger (2017) så også på korrekthet i bruk av bestemthetsmarkører, men studiene er ikke direkte sammenlignbare ettersom Nordanger hadde delt inn i ulike referensielle kontekster og i tillegg så på ubestemte former, mens mitt materiale bare har obligatoriske kontekster for bestemt form og dessuten slår sammen de to referensielle kontekstene som kan kreve bestemte former, nemlig definitte og generisk. Nordangers funn viste spesielt at russiske innlærere (-GB) hadde utfordringer med ubestemt form i indefinitte kontekster, og at engelske innlærere (+GB) hadde problemer med bestemte former i definitte kontekster (Nordanger 2017: 259).

6.5.2 Transfereffekter: morsmålssamsvar og intralingval kontrast

At gruppen som ikke kjenner bestemthet fra morsmålet, markerer signifikant færre kontekster og har færre tekster med 100 % markering enn gruppen av innlærere som kjenner bestemthet fra morsmålet, tyder på at det er en viss *heterogenitet* mellom de to gruppene, altså at de skiller seg fra hverandre som grupper tatt i betraktning. De statistiske testene som ble brukt, tar som sagt hensyn til både gjennomsnittet i gruppene og de individuelle avvikene fra dette gjennomsnittet, og kan dermed også vise at det er *homogenitet innad* i gruppene. Det var en god del variasjon innad i gruppene hva markeringsprosent angår, men denne variasjonen gjaldt begge gruppene og graviterte rundt gruppenes respektive sentralverdier.

På denne måten kan vi altså si at jeg har undersøkt og funnet mønster som er i tråd med de to gruppebaserte transfereffektene *intergruppe-heterogenitet* og *intragruppe-homogenitet*. For det første har jeg funnet at de språklige mønstrene er spesifikke, og for det andre at de er representative for hver av gruppene. Det er imidlertid to effekter til som kan (og bør) støtte oppunder et argument for transfer, nemlig *morsmålssamsvar* og *intralingval kontrast*. Disse vil jeg ta for meg under.

²⁸ «I don't know if this type of change would have to be instituted by journal editors in order to be adopted, or if individual authors could just quote people who advocate such an increase in the alpha level, but Why don't you try it? Quote me, and quote also Kline (2004) and Murphy and Myors (2004), who have argued that the alpha level should be set to $\alpha = .10$ in the social sciences.» (Larsen-Hall 2010: 102).

6.5.2.1 *Morsmålssamsvar*

Selv om jeg har vist at de to gruppene skiller seg fra hverandre, så gjenstår det fremdeles å vise at det er en sammenheng mellom morsmålet og innlærerspråket. Siden det jeg har sett på, er *markeringsprosent*, altså et kvantitativt uttrykk for hvor ofte man bøyer substantiv i obligatoriske kontekster, har jeg ikke sett på noe spesifikt språklig samsvar mellom morsmålet og innlærerspråket. I kapittel 7.1.4 nevnte jeg fravær og tilstedeværelse av grammatisk bestemthet i morsmålet som foreløpige kandidater til morsmålssamsvar for de to innlærergruppene: Mangel på kategorien i morsmålet, fører til mindre markering i obligatoriske kontekster i innlærerspråket, mens tilstedeværelse av kategorien i morsmålet fører til mer markering. Dette er imidlertid en noe upresis form for morsmålssamsvar, og om transfer skal være en form for kunnskap fra morsmålet som overføres og tas i bruk på andrespråket (jf. definisjonen i 2.1), kan en definisjon av morsmålssamsvar som er mer knyttet til *hva* som blir overført fra morsmålet, være relevant.

For +GB-innlærere er dette kanskje ukomplisert. Det er kunnskapen om grammatisk bestemthet som overføres til innlærerspråket, og siden morsmål og målsspråk i stor grad er sammenfallende, fører det til at det som regel brukes rett. Vi kan se for oss at det er snakk om en ren overføring av den distribusjonen av former som fins i morsmålet, og at dette bidrar til at markeringen stort sett blir korrekt i norsk hos +GB-innlærere. Nordanger (2017: 375) viser også hvordan det å ha engelsk som morsmål virker fasiliterende på bruken av bestemte former, selv om hun mener at det kan være en utfordring tidlig i innlæringen på grunn av de strukturelle forskjellene mellom engelsk og norsk, altså at engelsk har bestemte artikler, mens norsk har bøyning.

For -GB-innlærere er det derimot mer komplisert å skulle knytte mønstrene i innlærerspråket til overføring av spesifikke språklige forhold i morsmålet. Er den lavere markeringsgraden kun et generelt kvantitativt utslag av *generelle vansker* med bruk av bestemte former ettersom morsmålet ikke har kategorien, eller kan det tenkes at det avspeiler en overføring av spesifikke språklige forhold fra morsmålet som påvirker *distribusjonen av former* i innlærerspråket. I så fall er de kvantitative utslagene vi har sett, bare en sekundæreffekt av den faktiske morsmålspåvirkningen, altså det jeg i 2.2.2 omtalte som *kumulative effekter* (Odlin 1989: 38).

Dette spørsmålet kan ikke besvares gjennom undersøkelse 1, men i undersøkelse 2 er målet å se på hvordan mer spesifikke forhold fra morsmålet til -GB-innlærere kan påvirke formvalgene i ulike typer språklige kontekster i innlærerspråket. Om det viser seg at disse hypotesene bidrar til å avdekke et mer spesifikt morsmålssamsvar hos de polske innlærerne, kan det være aktuelt å diskutere om det er denne morsmålpåvirkningen som på et eller annet vis bidrar til forekomsten av lavere markeringsgrad i -GB-gruppen i undersøkelse 1. Dette komme jeg tilbake til i avslutningskapitlet (11.3.8.1).

6.5.2.2 *Intralingval kontrast*

Den fjerde transfereffekten, *intralingval kontrast*, viser til at det vil være ulike utslag av transfer etter om morsmål og målspråk er like eller ulike på spesifikke områder (se 2.4.3.2). Denne effekten kan undersøkelse gjennom en sammenligning av ulike kontrastive relasjoner mellom morsmål og målspråk, men på grunn av måten undersøkelse 1 er lagt opp på, har ikke denne transfereffekten blitt undersøkt. Det er for det første av praktiske årsaker: På det tidspunktet undersøkelse 1 ble planlagt og gjennomført, var ikke denne fjerde effekten lansert av Jarvis enda, den kom i en artikkel fra 2010. Et annet problem med intralingval kontrast for undersøkelse 1, er at denne effekten henger sammen med hva som defineres som morsmålssamsvar, og uten at dette er klart definert, er det også vanskelig å undersøke denne fjerde effekten. En løsning som kunne passet til prediksjonene for undersøkelse 1, er å sammenligne markeringsprosenten for bestemt form med markeringsprosenten til en annen kategori, som i motsetning til grammatisk bestemthet er felles for norsk og -GB-språk, for eksempel kategoriene numerus og genus. For +GB-gruppen kunne jeg sammenligne markeringsgraden i kontekster der morsmålet og norsk er like og ikke like innad i bestemthetskategorien, men dette kan være omfattende når det er snakk om tre ulike +GB-språk i materialet. I undersøkelse 2 vil jeg derimot ha en klarere ide om hva som er det forventede morsmålssamsvaret, og dermed vil jeg også lettere kunne se på intralingval kontrast, noe jeg vil komme tilbake til i kapittel 8.

6.5.3 Forbehold ved utvalget av substantivfraser

Jeg vil nå se på noen forbehold ved undersøkelse 1 hva angår utvalget av substantivfraser. Siden jeg har søkt på *alle substantiv* i ASK uavhengig av hva slags type frase de er kjerne i, og dessuten bare substantiv som har bøyning i bestemt form eller skulle hatt det, vil en stor del av substantivene som i undersøkelsen er regnet som *umarkerte*, likevel kunne ha markering andre steder i frasen, for eksempel i form av et foranstilt determinativ, som i ‘den røde bil’ og ‘den bil’. Selv om feilkodingen og korreksjonen som er lagt inn i ASK-korpuset, tilsier at frasen ‘den røde bil’ skulle ha hatt formen ‘den røde bilen’ og at substantivet derfor har gal form, er det problematisk å si at frasen ikke har markert bestemthet. Bestemtheten er nemlig markert, men ikke helt i tråd med målspråkets måte å markere den på, siden selve substantivet har ubestemt form.²⁹ På grunn av dette er det viktig å understreke at undersøkelse 1 ser på bruk av bestemt form av substantivet, og ikke på markering av bestemthet i substantivfrasen som helhet.

Det er et omfattende arbeid å kartlegge omfanget av fraser med andre bestemthetsmarkører i et korpus på 1200 tekster, i hvert fall om en skal gå ned på individnivå. I Johansen (2010: 170) gjorde jeg likevel noen ekstra søk i ASK på *gruppenivå* for å forsøke å anslå hvor stor andel av de umarkerte substantivene som hadde slike foranstilte determinativer, enten direkte foran seg eller adskilt av opptil to ord, som regel adjektiv. Dette er fraser som ‘de foreldre’, ‘den gode bok’, eller ‘den lille, røde hatt’. Det viste seg at i +GB-gruppen hadde 32 % av de substantivene som ble regnet som umarkerte, foranstilt determinativ, mens dette gjaldt bare 16 % i -GB-gruppen. Bruk av foranstilte determinativer uten å samtidig bøye substantivet, er altså mer utbredt i +GB-gruppen enn i -GB-gruppen. Siden dette øker antall markerte *substantivfraser* i ulik grad i de to gruppene, ville altså markeringsprosenten ha gått opp i +GB-gruppen, men ikke fullt så mye opp i -GB-gruppen om jeg hadde justert for markering andre steder i frasen. Forskjellen mellom de to gruppene ville dermed ha vært

²⁹ I norsk kan man skille mellom fraser med *enkel bestemthet* (determinativ + ubestemt form) og *dobbel bestemthet* (determinativ + bestemt form). I ASK vil noen fraser med enkel bestemthet ha blitt korrigeret til bestemt form, mens andre nok ikke har blitt det, siden enkel bestemthet i mange tilfeller oppleves som korrekt. Slike fraser vil ikke være inkludert i materialet mitt siden ubestemt form av substantivet bare er inkludert om det har feilkoden F.

større og potensielt enda mer i tråd med prediksjonene enn det vi har sett i min undersøkelse.

Om en tar i betraktning at de innlærerne som utgjør +GB-gruppen, alle har morsmål som markerer bestemthet med foranstilt artikkel (engelsk, nederlandsk og tysk), så kan det også være en potensielt morsmålsrelatert effekt at disse innlærerne oftere bruker foranstilt determinativ enn substantivbøyning. Dette er imidlertid kun en spekulasjon basert på søk i materialet uten å ta hensyn til individuelle forskjeller, og denne ekstra undersøkelsen på gruppenivå fanget nok heller ikke opp alle tilfeller av slike andre markeringer, bare fraser som har determinativ opptil tre plasser fra substantivet. Det kan likevel være en indikasjon på en generell og interessant tendens.

Måten undersøkelse 1 er gjennomført på, er altså først og fremst rettet mot å undersøke om bestemt form av substantivet brukes i obligatoriske kontekster, ikke om substantivfrasen har bestemthetsmarkør, og undersøkelsen skiller dermed ikke tydelig mellom tilfeller der substantivet krever bestemt form av strukturelle grunner og tilfeller der substantivet krever bestemt form av referensielle grunner. I undersøkelse 2 vil jeg derimot kun se på enkle, umodifiserte substantivfraser og dermed vil strukturelle forhold i større grad være utelukket.

6.5.4 Manglende differensiering av referensielle typer

Et annet forbehold med de søkene jeg har gjort i ASK, handler ikke direkte om utvalget av fraser de gir som treff, men om hvilken lingvistisk informasjon som er tilgjengelig i ASK. ASK inneholder ikke informasjon om hvilken type referanse de ulike substantivfrasene har, og dermed skiller ikke undersøkelse 1 mellom de to referensielle typene som i hovedsak utgjør obligatoriske kontekster for bestemt form, nemlig *definit* og *generisk* referanse.

Definit og generisk referanse styrer valg av bestemt form på ulike måter, for eksempel med at bestemt form (eller annen bestemthetsmarkering) som regel er obligatorisk ved definit referanse, mens man ved generisk referanse bare bruker bestemte former ved enkelte generiske undertyper, og veldig ofte kan bruke ubestemt og bestemt form om hverandre uten at den generiske fortolkningen forsvinner (se 7.2.4.2).

Denne forskjellen mellom definitte og generisk referanse kan ha hatt konsekvenser for hvilke typer substantivfraser som er inkludert i materialet, på den måten at de generiske forekomstene stort sett bidrar til å øke antall obligatoriske kontekster, men sjeldnere bidrar til andelen av umarkerte substantiv. Dette kan bety at mens materialet som helhet består av begge referansetyper, så er det kanskje først og fremst de definitive forekomstene som har feil og bidrar til å redusere markeringsprosenten. Om det i tillegg fins morsmålsbaserte forskjeller knyttet til bruk av ulike referensielle typer eller generiske undertyper, kan dette altså ha påvirket markeringsprosentene til de to gruppene på ulike måter.

Dette viser at å kunne skille mellom ulike referensielle typer er viktig for å kunne gi et mer presist mål på markeringsgraden i gruppene, men dette krever en omfattende manuell annotering av materialet, ettersom referansetype ikke allerede ligger inne i ASK. I undersøkelse 2 har jeg derimot et mer begrenset utvalg av tekster, så der har jeg hatt mulighet til å klassifisere substantivfrasene etter referensiell type, noe jeg kommer tilbake til i kapittel 8.

6.5.5 Konklusjon for undersøkelse 1

Den overordnede konklusjonen av undersøkelse 1 er at det *er* forskjeller mellom innlærere med og uten grammatisk bestemthet i morsmålet når det gjelder bruk av bestemt form i obligatoriske kontekster. Statistisk signifikante resultater fra statistiske tester av de to transfereffektene *intragruppe-homogenitet* og *intergruppe-heterogenitet*, tyder på at det foreligger morsmålspåvirkning, men det kommer ikke fram gjennom undersøkelsen om dette gjelder både -GB-gruppen og +GB-gruppen, eller bare én av gruppene. Funnene er likevel i tråd med den store mengden tidligere studier som tyder på at -GB-innlærere har større utfordringer med bestemthetsmarkering enn +GB-innlærere. Hvorvidt disse utfordringene kun er generelle i den forstand at morsmålet bare bidrar til at bestemt form har en tendens til å utelates på en usystematisk måte, eller om det ligger en morsmålspåvirket systematikk bak når formene brukes og ikke brukes, kan ikke denne undersøkelsen svare på, og heller ikke hva som bidrar til at bestemte former faktisk brukes eller i hvilke tilfeller de *ikke* brukes. Tendensene er at det er relativt høy markeringsprosent både hos -GB- og +GB-innlærere, selv om -GB-gruppene ligger signifikant lavere enn +GB-gruppene.

Jeg har også pekt på noen forbehold ved undersøkelse 1. For det første om transfereffekten *morsmålsansvar* er tydelig nok definert til at man kan argumentere for at forskjellene mellom de to innlærergruppene faktisk skyldes transfer. Det er vanskelig å definere klart hva i morsmålet som motiverer lavere grad av markering hos -GB-innlærere, siden jeg bare ser på kvantitative forhold og ikke direkte på de språklige valgene som er gjort i ulike språklige kontekster. For det andre påpekte jeg forbehold ved utvalget av substantivfraser slik de er søkt fram i ASK, og at undersøkelsen bare kunne si noe om hvorvidt substantivet hadde bestemt form, og ikke om andre bestemthetsmarkører har blitt brukt. Dette gjør det vanskelig å si om mangel på markering skyldes manglende kunnskap om at substantivfrasen skal ha bestemthetsmarkering, eller om det skyldes manglende kunnskap om hvordan dette uttrykkes i norsk. Til sist pekte jeg på at ASK ikke legger til rette for å kunne skille mellom substantiv med definitiv og generisk referanse. Både strukturelle og referensielle forhold kan være viktig å ha kontroll på i en undersøkelse som dette. Dette er både vesentlig med tanke på at det i norsk eksisterer ulike regler og ulik grad av obligatoriskhet for bestemt form i ulike syntaktiske og referensielle kontekster, men også med tanke på at dette er områder hvor morsmålet kan slå ulikt ut både hos -GB-innlærere og +GB-innlærere. Som vi så, var for eksempel bruk av foranstilt determinativ langt mer utbredt i +GB-gruppen enn det var i -GB-gruppen, noe som kan påvirke sammenlignbarheten mellom gruppene.

6.5.6 Neste delundersøkelse

I de neste kapitlene vil jeg gå over til undersøkelse 2. Først vil jeg se på relevante forhold knyttet til substantiv og substantivfraser i norsk, polsk og engelsk (kapittel 7), deretter presenterer jeg bakgrunnen for de to hypotesene om polsk morsmålspåvirkning, utvalget av datamateriale fra ASK og metode (kapittel 8), og til slutt presenterer jeg resultatene for de to hypotesene (kapittel 9 og 10). Kapittel 11 oppsummerer hele avhandlingsstudien.

Undersøkelse 2 vil ikke se på markeringsprosenter og skal besvare et annet forskningsspørsmål enn undersøkelse 1, men jeg har valgt å tilpasse meg noen av de metodiske problemstillingene som er nevnt over. Jeg har blant annet begrenset datamaterialet til å bare inneholde enkle, umodifiserte substantivfraser, men jeg har

utvidet materialet til å gjelde alle forekomster av bestemte og ubestemte former, og jeg har dessuten klassifisert alle forekomstene etter hvilken type referanse de har. Dette vil jeg komme tilbake til i kapittel 8.

7 SUBSTANTIV OG REFERANSE I NORSK, POLSK OG ENGELSK

I dette kapitlet vil jeg presentere de tre språkene som er aktuelle for undersøkelse 2, altså norsk, polsk og engelsk, med tanke på forholdet mellom substantivets form og referanse. Formålet er å beskrive noen sentrale likheter og forskjeller mellom disse språkene, som videre danner grunnlaget for hypotesene om polsk morsmålspåvirkning og måten transfereffektene for disse hypotesene undersøkes på.

En hovedforskjell mellom språkene som jeg allerede har vært inne på flere steder, er at norsk og engelsk har grammatisk bestemthet, noe polsk ikke har, men polsk har derimot kasus som en sentral kategori ved substantiv. Denne forskjellen har jeg tidligere beskrevet som en *nullkontrast* mellom morsmål og målspråk. Selv om grammatisk bestemthet blant annet brukes til å skille mellom ulike typer referanse, er referanse også et begrep innen polsk, men uten å være avgjørende for valg av substantivform, slik det er i norsk og engelsk. De referensielle hovedtypene som ble omtalt i kapittel 3 (generisk, indefinit, definit og ikke-referensiell) er altså til stede i alle de tre språkene siden substantivfraser kan brukes for å referere til individuelle entiteter, generiske klasser eller for å beskrive referenter i alle språkene. Dette gjør at referensiell type kan brukes som et felles utgangspunkt for de kontrastive beskrivelsene i dette kapitlet. I tillegg vil jeg komme inn på bestemthetsekvivalenter og kasus i polsk.

Som nevnt i kapittel 6, er materialet i undersøkelse 2 avgrenset til umodifiserte substantivfraser, og derfor har jeg for det meste også begrenset den kontrastive beskrivelsen til denne typen enkle substantivfraser, og stort sett utelatt forhold som gjelder andre typer substantivfraser i de tre språkene.

Kapitlet har fire hoveddeler. Først ser jeg kort på begrepet enkle substantivfraser (7.1). Deretter presenterer jeg norske, polske og engelske substantivfraser, og hvordan substantivenes form knyttes til referansetype i de ulike språkene (7.2 til 7.4) Til slutt oppsummeres de viktigste likhetene og forskjellene mellom språkene i 7.5.

7.1 Enkle substantivfraser

Enkle substantivfraser vil her regnes som substantivfraser som ikke er modifisert med foranstilte og etterstilte adledd, bortsett fra kvantorene *en/ei/et*, altså ubestemt artikkel.

Dette vil si substantivfraser som består av enten bare et substantiv, eller av ubestemt artikkel + substantiv.

I slutten av kapittel 6 var jeg inne på hvorfor jeg ville avgrense meg til å se på disse frasene, nemlig for å unngå problematikken rundt at bestemthet kan være uttrykt andre steder i frasen, men også at forholdet mellom innhold og form er mer tydelig ved enkle substantivfraser ettersom det ikke er strukturelle, men referensielle forhold som avgjør formen på substantivet. Jeg har av samme grunn også valgt å utelate substantivfraser som står som utfylling i preposisjonsfraser.

Enkle substantivfraser i denne avhandlingen vil dermed for norsk være de fem formene *ubestemt entall*, *ubestemt flertall*, *bestemt entall*, *bestemt flertall* og det jeg har valgt å omtale som *artikkelfrase*, altså substantivfraser som består av ubestemt artikkel og substantiv i norsk.

7.2 Norsk

Siden tekstene i datamaterialet mitt er skrevet på norsk bokmål, er det dette som vil bli beskrevet her og vist til som *norsk*. Hovedkildene mine til beskrivelsen av norsk er Norsk referansegrammatikk (Faarlund, Lie & Vannebo 1997) og Norsk grammatikk for andrespråklærere (Hagen 1998).

7.2.1 Norske substantiv

De grammatiske kategoriene for norske substantiv er *genus*, *tall* og *bestemthet*. Det er fire hovedformer av substantivet: *ubestemt entall*, *ubestemt flertall*, *bestemt entall* og *bestemt flertall*. Ubestemt entall forekommer dessuten ofte sammen med delkvantorene 'en'/'ei'/'et', altså det som ofte omtales som «den ubestemte artikkelen» (Hagen 1998: 226), en term jeg også har brukt og vil bruke videre i denne avhandlingen. Siden den ubestemte artikkelen opererer i samspill med de ulike substantivformene, vil også substantiv med artikkel være del av beskrivelsen av norske substantiv og regnes som en enkel substantivfrase i denne avhandlingen.

Norsk er et flekterende språk, og har dermed ingen klare grenser mellom de ulike bøyingsmorfemene på substantivet, som uttrykker kombinasjoner av trekk for kategoriene tall og bestemthet (Hagen 1998: 33). De mest vanlige paradigmatisk

formene brukt i skrevet norsk, er framstilt i tabellen under. Substantivfrase med ubestemt artikkel er tatt med som en egen form.

Tabell 7-1 Det norske bøyningssystemet for substantiv

	Ubestemt entall	Ub. artikkel + entall	Ubestemt flertall	Bestemt entall	Bestemt flertall
Hankjønn	Bil	En	En bil		Biler
Hunkjønn	Jente	ei	Ei jente	<i>-er</i>	jenter
Intetkjønn	Eple	et	Et eple		Epler
				<i>-en</i>	Bilen
				<i>-a</i>	Jenta
				<i>-et</i>	Eplet
					<i>-ene</i>
					Bilene
					Jentene
					Eplene/epla

Genus er en inherent egenskap ved substantivene, og det er tre ulike kjønn i norsk: *hankjønn*, *hunkjønn* og *intetkjønn* (Faarlund et al. 1997: 149). Genus kommer i norsk bokmål som regel bare til syne på entallsformen, ettersom det er full synkretisme for flertallsformene, men det er ikke uvanlig at man har endelsen *-a* for intetkjønn i bestemt form flertall, og dermed også har et skille mellom intetkjønn og hankjønn/hunkjønn (Hagen 1998: 65). Verken genus eller formvariasjonen innen hver av de morfosyntaktiske kategoriene kommer til å være sentral for undersøkelse 2, så dette kommer jeg ikke til å gå nærmere inn i.

7.2.2 Generelle forhold knyttet til substantivets form

Den generelle regelen for bestemthetsbøyning av norske substantiv er at bestemt form brukes når substantivet har definnitt referanse, og i noen tilfeller ved generisk referanse, men *aldri* ved indefinnitt referanse (Hagen 1998: 55). Ubestemte former er typisk forbundet med indefinnitt referanse, men kan likevel også brukes både ved generisk og definnitt referanse. At ubestemt form av substantivet kan brukes selv om det er snakk om definnitt referanse, skyldes at bestemthet også kan uttrykkes med andre elementer i substantivfrasen, noe som av og til gjør det mulig å utelate den bestemte endelsen på substantivet (Faarlund et al. 1997: 283; Hagen 1998: 55).

Når man har et adjektiv i substantivfrasen vil foranstilt demonstrativ determinativ ('den'/'det'/'de') normalt sett være obligatorisk i norsk om referansen er definnitt, som i 7.2-a under. I slike tilfeller omtaler man gjerne determinativet som «adjektivets bestemte artikkel» (Hagen 1998: 212). Adjektivets bestemte artikkel forekommer som regel *sammen* med bestemt form av substantivet, og siden både artikkelen og endelsen

uttrykker bestemthet, kaller man dette *dobbel bestemmelse* (Hagen 1998: 55).³⁰ I noen tilfeller kan artikkelen imidlertid opptre sammen med ubestemt form, noe som kalles «enkel bestemmelse», og dette skjer ofte av stilistiske årsaker, men dobbelt og enkelt bestemmelse kan i noen tilfeller også skille mellom definitiv og generisk referanse, som i 7.2-b og 7.2-c (Hagen 1998: 56).

- 7.2-a Den grønne sykkelen (Dobbel bestemmelse)
- 7.2-b Den hvite mann (Enkel bestemmelse/generisk)
- 7.2-c Den hviten mannen (Dobbel bestemmelse/definitiv)
- 7.2-d Den sykkelen er ikke min. (Påpekende)

Demonstrativer kan også plasseres direkte foran substantivet, men denne bruken er utelukkende demonstrativ (påpekende), og vil nesten alltid kreve bestemt form av substantivet, som i 7.2-d over (Hagen 1998: 212).

Også andre determinativer kan stå sammen med ubestemt form selv om referansen er definitiv. Possessive determinativer kan for eksempel stå både før og etter substantivet, noe som påvirker om substantivet skal ha bestemt eller ubestemt form: Foranstilt possessiv krever ubestemt form av substantivet, som i 7.2-e, mens etterstilt possessiv krever bestemt form, som i 7.2-f (Hagen 1998: 57):

- 7.2-e Min sykkel er ikke grønn.
- 7.2-f Sykkelen min er ikke grønn.

Samspeillet mellom demonstrativer og bestemthetsendelse i norsk gjør altså at ubestemt form av substantivet av og til kan brukes selv om det er definitiv referanse, men dette vil sjelden skje i *enkle* substantivfraser i korrekt norsk. Hagen (1998: 55) skriver at «Når substantivfrasen mangler foranstilte ledd, slik at kjernen blir det første (eller eneste) ledd i substantivfrasen, er bestemt form obligatorisk dersom definitiv referanse er intendert.»

7.2.3 Tall og tellelighet

I norsk substantivbøyning skiller man også mellom *entall* og *flertall*, og kategorien tall er som jeg var inne på i 3.4, en relativt gjennomiktig kategori innen indefinitiv og definitiv

³⁰ En mer presis term kunne kanskje ha vært «trippel bestemmelse» siden bestemthet også uttrykkes på adjektivet i disse frasene. Lundeby (1965: 36) bruker derfor heller termen «overbestemt».

referanse, men ved generisk referanse er valget mellom entall og flertall styrt av andre forhold enn antall, ettersom man ved generisk referanse alltid refererer til mengder.

Som andre språk skiller også norsk mellom *tellelige* og *ikke-tellelige* substantiv (Faarlund et al. 1997: 143-144). Tellelige substantiv viser til individuelle enheter som kan telles, og disse vil ha ubestemt artikkel ved entall, og flertallsendelse ved flertall, som i setning 7.2-g (under). Ikke-tellelige substantiv viser til abstrakte begreper eller massesubstantiv, og vil typisk ikke ha ubestemt artikkel eller flertallsendelse, men heller ubestemt entallsform, som i 7.2-h. Selv om ubestemt entallsform er mest vanlig ved ikke-tellelige substantiv, kan denne formen også enkelte ganger brukes ved tellelige substantiv, som da får en ikke-tellelig fortolkning (7.2-i) (Faarlund et al. 1997: 148). I tillegg kan ikke-tellelige substantiv stå med flertallsendelse og dermed få en individualisert betydning (7.2-j) (Faarlund et al. 1997: 145).

7.2-g Frode så en elg/elger i hagen.

7.2-h Ine så sukker i skapet.

7.2-i Det var elg i hagen.

7.2-j Det fins mange forskjellige intelligenser.

Vekslingen mellom ubestemt entall med og uten ubestemt artikkel knyttes særlig til forskjellen mellom tellelige og ikke-tellelige substantiv, men ubestemt entall uten artikkel kan i norsk altså også brukes ved tellelige substantiv, noe jeg vil komme tilbake til under.

7.2.4 Referanse og valg av bestemte og ubestemte former

Jeg vil videre beskrive forholdet mellom form og referanse slik det typisk forekommer i norsk, først for individuell referanse (definit/indefinit), deretter generisk referanse og ikke-referensiell bruk.

7.2.4.1 Individuell referanse

Som nevnt i 3.2 er individuell referanse i denne avhandlingen brukt som en fellesterm for definit og indefinit referanse, og som vi så over, er det bestemte former som er det typiske formvalget ved definit referanse, mens indefinit referanse typisk har ubestemte former. Innen individuell referanse har jeg her også inkludert tilfeller av *svak referanse*, altså substantivfraser som grenser til å være ikke-referensielle, men likevel også introduserer en diskursreferent (se 3.2.1). Dette gjelder *ikke-spesifikke* og *attributive*

substantivfraser, men det gjelder også framheving av *type* heller enn *eksemplar* ved indefinitt referanse, som jeg vil omtale under.

Norsk markerer ikke forskjellen mellom *spesifikk* og *ikke-spesifikk* referanse, det vil si om referansen er til én spesifikk referent som man har i tankene, eller til en hvilken som helst referent som passer til beskrivelsen (Hagen 1998: 58). Norsk bruker ubestemt form i begge tilfeller, som vist i setning 7.2-k og 7.2-l under. Det samme gjelder forskjellen mellom definitt referanse (som alltid er spesifikk) og *attributiv bruk*, hvor referansen er til en spesifikk referent, men hvor denne referenten ikke er kjent for senderen av ytringen (Faarlund et al. 1997: 51). I begge tilfeller vil norsk bruke bestemt form, som i 7.2-m og 7.2-n.

7.2-k John Inge er gift med en lærer. (Spesifikk)

7.2-l John Inge vil gifte seg med en lærer. (Ikke-spesifikk)

7.2-m Læreren som sto foran meg, var sint. (Definitt)

7.2-n Læreren som fortalte deg dette, tok feil. (Attributiv)

Et annet aspekt av individuell referanse i norsk er muligheten til å skille mellom fokus på *type* og fokus på *eksemplar* (token) ved indefinitt referanse (Borthen 2003: 22, 51; Vannebo 2010: 231). Selv om det vanlige er å forbeholde ubestemt entallsform til ikke-tellelige substantiv (7.2-o, under), kan man i objektsposisjon også ofte bruke denne formen ved tellelige substantiv, noe som av flere omtales som *naken substantivfrase*.³¹ Forskjellen på et tellelig substantiv med og uten artikkel er at man *med* artikkel framhever referenten i seg selv, altså eksemplaret (7.2-p), mens man *uten* artikkel framhever aktiviteten i setningen og typen av referent (7.2-q). Fokus på type gjennom entallsform kan også gjøres når det er snakk om en *mengde* av referenter (Halmøy 2010: 82), noe vi ser av substantivet «kake» i setning 7.2-q, som ikke sier noe om hvorvidt det er én eller flere kaker som har blitt kjøpt.

³¹ Flere forbeholder i beskrivelsen av norsk benevnelsen «naken substantivfrase» til bruk av ubestemt entall på tellelige substantiv, og bruker det ikke om samme formen brukt ved ikke-tellelige substantiv (Borthen 2003: 10; Vannebo 2010: 230). Det understrekede substantivet i ‘Ine har kjøpt kake’ er dermed en naken substantivfrase, mens det ikke er det i setningen ‘Ine har kjøpt sukker’.

- 7.2-o Ine har kjøpt sukker. (Ikke-tellelig)
7.2-p Ine har kjøpt en kake. (Eksemplar-fokus)
7.2-q Ine har kjøpt kake. (Type-fokus)

Som jeg kommer tilbake til, er dette et av de områdene innen grammatisk bestemthet hvor norsk og engelsk skiller seg mest fra hverandre.

7.2.4.2 *Generisk referanse*

I norsk kan både bestemte og ubestemte former brukes ved generisk referanse, faktisk kan en hvilken som helst form av substantivet brukes (Faarlund et al. 1997: 292). Valg av form ved generisk referanse styres av både syntaktiske og semantiske forhold. Syntaktisk sett er bestemte entallsformer mer akseptert enn de andre formene når den generiske frasen ikke er et subjekt (Hagen 1998: 61). Man kan videre til en viss grad knytte valg av form til ulike undertyper av generisk referanse: *taksonomisk*, *generaliserende* og *referanse til klasse* (se 3.2.1.1).

For taksonomiske generika er artikkelfrase det vanlige i norsk, men også flertallsformen og ubestemt entall kan brukes, og det både med tellelige og ikke-tellelige substantiv (Faarlund et al. 1997: 292; Hagen 1998: 60). For generaliserende generika er det mest vanlig å bruke ubestemt flertall, som i setning 7.2-r under, mens referanse til klasse vanligvis uttrykkes med bestemt entall eller flertall, som i 7.2-s (Halmøy 2010: 69). Halmøy hevder også at ubestemt flertall ikke kan brukes for referanse til klasser i norsk, på den måten at «poteter» i setning 7.2-t ikke kan tolkes som generisk referanse, men må tolkes som individuell, indefinit referanse (Halmøy 2010: 61-62, eksempler basert på Halmøy).

- 7.2-r Poteter inneholder proteiner. (Generaliserende)
7.2-s Poteten kom fra Peru til Europa med båt. (Klasse)
7.2-t Poteter kom fra Peru til Europa med båt. (ikke-generisk, indefinit)

Ved generisk referanse i norsk er det altså mulig å bruke en hvilken som helst form av substantivet og valget styres i stor grad av generiske undertyper, men det er også en god del fleksibilitet i hvilke former som kan brukes. Selv om det vanlige i norsk ser ut til å være artikkelfrase ved taksonomi, ubestemt flertall ved generaliserende, og bestemt form entall ved referanse til klasse, så kan andre former også brukes i disse generiske kontekstene.

7.2.4.3 *Ikke-referensiell bruk*

Siden svak referanse, som grenser til å være ikke-referensiell, i denne avhandlingen regnes som referensielle, er det først og fremst predikative substantivfraser som inngår som ikke-referensielle, da disse helt klart beskriver framfor å referere (se 3.2). Norske predikative substantivfraser kan ha både bestemt og ubestemt form, og stå med eller uten ubestemt artikkel. Valget av ubestemt form med eller uten artikkel er semantisk styrt, knyttet til hvorvidt egenskapene predikativet viser til er ment å være *klassifiserende* eller *karakteriserende* (Halmøy 2010: 77; Vannebo 2010: 231). Mens klassifisering uttrykkes med naken form eller flertall, avhengig av om det er entall eller flertallssubjekt (7.2-u), er karakteriseringer uttrykt med artikkelfrase ved entall, og *noen* + ubestemt flertall ved flertall (7.2-v).

7.2-u Jeg er lærer/De er lærere. (Klassifiserende)³²

7.2-v Jeg er en djevel/De er noen djevler. (Karakteriserende)

Det er også andre typer predikativer, som kan ta bestemte former, men de fleste av disse går inn under svak, og dermed individuell, referanse, slik jeg har organisert denne framstillingen.

7.2.4.4 *Oppsummering av substantivfraser og referanse i norsk*

Som vi har sett, er bruk av bestemt og ubestemt form i norsk styrt av ulike semantiske forhold som hører inn under den grammatiske kategorien bestemthet. For det første er det bruk av bestemt og ubestemt form for definit og indefinit referanse, som hovedsakelig handler om *identifiserbarhet*, men hvor også type/eksemplar-fokus kan uttrykkes innen indefinit referanse. For det andre har vi de generiske bruksområdene, hvor valg av form er mer fleksibelt og styres delvis av syntaks og delvis av ulike generiske undertyper, og hvor valget mellom entall og flertall ikke direkte reflekterer antall i seg selv siden generisk referanse alltid vil vise til mengder. For det tredje har vi den ikke-referensielle bruken hvor man i norsk skiller mellom klassifisering og karakterisering, noe som først og fremst er relevant for valget mellom ubestemt flertall og artikkelfrase på den ene siden, og naken form av substantivet på den andre.

³² Her er det verdt å merke seg at man i norsk kan se ut til å ha en språkendring på gang, der det blir vanligere å bruke artikkel også ved klassifiserende predikativ.

Typiske formvalg ved enkle substantivfraser i norsk vil dermed være følgende:

Definit referanse	Bestemte former
Indefinit referanse	Ubestemte former (artikkel frase, ubestemt entall/flertall) a) Artikkel frase ved tellelige entallssubstantiv (fokus på eksemplar) b) Ubestemt flertall ved tellelige flertallssubstantiv c) Ubestemt entall ved ikke-tellelige substantiv d) Ubestemt entall ved tellelige substantiv (fokus på type)
Generisk referanse	Ubestemte og bestemte former a) Artikkel frase ved taksonomisk b) Ubestemt flertall ved taksonomisk og generaliserende c) Bestemte former ved referanse til klasse d) Ubestemt entall ved ikke-tellelige substantiv
Ikke-referensiell	Ubestemte (og av og til bestemte former) a) ubestemt entall og flertall ved klassifisering b) artikkel frase og flertall ved karakterisering

Jeg vil nå gå videre til å presentere substantiv i polsk og engelsk, og i slutten av kapittelet vil jeg oppsummere likheter og forskjeller mellom de tre språkene.

7.3 Polsk

I dette underkapittelet om polsk vil jeg presentere en generell beskrivelse av relevante aspekter av polsk substantivbøyning med vekt på kasus, i tillegg til noen foreslåtte bestemthetsekvivalenter i polsk. Hovedkildene til beskrivelsen av polsk språk er Swans *A Grammar of Contemporary Polish* (2002), supplert av Wiese (2011) og Feldstein (2001), men også noen kontrastive arbeider som tar for seg polsk og engelsk (Szwedek 1976; Król-Markefka 2008; Venneri 2011).

7.3.1 Polske substantiv og morfologiske kategorier

Polsk substantivbøyning uttrykker morfologisk kategoriene tall, genus og de syv kasusene *nominativ*, *akkusativ*, *genitiv*, *dativ*, *lokativ*, *instrumentalis* og *vokativ*. Jeg vil her kort gå igjennom noen grunnleggende trekk ved kasus i polsk, og jeg vil først se på funksjonssiden og deretter si litt om det morfologiske systemet. Alle eksempelsetninger i omtalen av polsk er hentet fra Swan (2002).

7.3.1.1 *De ulike kasusene i polsk*

Kasus uttrykkes i polsk kun ved bøyningssendelser, eller ved en kombinasjon av bøyning og preposisjoner, og Swan (2002: 327) skiller mellom tre bruksområder for kasus: 1) grunnleggende syntaktisk bruk, 2) bruk styrt av spesifikke verb eller preposisjoner og 3) idiomatisk bruk.

Nominativ form er ifølge Swan (2002: 327) oppslagsformen for polske substantiv. *Nominativ* er først og fremst assosiert med subjektet (7.3-a), men kan også forekomme i andre funksjoner, for eksempel som spesifiserende substantivfrase etter kopulaverb, som deler referanse med subjektet og derfor kongruere med det (7.3-b).

7.3-a Pada deszcz. (Regn-NOM faller)

7.3-b To jest mój brat. (Det-NOM er min bror-NOM)

Akkusativ brukes først og fremst ved direkte objekt (7.3-c), etter visse preposisjoner og blir dessuten ofte brukt idiomatisk og for å uttrykke tidsperioder og mengde (7.3-d) (Swan 2002: 354).

7.3-c Ogladam telewizje. (Jeg ser på TV-AKK)

7.3-d Czekam tu godzinę. (Jeg har ventet her flere timer-AKK)

Genitiv er en av de mest brukte kasusformene i tillegg til *nominativ* og *akkusativ*, og har mange ulike bruksområder i tillegg til den grunnleggende possessive bruken, som er vist i setning 7.3-e under (Swan 2002: 329). *Genitiv* brukes også ved mengde og etter tallord fra fem og oppover (7.3-f), og formen kan dessuten brukes partitivt (7.3-g). Et siste bruksområde er den såkalte «genitive of negation», hvor *genitiv* brukes i stedet for *akkusativ* for objekter i negerte setninger, som i 7.3-h (Swan 2002: 333).

7.3-e Książka studenta (studentens-GEN bok)

7.3-f Piec jabłek (fem epler-GEN)

7.3-g Dodac cukru. (Skal jeg legge til litt sukker-GEN)

7.3-h Nie ogladam telewizji. (Jeg ser ikke på TV-GEN)

Instrumentalis blir av Swan (2002: 360) beskrevet som det av kasusene som har mest variert bruk, men brukes i hovedsak til å beskrive *måten* noe gjøres på i tillegg til en del idiomatisk bruk. *Nominale* predikativer står dessuten nesten uten unntak i *instrumentalis*, som i 7.3-i (Swan 2002: 363).

7.3-i On jest polakeim. (Han er polakk-INSTR)

Dativ brukes typisk ved indirekte objekt (7.3-j), men slike objekter er vanligvis uttrykt med pronomer og navn, og sjeldent av substantiv. Swan (2002: 344) hevder derfor at dette er det minst utbredte av de polske substantivkasusene, i tillegg til vokativ.

7.3-j Dalem pasze koniowi. (Jeg ga fôr til hesten-DAT)

De to siste kasusene er lokativ og vokativ. *Lokativ* opptrer alltid sammen med preposisjoner, som i setning 7.3-k (Swan 2002: 368-369), mens *vokativ* er et relativt sjeldent kasus brukt for å henvende seg direkte til noen, sammen med titler eller navn (Swan 2002: 344, 371).

7.3-k Na basenie (i svømmebassenget-LOK)

7.3-l Kochana Siostró (kjære søster-VOK)

Dette var en kort beskrivelse av typiske bruksområder for polske kasus, og jeg vil nå gi en grunnleggende oversikt over hvordan de ulike kasusene kommer til uttrykk gjennom substantivmorfologi.

7.3.1.2 *Kasusenes morfologi i polsk*

Formene på de polske substantivene uttrykker både kasus og tall, og valg av endelse er dels styrt av hvilket kjønn substantivet har (hunkjønn, hankjønn eller intetkjønn), og dels av hvilken konsonant stammen ender på. Her skiller man mellom *harde* og *myke* konsonanter i stammen, som i praksis er et skille mellom palatale eller palataliserte konsonanter og de øvrige (Feldstein 2001: 39; Wiese 2011: 104-105). For maskuline substantiv i akkusativ vil dessuten kategoriene *animat* og/eller *viril* være relevant for formvalget, altså for det første om referenten er *levende* eller ikke, og for det andre om en levende hankjønnreferent er *menneskelig* eller ikke (Feldstein 2001: 41).

Wiese (2011: 107-108) beskriver det polske substantivsystemet som en blanding av flekterende og agglutinerende, og skriver at bøyningssystemet er preget av mange ikke-unike forhold mellom form og funksjon, der flere likelydende endelser opptrer i ulike kasus og i ulike bøyningssklasser. På grunn av dette foreslår Wiese (2011: 111) at endelsene med lik form ikke bare skal ses på som homonymer, men som de samme endelsene. Konsekvensen av dette er at ingen endelser er unikt assosiert med ett bestemt kasus (Wiese 2011: 107).

Tabellene under viser de mest sentrale endelsene på polske substantiv i entall (tabell 7-2) og flertall (tabell 7-3).³³ Der formene varierer etter stamme innen de ulike genusklassene, har jeg tatt med flere endelser, men uten å presisere hvilken stamme de brukes ved, da jeg ikke fant dette nødvendig for mitt formål.

Tabell 7-2 Bøyingssystemet for polske substantiv i entall

Entall	Hunkjønn	Hankjønn	intetkjønn
Nominativ	-a /-i/Ø	Ø	-o/-e
Akkusativ	ę/Ø	Ø/a	-o/-e
Genitiv	-i	-u/a	-a
Instrumentalis	-ą	- em	
Dativ	-e/-i	-owi	-u
Lokativ	-e/-i	-e/-u	
Vokativ	-o/-i	-e/-u	-o/-e

Tabell 7-3 Bøyingssystemet for polske substantiv i flertall

Flertall	Hunkjønn	Hankjønn	intetkjønn
Nominativ/vokativ	-i/-e	-i/-owie/-e	-a
Akkusativ	-i/-e	-e/-i/-ow	-a
Genitiv	Ø/-i	-ów	Ø/ ow/i
Instrumentalis	-ami		
Dativ	-om		
Lokativ	-ach		

Det som kommer til syne her, er altså stor grad av sammenfall mellom formene for de ulike kasusene og innen de ulike genusklassene. En kan også merke seg at *mangel på bøyningsendelse* (nullform) kun forekommer i nominativ, genitiv og akkusativ, men aldri innen dativ, lokativ, instrumentalis og vokativ. Ofte er det nominativformen i kasusspråk som er både funksjonelt og morfologisk *umarkert*, altså både representerer den nøytrale bruken og mangler endelse (Falk 1999: 2), men som vi ser av tabellene over, er dette ikke tilfellet i polsk, da nullformen for det første bare opptrer ved visse nominativer (alle hankjønn og noen hunkjønnord), og for det andre både i akkusativ og i genitiv flertall for enkelte ord. Som jeg skal komme tilbake til, kan likevel nominativ regnes som det umarkerte kasuset i polsk, ettersom morfologisk markerthet kun er sekundært til funksjonell markerthet (Falk 1999: 2).

³³ Framstillingen i tabellene er basert på både Feldstein (2001) og Wiese (2011).

Sammenlignet med norsk er det altså en langt mer variert substantivmorfologi i polsk, men systemene deler det at de uttrykker kategorien tall og at det er tre ulike genusklasser, selv om det kan være forskjellig hvilket genus spesifikke ord i de to språkene tilhører. Norsk og polsk har altså det til felles at de har genus og substantivmorfologi som uttrykker mer enn kategorien tall, noe som ikke gjelder engelsk, som bare bøyer substantiv i tall og ikke har genus.

7.3.2 Bestemthetsekvivalenter i polsk

I kapittel 3.3 var jeg inne på fenomenet bestemthetsekvivalenter i -GB-språk, hvor det blant annet ble trukket fram en kobling mellom pragmatisk ordstilling og grammatisk bestemthet knyttet til at begge deler kan sies å signalisere identifiserbarhet, men på ulike måter og bare delvis overlappende når det kommer til hvor referentene kan identifiseres: Pragmatisk ordstilling signaliserer først og fremst om noe er nevnt før eller ikke (tekstuell identifiserbarhet), mens grammatisk bestemthet inkluderer både tekstuell og utenom-tekstlig identifiserbarhet. Som jeg også var inne på i kapittel 3, omtalte Trenkic (2002: 114) denne forskjellen som en forskjell mellom identifiserbarhet som en *bevissthetskategori* ved pragmatisk ordstilling, og identifiserbarhet som en *kunnskapskategori* ved grammatisk bestemthet. Dette vil i så fall si at forskjellen mellom disse to uttrykkene for identifiserbarhet ikke bare handler om når en substantivfrase er *kjent* og *ny* i de ulike språkene, men kanskje også om en begrepsmessig forskjell, nemlig at brukere av -GB-språk og +GB-språk har ulike identifiserbarhetsbegreper.

Vitenskapelige arbeider som ser på bestemthetsekvivalenter spesifikt i polsk, nevner også pragmatisk ordstilling, i tillegg til bruk av demonstrativer og intonasjon, som ekvivalenter til bestemthet (Szwedek 1976: 203; Król-Markefka 2008: 106; Venneri 2011: 207). Som jeg var inne på i kapittel 3.3 mener Szwedek (1976: 206) at ordstilling i polsk er involvert i anafori og derfor uttrykker en lignende funksjon som grammatisk bestemthet: «[...] contrary to widespread and advocated beliefs, word order in Polish is not free and that it plays a crucial role in anaphoric processes.» Szwedek viser også til samspillet mellom ordstilling og demonstrativer, der demonstrativer kan spille en rolle i å signalisere identifiserbarhet når ordstillingen ikke følger tema/remainprinsippet: Setningsinitiale og setningsfinale substantivfraser blir i polsk fortolket som henholdsvis definte og indefinte med mindre det er eksplisitt markert at de ikke er

det, og Szwedek (1976: 208) trekker her fram det polske pronomenet ‘ten’ som en ekvivalent til bestemt artikkel i engelsk ettersom ‘ten’ er obligatorisk når definite substantiv står finalt i setningen:³⁴

[...] nouns with indefinite interpretation appear in sentence final position only (unless explicitly marked indefinite in some other way). This is why the pronoun *ten* is obligatory with a noun in this position if the noun is to be interpreted as definite. Nouns with definite interpretation appear in non-final positions (again, unless explicitly marked otherwise).

Król-Markefka (2008: 108) viser imidlertid til at ‘ten’ bare delvis overlapper med bestemt artikkel i engelsk, og at overlappingen er knyttet til *demonstrativ* bruk av artikler, ikke definitt. Axelsson (1994: 22) mener også at ‘ten’ først og fremst er en ekvivalent til demonstrativer i svensk.

Andre uttrykk i polsk som har blitt vurdert som ekvivalenter til engelske artikler, er først og fremst pronomenet ‘jakis’ og tallordet ‘jeden’, som skal ha likhetstrekk med henholdsvis ‘a(n)’ og ‘one’ i engelsk (Król-Markefka 2008: 107-108; Venneri 2011: 207). Siden vi i norsk ikke skiller like tydelig mellom tallord og artikkel som i engelsk, altså at de determinative kvantorene ‘en’/’ei’/’et’ fungerer som begge deler i norsk, kan man altså si at både ‘jakis’ og ‘jeden’ tilsvarer ‘en’/’ei’/’et’ i norsk. Król-Markefka (2008: 107-108) mener likevel at det også for begge disse to polske markørene bare er snakk om en delvis overlapping med engelsk, da bruk av både ‘jakis’ og ‘jeden’ er begrenset til spesifikk referanse, og hun viser også til at ‘jakis’ ikke kan brukes ved generisk referanse. Ubestemte artikler i engelsk er derimot ikke begrenset til å være spesifikke, og som nevnt i beskrivelsen over, gjelder dette også norsk ettersom ubestemt artikkel brukes både ved spesifikk og ikke-spesifikk referanse (se 7.2.4.1).

Król-Markefka (2008: 108-109) konkluderer sin framstilling av forholdet mellom polsk og engelsk med at det fins noen delvise ekvivalenter, men at verken ordstilling eller demonstrativer sammenfaller fullt ut med artikler i engelsk:

[...] it might be stated that each of the aspects of the Polish language usually presented as equivalent to *the* or *a/an* only partially overlaps with the definite/indefinite distinction: word order in some cases reflects the juxtaposition of given (familiar) and new (unfamiliar)

³⁴ Merk at Szwedek (1976) omtaler ‘ten’ som pronomen, og det samme gjelder Swan (2002), mens Król-Markefka (2008) og Venneri (2011) omtaler det som en demonstrativ.

information; indefinite pronouns refer to specific entities which are not identifiable by the speaker, and the demonstrative *ten* in some contexts is used as the definite article.

Gjennomgangen tyder på at det er en tendens til at samtlige av disse foreslåtte ekvivalentene til bestemthet kun regnes som *delvis* overlappende med både engelske og norske bestemthetsmarkører, men denne delvise overlappingen er likevel interessant, og jeg vil komme tilbake til noe av dette i forbindelse med hypotese 1 i kapittel 8.

7.3.3 Referensielle typer i polsk

Selv om polsk ikke har eksplisitte markører for å uttrykke de ulike referensielle typene, vil jeg nå beskrive hvilke former som typisk brukes i polsk ved ulike typer referanse.

7.3.3.1 Individuell referanse

En av hovedforskjellene mellom språk med og uten bestemthet er den grammatikaliserte markeringen av identifiserbarhet, som først og fremst er relevant for kontekster med individuell referanse (definit og indefinit). Som jeg var inne på over, kan identifiserbarhet ha en sammenheng med hvilken setningsplassering substantivfrasen har i polsk, og i noen tilfeller er identifiserbarhet også relevant for bruk av demonstrativer og pronomen i polsk. Sett bort fra dette, er polske substantivfraser i seg selv *tvetydige* med tanke på identifiserbarheten til referenten, denne går heller fram av ordstillingen, og valg av form er heller basert på kategoriene kasus og tall. Ved individuell referanse i polsk, vil derfor formen i stor grad variere med *syntaktisk funksjon* og andre forhold som dekkes av kasusformer, heller enn identifiserbarhet. Der norsk knytter ubestemte former til indefinit referanse og bestemte former til definit referanse, vil det altså i polsk være slik at substantivenes form heller varierer med kasus både ved definit og indefinit referanse.

7.3.3.2 Generisk referanse

Polsk har ingen spesifikk markør for generisk referanse, men Smólska og Rusiecki (1980: 55) viser at generisk referanse er psykologisk reelt i polsk til tross for at det ikke uttrykkes eksplisitt. C. Lyons (1999: 185) hevder at forskjellen mellom entalls- og flertallsgenerika foreligger også i språk uten grammatisk bestemthet, der entallssubstantiv brukes taksonomisk, altså for å uttrykke en essensiell egenskap ved en

klasse. Smólska og Rusiecki (1980: 55) mener imidlertid at forskjellen mellom entall og flertall ved generisk referanse i polsk først og fremst er stilistisk, selv om de også viser til at det fins visse generiske substantivfraser som aldri kan stå i flertall. Utover dette er det ved generisk referanse også kasussystemet som styrer hvilken form substantivet har, slik det er ved definitiv og indefinitiv referanse. Dette betyr at det ikke er noen vesentlig forskjell på hva som styrer formvalget innen de tre referensielle typene definitiv, indefinitiv og generisk i polsk.

7.3.3.3 Ikke-referensiell bruk

Som nevnt i omtalen av polske kasus, vil substantiv som fungerer som predikativ, så å si alltid ha instrumentalis i polsk (Swan 2002: 363). Polsk ser dermed ikke ut til å uttrykke den forskjellen mellom klassifiserende og karakteriserende predikativer som ble omtalt for norsk (se 7.2.4.3). Predikative substantivfraser vil imidlertid i polsk skilles fra *spesifiserende* substantivfraser, altså postkopulære ledd som deler referanse med subjektet. Ifølge Swan (2002: 378) bøyes disse i samsvar med subjektet og tar derfor nominativ (7.3-m), i motsetning til de ikke-referensielle predikativene, som altså tar instrumentalis (7.3-n):

7.3-m To jest mój brat. (Det-NOM er min bror-NOM)

7.3-n On jest polakeim. (Han er polakk-INSTR)

Spesifiserende fraser er imidlertid i denne avhandlingen regnet som svak, individuell referanse, og ikke som ikke-referensielle, så i all hovedsak kan man si at ikke-referensielle, enkle substantivfraser, slik de er definert her, vil ha instrumentalis i polsk.

7.3.4 Oppsummering av beskrivelsen av polsk

Sammenlignet med norsk er kanskje den tydeligste forskjellen at polsk ikke har grammatisk bestemthet, men grammatisk kasus, og i denne beskrivelsen av polsk har jeg sett på noen grunnleggende trekk ved polsk kasus og substantivmorfologi, samt andre språklige forhold som det har blitt hevdet at uttrykker noen av de samme funksjonene som bestemthetsmarkører i +GB-språk. Selv om det er visse overlappinger mellom ordstilling eller enkelte demonstrativer og pronomener i polsk og artikler i

engelsk, ser det altså fra litteraturen ut til at det bare er snakk om en delvis overlapping, noe som også må gjelde norsk.

Et annet viktig moment er at ulike referansetyper ikke har noen særlig påvirkning på valg av form i polsk, og kasus vil derfor operere likt på tvers av referensiell type: Formvalget både ved indefinit, definit og generisk referanse er styrt av kasus i polsk, typisk nominativ, akkusativ, dativ og genitiv, mens ikke-referensielle predikativer har instrumentalis.

Under er en oppsummerende oversikt over hvilke former som *typisk* vil bli brukt innen de ulike referansetyperne i polsk:

Definit referanse		Ulike former avhengig av kasus:
Indefinit referanse		
Generisk referanse		
		a) Nominativ ved subjekter
		b) Akkusativ og genitiv ved objekter
		c) Dativ ved indirekte objekter
		d) Lokativ ved adverbialer
Ikke-referensiell		Instrumentalis

Etter å ha sett på engelske substantivfraser, vil jeg komme tilbake til en oppsummerende sammenligning av polsk, norsk og engelsk.

7.4 Engelsk

Siden engelsk i likhet med norsk har grammatisk bestemthet, deler engelsk mange av de formelle og semantiske trekkene som er beskrevet for norsk, men det er også noen viktige forskjeller. I dette underkapittelet vil jeg kort presentere engelske substantivfraser med noen eksempler, og peke på noen sentrale forskjeller fra norsk, spesielt forskjeller knyttet til bruk av artikkelfrase og naken form av substantivet, i tillegg til valg av form ved generisk referanse.

7.4.1 Engelske substantiv og substantivfraser

Engelske substantiv bøyes kun i tall, mens bestemthet er markert med foranstilte bestemte og ubestemte artikler ('the' / 'a(n)'). Den engelske umodifiserte

substantivfrasen har fem ulike former, som korresponderer direkte til de fem typene enkle substantivfrasene som ble omtalt for norsk (se 7.1). Formelt sett ligger forskjellen mellom de to språkene hovedsakelig i hvordan trekket bestemt realiseres: Norsk har bøyning, engelsk har artikkel. En annen forskjell fra norsk er at den ubestemte artikkelen ‘a(n)’ i engelsk ikke er homonym med tallordet ‘one’, slik det er i norsk (‘en’/’ei’/’et’). Tabellen under viser hovedsystemet for bestemte og ubestemte enkle substantivfraser i engelsk.

Tabell 7-4 Det engelske systemet for substantiv

Naken ubestemt	Ubestemt entall		Ubestemt flertall		Bestemt entall		Bestemt flertall	
	a(n)	a car	-s	cars	The	the car	the -s	the cars
Car	a(n)	a car	-s	cars	The	the car	the -s	the cars

Forskjellen i realiseringen av trekket bestemt i engelsk og norsk har noe påvirkning på substantivfrasen for øvrig, for eksempel at komplekse determinativer som uttrykker en kombinasjon av bestemthet og possessivt eller demonstrativt innhold, *erstatter* den bestemte artikkelen i engelsk (7.4-a), mens slike ofte forekommer *sammen* med bestemt form i norsk (7.4-b) (C. Lyons 1999: 118; Hagen 1998: 55).

7.4-a He chose both cars.

7.4-b Han valgte begge bilene.

Til tross for disse forskjellene korresponderer engelske og norske former i stor grad direkte, men de korresponderende formene utviser forskjeller når det kommer til hvordan de vil fortolkes i spesifikke kontekster og med tanke på hvor frekvente de er (Halmøy 2010: 54). Et eksempel på dette er at bruk av ubestemt entallsform uten artikkel har et mye mer begrenset bruksområde i engelsk enn i norsk: I engelsk er naken entallsform begrenset til *ikke-tellelige* substantiv, i tillegg til noen få andre tilfeller av bruk med tellelige substantiv. De Swart og Zwarts (2009: 280) nevner fem kontekster der naken entallsform (eng. ‘bare noun’) kan brukes ved tellelige substantiv: *bare location* (7.4-c), *bare coordination* (7.4-d), *bare predication* (7.4-e), *bare reduplication* (7.4-f) og *bare incorporation* (7.4-g). Eksemplene er fra de Swart og Zwarts (2009).

7.4-c John is at school.

7.4-d The way to use knife and fork.

7.4-e Mary is chair of the department.

7.4-f She went through door after door.

7.4-g John is playing piano.

Som jeg var inne på i omtalen av norsk, kan den ubestemte entallsformen ofte brukes også ved tellelige substantiv i norsk, men i engelsk vil man altså som regel foretrekke artikkelfrase eller ubestemt flertallsform ved tellelige substantiv. Forskjellene mellom norsk og engelsk slår spesielt ut ved tellelige objekter og ved predikativ, noe jeg vil komme tilbake til under.

7.4.2 Referanse og valg av artikkel i engelsk

Den følgende delen beskriver noen forskjeller og likheter mellom engelsk og norsk med hensyn til hvordan de bestemte og ubestemte formene er distribuert, og hva som uttrykkes utover kjernedistinksjonen mellom definitiv og indefinitiv referanse. Beskrivelsen er også her delt inn etter referensielle typer, og forskjellene som vil bli vektlagt, er for det første formvalg ved generisk referanse og for det andre valget mellom naken substantivfrase og artikkelfrase ved indefinitiv objekter og ved predikativer.

7.4.2.1 Individuell referanse

Ettersom individuell referanse representerer kjernen i bestemthetsinnholdet, er dette et område hvor engelsk og norsk er mer eller mindre like i valg av form: Ubestemt form assosieres med indefinitiv referenter, og bestemt form med definitive. Heller ikke hoveddomenene for å etablere definitiv referanse (*anaforisk, situasjonell, unik og assosiativ*) er nevneverdig ulike i de to språkene (Chesterman 1991: 52; C. Lyons 1999: 4).

I beskrivelsen av norsk presenterte jeg tilfeller av svak referanse, og også her er norsk og engelsk langt på vei like. Distinksjonen spesifikk/ikke-spesifikk, altså om senderen av ytringen har en spesifikk referent i tankene eller ikke, er heller ikke uttrykt eksplisitt i engelsk, slik vi ser i setning 7.4-h og 7.4-i under (Declerk 1986: 27). På samme måte er også engelske substantivfraser tvetydige med tanke på om de er attributive eller definitive, som vil si at bestemt artikkel brukes både når referenten er kjent og ukjent for senderen av ytringen (7.4-j og 7.4-k).

- 7.4-h I'm married to a teacher. (Spesifikk)
 7.4-i I want to marry a teacher. (Ikke-spesifikk)
 7.4-j The teacher in front of me was angry. (Definitiv)
 7.4-k Whoever he is, the teacher who told you this, was wrong. (Attributiv)

Når det gjelder den tredje typen svak referanse som ble omtalt i beskrivelsen av norsk (fokus på type), så er dette et av de områdene hvor norsk og engelsk skiller seg tydeligst fra hverandre. I norsk så vi at ubestemt entall med og uten ubestemt artikkel kunne uttrykke et skille mellom fokus på eksemplar eller fokus på type, noe som ikke gjelder engelsk siden naken entallsform stort sett er forbeholdt ikke-tellelige substantiv. Ubestemt artikkel kan dermed ikke utelates ved tellelige substantiv i engelsk slik det er mulig å gjøre i norsk, noe som kommer fram av at 7.4-l er ugrammatisk, mens 7.4-m ikke er det. Artikkel må derimot utelates ved ikke-tellelige substantiv i engelsk (7.4-n).

- 7.4-l *I have bought cake. (Tellelig)
 7.4-m I have bought a cake. (Tellelig)
 7.4-n I love sugar. (Ikke-tellelig)

Mens det er store likheter med tanke på hvilke former som brukes ved definitiv referanse i engelsk og norsk, er det altså noen forskjeller innen indefinitiv referanse og i valget mellom de ulike ubestemte formene. Dette slår spesielt ut ved objekter, som er det typiske området for å skille mellom fokus på type og eksemplar i norsk.

7.4.2.2 *Generisk referanse*

I prinsippet kan alle engelske substantivformer brukes generisk, som i norsk, men det er forskjeller mellom de to språkene når det gjelder generisk bruk av flertallsformer, da norsk tillater både ubestemt og bestemt flertall, mens bestemt flertall har veldig begrenset bruk i engelsk. Disse formene brukes kun i kjerneløse substantivfraser med nasjonaladjektiv, for eksempel *The English* (Smólska & Rusiecki 1980: 40; C. Lyons 1999: 181).

I engelsk er det ifølge Halmøy (2010: 58) *ubestemt flertall* som er den mest vanlige formen ved generisk referanse, og det er også den eneste formen som kan brukes både taksonomisk, generaliserende og ved referanse til klasse (hhv. 7.4-p, 7.4-q og 7.4-r). Ubestemt artikkel kan kun brukes taksonomisk (7.4-o), mens bestemt entall kun kan brukes ved referanse til klasse (7.4-s), dog sjeldnere enn i norsk, og bestemt flertall brukes sjeldent i engelsk (Halmøy 2010: 71). Ved generisk referanse til ikke-tellelige

massesubstantiv brukes naken entallsform, som i norsk (7.4-t) (Eksempler basert på Halmøy 2010).

7.4-o A lion is a big cat. (Taksonomisk)

7.4-p Lions are big cats. (Taksonomisk)

7.4-q Lions eat other animals. (Generaliserende)

7.4-r Lions are not becoming extinct. (Klasse)

7.4-s The lion is not becoming extinct. (Klasse)

7.4-t Electricity was invented before year 1900. (Ikke-tellelig)

7.4.2.3 *Ikke-referensiell bruk*

Predikativer er et annet område hvor forskjellen mellom engelsk og norsk i bruk av naken entallsform kommer til syne. I norsk skilles det som nevnt, mellom klassifiserende og karakteriserende substantivpredikativer i entall, mens man i engelsk bruker ubestemt artikkel i begge tilfeller med mindre det er snakk om ikke-tellelige substantiv (Halmøy 2010: 77-78). For flertallspredikativer er det ingen forskjell på språkene, siden begge bruker flertallsform. Eksempelene under viser henholdsvis klassifisering og karakterisering i både entall og flertall i engelsk.

7.4-u He is a teacher. / They are teachers. (Klassifiserende)

7.4-v He is a devil. / They are (some) devils. (Karakteriserende)

7.4.3 *Oppsummering engelsk*

Som vi ser, er engelsk og norsk ganske like i distribusjonen av bestemte og ubestemte former, spesielt innen definit referanse, men det foreligger noen forskjeller innen de øvrige referensielle typene, spesielt når det gjelder valget mellom de ulike ubestemte formene, hvor naken ubestemt entallsform er mye mindre brukt i engelsk enn i norsk både ved objekter og predikativer, noe som skyldes at denne formen er forbeholdt ikke-tellelige substantiv i engelsk. I tillegg er det noen forskjeller innen generisk referanse, spesielt knyttet til mer utstrakt bruk av ubestemt flertall i engelsk, og det at man i norsk kan bruke bestemt flertallsform.

Typiske formvalg ved enkle substantivfraser i engelsk vil dermed være følgende:

Definit referanse	Bestemte former
Indefinit referanse	Ubestemte former (artikkelfrase, ubestemt entall/flertall)
	a) Artikkelfrase ved tellelige entallssubstantiv
	b) Ubestemt flertall ved tellelige flertallssubstantiv
	c) Ubestemt entall ved ikke-tellelige substantiv
Generisk referanse	Ubestemte og bestemte former
	a) Artikkelfrase ved taksonomisk
	b) Ubestemt flertall ved taksonomisk, generaliserende og klasse
	c) Bestemte entall ved referanse til klasse
	d) Ubestemt entall ved ikke-tellelige substantiv

7.5 Oppsummering av språklige forskjeller og likheter

Dette kapittelet har beskrevet nærmere forskjeller og likheter mellom språkene norsk, polsk og engelsk når det gjelder substantivets form og referanse, utover den grunnleggende nullkontrasten som har blitt beskrevet tidligere i avhandlingen.

Som vi har sett, skiller polsk seg fra både norsk og engelsk ved å ha den grammatiske kategorien kasus heller enn grammatisk bestemthet. I tillegg har kapittelet beskrevet noen bestemthetskvivalenter i polsk, men som vi har sett, er disse bare delvis sammenfallende med bestemthetsmarkører i +GB-språk. For eksempel så vi at identifiserbarhet gjennom pragmatisk ordstilling først og fremst handler om en form for anaforisk identifiserbarhet, og at polske funksjonsord som ‘ten’, ‘jakis’ og ‘jedem’ bare delvis har samme bruksmønster som de ulike bestemthetsmarkørene i engelsk og norsk.

Framstillingen av polsk innen de ulike referensielle typene, tydeliggjorde en viktig side av nullkontrasten mellom norsk og polsk, nemlig at valg av de ulike bestemte og ubestemte formene i norsk hovedsakelig avspeiler referensielle forhold, mens referanse ikke markeres eksplisitt i polsk, og kasusmarkering i stedet vil være styrende for formvalg innen *alle* de referensielle typene: Mens ikke-referensielle predikativer typisk vil ha instrumentalis, vil det veksles mellom (hovedsakelig) nominativ, akkusativ, dativ og genitiv ved indefinit, definit og generisk referanse.

Selv om den strukturelle korrespondansen mellom norsk og polsk i enkle substantivfraser framstår som mer komplisert og omfattende enn mellom engelsk og norsk, siden det polske bøyningssystemet har langt flere former enn norsk, er polsk og norsk likevel like på den måten at de begge har substantivmorfologi i større grad enn

engelsk. Denne strukturelle likheten mellom polsk og norsk kan potensielt utgjøre en vesentlig forskjell mellom norsk og engelsk som målspråk for polske innlærere. I tillegg innebærer det også en mulig forskjell mellom polske og engelske innlærere av norsk ettersom de polske kan ha en fordel av å ha mer erfaring med substantivmorfologi enn de engelske. Dette er et poeng både Nordanger (2017) og Axelsson (1994) også har framhevet i sine studier av henholdsvis russiske norskinnlærere og polske svenskinnlærere.

Når det gjelder det engelske systemet for bestemthet, korresponderer det i stor grad med det norske både strukturelt og innholdsmessig, men som vi har sett er det forskjeller også. Strukturelt sett har både engelsk og norsk korrespondanse mellom de fem enkle substantivfrasene som blir studert i undersøkelse 2, og den mest framtrede strukturelle forskjellen er at bestemte substantiv har artikkel i engelsk, men bøyning i norsk.

Når det gjelder distribusjonen av formene, er det også stor likhet mellom engelsk og norsk, men noen ulikheter har også kommet fram. De er ikke så framtrede ved definitt referanse, hvor bestemt form er vanlig i både norsk og engelsk, men vi har sett at det innen generisk referanse kan være noen forskjeller i hvilke former som foretrekkes, for eksempel at bestemt flertall har veldig begrenset bruk i engelsk, og at det er forskjeller i hvilke generiske undertyper de ulike formene brukes i. En annen framtrede forskjell som ble nevnt, er knyttet til at bruk av naken ubestemt entallsform er mer begrenset i engelsk enn i norsk, ettersom man i norsk ikke i så stor grad som i engelsk forbeholder ubestemt entallsform til ikke-tellelige substantiv. Dette innebærer også at artikkelfrase er mer brukt i engelsk enn i norsk i visse kontekster, siden norsk kan velge å utelate ubestemt artikkel for å framheve type heller enn eksemplar ved objekter, eller for å uttrykke karakterisering framfor klassifisering ved predikativ.

I hypotesene som presenteres i neste kapittel, er det noen av disse forholdene mellom språkene som vil være mer sentrale: For det første nullkontrasten mellom norsk og polsk, for det andre at substantivmorfologi i polsk brukes til å uttrykke kasus. De kontrastive beskrivelsene fra dette kapitlet er imidlertid også relevante for måten transfereffektene vil bli undersøkt på, spesielt transfereffekten *intergruppe-*

heterogenitet, som skal avdekke forskjeller mellom den polske og den engelske gruppen, og *intralingval kontrast*, som skal skille mellom morsmålsamsvar og målspråkslikhet.

8 UNDERSØKELSE 2: HYPOTESER OG DATAMATERIALE

Undersøkelse 2 har som formål å granske om morsmålet påvirker distribusjonen av bestemte og ubestemte former i innlærerspråket til polske innlærere. I den forbindelse har jeg satt opp to hypoteser som begge sier noe om hvilke språklige forhold i polsk som antas å bli overført til innlærerspråket, og hvordan dette påvirker hvilke former som brukes i hvilke språklige kontekster. Undersøkelse 2 har altså et forskningsspørsmål som retter seg mer mot kvalitative sider av morsmålspåvirkningen enn det undersøkelse 1 hadde: *Påvirker morsmålet distribusjonen av bestemte og ubestemte former i innlærerspråket til polske innlærere?*

De to hypotesene om polsk morsmålspåvirkning er utviklet med utgangspunkt i tidligere forskning på bestemthet og/eller transfer, i tillegg til noen av de kontrastive forholdene mellom norsk og polsk som ble beskrevet i kapittel 7. Begge hypotesene tar utgangspunkt i at det foreligger en nullkontrast mellom polsk og norsk, og at språkene derfor er for ulike til at innlæreren kan foreta noen tverrspråklig identifikasjon mellom strukturer i morsmålet og målspråket, noe som videre fører til en antakelse hos innlærerne om at målspråket fungerer på samme måte som morsmålet (2.3.6). Hva mer spesifikt som antas å bli overført fra morsmålet, og hvordan det manifesterer seg i innlærerspråket, er derimot ulikt for de to hypotesene: For hypotese 1 antas en overføring av pragmatisk ordstilling som signal om identifiserbarhet og en oppfatning av at markering av identifiserbarhet er redundant så lenge ordstillingen følger tema/rema-prinsippet, mens det for hypotese 2 antas en overføring av substantivmorfologi som uttrykk for kasusinnhold. Begge deler antas å ha konsekvenser for hvordan bestemte og ubestemte former brukes i innlærerspråket.

I dette kapittelet vil jeg presentere hypotesene, datamaterialet og den generelle metoden for undersøkelse 2. Først gjør jeg rede for bakgrunnen for de to hypotesene om morsmålspåvirkning hos polske innlærere (8.1). Videre gjør jeg rede for utvalget av tekster fra ASK-korpuset (8.2), utvalget av substantivfraser og hvordan de er søkt fram i korpuset (8.3), og koding av disse frasene (8.4). Metoden for å undersøke hypotesene, for eksempel operasjonaliseringen av de fire transfereffektene, blir til slutt presentert i 8.5. Resultater og analyse for hver av de to hypotesene legges fram i kapittel 9 og 10.

8.1 Bakgrunnen for de to hypotesene om polsk morsmålpåvirkning

Så langt i avhandlingen har jeg presentert og gjort rede for en rekke forhold som kan bidra til utformingen av hypoteser om morsmålpåvirkning hos polske innlærere. Det er både trekk ved polsk i seg selv og forholdet mellom polsk og norsk som er relevante, og jeg vil her gjøre rede for bakgrunnen for begge hypotesene.

8.1.1 Bakgrunn for hypotesene

I kapittel 7 gikk jeg igjennom noen trekk ved substantivfraser og referanse i polsk sammenlignet med norsk. Det ble der presentert noen språklige forhold som er relevante for utformingen av hypoteser om morsmålpåvirkning fra polsk: For det første har vi den generelle og «objektive» nullkontrasten mellom de to språkene, nemlig at norsk har en grammatisk substantivkategori som polsk ikke har, samtidig som polsk har en kategori som norsk ikke har, nemlig kasus. I tillegg til denne grunnleggende forskjellen har jeg vært inne på en *likhet* mellom de to språkene, nemlig at begge legger til rette for fortolkning av substantivfrasers identifiserbarhetsstatus (kjent/ny), selv om de gjør det på ulike måter. Mens norsk uttrykker identifiserbarhet gjennom grammatisk bestemthet, signaliserer polsk det først og fremst gjennom pragmatisk ordstilling, men også delvis gjennom bruk av demonstrativer og pronomener som i noen tilfeller kan brukes på en lignende måte som bestemthetsmarkører i norsk (se 7.3.2).

I kapittel 2.3.3 beskrev jeg forholdet mellom tverrspråklige likheter/forskjeller og transfer, og der omtalte jeg tre kontrastive relasjoner som kan transfer kan være utfall av: *likhetsrelasjoner*, *ulikhetsrelasjoner* og *nullrelasjoner*. En hovedforskjell mellom disse tre er hvorvidt transfer er et resultat av *tverrspråklig identifikasjon*, eller ikke. Om innlærere oppfatter noe i målspråksinnputt som fører til en antakelse om at det er enten likheter eller forskjeller mellom språkene, skjer det en tverrspråklig identifikasjon (likhets- eller ulikhetsrelasjon). Om forskjellene derimot er for store eller likhetene for abstrakte, vil det være mindre sannsynlig at en slik identifikasjon finner sted (nullrelasjon). En mulig konsekvens av at forholdet mellom språkene oppfattes som en nullrelasjon, er at innlærerne ender opp med en generell antakelse om at målspråket fungerer på samme måte som morsmålet (Ringbom & Jarvis 2009: 108).

Om vi ser på forskjeller og likheter mellom språkene norsk og polsk i lys av disse ulike kontrastive relasjonene, mener jeg det er grunnlag for å si at innlærerne kan ende opp med å se på forholdet som en nullrelasjon.³⁵ For det første foreligger det en nullkontrast mellom språkene, der norsk har grammatisk bestemthet, mens polsk har kasus som kategori ved substantiv, i tillegg til numerus, som fins i begge språkene og derfor ses bort fra her. Forskjellene mellom språkene kan antas å være *for store* til at det kan skje noen tverrspråklig identifikasjon mellom kategoriene. For det andre foreligger det en likhet mellom pragmatisk ordstilling og grammatisk bestemthet med tanke på at begge er involvert i å signalisere identifiserbarhet, men det er rimelig å anta at dette framstår som en *for abstrakt* likhet til at det kan skje en tverrspråklig identifikasjon. Dette er særlig fordi likhetene knytter seg til ulike språklige nivå i de to språkene, altså syntaks og morfologi.

Med tanke på at forskjellene virker for store og likhetene for abstrakte, mener jeg at det kan være grunn til å anta at innlærerne oppfatter dette som en nullrelasjon. I slike tilfeller er det en mulighet for at transfer ikke skjer, men det er også, som nevnt over, en mulighet for at innlærerne sitter igjen med en umotivert antakelse om at morsmål og målspråk fungerer på samme måte. Denne typen transfer har jeg tidligere (2.3.6.1) omtalt som *transfer to nowhere* (Kellerman 1995) og *transfer som inaktivt utfall* (Jarvis 2000: 250).

Transfer som inaktivt utfall omtales av Jarvis (2000: 250) som ikke-dynamisk og ikke-reaktiv transfer, og er en type transfer som oppstår som resultat av mangel på tverrspråklig identifikasjon. Etter det jeg forstår, innebærer ikke denne typen transfer noen form for «aktiv» overføring, men heller at noe som ligger til grunn for morsmålet fortsetter å gjøre det på andrespråket, til tross for at de to språkene er ulike. Jarvis (2000:250) omtaler dette som resultat av at det samme begrepssystemet ligger til grunn for både morsmål og innlærerspråk. Denne typen transfer virker å være relativt sammenfallende med Kellermans (1995) *transfer to nowhere*, som også beskriver transfer når forskjellene mellom morsmål og målspråk er for store. De store forskjellene har ifølge Kellerman (1995: 141) to konsekvenser: For det første at måten målspråket

³⁵ Jeg har tidligere forklart at jeg skiller mellom «nullkontrast», som viser til en relativt objektiv sammenligning mellom språkene, og «nullrelasjon», som viser til innlæreres subjektive forståelse av forholdet mellom språkene (se 2.3.3.)

fungerer på, ikke blir oppfattet av innlærerne, og for det andre at innlærerne ikke blir bevisst på at språklige forhold som er sentrale for morsmålet, kan være språkspesifikke og mangle i målspråket. Disse tingene bidrar til at innlærerne antar at målspråket fungerer på samme måte som morsmålet, og kategorier eller perspektiv fra morsmålet fortsetter dermed å gjelde selv om man formulerer seg på et annet språk. Dette skjer gjennom at innlærerne leter etter språklige uttrykk i målspråket som kan opprettholde morsmålet perspektiv eller språklige innhold:

[...] when it comes to verbalizing events in a second language, learners may not look for the perspectives peculiar to that language; instead, they may seek the *linguistic* tools which will permit them to maintain their L1 perspectives (Kellerman 1995: 141).

I sammenheng med nullkontrasten mellom norsk og polsk kan dette bety at det ikke skjer noen tverrspråklig identifikasjon, og at måten grammatisk bestemthet fungerer på, ikke blir lagt merke til. De polske innlærerne vil derimot anta at norsk fungerer på samme måte som polsk, og dermed lete etter språklige midler i norsk som kan uttrykke språklig innhold fra morsmålet.

Om det er slik at innlærerne antar at morsmål og målspråk fungerer på lik måte, er det nødvendig å finne ut *hva* i morsmålet som «fortsetter å virke» på innlærerspråket, og hvilke språklige utslag dette har.³⁶ Både pragmatisk ordstilling og grammatisk kasus er språklige forhold som er sentrale og obligatoriske i polsk, og begge disse kan ses på som kandidater til å bli overført som resultat av en nullrelasjon eller *transfer to nowhere*. Jeg har derfor valgt å koble begge disse språklige trekkene fra polsk til hver sin hypotese om morsmålpåvirkning fra polsk: Den første hypotesen går ut på at polske innlærere overfører et morsmålsperspektiv der pragmatisk ordstilling fungerer som signal om identifiserbarhet, og at det dermed oppfattes som unødvendig å markere identifiserbarhet når pragmatisk ordstilling brukes. Den andre hypotesen handler om at de polske innlærerne overfører et morsmålsperspektiv knyttet til kasus, der kasusinnhold fra morsmålet kommer til uttrykk gjennom norsk substantivmorfologi.

³⁶ Det ligger i kjernen av *transfer to nowhere* at noe «fortsetter å virke»: «There is *transfer to nowhere*, since the blueprint established for the verbal expression of experience **continues to function** regardless» (Kellerman 1995: 141, min utheving). Denne typen transfer er, slik jeg oppfatter det, utenfor det som kan kalles en *overføring* fra morsmål til innlærerspråk, og er kanskje heller en *videreføring*. Jeg vil likevel omtale denne typen transfer som en *overføring*. Kellerman bruker jo også selv uttrykket «transfer».

Morsmålpåvirkningen antas for begge hypoteser å komme til uttrykk i innlærerspråket gjennom at de norske bestemthetsmarkørene tas i bruk, enten for å signalisere brudd på tema/rema-prinsippet, eller for å uttrykke kasusdistinksjoner. Hypotese 1 kobler transfer til fenomenet *kommunikativ redundans*, som jeg var inne på i omtalen av tidligere studier i kapittel 4, og som jeg vil komme tilbake til under. De to hypotesene vil bli nærmere gjort rede for i den videre teksten.

8.1.2 Hypotese 1

I kapittel 4 var jeg inne på flere studier som har berørt fenomenet *kommunikativ redundans* som forklaring på utelatelse av artikler i engelsk innlærerspråk (Young 1996; Jarvis 2002; Trenkic 2009; Trenkic & Pongpairoj 2013). Jeg beskrev i 4.3 hvordan disse studiene så ut til å legge et noe ulikt innhold i fenomenet redundans, og at redundans ble antatt å interagere på ulike måter med transfer. For Trenkic (2009) og Trenkic og Pongpairoj (2013) så det ut til å være snakk om et *målspråksfenomen*, der artikkelbruken i målspråket engelsk er innholdsmessig redundant og pleonastisk, i tillegg til at det er et generelt kommunikativt prinsipp å unngå redundans, altså at man ikke bruker en mer informativ form enn nødvendig (Grice 1975). I disse forskernes tilnærming kommer dessuten redundans kun *i tillegg* til en forklaring om morsmålsbasert preferanse for bruk av naken form, men redundans er ikke i seg selv en form for morsmålpåvirkning (Trenkic 2009: 12; Trenkic & Pongpairoj: 163). Når det gjaldt Young (1996) og Jarvis (2002), hevdet jeg derimot at redundans blir sett på mer som et *innlærerspråksfenomen* som ikke er like tydelig adskilt fra transfer. Morsmålet blir nemlig antatt å påvirke oppfatningen av hvilke språklige kontekster bruk av bestemthetsmarkører er redundant og ikke redundant i, noe som gjør at markering av identifiserbarhet oppleves som redundant når dette går fram av ordstillingen, men ikke ellers.

En sentral konsekvens av ulikheten mellom disse tilnærmingene er at Trenkic (2009) og Trenkic og Pongpairoj (2013) mener at redundansfenomenet bare kan forklare *utelatelse* av bestemthetsmarkører i innlærerspråk, siden markørene alltid er redundante, mens Young (1996) mener redundansfenomenet kan forklare en *variabel* bruk av markører, der de noen ganger utelates og andre ganger brukes. Youngs studie er dermed et mer egnet utgangspunkt for en hypotese om distribusjon av former der både bruk og ikke bruk av markører antas å styres av transfer fra morsmålet. I tillegg er også Youngs

informanter innlærere med et slavisk språk som morsmål, nemlig tsjekkisk og slovakisk. Jeg har derfor valgt å basere meg på Young (1996) for den første hypotesen min, men med noen forskjeller som jeg vil komme tilbake til under.

Hypotese 1 tar altså utgangspunkt i at polske innlærere overfører pragmatisk ordstilling som signal om identifiserbarhet, og med dette også en oppfatning av at identifiserbarhet allerede er signalisert når ordstillingen følger tema/rema-prinsippet. Det antas at innlærerne benytter de språklige elementene som er tilgjengelige i norsk, nemlig bestemthetsmarkører, for å markere eventuelle *brudd* på tema/rema-prinsippet. Dette antas videre å manifestere seg i innlærerspråket gjennom at bestemte former og ubestemt artikkel spesielt blir utelatt i språklige kontekster hvor substantivfrasen er plassert i tråd med tema/rema-prinsippet, mens bestemt form og artikkel derimot vil tas i bruk i språklige kontekster hvor tema/rema-prinsippet brytes. Hypotese 1 er dermed følgende:

De polske innlærerne vil overføre morsmålets bruk av pragmatisk ordstilling som signal om tekstuell identifiserbarhet og en oppfatning om at markering av identifiserbarhet er overflødig når pragmatisk ordstilling brukes. De norske bestemthetsmarkørene vil bli tatt i bruk når identifiserbarhet ikke går fram av ordstillingen.

I tillegg til overføring av ordstilling som identifiserbarhetssignal, innebærer denne typen transfer også nødvendigvis et *morsmålpåvirket identifiserbarhetsbegrep*. Som jeg var inne på i kapittel 3, er det forskjell på hvor skillet går mellom kjente og nye referenter (identifiserbare og ikke identifiserbare) ved grammatisk bestemthet og pragmatisk ordstilling. Denne forskjellen handler i hovedsak om at pragmatisk ordstilling signaliserer om referenten er identifiserbar *gjennom teksten* eller ikke, mens referenten ved grammatisk bestemthet kan være identifiserbar også gjennom utenom-tekstlige forhold (se 3.3 og 7.3.2). Det er naturlig å anta at om polsk morsmål påvirker innlærerspråket på dette området, så er det snakk om identifiserbarhet slik det framkommer ved pragmatisk ordstilling, altså *tekstuell identifiserbarhet*. Når hypotesen over nevner «signalisering av tekstuell identifiserbarhet», er det altså snakk om signalisering av om referenten er nevnt tidligere i teksten eller ikke.

Som nevnt ovenfor, er det viktig å understreke at selv om hypotesen sier at det er *pragmatisk ordstilling* som blir overført fra morsmålet til innlærerspråket, så er ikke

hypotesen rettet mot å undersøke ordstillingen i innlærertekstene. Det er ikke pragmatisk ordstilling i seg selv som er det interessante her, men hvordan pragmatisk ordstilling, når det brukes som et signal om identifiserbarhet, påvirker markeringen av identifiserbarhet ellers i innlærerspråket.³⁷

8.1.2.1 *Hypotese 1: operasjonalisering og prediksjoner*

Hypotese 1 bygger altså i stor grad på Young (1996), men med noen forskjeller som vil gjøres rede for her.

Som jeg var inne på i 4.2.2.1, undersøkte Young (1996) om bruk av pragmatisk ordstilling ble overført fra tsjekkisk til engelsk innlærerspråk, og han undersøkte blant annet om engelske artikler kun ble brukt når plasseringen av en substantivfrase brøt med tema/rema-prinsippet, og om de ble utelatt når identifiserbarheten gikk fram av konteksten gjennom ordstillingen. Young hadde fire delhypoteser om artikkelvalg.³⁸

- 1) **Finale rema** fremmer bruk av nakent substantiv
- 2) **Initiale tema** fremmer bruk av nakent substantiv
- 3) **Finale tema** fremmer bruk av bestemt artikkel
- 4) **Initiale rema** fremmer bruk av ubestemt artikkel

Undersøkelsen til Young viste at dette til en viss grad stemte, men ikke helt. Analysen viste at hypotesene stemte godt med tanke på bruk av nakent substantiv versus substantiv med artikkel, men at den interne fordelingen av ubestemte og bestemte artikler ikke var helt i tråd med hypotesene (Young 1996: 167). *Bestemte* artikler ble brukt generelt ved tematiske substantiv, uavhengig av plassering i setningen, og *ubestemte* artikler så ikke ut til å ha noe å gjøre med verken plassering eller tema/rema (Young 1996: 168). Til tross for at funnene til Young ikke var entydig positive, ser konklusjonen ut til å være at innlærerne har transfer på dette området, og jeg vil derfor basere meg på Youngs hypoteser, men med en noe annen operasjonalisering av *tema/rema* og *setningsposisjon*.

³⁷ Tema/rema-prinsippet er også relevant for ordstillingen i norsk, selv om grammatisk ordstilling står sterkere (Hagen 1998: 260). Om polske innlærere har sterkere tendenser til å organisere setningen i innlærerspråket etter tema/rema enn norske morsmålsbrukere, er interessant i seg selv, men ligger utenfor denne studien.

³⁸ Merk at Young opererer med fire *delhypoteser* som omhandler *finale rema*, *finale tema*, *initiale tema* og *initiale rema* hver for seg (Young 1996: 156), mens jeg opererer med *delprediksjoner* for hver av disse og en overordnet og mer generell hypotese.

Tema/rema

Young (1996: 150) definerer tema og rema (identifiserbarhetsstatus) med utgangspunkt i Bickertons (1975) identifiserbarhetskategori *Assumed Hearer Knowledge* (\pm HK), altså hvorvidt taleren (senderen) mener at tilhøreren kan identifisere referenten (se 4.1). *Tema* defineres som identifiserbar (+HK), noe som innebærer definitte og generiske substantivfraser, mens *rema* defineres som ikke identifiserbar (-HK), som innebærer indefinitte og ikke-referensielle fraser. Dette betyr, så vidt jeg har forstått, at en substantivfrases identifiserbarhetsstatus hos Young er basert på identifiserbarhet slik det framkommer gjennom grammatisk bestemthet. Som jeg var inne på over, ser jeg på det som mer naturlig at innlærerne vil ha en morsmålspåvirket forståelse av identifiserbarhet, og at det dermed er mer relevant å definere tema og rema med utgangspunkt i hvordan det framkommer ved pragmatisk ordstilling, altså om noe er nevnt før eller ikke (tekstuell identifiserbarhet). *Tema* vil dermed være et smalere begrep i min undersøkelse, siden det bare omfatter referenter som er *tekstuelte aktive*, altså nevnt tidligere, mens *rema* får større rekkevidde, siden det inkluderer alt som ikke er tekstuelte aktivert. I tillegg vil generisk referanse i min undersøkelse regnes som aktivert (tema) om det er nevnt før, men som ikke-aktivert (rema) om det ikke er nevnt før. Hos Young var generisk referanse tema uansett siden det regnes som +HK (Young 1996: 150). Jeg vil heller ikke bruke termparet «tema/rema» i særlig grad, men heller omtale det som «identifiserbar/ikke identifiserbar» eller «aktivert/ikke-aktivert». Jeg vil komme tilbake til dette i 8.4.3.

Setningsposisjon

Det andre forholdet som er noe annerledes definert her enn hos Young (1996), er *setningsposisjon*. Young delte i utgangspunktet substantiv inn etter om de sto *initialt*, *medialt* eller *finalt* i setningen (Young 1996: 151), men analysen viste at det ikke var noen forskjell på initial og medial, så disse to ble slått sammen til *ikke-final* (Young 1996: 167).³⁹

For å bestemme hva som kan regnes som initialt, medialt og finalt i min egen undersøkelse, har jeg tatt utgangspunkt i Hagens (1998) feltskjema for hovedsetninger

³⁹ Dette er koblet til at Young bruker VARBRUL, et verktøy for statistisk analyse som vil slå sammen faktorer som ser ut til å ikke påvirke hver for seg (Young 1996: 168).

og leddsetninger. Norske hovedsetninger og leddsetninger har til dels ulik oppbygging, der en sentral forskjell er at subjektet kan stå både før og etter det finitte verbet i hovedsetninger. Dette er knyttet til at norsk er et V2-språk, som innebærer at finitt verb må stå på andre plass i hovedsetninger. Når et annet ledd enn subjektet står i starten av setningen, må derfor subjektet stå *etter* det finitte verbet. I leddsetninger, som ikke har noen V2-regel, vil subjektet stå før det finitte verbet.

Setningene under har alle tre substantivfrasen «*elevene*» som subjekt, i tillegg til de to objektene «*lærerne*» og «*blomster*».

8.1-a Elevene har gitt lærerne blomster.

8.1-b Lærerne har elevene gitt blomster.

8.1-c Hun fortalte at elevene har gitt lærerne blomster.

I disse setningene ser vi hvordan subjektet kan stå *både før og etter* det finitte verbet i hovedsetninger (8.1-a og 8.1-b), mens det må stå *før* i leddsetninger (8.1-c). Objektene «*lærerne*» og «*blomster*» kan enten stå før det finitte verbet, slik «*lærerne*» gjør i 8.1-b, eller *etter* alle verb, som de øvrige objektene i setningene over. I leddsetninger vil objekter ikke kunne stå før det finitte verbet. Samlet sett kan altså nominale ledd som subjekter og objekter stå enten *før* finitte verb, *etter* finitte verb, eller *mellom* finitte verb og eventuelle infinitive verb i hovedsetninger, mens de i leddsetninger enten står før eller etter *alle* verb.

Hagen (1998: 263) opererer i sitt feltskjema med tre ulike setningsposisjoner som nominale ledd i hovedsetninger kan stå i: *forfelt*, *midtfelt* og *slutfelt*.⁴⁰ Forfeltet er posisjonen *før* det finitte verbet, midtfelt er posisjonen *mellom* finitte verb og eventuelle infinitive verb, og slutfeltet er posisjonen *etter* alle verb (Hagen 1998: 263). Hagens feltskjema for hovedsetninger er illustrert under i figur 8-1. Feltskjemaet plasserer egentlig alle typer setningsledd, men er her forenklet for vår sammenheng, og viser derfor bare de nominale leddene subjekt og objekt, i tillegg til verbene.

Figur 8-1 Forenklet fremstilling av Hagens feltskjema for hovedsetninger (1998: 264)

Forfelt	v-plass	midtfelt	V-plass	Slutfelt
subjekt objekt	finitt verb	subjekt	infinitive verb	objekt

⁴⁰ Slutfeltet hos Hagen deles videre inn i en «komplementær plass» og en «friplass» for henholdsvis valensbundne og valensuavhengige setningsledd (Hagen 1998:264). Dette er felles for hovedsetninger og leddsetninger.

Forskjellen på et «felt» og en «plass» er at det første kan ha varierende innhold, for eksempel enten subjekt eller objekt, mens det andre kun kan ha én type innhold, for eksempel finitt verb.

Lesere som er kjent med feltskjemaet til Diderichsen (1946), som Hagens feltskjema er basert på, vil legge merke til at navnene på de ulike setningsposisjonene brukes på en noe annen måte hos Hagen. Jeg vil ikke gå inn i en sammenligning mellom dem her, men vi understreke at termer som «forfelt», «midtfelt» og «slutfelt», som er de relevante posisjonene i denne sammenhengen, skal forstås utelukkende ut fra Hagens feltskjema slik det blir beskrevet her.

I Hagens feltskjema for leddsetninger er de mulige setningsposisjonene for nominale ledd annerledes: Leddsetninger har verken forfelt eller midtfelt, men derimot en preverbal *subjektplass* som kun kan ha subjekt, og et slutfelt hvor øvrige nominale ledd kan stå (Hagen 1998: 269-270).⁴¹ Siden nominale ledd, eller andre ledd, aldri kan stå mellom verbene i leddsetninger, er det ikke noen egen posisjon der tilsvarende hovedsetningens midtfelt.

Figur 8-2 Forenklet fremstilling av Hagens feltskjema for leddsetninger (1998: 269)

Subjunk.plass	subjektplass	midtplass	v-plass	V-plass	slutfelt
(subjunksjon)	subjekt	(setn.adverbial)	finitt verb	infinitverb	objekt

En formålstjenlig måte å skille mellom initiale og finale substantivfraser på er å vurdere om de utgjør *preverbale* eller *postverbale* setningsledd. Substantivfraser som står før det finitte verbet i hovedsetninger og leddsetninger, vil da regnes som initiale, mens substantivfraser som står etter finitte verb og eventuelle infinitte verb, vil regnes som finale.

Siden substantivfraser i leddsetninger enten står før det finitte verbet, eller etter alle verb, vil dette gi en enkel todeling av substantivfrasers plassering i leddsetninger: De kan regnes som initiale (preverbale) eller finale (postverbale). I hovedsetninger er

⁴¹ De tre første posisjonene i leddsetninger omtales riktignok ett sted som «forfelt» av Hagen (1998:270), men denne forståelsen av «forfelt» kommer ikke til syne i de faktiske skjemaene som er brukt i boken. Når jeg bruker «forfelt» om leddsetninger er det kun SUBJ-plassen jeg mener. Dette er beklageligvis noe forvirrende, og en problemstilling som har dukke opp fordi det er noe utydelig i Hagens fremstilling om leddsetninger har forfelt eller ikke, og jeg har valgt å tolke det som at de ikke har det, med utgangspunkt i det faktiske skjemaet som presenteres i boken.

saken mer komplisert fordi en også kan ha substantivfraser i midtfeltet, altså mellom finitt verb og eventuelle infinitte verb. Substantivfraser i midtfeltet vil kunne oppfattes som *mediale* eller *finale*, alt avhengig av hvilken type setning de står i. Eksemplene under illustrerer dette.

8.1-d I går sov elevene.

8.1-e I hele dag har elevene danset vals.

I begge disse setningene er «elevene» subjekter, og begge disse substantivfrasene er plassert i midtfeltet, men mens «elevene» i den første setningen (8.1-d) tydelig er plassert finalt, er «elevene» i den andre setningen (8.1-e) mer *medialt* plassert, i den forstand at den står mellom de to verbene. Denne variasjonen for midtfelt gjør at jeg har valgt å utelate substantivfraser som står i hovedsetningers midtfelt fra undersøkelsen av hypotese 1. Dette gjelder uavhengig av om de har et infinitt verb bak seg eller ikke, og gjelder i praksis alle subjekter som er plassert bak det finite verbet.

For setningsposisjon opererer jeg dermed med posisjonene *forfelt* og *slutfelt*, som representerer henholdsvis *initiale* og *finale* substantivfraser. Selv om leddsetninger ikke har noe forfelt, har jeg likevel funnet det hensiktsmessig å bruke benevnelsen «forfelt» både om hovedsetningens forfelt og leddsetningens *subjekts plass*. For mer om inndelingen i forfelt og slutfelt, se 8.4.2.

Prediksjoner for hypotese 1

Der Young ser på tema/rema og ikke-final/final, vil jeg altså forholde meg til distinksjonene *aktivert/ikke-aktivert* og *forfelt/slutfelt*, og jeg vil også bruke disse termene videre gjennom undersøkelsen.

Når det gjelder den formelle siden av substantivfrasene, skiller Young i sin studie mellom fraser med bestemt artikkel, ubestemt artikkel og nullartikkel (Young 1996: 44). Denne inndelingen passer for engelsk som målspråk, men ikke for norsk, som også har bøyning, derfor har jeg laget tilsvarende inndeling tilpasset norsk som målspråk, der jeg skiller mellom tre typer fraser etter form:

- 1) Ubestemt entall/flertall
- 2) Bestemt entall/flertall
- 3) Artikkel frase

Det er altså tilstedeværelse eller ikke tilstedeværelse av bestemt form eller ubestemt artikkel som styrer denne inndelingen, på samme måte som hos Young. Det er ikke skilt mellom entall og flertall, ettersom det ikke er ventet at verken den samlede oppfatningen av redundans eller aktivering og setningsposisjon hver for seg, skal virke ulikt på om det blir valgt entall- eller flertallsform.

Prediksjonene for hypotese 1 er at de polske innlærerne vil ha følgende assosiasjoner mellom form og kontekst:

- 1. Ubestemt entall/flertall og**
 - a) substantivfraser i forfeltet som har aktivert referent
 - b) substantivfraser i sluttfeltet som har ikke-aktivert referent
- 2. Bestemte entall/flertall og**
substantivfraser i sluttfeltet som har aktivert referent
- 3. Artikkelfrase og**
substantivfraser i forfeltet som har ikke-aktivert referent

I undersøkelsen av hypotese 1 vil jeg se på disse tre delprediksjonene i tekstene skrevet av den polske gruppen, med fokus på at det skal være en *gjensidig* assosiasjon mellom formen og den predikerte konteksten. Hvordan dette gjennomføres vil jeg komme tilbake til i 8.5.1.

8.1.3 Hypotese 2

Hypotese 2 er ikke i like stor grad som hypotese 1 basert på tidligere studier av bestemthet i innlærerspråk, men vokser mer direkte ut av tidligere omtalte tilnæringer til morsmålpåvirkning som berører transfer ved nullrelasjon, altså *transfer to nowhere*, og dessuten elementer av Slobins (1996) hypotese om *thinking for speaking*. Hypotese 2 retter seg inn mot den grammatiske kategorien kasus og denne kategoriens status som obligatorisk i morsmålet, noe som innebærer at den representerer det man kan kalle et grammatisk uttrykt perspektiv (Slobin 1996: 89). Hypotese 2 handler i korte trekk om at polske innlærere overfører dette grammatiske perspektivet til innlærerspråket, og at kasusinnhold vil bli uttrykt ved hjelp av norske bestemthetsmarkører.

Ifølge Lakoff (1987: 320) har grammatiske kategorier en annen status i språk enn kategorier som bare uttrykkes leksikalsk: De brukes automatisk og ubevisst, og er i større grad obligatoriske, og de er en del av vår tenkning på en annen måte enn kategorier

som ikke er grammatikaliserte: «Concepts that are automatic and unconscious are used in thinking and understanding; they are not merely pondered as objects of thought» (Lakoff 197: 320). Slobin (1996: 89) hevder at morsmålet vårt styrer oss til å fokusere på de aspektene av en situasjon som grammatikken framhever, og at dette morsmållspesifikke perspektivet, som han kaller *thinking for speaking*, styrer hva vi trekker ut av en situasjon når vi skal formidle den. Brukere av ulike språk retter i større grad blikket mot og husker bedre elementer som er sentrale for de grammatiske kategoriene morsmålet har, og Slobin (1996: 89) mener at dette *språkpåførte* perspektivet er vanskelig å endre eller legge fra seg når man skal lære et nytt språk:

In brief, each native language has trained its speakers to pay different kinds of attention to events and experiences when talking about them. This training is carried out in childhood and is exceptionally resistant to restructuring in adult second-language acquisition.

Slobin (1996: 91) trekker videre fram visse typer grammatiske kategorier som har den egenskapen at de eksisterer kun *i* og *for* språket, og disse kaller han *thinking for speaking*-kategorier. Eksempler på dette er aspekt, bestemthet og diatese, og Slobin (1996: 89) hevder at slike kategorier kun kan læres gjennom språket, siden de ikke viser til noe utenom-språklig, til forskjell fra for eksempel tall-kategorien.

Slobins utsagn om slike kategorier handler om to sider av språklæringen, da det både sier noe om kategorier i målspråket, som er spesielt vanskelig å lære, og om kategorier i morsmålet, som kan være spesielt vanskelig å endre eller *la være å uttrykke* når man skal lære et nytt språk. Det første gjelder grammatisk bestemthet i norsk, og det andre gjelder kasus i polsk. Begge disse kategoriene har en framtrædende status i ytringsprosessen, da de har å gjøre med forhold som man er nødt til å ta stilling til siden de er obligatoriske.

Tanken bak hypotese 2 er at kasus på grunn av sin obligatoriske og automatiserte status i morsmålet vil overføres til innlærerspråket på den måten at det morsmålpåvirkede grammatiske perspektivet vil fortsette å gjelde når polske innlærere skal formulere seg på norsk. Dette antas å manifestere seg gjennom at de polske innlærerne tar i bruk den substantivmorfologien som er tilgjengelig i norsk, altså bestemte og ubestemte former, for å uttrykke kasusinnhold. Hypotese 2 er dermed formulert som følgende:

De polske innlærerne vil overføre kasus fra morsmålet, og norsk substantivmorfologi vil bli tatt i bruk for å uttrykke kasusdistinksjoner.

Jeg vil videre gå inn i hvordan denne hypotesen kan operasjonaliseres.

8.1.3.1 *Hypotese 2: operasjonalisering og prediksjoner*

Siden hypotese 2 innebærer at språklig innhold som er obligatorisk uttrykt i morsmålet, uttrykkes ved hjelp av den substantivmorfologien som er tilgjengelig i målspråket, er det nødvendig med en presisering av *hvilke funksjoner* i morsmålet som antas å overføres til *hvilke former* i innlærerspråket. Dette er ikke en rett fram sak ettersom bøyningssystemet i polsk er langt mer omfattende enn i norsk.

I kapittel 7.3 kom det fram at polsk har syv ulike kasus, og at formene varierer med en rekke ulike faktorer, for eksempel kjønn og substantivets stamme. Det er også en rekke ulike funksjoner som uttrykkes gjennom de ulike kasusformene, og de fleste kasusformene uttrykker flere innholdsnyanser (Swan 2002: 327). Vi så dessuten at det er stor grad av sammenfall mellom endelsene i de ulike kasusene (Wiese 2011: 105), og at mens kasusene dativ, instrumentalis, lokativ og vokativ gjerne bare har én entallsform og én flertallsform hver, er formvalgene langt mer varierende for nominativ, akkusativ og genitiv. Disse kasusene har i tillegg stor grad av synkretisme og dermed uttrykkes de ofte med de samme endelsene (Wiese 2011: 106).

Substantivmorfologien i norsk uttrykker på sin side bare to kategorier, nemlig tall og bestemthet, og det er snakk om fire ulike substantivformer, som varierer noe med genus (se 7.2.1). Om en polsk innlærer skal uttrykke kasus gjennom de morfologiske markørene i norsk, sier det seg selv at det er altfor få ulike markører tilgjengelig til at det er mulig å dekke hele det innholdsområdet som uttrykkes gjennom polsk kasus. I tillegg er utgangspunktet for hypotese 2 at transfer ikke skjer som resultat av tverrspråklig identifikasjon mellom strukturer i polsk og norsk, men heller at morsmåletts innholdselementer også ligger til grunn for innlærerspråket og dermed finner veien til tilgjengelige språklige midler i norsk. Begge disse forholdene gjør det nødvendig å finne en formålstjenlig todeling av kasusene i polsk, som ikke tar utgangspunkt i form. Jeg har her valgt å ta utgangspunkt i begrepet *markert*, altså skillet mellom *markerte* og *umarkerte* språklige trekk.

Begrepet markerthet ble introdusert på 30-tallet av Praha-strukturalistene Trubetzkoy og Jakobson, men har i ettertid blitt brukt med mange ulike betydninger og formål, og etter hvert også teoriuavhengig (Haspelmath 2006: 25-27). Hva som ligger i distinksjonen umarkert/markert, varierer veldig, og Haspelmath (2006: 27) beskriver hele tolv ulike varianter, for eksempel fonologisk, semantisk, formell, distribusjonell og typologisk markerthet. På grunn av denne variasjonen er det nødvendig å presisere både formålet med og valget av markerthetsvariant for min undersøkelse.

Markerthetens rolle i denne undersøkelsen svarer mest til det som Haspelmath (2006: 42) kaller «metagrammatisk».⁴² Begrepet er her ikke ment å forklare hvorfor og hvordan transfer forekommer, men fungerer først og fremst som et verktøy til å skape en formålstjenlig lingvistisk inndeling av ulike språklige trekk, i dette tilfellet kasus. Markerthet skal bidra til å trekke et hovedskille mellom *markerte* og *umarkerte* kasus, slik at jeg kan formulere et sett med prediksjoner som ikke nødvendigvis gjelder alt innhold som uttrykkes av kasus, men som likevel vil kunne være egnet til å fange opp et eventuelt utslag av slik transfer som hypotesen handler om.

Den markerthetsvarianten som ser ut til å passe best til dette formålet, er det som kalles *funksjonell markerthet* (Dixon 1979).⁴³ Funksjonell markerthet retter seg mot hvilke kontekster ulike kategorier brukes i, og ikke mot formen som brukes. Denne typen markerthet er dermed uavhengig av *formell* markerthet, som viser til om en form er morfologisk umarkert eller markert (Haspelmath 2006: 29).

Et sentralt kriterium knyttet til funksjonell markerthet, er *kontekstuell nøytralisering* (Haspelmath 2006: 39), altså at den umarkerte varianten opptrer i kontekster hvor skillet mellom umarkerte og markerte varianter er opphevet, altså nøytralisert (Greenberg 1966: 29; Haspelmath 2006: 39; Croft 1990: 100). Eksempler på slike kontekster er oppslagsord eller ekstraposisjoner (Falk 1999: 2).

Når det gjelder kasus, er det ifølge Falk (1999: 2, 5), som bygger på Dixon (1979: 72), *nominativ* som er det funksjonelt umarkerte kaset, og her viser både Falk og Dixon til at det er denne formen som brukes i nøytrale kontekster. Funksjonell

⁴² Mer presist omtaler Haspelmath (2006: 42) det som «[...] a metagrammatical concept used by linguists for their convenience.» Det ligger et snev av kritikk i denne formuleringen, slik jeg tolker det, da Haspelmath (2006: 62) mener at termen markerthet er altfor generell og lite gjennomsluktig i en slik sammenheng.

⁴³ Falk (1999) omtaler funksjonell markerthet som «featural», mens Haspelmath (2006) kaller det «distributional».

markerthet er ifølge Falk (1999: 2) i prinsippet uavhengig av om trekket er *formelt* markert eller umarkert, altså om det har endelse eller om det har en såkalt *nullform*. I mange språk er det likevel slik at det er de umarkerte funksjonene som kommer til uttrykk gjennom nullform, og dermed også er *formelt* umarkerte. Dette er tilfelle i for eksempel tyrkisk, der nominativ har nullform, mens øvrige kasus har endelse (Wiese 2011: 102; Nistov 2001: 222). Dette er derimot *ikke* tilfelle i polsk, hvor nullformen opptrer både innen nominativ, akkusativ og genitiv, og nominativ både kan ha nullform og bøyning (se 7.3.1.2). Som Falk (1999: 2) understreker, er det heller ikke et krav at noe er formelt umarkert for at det skal anses å ha umarkert funksjon, og den funksjonelle markertheten må ses på som primær i forhold til den formelle. Han nevner også at dette er spesielt viktig og relevant når man har å gjøre med språk med rike kasussystem (Falk 1999: 2). Nominativ kan dermed regnes som funksjonelt umarkert i polsk og de fleste andre nominativ/akkusativ-språk, mens de øvrige kasusene regnes som markerte (Falk 1999: 5-6).

Basert på denne inndelingen i markerte og umarkerte kasus, vil jeg ta utgangspunkt i skillet mellom *nominativ* og *ikke-nominativ* for å formulere prediksjoner for hypotese 2. Om hypotesen stemmer, vil bestemthetsmarkørene i norsk bli tatt i bruk for å markere dette skillet innen kasus: Ubestemte former vil brukes i språklige kontekster som er assosiert med nominativ i morsmålet, og bestemte former vil tas i bruk i kontekster som er assosiert med ikke-nominativ. Siden nominativ stort sett brukes ved subjekter i polsk, med noen unntak som jeg ser bort ifra her, gir dette en foreløpig prediksjon om at subjekter vil ha ubestemt form, mens ikke-subjekter vil ha bestemt form i innlærerspråket til de polske innlærerne.

Subjekt og ikke-subjekt i tidligere studier

At bestemthetsmarkører brukes til å skille mellom subjekter og ikke-subjekter, er noe som tidligere har vært trukket fram av Parrish (1987) og Nistov (2001), og helt kort også av Jarvis (2002). Parrish (1987) studerte språket til én japansk innlærer og fant at innlæreren i en tidlig periode assosierte defintitte subjekter med naken form (ø), mens defintitte ikke-subjekter ble assosiert med bestemt form:

There does, indeed, appear to be an initial tendency to use \emptyset for subject position [+SR][+HK] NPs and to use *the* for predicate position [+SR][HK] NPs, a tendency that gradually changes until Time 8, when it essentially disappears. (Parrish 1987: 368)

Nistov (2001) har i en mindre delundersøkelse sett på det referensielle uttrykket til subjekter og ikke-subjekter, mer spesifikt objekter, i tekstene til tyrkiske norskinnlærere. Utgangspunkt var en antakelse om at disse innlærerne på grunn av morsmålet ville ha mer bruk av bestemt form ved objekter enn ved subjekter. Tyrkisk er som nevnt over et av de språkene hvor nominativformen faktisk har nullform, mens de øvrige kasusene har bøyninger. Nistov (2001: 225) finner at 8 av 9 forekomster av nakent substantiv står som subjekt, mens bare én forekomst er objekt, noe hun mener kan indikere at syntaktisk funksjon spiller en rolle for valg av form, men hun understreker at det ikke er noe sterkt bevis, og at bestemte former gjerne brukes ved både subjekter og objekter.

Nistov (2001: 222) viser i forbindelse med dette til Slobins *thinking for speaking*, og forklarer følgende:

[...] there would be greater probability of transfer from the learners' L1 in categories serving only as "thinking for speaking", since these are more difficult to "retrain", than in categories that are directly reflected in one's dealings with the real world.

Nistov mener altså at det er større sjanse for transfer av morsmålskategorier som kan regnes som del av morsmålet *thinking for speaking*, siden disse kategoriene er spesielt vanskelig å endre på. Nistov (2001: 222) foreslår derfor at dette fører til at bestemthetsendelse oftere utelates ved subjekter, som har nullform i tyrkisk, mens endelsene oftere brukes ved objekter, som også har endelser i tyrkisk.⁴⁴

Også Jarvis (2002: 411) foreslår markering av subjekt versus ikke-subjekt som en alternativ forklaring til sine funn, fordi disse trekkene i hans materiale sammenfaller med kategorier som viste seg å være sentrale for formvalg, for eksempel emne/kommentar, men Jarvis går ikke nærmere inn på dette.

⁴⁴ Også Ekiert (2010: 147) har knyttet bruk av bestemthetsmarkører i innlærerspråk til *thinking for speaking*, men da i form av identifiserbarhet og informasjonsorganisering i setningen, og ikke kasus.

Avgrensning av materialet for hypotese 2: subjekter og objekter

I likhet med Nistov (2001) vil jeg begrense ikke-subjekter til objekter, og dermed undersøke hypotese 2 i et materiale som bare består av *subjekter* og *objekter*. Det er ulike grunner til dette, men hovedsakelig er det fordi en del av de polske kasusene som er assosiert med de andre syntaktiske funksjonene, allerede er ekskludert fra materialet av andre grunner. For det første har jeg ikke tatt med ikke-referensielle substantivfraser i materialet, og som jeg har vært inne på, dekker dette i stor grad predikativer, en syntaktisk funksjon som først og fremst markeres med *instrumentalis* i polsk (Swan 2002: 363). Jeg har dessuten ikke tatt med substantivfraser som er utfylling i preposisjonsfraser, noe som i stor grad gjør at lokativ heller ikke blir relevant. Jeg sitter da igjen med nominativ, akkusativ og genitiv, hvorav nominativ forbindes med subjekt og akkusativ med objekt, men ved negasjon brukes også genitiv ved objekter (Swan 2002: 333). Objekter forbindes dermed både med akkusativ og genitiv, men begge disse må regnes som funksjonelt markerte sammenlignet med nominativ.

Dette som er nevnt over, gjør at jeg i undersøkelsen av hypotese 2 vil avgrense *ikke-subjekter* til bare *objekter*. Subjekter og objekter utgjør for øvrig den aller største delen av de syntaktiske funksjonene som er representert i materialet, til sammen 83 % av alle substantivfrasene, og fordelingen av subjekter og objekter ligger jevnt rundt 40-45 % for begge funksjoner både for den polske, engelske og norske gruppen. Det at de er de aller mest frekvente funksjonene i materialet, er også en god grunn til å begrense seg til disse to funksjonene. De mer spesifikke prediksjonene for hypotese 2 er dermed at subjekter vil assosieres med ubestemte former (entall og flertall), mens objekter vil assosieres med bestemte former (entall og flertall). De polske innlærerne ventes altså å ha assosiasjoner mellom følgende former og språklige kontekster:

1. Ubestemt entall/flertall og subjektskontekster
2. Bestemte entall/flertall og objektskontekster

Hypotesen retter seg mot hva substantivmorfologi brukes til, og bruk av artikkelfraser er derfor ikke relevant, siden artikler ikke er morfologiske uttrykk for bestemthet. Artikkelfraser inngår dermed ikke i prediksjonene, og er også utelatt fra materialet for hypotese 2.

8.1.4 Hypoteser og prediksjoner oppsummert

Over har jeg presentert to hypoteser om hvordan morsmålet kan påvirke valg av substantivform i innlærerspråket til polske innlærere i ulike språklige kontekster. Hypotesene tar utgangspunkt i en transfer av typen *transfer to nowhere*, der språklige strukturer i norsk tas i bruk for å uttrykke et språklig innhold som er sentralt i morsmålet: Pragmatisk ordstilling vil på samme måte som i morsmålet signalisere om en referent er identifiserbar eller ikke, men bestemthetsmarkører vil tas i bruk når ordstillingen ikke følger tema/rema-prinsippet. Substantivmorfologi vil på samme måte som i morsmålet uttrykke kasisinnhold, og det er bestemthetsmarkørene som får denne rollen.

Prediksjonene for de to hypotesene er at det vil være assosiasjoner mellom visse former og kontekster i innlærerspråket til de polske innlærerne, og disse går fram av tabellen under.

Tabell 8-1 Form/kontekst-assosiasjoner for de to hypotesene

Hypotese 1	Form	Kontekst
	Ubestemt entall/flertall	Aktiverte forfelt/ ikke-aktiverede slutfelt ⁴⁵
	Bestemt entall/ flertall	Aktiverte slutfelt
	Artikkelfrase	Ikke-aktiverede forfelt
Hypotese 2	Form	Kontekst
	Ubestemt entall/flertall	Subjekter
	Bestemt entall/ flertall	Objekter

Hypotese 1 tilsier at bestemt form og ubestemt artikkel kun brukes når substantivfrasen ikke er plassert i tråd med tema/rema-prinsippet, mens de ellers vil utelates. Hypotese 2 tilsier at ubestemte former brukes i språklige kontekster som assosieres med nominativ i polsk, altså subjekter, mens bestemte former brukes i språklige kontekster for ikke-nominativ, her avgrenset til objekter.

Analyse og resultat for hypotese 1 og 2 vil bli presentert i henholdsvis kapittel 9 og 10. Videre i dette kapittelet vil jeg først se på datamaterialet og utvalg av tekster og substantivfraser for undersøkelse 2, og deretter presenterer jeg den generelle metoden for undersøkelsen, altså de metodiske grepene som er felles for begge hypotesene.

⁴⁵ Se fotnote 49 for omtale av formuleringsmåten «aktivert forfelt» o.l.

8.2 Datamateriale – utvalg av tekster fra ASK

ASK har tilgjengelig 200 tekster skrevet av polske og 200 av engelske innlærere, og alle disse var inkludert i undersøkelse 1, men jeg har for undersøkelse 2 avgrenset materialet for å kunne kontrollere for variablene *testtype* og *ferdighetsnivå*. Utvalget mitt har derfor kun tekster skrevet til Språkprøven, og av disse har jeg kun inkludert tekster som er vurdert til å ligge på B1-nivå. Det er i alt 100 tekster som fyller disse kriteriene; 41 av dem skrevet av polske innlærere, og 59 av engelske. For å gjøre antallet mer likt, har jeg redusert den engelske gruppen ytterligere ved hjelp av variabelen *hjemland*, slik at kun engelske innlærere med hjemland Storbritannia og USA er tatt med.⁴⁶ Denne avgrensningen gjorde at de to innlærergruppene ble like med hensyn til antall tekster, med 41 tekster hver. Tabellen under gir en oversikt over hvordan antall tekster ble redusert ved hjelp av de ulike variablene.

Tabell 8-2 Oversikt over antall tekster i engelsk og polsk delkorpus

Antall tekster:	S1= Engelsk	S1= Polsk	Totalt
Totalt i ASK	200	200	400
Testtype (Språkprøven)	100	100	200
+ Ferdighetsnivå (B1)	59	41	100
+ Hjemland	41	41	82

I tillegg til innlærertekstene har jeg brukt det norske kontrollkorpuset i ASK, som består av tekster skrevet av personer med norsk som morsmål. Også her har jeg redusert de 100 tilgjengelige tekstene fra Språkprøven til 41, men siden ferdighetsnivå ikke er en aktuell variabel for morsmålsbrukere, reduserte jeg antallet ved velge de 41 første tekstene etter nummereringen av tekster i ASK (tekstnr. s2001 - s2041).

Tabell 8-3 (neste side) gir en oversikt over størrelsen på delkorpuset for undersøkelse 2, både antall ord og tekstlengde, delt inn etter morsmålsbakgrunn. Dette delkorpuset består totalt av 33444 ord, tilnærmet jevnt fordelt på de tre gruppene (pol. 11642; eng. 10849; nor. 10953). Den gjennomsnittlige tekstlengden er det heller ingen stor forskjell på i de tre gruppene, men det polske gjennomsnittet ligger på cirka 20 ord mer enn det engelske og norske (pol. 284; eng. 265; nor. 268). Én innlærer i det

⁴⁶ Denne avgrensningen ble kun gjort fordi det ga nøyaktig det antallet tekster jeg hadde behov for å redusere med, ikke fordi hjemland er ventet å spille noen rolle for resultatet.

polske korpuset stikker seg ut ved å ha skrevet en tekst på kun 99 ord, ellers ligger tekstlengden i hele materialet på mellom 151 og 473 ord, altså et relativt stort spenn.

Tabell 8-3 Antall ord og gjennomsnittlig tekstlengde materialet for undersøkelse 2

Morsmål=	Polsk (N=41)	Engelsk (N=41)	Norsk (N=41)	Tot. (N=123)
Antall ord totalt	11642	10849	10953	33444
Tekstlengde				
Gjennomsnitt:	284 ord	265 ord	268 ord	272 ord
Min/maks:	(99)185-473 ⁴⁷	174-422	151-390	(99) 151-473

8.2.1 Tekstene

Siden alle tekstene er fra samme språktest, er tekstene skrevet under relativt like forhold, og man kan til en viss grad anta at alle har skrevet relativt like typer tekster, selv om det nok også er en god del variasjon med tanke på hvor de ligger sjangermessig (se 5.2.2.1). Også temaet for tekstene varierer en god del. Ved Språkprøven har kandidatene hatt et valg mellom tre ulike oppgavetema, men tekstene er hentet fra ulike testavviklinger, så kandidatene har skrevet om en rekke ulike tema, noe som avspeiles i mitt materiale. Det er kun noen få av disse temaene som er representert i både den engelske og den polske gruppen, nemlig «mobiltelefoner», «nyheter», «vennskap» og «viktige verdier i livet». Disse fire temaene utgjør totalt 37 av tekstene, altså litt under halvparten. Temaet «mobiltelefoner» er mest balansert i de to gruppene (pol. 10; eng. 9). Noen av temaene er bare representert i den ene gruppen, og her er det spesielt temaet «bomiljø» som skiller seg ut, siden det er skrevet av hele 20 av de engelske kandidatene, men ingen av de polske.

De norske kontrolltekstene er skrevet med utgangspunkt i oppgavesett som ligner, men som ikke har de samme temaene som de ekte testene. Av de 41 norske kandidatene er det fire ulike tema som er valgt, hvorav temaet «Sunnhet og helse» har vært det mest populære (21), etterfulgt av «Penger og andre verdier» (12). Vedlegg 2 gir en detaljert oversikt over hvordan de ulike temaene er fordelt i de tre gruppene.

⁴⁷ Teksten på kun 99 ord stikker seg ut fra resten av tekstene. Denne er derfor satt i parentes i tabellen.

8.2.2 Tekstskriverne

Personene som har skrevet tekstene i delkorpuset for undersøkelse 2, varierer en god del med tanke på personvariabler som *alder*, *utdanning*, *kjønn*, *oppholdstid* osv. Det utvalget jeg har gjort, vil til en viss grad avspeile hva slags engelske og polske innlærere som har tatt testen, og tendensen for mitt spesifikke delkorpus er at det er overvekt av innlærere som er *høyt utdannet* og som har relativt *kort oppholdstid* i Norge (mindre enn to år), men det fins også flere som har lavere utdanning og lengre oppholdstid. Det er dessuten en tydelig overvekt av kvinner i det polske delkorpuset. En fullstendig oversikt over personvariabler for de ulike tekstskriverne ligger i vedlegg 1.

Alderen strekker seg fra 22-48 år i den polske gruppen, 17-63 år i den engelske, og 17-69 år i den norske, men hoveddelen i alle de tre gruppene er mellom 20 og 40 år. Når det kommer til kjønn, er de to innlærergruppene mindre like. Den engelske gruppen har dobbelt så mange kvinner som menn (♀ 27; ♂ 14), men overvekten av kvinner er langt større i den polske gruppen (♀ 36; ♂ 4). Den norske gruppen har ganske lik kjønnsfordeling som den engelske (♀ 25; ♂ 16), altså har alle tre gruppene overvekt av kvinner. Med tanke på utdanning er gruppene ganske like, da alle tre gruppene har en klar overvekt av kandidater med utdanning fra høyskole eller universitet (pol. 32; eng. 27; nor. 29), og bare en liten andel har videregående skole som høyeste utdanning.

Størstedelen av begge innlærergruppene har tatt testen etter å ha vært mindre enn to år i Norge, og utover dette har en liten andel vært i landet mellom to og fem år, og bare syv stykker i mer enn fem år. Innlærergruppene er her ganske like i fordelingen, selv om det er en litt større andel av de engelske med oppholdstid på ett til to år, mens det for de polske er en høyere andel med oppholdstid på mindre enn ett år.

Det er ikke oppgitt hvilke andre andrespråk tekstskriverne i ASK har, men for de fleste polske innlærerne fins det informasjon om ferdighetsnivå for engelsk (tabell 8-4, neste side). Den største delen av de polske innlærerne oppgir å være på grunnivå (15), mens syv er på høyere nivå, åtte på mellomnivå og fire er på begynnernivå. Bare seks polske innlærere oppgir å ikke ha noen engelskkunnskaper. Dette er imidlertid en litt vanskelig variabel å fortolke ettersom det er snakk om selvrapporing, og dermed er det vanskelig å vite hva innlærerne har lagt i de ulike nivåene.

Tabell 8-4 Selvrapporterte engelskferdigheter hos de polske informantene

Oppgitte engelskferdigheter	Polsk
Høyere nivå	7
Mellomnivå	8
Grunn-nivå	15
Begynnernivå	4
Ikke noe	6
Ikke oppgitt	1
SUM	41

8.3 Datamateriale – utvalg av substantivfraser og søk i ASK

Jeg vil nå presentere den språklige avgrensningen av datamaterialet, altså *enkle substantivfraser*, og jeg vil se på både hva slags typer substantivfraser som er inkludert, og hvordan de er søkt fram i ASK. De typene enkle substantivfraser som er søkt fram fra korpuset, kan deles i fem ulike typer med tanke på form (se også 7.1):

- 1 Ubestemt entall ('bil')
- 2 Ubestemt flertall ('biler')
- 3 Bestemt entall ('bilen')
- 4 Bestemt flertall ('bilene')
- 5 Ubestemt artikkel + ubestemt entall ('en bil')

Type 5 (artikkel + ubestemt entall) vil fortsatt bli omtalt som «artikkelfrase», som tidligere i avhandlingen. Type 1, 2 og 5 vil noen ganger også omtales med fellesbenevnelsen «ubestemte former», mens type 3 og 4 vil omtales som «bestemte former», der det ikke er hensiktsmessig å være mer presis. Autentiske eksempler på slike substantivfraser fra ASK fins i tabell 8-6 lenger ned.

8.3.1 Søkemetoder: grunnsøk og modifisering

For å søke fram enkle substantivfraser har jeg måttet kombinere fem ulike korpussøk, ett for hver av de fem formtypene, og deretter manuelt gå igjennom treffene for å luke ut feiltreff og treff som ikke er interessante for undersøkelsen. De fem ulike søkene vil bli omtalt som *grunnsøk 1-5*.

Alle de fem grunnsøkene tar utgangspunkt i den automatiske taggingen som er gjort av ordklasser og grammatiske kategorier i ASK. Hovedstrukturen i grunnsøk 1-4 er å søke etter alle substantiv som er bestemte eller ubestemte, og har entall eller flertall, inkludert en spesifisering av hvilke grammatiske trekk som ikke skal forekomme *før* og *etter* substantivet, for eksempel adjektiv, preposisjon, determinativ. Dette er for å kunne

utelukke modifiserte substantivfraser i søkene. For grunnsøk 5 (artikkelfraser) har jeg søkt etter substantiv som har lemmaene ‘en’, ‘ei’ eller ‘et’ like foran seg.

Tabell 8-5 En forenklet framstilling av grunnsøk 1-5.

	[posisjon foran]	[treffposisjon]	[posisjon bak]
Grunnsøk 1	Skal ikke inneholde: Adjektiv, determinativ, preposisjon, preposisjons- utfylling, adverb eller substantiv	[subst ent be]	Skal ikke inneholde: Determinativ
Grunnsøk 2		[subst ent ub]	
Grunnsøk 3		[subst flt be]	
Grunnsøk 4		[subst flt ub]	
Grunnsøk 5		[subst]	

Tabell 8-5 (over) gir en forenklet framstilling av grunnsøk 1-5. Andre kolonne fra venstre viser hva som skal forekomme eller ikke forekomme i posisjonen *før* substantivet, tredje kolonne viser formen på selve substantivet (treffposisjon), og fjerde kolonne viser hva som ikke skal forekomme i posisjonen *etter* substantivet. Se vedlegg 3 for de fullstendige søkeuttrykkene til hvert og ett av grunnsøkene.

Disse fem grunnsøkene var imidlertid ikke tilstrekkelige for å søke fram utelukkende enkle substantivfraser, enten på grunn av at en av posisjonene inneholdt feilkategoriseringer fra den automatiske taggerens sin side, eller fordi søket ikke utelukket tilstrekkelig med modifiserende adledd. Konkordansene med treff inneholdt altså både treff som ikke var substantivfraser, og substantivfraser som ikke var umodifiserte. I tillegg ønsket jeg å utelate en del faste uttrykk og idiomatisk bruk av enkle substantivfraser, og det var også noen få substantiv som var feilkategorisert for bestemthet og tall. For å luke vekk uønskede treff, tok jeg i bruk funksjonen i ASK som tillater egen annotasjon av hvert treff i konkordansen. Her la jeg inn egne annotasjonskoder for de treffene som ikke var relevante, og utførte alle grunnsøkene på nytt, denne gangen med en ekstra spesifisering av at alle treff som var kodet som irrelevante i annotasjonsfeltet, skulle utelates. Etter dette ble alle treffene på grunnsøk 1-5, sammen med person- og tekstvariabler, lastet ned og bearbeidet i Excel og SPSS.

Setningene i tabell 8-6 (neste side) er faktiske eksempler på hver og en av formene som ble søkt fram. I parentes står identifikasjonsnummeret til teksten setningen er hentet fra, såkalt *tekst-id*.

Tabell 8-6 Eksempler på de fem typene enkle substantivfraser

Formtype	Eksempel
Ubestemt entall	Jeg har kjøpt <u>mobiltelefon</u> i 2 år. (s0763)
Ubestemt flertall	Det er mange fordeler som <u>mobiltelefoner</u> har ... (s0770)
Bestemt entall	Nesten alle mennesker har <u>mobiletelefonen</u> i lomma eller veska. (s0791)
Bestemt flertall	Da jeg tenker at vi trenger ikke <u>mobiltelefonene</u> ... (s0978)
Artikkelfrase	Jeg tror at det er bra å ha <u>en mobiltelefon</u> . (s0767)

I treffene for artikkelfraser var det noen få som strukturelt sett ikke var i tråd med målspråksnormen fordi de kombinerte ubestemt artikkel med bestemt form eller flertallsform av substantivet. Det var to slike tilfeller: «en velferdsstaten» og «en utledninger», men begge disse var ikke-referensielle predikativer, og er derfor ikke inkludert i materialet.

8.3.2 Andel av enkle substantivfraser

Antall og andel av substantiv og enkle substantivfraser (ESF) i delkorpuset for undersøkelse 2 er oppgitt i tabell 8-7 (neste side). I gjennomsnitt utgjør ordklassen substantiv rundt 20 % av alle ord i delkorpuset, uten nevneverdige forskjeller mellom de tre gruppene (pol. 18,5 %; eng. 19,4 %; nor. 21,3 %). Den relative andelen av enkle substantivfraser i forhold til antall substantiv totalt, er ganske jevn i de to innlærergruppene (pol. 19 %; eng. 18,2 %), mens man i de norske tekstene finner en noe større andel enkle substantivfraser (23,5 %).⁴⁸

Det er ikke forventet at fordelingen av former er lik i de tre gruppene, men det er noen tall som stikker seg ut her. Den polske gruppen stikker seg ut ved å ha en noe lavere andel bestemt entall enn de to andre gruppene (pol. 14 %; eng. 17 %; nor. 20 %), i tillegg til en høyere andel bestemt flertall (pol. 13 %; eng. 7 %; nor. 8 %). For ubestemte former, ser vi at det er den engelske gruppen som stikker seg mest ut, med en god del lavere andel av ubestemt entall (pol. 36 %; eng. 23 %; nor. 39 %), og en tilsvarende større andel artikkelfraser (pol. 10 %; eng. 21 %; nor. 8 %).

⁴⁸ Jeg har valgt å måle andelen av enkle substantivfraser i forhold til antall *substantiv* i delkorpuset, ikke antall *substantivfraser*. Dette er fordi tallene mine er basert på søk etter kun substantiv i delkorpuset, noe som i stor grad vil sammenfalle med antall substantivfraser, men ikke helt, ettersom noen substantivfraser kan være kjerneløse og mangle substantiv.

Tabell 8-7 Andelen av substantiv og enkle substantivfraser fordelt på de tre gruppene

S1=	Engelsk		Polsk		Norsk		Totalt	
Substantiv	2110	(19,4 %)	2153	(18,5 %)	2333	(21,3 %)	6596	
ESF	401	(18,2 %)	506	(19,0 %)	425	(23,5 %)	1332 (20,2 %)	
Bestemte former								
- entall	70	17,4 %	71	14,0 %	86	20,4 %	227	17,1 %
- flertall	27	6,7 %	64	12,6 %	34	8,1 %	125	9,4 %
Ubestemte former								
- entall	94	23,3 %	182	36,0 %	165	39,1 %	441	33,1 %
- flertall	129	32,0 %	140	27,7 %	102	24,2 %	371	27,9 %
- artikkel	83	20,6 %	49	9,7 %	35	8,3 %	167	12,5 %

8.4 Koding av kontekster

I omtalen av hypotesene tidligere i kapittelet går det fram at det er visse egenskaper substantivfrasene i materialet må klassifiseres etter for å undersøke de to hypotesene: *syntaktisk funksjon*, *setningsposisjon* og *tekstuell aktivering*. Jeg har i tillegg klassifisert substantivene etter *referensiell type* og *tellelighet*. Under beskrives hver av disse klassifiseringsområdene med eksempler. Med unntak av syntaktisk funksjon, som har vært direkte tilgjengelig gjennom ASK, er klassifiseringen gjort manuelt ved å legge inn egne koder i annoteringsfeltet i konkordansen.

8.4.1 Syntaktisk funksjon

Datamaterialet for undersøkelse 2 består hovedsakelig av subjekter og objekter, men til sammen 17 % av substantivfrasene har andre syntaktiske funksjoner eller egenskaper. Tabell 8-8 (under) viser eksempler på *subjekt*, *direkte objekt*, *indirekte objekt*, *predikativ* og *adverbial* fra materialet. I tillegg har jeg lagt til substantivfraser som står som *løs substantivfrase*, altså fraser som ikke er en del av en setning, eller som står i ekstraposisjon.

Tabell 8-8 Eksempler på de ulike syntaktiske funksjonene

Synt.funk.	Eksempel	
Subjekt	Mobiltelefoner tilbyr også sikkerhet i mange måter.	(s0329)
Objekt	Jeg liker mobiltelefoner , men jeg tror at folk bruker dem altfor mye.	(s0347)
Ind. obj	... det viktig å ha steder i nærheten som gir barn plass å leke	(s0302)
Predikativ	En av dem er murder allerede i hjemlandet sitt.	(s0793)
Adverbial	I @sted måtte jeg kjøre en time til jobb og det var alltid en kø.	(s0187)
Løs SF	Mobiltelefoner ; du kan ikke tenke om livet uten dem.	(s0785)

Klassifiseringen av syntaktisk funksjon har stort sett vært tilgjengelig gjennom ASK, men den syntaktiske taggeren til korpuset er ikke alltid like pålitelig, enten på grunn av feilklassifisering eller på grunn av flere alternative forslag. Slike feil og tvetydigheter har blitt manuelt rettet opp etter at treffene ble lastet ned.

Materialet består i hovedsak av subjekter og objekter, som til sammen utgjør 84 % av alle de enkle substantivfrasene. Den polske gruppen skiller seg ut med å ha litt høyere andel objekter enn de to andre gruppene (pol. 49 %; eng. 44 %; nor. 41 %). Merk at tallene ikke sier noe om antall objekter og subjekter generelt i tekstene, men antallet som er uttrykt med enkle substantivfraser.

Oversikten over viser at det i materialet fins enkle substantivfraser med ulike syntaktiske funksjoner, men ikke alle disse er inkludert i undersøkelsen. Det gjelder generelt for begge hypoteser at ikke-referensielle substantivfraser ikke er inkluderte, (predikativer og adverbialer), og det samme gjelder løse substantivfraser. For hypotese 2 ser jeg som jeg har vært inne på tidligere, bare på subjekter og direkte objekter, mens for hypotese 1 har jeg inkludert både subjekter, direkte og indirekte objekter.

8.4.2 Setningsposisjon

I forbindelse med hypotese 1 må substantivfrasene i materialet blant annet klassifiseres etter hvor i setningen de er plassert, altså initialt eller finalt, som sammen med *aktivering* er en del av vurderingen av om markering er redundant eller ikke. Som jeg var inne på tidligere (8.1.2.1), har jeg valgt å operasjonalisere dette som *forfelt (initial)* og *slutfelt (final)*, der «forfelt» i denne undersøkelsen viser til både hovedsetningers forfelt og leddsetningers subjekt plass, som begge er plassert før det finitte verbet.

I tillegg til dette har hovedsetninger *midtfelt*, en posisjon mellom finitt verb og eventuelle infinitte verb, men fordi det vil variere om midtfeltet i praksis står medialt eller finalt, har jeg som tidligere nevnt valgt å utelate disse substantivfrasene fra materialet for hypotese 1.

En ytterligere problemstilling knyttet til setningsposisjoner og V2-regelen i norsk er at det er et utbredt avvik i innlæreres norsk å plassere subjektet før det finitte verbet selv om det står et annet ledd i forfeltet. Dette fenomenet omtales ofte som *underinversjon*. Eksempel 8.4-a viser et eksempel på dette hentet fra materialet:

8.4-a Hvis jeg har bare litt uteplass **balkong** er nok. (s0295)

Slike substantivfraser forekommer også i materialet mitt, og disse har jeg valgt å klassifisere som *initiale*, siden de er preverbale. Subjekter som er korrekt plassert etter det finite verbet, vil derimot være utelatt siden disse står i midtfeltet.

Tabell 8-9 (under) viser eksempler på ulike setninger der en substantivfrase står initialt/forfelt eller finalt/slutfelt, i tillegg til et eksempel på midtfelt. Eksemplene viser kun hovedsetninger og leddsetninger, men jeg har også fulgt den samme praksisen for initial/final i andre innføyde setninger, for eksempel relativsetninger. Når det gjelder infinitivskonstruksjoner, har disse ikke noe subjekt og dermed ikke noe forfelt, men kun slutfelt.

Tabell 8-9 Eksempler på de ulike setningsposisjonene

Posisjon	Eksempel
Initial	<i>Forfelt i hovedsetning:</i> Mobiltelefoner tilbyr også sikkerhet i mange måter. (s0329)
	<i>Subjekt i hovedsetning med underinversjon:</i> Hvis jeg har bare litt uteplass balkong er nok. (s0295)
	<i>Subjekt plass i leddsetning:</i> Det er veldig hyggelig når naboer tar en kopp kaffe sammen, ... (s0189)
Finalt	<i>Slutfelt i hovedsetning:</i> Jeg liker mobiltelefoner , men jeg tror at folk bruker dem altfor mye. (s0347)
	<i>Slutfelt i leddsetning:</i> Selv om jeg eide en mobiltelefon for 8 år siden, ... (s0738)
	<i>Midtfelt i hovedsetning:</i> Nå er mobiltelefoner bedre enn før flere år siden (s0749)

I gjennomgangen av undersøkelse 2 vil det bli gitt flere eksempler på forfelt og slutfelt innen ulike typer referanse, da også vurdert etter om substantivfrasenes referenter er tekstuelte aktivert eller ikke.

8.4.3 Tekstuell aktivering

I forbindelse med hypotese 1 er også klassifisering av tekstuell aktivering relevant for vurderingen av kommunikativ redundans. Tekstuell aktivering er en operasjonalisering av om en referent er identifiserbar eller ikke (tema/rema). Dette er fordi det er tekstuell identifiserbarhet som ligger til grunn for hva som er tema og rema ved pragmatisk ordstilling, som igjen er styrende for om det går fram av ordstillingen om referenten er identifiserbar eller ikke for hypotese 1.

Alle referensielle fraser har blitt kodet for om referentene er tekstuert aktiverte eller ikke, altså hvorvidt referenten er nevnt tidligere i teksten. I vurderingen av tekstuell aktivering er det hovedsakelig to forhold jeg har vurdert: For det første det Givon (1983: 13) kaller *referensiell distanse til venstre*, altså avstanden mellom det referensielle uttrykket (substantivfrasen) og forrige referensielle uttrykk med samme referanse. Her har jeg som kriterium at referenten må være nevnt tidligere i samme eller foregående hovedsetning. Om den ikke er det, er referenten ikke-aktivert. Dette en relativt kort referensiell distanse, da det kun inkluderer de nærmeste setningene, og i tillegg ikke tar hensyn til om det er andre referenter i de samme setningene som er *mer fremtredende* (*potential interference* (Givon 1983: 13)).

Det andre forholdet jeg har vurdert, er *samreferanse*, altså om den aktuelle substantivfrasen viser til akkurat samme referent som det tidligere referensielle uttrykket eller ikke. Jeg har valgt å regne en referent som aktivert om *den samme referenten* eller *en referent av samme type* er nevnt tidligere i samme eller foregående hovedsetning. Det er med andre ord ikke noe krav om at de to uttrykkene er samreferensielle, det holder at de viser til samme type referent og har samme substantiv som kjerne i frasen. Med referent av samme type mener jeg for eksempel en annen referent som tilhører samme klasse av referenter, eventuelt generisk referanse til klassen referenten tilhører. De følgende eksemplene viser forskjellen mellom disse.

8.4-b Det var min katten som var påkjørt foran mitt hus. **Katten** gikk tilbake fra skogen hjem og ble påkjørt på veien. (s0940)

Vi ser over et eksempel med substantivfrasen «katten», som har blitt klassifisert som *aktivert* ettersom referenten er nevnt i setningen før. Det er her snakk om samreferanse siden det er samme katt som omtales i begge setningene. I eksemplet under (8.4-c) kan

man også si at det er snakk om aktivering gjennom en form for samreferanse, men det er snakk om to fraser, «kvinner» og «kvinnen», som begge har samme *generiske* referanse. Den understrekede forekomsten «kvinner» er altså klassifisert som aktivert.

8.4-c Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at **kvinner** fikk stemmerett. (s0654)

Vi ser også av setningen over at det uttrykket som utløser aktivering, ikke selv trenger å være et selvstendig setningsledd, men kan være utfylling til et annet substantiv, som «kvinnen» er i frasen «kvinnen possision».

I de neste eksemplene er det *ikke* snakk om samreferanse mellom den understrekede frasen og det tidligere uttrykket. I eksempel 8.4-d er mobiltelefoner nevnt to ganger, der det først vises generisk til alle mobiltelefoner, mens det i påfølgende setning vises til individuelle mobiltelefoner som folk eier. Den understrekede forekomsten «mobiltelefonen» er her regnet som aktivert fordi samme type av referent har vært nevnt før, selv om det ikke kan sies å være samme referent.

8.4-d Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

Et lignende tilfelle finner vi i eksempel 8.4-e under, der også mobiltelefoner nevnes to ganger. Den første gangen vises det til en spesifikk mobiltelefon som skribenten eier, mens det andre gangen vises til en ikke-spesifikk mobiltelefon. Den understrekede forekomsten «en mobiltelefon» er altså klassifisert som aktivert fordi et referensielt uttrykk med samme substantiv og samme type referent er nevnt tidligere, selv om det ikke er snakk om samme referent.

8.4-e Jeg reiser veldig ofte og det er mye lettere for meg hvis jeg kan bruke mobiltelefonen min. Det skjer også mange rare situasjoner i vår liv når det er bra å ha **en mobiltelefon** med – for eksempel bilulykker. (s0983)

Det siste eksemplet (8.4-f) viser substantivfrasen «en by» som er klassifisert som *ikke-aktivert* fordi referenten eller en referent av samme type ikke har vært nevnt i samme eller foregående hovedsetning.

8.4-f Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan **en by** passer for barn. (s0295)

Både kriteriet for referensiell distanse og valget om å ikke kreve samreferanse har blitt utformet med tanke på effektivisering av klassifiseringen, siden dette i utgangspunktet er en korpusbasert undersøkelse, noe som i mindre grad har åpnet for nærlesning av hele tekster i klassifiseringsprosessen. Klassifiseringen har nemlig blitt gjort i annoteringsfeltet i selve konkordansen, derfor har jeg latt meg styre av hvor stor mengde kontekst før treffordet som er mulig å se der. Illustrasjon 8-1 under viser et skjermbilde av en konkordans med maksimal venstrekontekst, som er 700 piksler. I noen tilfeller åpner venstrekonteksten for lengre referensiell distanse, men det var også ønske om at distansen skulle være lik for alle forekomstene, derfor har jeg valgt en minste felles distanse.

Illustrasjon 8-1 Skjermbilde av konkordans med maksimal venstrekontekst i ASK.

The screenshot shows the ASK Hovedkorpus Konkordans interface. At the top, it displays the search query: `[annotation = "s3.*"] \ <> \ \ :: language != "norsk"`. Below the query, there are buttons for 'Søk', 'Raffiner', 'vindu: sentence', 'Stop', 'Lagrede søk ...', and 'Lagre søket'. The interface also shows 'Done. Running time: 0.22 sec. (0.01 CPU sec.)' and various filters like 'Type: kwic', 'Vis annotasjon', 'Vis linjefilter', 'Attributter ...', 'Strukturer ...', 'Linjer per side', and 'Kontekststørrelse'. The search results are displayed in a table with columns for 'Tekst-id', the text snippet, and 'treff'. The table shows several rows of results, each with a unique ID and a corresponding text snippet and search hit.

Tekst-id	treff
s0011	er trivelig å bo på hvis du har noen venner der også. Man kan bo i en spennende by, men uten venner, byen vil være veld
s0219	ennende. Jeg studerte ham på skolen og jeg trodde ikke at jeg ville kunne se det bildet ansikt til ansikt. Bildet er mye mind
s0219	eg trodde! Jeg har også gått til parken hvor står mange mennesker som er lagte av stein. Jeg har glemt navnet unnskyld! De
s0349	o i mobiltelefonbruk. Jeg håper at det er mulig å finne en avlesning snart. Et annet negativt poeng er at melding muligheten t
s0645	Før og nå. Det er mange forskjeller i samfunn før og nå. Spørsmålet er hvilken pe
s0015	Sentrum. Det ligger ved siden av @sted i en gammel naboheit. Men vi tenker å flytter oss i det neste år. Miljøet er deilig, me
s0015	jeg kan også treffe andre mor ute på gate eller en lekeplass. Hos oss er der ganske ensomme av og til. Huset ligger langt t
s0018	nå man være fornøyd. For dem som liker å gå på tur eller på ski, er Norge det beste mulig land å bo på. Naturen er vakker og
s0018	sine barn, og mine trives på skolen og barnehagen. De får frisk luft, mye mosjon, og et trygt miljø. Men været er også viktil
s0078	tesker som vi må også respektere. Det er en stor forskjell mellom barna fra gamle dager og våre dager. Forskjellen ligger i oppf
s0163	svare. Det er også skolemiljø som påvirker til barna. Der lærer barna hvordan og hvor kan de skaffe seg kunnskapet Hvis de leser
s0163	ljø og ekologi. Barnaoppdragelse er vanskelig men når de er voksne te kan foreldrene være glade i dem. Resultaten er at de er g
s0187	det viktig å ha familie eller venner for å støtte deg. En annen ting som kan gjøre et sted trivelig å bo er været . Det er man
s0189	ig med en fin utsikt, for eksempel, vann, skog eller fjell. Jeg er veldig fornøyd med min leilighet. Jeg har utsikten ut over fjord
s0191	n i samfunnet. Huis man flytter til et sted hvor de snakker et forskjellig språk, er det viktig at han lærer språket . Huis man e
s0191	ted hvor de snakker et forskjellig språk, er det viktig at han lærer språket. Huis man er ikke villig å lære språket , da skiller h
s0192	: kastet ut hvor som helst men at det er en plass på soppel. Frisk luft er også viktig så hvis vi kan bruke bussen det er bedre

Tabellen på neste viser flere eksempler på aktiverte og ikke-aktiverte referenter, og eksemplene over er også lagt inn for oversiktens del.

Tabell 8-10 Eksempler på tekstuelt aktivert og ikke-aktivert referent

Aktivering	Eksempel
Ikke aktivert	Jeg har <u>ei datter</u> og hun mener alt til meg. (s0351) Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan <u>en by</u> passer for barn. (s0295)
Aktivert v/ samreferanse	Man kan bo i en spennende by, men uten venner, <u>byen</u> vil være veldig kjedelig. (s0011) Det var min katten som var påkjørt foran mitt hus. <u>Katten</u> gikk tilbake fra skogen hjem og ble påkjørt på veien. (s0940) Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at <u>kvinner</u> fikk stemmerett. (s0654)
Aktivert gjennom samme type referent	Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har <u>mobiltelefonen</u> i lomma eller i veska. (s0791) Jeg reiser veldig ofte og det er mye lettere for meg hvis jeg kan bruke mobiltelefonen min. Det skjer også mange rare situasjoner i vår liv når det er bra å ha <u>en mobiltelefon</u> med – for eksempel bilulykker. (s0983)

8.4.4 Referensiell type

Referensielle typer har jeg vært inne på både i kapittel 3.2 og i kapittel 7, og jeg opererte da med et skille mellom *definitte*, *indefinitte*, *generiske* og *ikke-referensielle* substantivfraser. Klassifiseringen i referensielle typer for undersøkelse 2 har fulgt disse fire, og er basert på de beskrivelsene jeg har gitt tidligere i kapittel 3. Jeg vil først gi en kort oppsummering av hovedforskjellene mellom dem.

Indefinit og *definit* referanse er like med tanke på at man i begge tilfeller viser til én eller flere individuelle referenter, noe som skiller dem fra *generisk* referanse, hvor man viser til klasser av referenter (se 3.2.1). Indefinit og definit er derimot ulike i at man ved definit referanse viser til referenter som av senderen er antatt å være identifiserbar for mottageren av ytringen, mens referentene antas å *ikke* være identifiserbare ved indefinit referanse. Hovedområdene for etablering av identifiserbarhet ved definit referanse er: *anaforti*, *situasjon*, *kunnskap om verden* og gjennom *assosiasjon til en annen referent* (se 3.1). Den siste av de fire referensielle

typene er *ikke-referensielle* substantivfraser, som ikke refererer, men heller *beskriver* (se 3.2.2).

I klassifiseringen av substantivfrasene i materialet har jeg vurdert følgende forhold:

1) Er frasen referensiell eller ikke?

Den mest åpenbare ikke-referensielle bruken av substantivfraser er predikativer, som beskriver heller enn å referere, som i eksempel 8.4-g under. I kapittel 3.2.2. og 7.2.4 omtalte jeg dessuten noen grensetilfeller som bare er *svakt* referensielle. Svakt referensielle fraser kan for eksempel være fraser hvor senderen ikke har noen spesifikke referenter i tankene, altså *ikke-spesifikk* referanse, som vi ser i eksempel 8.4-h. En annen type svak referanse kan være når man i ytringen velger å fokusere mot *typen referent* heller enn eksemplaret (Borthen 2003: 22, 51; Vannebo 2010: 231), noe jeg har omtalt i 7.2. Et mulig eksempel på dette er «slips» i eksempel 8.4-i, men som jeg vil komme tilbake til lenger ned, er det spesielt vanskelig å trekke en grense mellom referanse og svak referanse i innlærerspråk.

8.4-g Jeg er **polakk** og jeg er oppvokset i Polen. (s0771)

8.4-h Det skjer også mange rare situasjoner i vår liv når det er bra å ha **en mobiltelefon** med –for eksempel bilulykker. (s0983)

8.4-i Om kvelden er det fremdeles to programleder, imidlertid sofaen er byttet på et bord, og mannen har **slips**.

Fraser med svak referanse har jeg har valgt å regne som referensielle, mens jeg som ikke-referensielle regner predikativer og noen andre typer fraser hvor substantivet ikke har noen egentlig referanse, for eksempel adverbialer som «en time». Denne typen fraser har blitt utelatt fra materialet for undersøkelse 2, mens de referensielle substantivfrasene har blitt videre vurdert for spørsmål 2.

2) Viser frasen til individuelle referenter eller til en klasse av referenter ?

For referensielle fraser skiller dette spørsmålet de definite og indefinite tilfellene fra de generiske. Ved referanse til klasser av referenter (generisk) viser man enten til alle medlemmer av en klasse eller til klassen som helhet, men ikke nødvendigvis alle medlemmene (se 3.2.1.1). Ved ikke-generisk referanse viser man bare til visse

medlemmer av klassen, altså én eller flere individuelle referenter. De to eksemplene under viser generiske substantivfraser. Frasen «nyheter» i 8.4-j viser generelt til alle nyheter og er en påstand om alle nyheter fordi det er ment som en forklaring på hva nyheter er. Frasen «læreren» i 8.4-k viser generelt til yrkesgruppen lærere, men ikke nødvendigvis til samtlige medlemmer av denne gruppen.

8.4-j **Nyheter** er informasjon som er levert til en person eller til mange mennesker. (s0355)

8.4-k I Polen har **læreren** stor autoritet. (s0652)

Ulike typer generisitet spiller ingen større rolle i undersøkelsen, men er ment som en konkretisering av hva som har blitt klassifisert som generisk. Om substantivfrasene ikke er generiske, er de videre vurdert etter spørsmål 3.

3) *Er referentene identifiserbare gjennom et av følgende domener:*

Anafori, situasjon, kunnskap om verden eller assosiasjon til en annen referent?

For å skille mellom definititt og indefinititt referanse, ble det vurdert om referentene var identifiserbare gjennom et av domenene som etablerer definititt referanse: gjennom å være nevnt tidligere i teksten (*anafori*), gjennom den umiddelbare eller utvidede ekstralingvistiske konteksten for ytringen (*situasjon*), gjennom *kunnskap om verden* eller gjennom *assosiasjon* til en annen referent. Disse er alle omtalt i 3.1, og under er eksempler på disse fire hovedtypene.

I det første eksemplet under (8.4-l) ser vi to tilfeller av anafori, hvor både «sofaen» og «mannen» er definitive fordi de har blitt introdusert i en tidligere setning. Første gangen disse to referentene nevnes, gjennom frasene «en sofa» og «en mann», er de derimot *indefinitte* ettersom de ikke er identifiserbare enda.

8.4-l Jeg vet ikke hvorfor, med som i storbritannia vannligvis er det to programledere, en mann og en kvinne som sitter på en sofa og pratter. Om kvelden er det fremdeles to programledere, imidlertid **sofaen** er byttet på et bord, og **mannen** har slips. (s0328)

Eksempel 8.4-m viser referenten «legene» som er definititt gjennom assosiasjon til en annen referent, nemlig sykehuset.

8.4-m Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikshospital sykehuset. **Legene** sa at operasjon gikk bra. (s0745)

Det neste eksemplet viser et tilfelle der referenten er identifiserbar gjennom den ekstralingvistiske konteksten, i dette tilfellet den geografiske eller kulturelle konteksten for ytringen. Vi vet at Norge bare har en konge, og dermed er «kongen» identifiserbar for mottageren av ytringen.

8.4-n I går hørte jeg på TV-en at **kongen** ble operert og at han ble lagt på en intensivavdeling. (s0756)

Eksemplet under (8.4-o) viser et av få tilfeller som kommer inn under kunnskap om verden, hvor identifiserbarheten gjerne er knyttet til at referenten er *absolutt unik*, noe som står i motsetning til den mer begrensede situasjonelle unikheten som ble omtalt i eksemplet med «kongen» over. I eksemplet under er det *våren*, her uttrykt som «vår», som jeg oppfatter som absolutt unik.

8.4-o **Vår** kommer litt tidligere til @sted, og mars gjennom mai kan bli veldig fint å være ute.

Som jeg var inne på i 3.1, kan referenter i tillegg være identifiserbare gjennom sitt beskrivende innhold om substantivfrasen er informativ nok til å peke mot en unik referent som passer til denne beskrivelsen. Siden mitt materiale består av kun enkle substantivfraser uten modifierende ledd, er ikke denne typen identifiserbarhet relevant. En type enkle substantivfraser som likevel kan komme inn under dette, er henvisninger til størrelser som «kongen» eller «statsministeren», men disse har jeg regnet som situasjonelt identifiserbare. Egennavn er et annet eksempel, men disse er ikke tatt med i materialet siden de normalt sett ikke bøyes i norsk, og fordi de som regel er anonymisert med en felles kode i korpuset (@navn).

Jeg var også inne på at referenter dessuten kunne regnes som identifiserbare gjennom å tilhøre en persons «personlige sfære», for eksempel kroppsdeler, slektninger eller andre ting som typisk knyttes til en person eller referent (se 3.1, s. 60). Denne typen ser jeg på som relatert til *assosiasjon*, altså at noe er assosiert med sin «eier» så lenge denne er nevnt. Et eksempel på dette fins i 8.4-p, hvor «mobiltelefonen» fremstår som en del av menneskers faste eiendeler og det på sett og vis kan ses på som at det foreligger en underforstått possessiv der: «mobiltelefonene sine».

8.4-p Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

De ulike områdene for etablering av definitte referanse er ikke i seg selv relevante for undersøkelsen, annet enn som en konkretisering av når en referent er blitt klassifisert som definitte.

Om referentene *ikke* kan sies å være identifiserbare gjennom et av domenene over, har de blitt regnet som *indefinitte*. Som indefinit referanse har jeg også regnet frasene med *svak referanse* som ble nevnt under punkt 1 over.

Et eksempel på indefinit referanse ble allerede omtalt over i forbindelse med eksempel 8.4-l, hvor «en sofa» og «mann» introduserte nye referenter som ved neste gangs nevning var definitive. Disse referentene var ved første gangs nevning indefinitte og *spesifikke*, på den måten at det vises til en spesifikk mann og en spesifikk sofa, ikke til en hvilken som helst mann og sofa.

Et eksempel på indefinit *ikke-spesifikk* referanse så vi over i 8.4-h, der det var snakk om en mobiltelefon, men uten at det var en spesifikk mobiltelefon. Lenger opp så vi også et eksempel på indefinit referanse som kunne tolkes å ha fokus på type heller enn eksemplar, nemlig «slips» 8.4-i (gjentatt under).

8.4-i Om kvelden er det fremdeles to programledere, imidlertid sofaen er byttet på et bord, og mannen har slips.

Når det står at mannen «har slips», heller enn «har et slips», kan dette tolkes som at det fokuseres mot type heller enn eksemplar, noe som er en vanlig i norsk i visse tilfeller der ubestemt entall uten artikkel brukes som objekt (Borthen 2003: 22, 51; Vannebo 2010: 231).

Grensen mellom referensiell og *svakt referensiell* er i seg selv en diskusjon, der noen regner som svakt referensielt det som andre regner som ikke-referensielt (se 3.2.2). I innlærerspråk kan intensjonene og systemet som ligger bak formvalgene, variere veldig og skille seg fra målspråket, noe som gjør det enda vanskeligere å vurdere grensen mellom referensiell og svakt referensiell i slike varieteter. Dette er i seg selv en god grunn til å inkludere fraser med potensielt svak referanse i det indefinitte datamaterialet.

For å oppsummere gjennomgangen av referensielle typer har altså substantivfraser blitt klassifisert som *definitte* om de viser til individuelle referenter som er identifiserbare gjennom et av domenene for etablering av definit referanse, *indefinitte* om de ikke er identifiserbare gjennom disse domenene, og *generiske* om de viser til

klasser av referenter heller enn individuelle referenter. Tabellen under viser eksempler på de fire referensielle typene, og eksemplene som er omtalt over, er inkludert for oversiktens del.

Tabell 8-11 Eksempler på de ulike referensielle typene hentet fra ASK

Ref.type	Eksempel
Indefinit	<p>Vi måtte ta en leiebil og kjøre rett over til andre siden av øyet, ... (s0328)</p> <p>Jeg vet ikke hvorfor, med som i storbritannia vannligvis er det to programledere, en mann og en kvinne som sitter på en sofa og pratter. Om kvelden er det fremdeles to programledere, imidlertid sofaen er byttet på et bord, og mannen har slips. (s0328)</p> <p>Flere etter bilulikker døde, flere er for hele livet ødelagt i kroppen, for eksempel de må bruke rullestol. (s0940)</p> <p>Ofta biler krasjer fordi det er for trangt i boligområde og mennesker kjøre for fort. (s0940)</p>
Definit	<p>Jeg vet ikke hvorfor, med som i storbritannia vannligvis er det to programledere, en mann og en kvinne som sitter på en sofa og pratter. Om kvelden er det fremdeles to programledere, imidlertid sofaen er byttet på et bord, og mannen har slips. (s0328)</p> <p>Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikshospital sykehuset. Legene sa at operasjon gikk bra. (s0745)</p> <p>Nesten alle mennesker har mobillefonen i lomma eller i veska. (s0791)</p> <p>I går hørte jeg på TV-en at kongen ble operert og at han ble lagt på en intensivavdeling. (s0756)</p> <p>Vår kommer litt tidligere til @sted, og mars gjennom mai kan bli veldig fint å være ute. (s0311)</p>
Generisk	<p>Vennskap er en av de viktigste tingene man kan skaffe i livet. (s0218)</p> <p>Mobillefoner ble veldig populært i de siste årene. (s0791)</p> <p>Det er ikke så veldig lenge siden at kvinner fikk stemmerett. (s0654)</p>
Ikke-referensiell	<p>Jeg er polakk og jeg er oppvokst i Polen. (s0771)</p>

De tre gruppene har en relativt lik andel av definte fraser (ca. 25 %) og ikke-referensielle fraser (ca. 3 %). Det er derimot noe mer ulikt når det gjelder andelen av indefinite og generiske fraser, hvor den engelske gruppen har en litt større andel indefinite enn de to andre gruppene (pol. 52 %; eng. 57 %; nor. 48 %), og en litt lavere andel generiske (pol. 20 %; eng. 16 %; nor. 22 %). Forskjellene er likevel ikke veldig store.

I undersøkelse 2 vil referensiell type være sentral for å dele opp materialet slik at jeg bedre kan undersøke forskjeller og likheter mellom den polske gruppen og de to andre gruppene med tanke på transfereffektene *intergruppe-heterogenitet* og *intringval kontrast*. Dette vil jeg komme nærmere tilbake til under, i 8.5.3.5.

8.4.5 Tellelighet

Den siste klassifiseringskategorien jeg har tatt med, er *tellelighet*, som ikke er direkte relevant for hypotesene, men som gjør det mulig å utelukke substantivfraser med ikke-tellelige referenter, noe som er ønskelig for hypotese 1. Tellelighet har i utgangspunktet to trekk: *tellelig* og *ikke-tellelig*, men jeg var også innstilt på å kode eventuelle *kollektiver*. Som nevnt tidligere, kan det være problematisk å snakke om tellelighet som en egenskap ved ordet i seg selv, og det kan være bedre å se på det som en kontekstuell egenskap (se 3.4). Dette er fordi vi kan individualisere ikke-tellelige referenter (Faarlund et al. 1997: 144), eller vi kan bruke tellelige substantiv med ikke-tellelig betydning (Faarlund et al. 1997: 149). Jeg har likevel valgt å se på tellelighet som en egenskap ved ordet, siden jeg da unngår at bøyingsformen til ordet påvirker hvilken klassifisering jeg har gitt. Tabellen under gir eksempler på tellelige, ikke-tellelige og kollektive substantiv fra materialet.

Tabell 8-12 Eksempel på tellelig, ikke-tellelig og kollektivt substantiv hentet fra ASK.

Tellelighet	Eksempler
Tellelig	Det er viktig å ha kinoer , museumer, caféer, idrettsbaner blant annet. (s0320)
Ikke-tellelig	De andre store ulempene er forurensing og sikkrihet. (s0304)
Kollektiv	Folk trenger ikke mobiltelefoner på arbeidet. (s0329)

Det er relativt få slike kollektiver i tekstene (38 stk.), og i all hovedsak er det snakk om ordet «folk». Utover dette har alle de tre gruppene flest tellelige substantiv, selv om

fordelingen av tellelige og ikke-tellelig varierer noe mellom gruppene: Den norske gruppen har lavest andel tellelige av de tre gruppene (pol. 75 %; eng. 79 %; nor. 63 %), og dermed en større andel ikke-tellelige. Dette kan blant annet skyldes at temaene for tekstene legger ulike føringer på bruk av ikke-tellelige substantiv. Veldig mange i den norske gruppen hadde for eksempel skrevet om «sunnhet og helse», og det kan tenkes at dette har bidratt til en overvekt av ikke-tellelige ord, for eksempel for matvarer.

8.5 Generelt om gjennomføringen av undersøkelse 2

Jeg vil i resten av dette kapittelet gjøre rede for de sidene av metoden som er felles for de to hypotesene. For det første prediksjoner om form/kontekst-assosiasjoner (8.5.1), for det andre undersøkelser på både gruppenivå og i individuelle profiler (8.5.2), for det tredje de fire transfereffektene (8.5.3), og til slutt klassifiseringen av individuelle profiler (8.5.4). De mer spesifikke forholdene som gjelder for hver av hypotesene, vil bli tatt underveis i kapittel 9 og 10.

8.5.1 Prediksjoner om assosiasjon mellom form og kontekst

Hypotesene i undersøkelse 2 handler begge om hvilke språklige forhold i morsmålet som kan tenkes å påvirke språklige valg i innlærerspråket, og begge hypotesene har prediksjoner om at visse former vil bli brukt i visse språklige kontekster, videre omtalt som *form/kontekst-assosiasjoner*. Tabell 8-13 (under) gir en oversikt over form/kontekst-assosiasjonene som er predikert for hver av de to hypotesene.

Tabell 8-13 Oversikt over form/kontekst-assosiasjoner som er predikert av hypotesene

Hypotese 1	Form	Kontekst
	Ubestemt entall/flertall	Aktiverte forfelt eller ikke-aktiverte slutfelt
	Bestemt entall/ flertall	Aktiverte slutfelt
	Artikkelfrase	Ikke-aktiverte forfelt
Hypotese 2	Form	Kontekst
	Ubestemt entall/flertall	Subjekter
	Bestemt entall/ flertall	Objekter

For hypotese 1 er det tre ulike form/kontekst-assosiasjoner som undersøkes, mens for hypotese 2 er det to, og det er disse form/kontekst-assosiasjonene som er hovedfokus for undersøkelsen av morsmålspåvirkning hos de polske innlærerne.

For hver av hypotesene har jeg hentet ut statistisk informasjon om hvilke former som er brukt i de ulike konteksttypene, både for *hele gruppen samlet*, og for *hver enkelt individuelle tekst*. På denne måten har jeg fått oversikt over både tall på gruppenivå og hver enkelt individs individuelle profil for formvalg i de ulike konteksttypene.

I analysen av form/kontekst-assosiasjonene på gruppenivå tar jeg i betraktning to sider av assosiasjonen: For det første undersøker jeg hvor ofte en form er brukt i sin predikerte konteksttype i forhold til andre former, og for det andre hvor ofte en form er brukt i sin predikerte konteksttype i forhold til hvor ofte den er brukt i andre konteksttyper. Disse to sidene vil jeg omtale som henholdsvis «formens andel innen kontekst» og «kontekstens andel innen form», eller bare «innen kontekst» og «innen form».

De sidene belyser assosiasjonen mellom form og konteksttype på ulike måter, og til sammen utgjør de et mål for hvor sterk assosiasjonen mellom form og kontekst er. En kategorisk og sterk assosiasjon foreligger når en form kun brukes i sin predikerte konteksttype, og når dette i tillegg er den eneste konteksttypen formen brukes i (dvs. 100 % innen kontekst og 100 % innen form). I tilfeller der det ikke er en slik kategorisk assosiasjon, vil styrken til assosiasjonen belyses av hvorvidt en konteksttype først og fremst har denne formen framfor andre former, og hvorvidt denne formen stort sett bare brukes i denne konteksttypen, og ikke i andre. Om det er andre former enn den predikerte som er mer brukt i konteksttypen, eller formen oftere brukes i andre konteksttyper enn den predikerte, regnes assosiasjonen som svak.

På gruppenivå har jeg ikke hatt mulighet til å teste disse assosiasjonene statistisk, derfor er vurderingen av forholdet mellom form og kontekst en skjønnsmessig vurdering, der jeg vurderer prosentandeler som ligger under 50 % innen form og/eller kontekst som et svakt funn. Statistisk testing vil i stedet bli utført på en opptelling av individuelle profiler som stemmer med prediksjonene for hypotesen, noe jeg vil komme tilbake til under.

Form/kontekst-assosiasjonene i den polske innlærergruppen blir for begge hypotesene sammenlignet med formvalgene i den engelske innlærergruppen og den norske kontrollgruppen, og jeg vil vise til disse gruppene som *den polske gruppen*, *den engelske gruppen* og *den norske gruppen*, alternativt bruker jeg noen steder

formuleringene *det polske materialet, det engelske materialet* eller *det norske materialet* om de samme gruppene uten at dette innebærer noen spesiell innholdsmessig forskjell. Når jeg derimot omtaler forhold i de *tre språkene i seg selv*, bruker jeg benevnelsene *polsk, norsk og engelsk*.

Jeg har over omtalt fire sentrale metodiske grep for undersøkelse 2:

- 1) prediksjoner om form/kontekst-assosiasjoner
- 2) undersøkelse av prediksjoner både *innen kontekst* og *innen form*
- 3) undersøkelse av prediksjoner både på gruppenivå og i individuelle profiler
- 4) sammenligning av den polske gruppen med både den engelske og norske gruppen

Ved hjelp av disse grepene, vil jeg forsøke å belyse de fire transfereffektene som er omtalt i kapittel 2, med det formål å avgjøre om språklig kontekst er styrende for formvalgene til de polske innlærerne på den måten hypotesene sier.

Jeg vil videre gjøre rede for metode for undersøkelsen på gruppe- og individnivå, for undersøkelsen av transfereffekter og for prosedyren for klassifisering av individuelle profiler.

8.5.2 Undersøkelse på gruppenivå og individnivå

For begge hypotesene om morsmålspåvirkning fra polsk sammenligner jeg den polske innlærergruppen med både den engelske innlærergruppen og den norske kontrollgruppen, og dette innebærer en sammenligning både av tall som gjelder for gruppen som helhet, og av de individuelle profilene.

Som tidligere nevnt, er materialet for undersøkelse 2 en god del mindre enn det jeg hadde i undersøkelse 1, og det er derfor begrenset mulighet for å bruke statistikk på samme måten. Dette skyldes at antall kontekster per individuelle tekst er så lavt at det vanskeliggjør bruk av både statistiske tester som baserer seg på gjennomsnitt og statistiske tester som baserer seg på rang, dvs. både parametriske og ikke-parametriske tester. Løsningen jeg har valgt, er å supplere en skjønsmessig vurdering av data på gruppenivå, som ikke tar i betraktning individuell variasjon, med en klassifisering av hver individuelle tekst etter om formvalgene stemmer med prediksjonene for den aktuelle hypotesen, eller ikke. Denne metoden er valgt for å både kunne beskrive tendenser i formvalg som kommer til uttrykk på gruppenivå, og for å kunne gjøre rede for individuell variasjon.

Selv om materialet ikke egner seg for slike statistiske tester som ble nevnt over, gjør en inndeling i individuelle profiler det mulig å bruke en annen statistisk test, nemlig *kjikkvadrat-test*, for å sammenligne gruppene med hensyn til hvor mange individuelle profiler som innfrir og ikke innfrir prediksjonene for polsk morsmålspåvirkning. Hver del av undersøkelsen vil derfor følge samme mal: Først en skjønnsmessig vurdering av gruppestatistikk hvor jeg sammenligner form/kontekst-assosiasjoner i den polske gruppen med de andre to gruppene. Deretter en opptelling og statistisk testing av antall individuelle profiler som er i samsvar med prediksjonene i hver av gruppene. Hvordan jeg gjennomfører undersøkelsen av transfereffektene og klassifiserer de individuelle profilene, vil jeg gjøre rede for under.

8.5.3 Undersøkelse av transfereffekter

Som nevnt vil jeg også i undersøkelse 2 basere meg på Jarvis' kriterier for påvisning av transfer, altså forsøke å finne evidens for form av *morsmålssamsvar*, *intragruppe-homogenitet*, *intergruppe-heterogenitet* og *intralingval kontrast* (Jarvis 2000; 2010). Jeg vil for hver av hypotesene formulere spesifikt hvordan de fire transfereffektene vil komme til uttrykk i materialet, men siden metoden er lik for de to hypotesene, vil jeg beskrive de generelle trekkene ved metoden her. De mer hypotesespesifikke detaljene blir redegjort for i kapittel 9.2 og 10.2.

8.5.3.1 Morsmålssamsvar: prediksjoner om form/kontekst-assosiasjoner

Transfereffekten *morsmålssamsvar* skal vise at det er noe i morsmålet som motiverer formvalget i innlærerspråket (Jarvis 2000: 255), og denne effekten vil for begge hypotesene undersøkes ved hjelp av prediksjoner om form/kontekst-assosiasjoner for den polske gruppen. Dette er de prediksjonene som ble presentert på bakgrunn av hver av hypotesene tidligere i dette kapittelet. Morsmålssamsvar undersøkes både på gruppenivå og gjennom individuelle profiler, altså om gruppen som helhet følger prediksjonen og om en større andel individuelle profiler har det predikerte formvalget. Hvilke prediksjoner det er snakk om vil jeg komme tilbake til for hver av hypotesene.

8.5.3.2 Intragruppe-homogenitet: likheter innad i den polske gruppen

Transfereffekten *intragruppe-homogenitet*, eller *likheter innad i gruppen*, skal vise at språklige valg hos de polske innlærerne er representative for denne gruppen, altså at de

predikerte form/kontekst-assosiasjonene er utbredt i tekstene skrevet av denne gruppen (Jarvis 2000: 254). Denne effekten forutsetter at man ser på individuelle formvalg, derfor vil den kun undersøkes gjennom opptelling av individuelle profiler som følger prediksjonene om morsmålssamsvar. Effekten er kun relevant om det faktisk foreligger tendenser til morsmålssamsvar, ettersom det er nettopp de predikerte form/kontekst-assosiasjonene som forventes å være representative for gruppen, og ikke en annen distribusjon av formene.

8.5.3.3 *Intergruppe-heterogenitet: forskjeller mellom innlærergruppene*

Effekten *intergruppe-heterogenitet*, eller forskjeller mellom innlærergruppene, skal vise at form/kontekst-assosiasjoner i den polske gruppen er spesifikke for denne gruppen sammenlignet med en annen innlærergruppe (Jarvis 2000: 254). I denne sammenhengen innebærer det at den polske gruppen skal ha spesifikke form/kontekst-assosiasjoner som skiller seg fra det en finner hos engelske innlærere.

Sammenligningen av den polske og den engelske gruppen vil bli gjort både på gruppenivå og individnivå, men jeg har valgt å skille mellom språklige kontekster hvor disse to gruppene er forventet å være *heterogene* (ulike) og kontekster hvor gruppene er forventet å *ikke* være det, altså hvor de tvert imot er forventet å være *homogene* (like). Selv om polske og engelske innlærere ut fra hypotesen generelt sett er forventet å ha ulike formvalg, ettersom polsk og engelsk som morsmål kan ventes å påvirke innlærerspråket på ulike måter, er det også forventet at det fins noen språklige kontekster hvor de to morsmålene tilfeldigvis har *like* implikasjoner for formvalg i innlærerspråket (Jarvis 2000: 257). En inndeling av kontekster etter om de impliserer likhet eller ulikhet mellom gruppene, er nødvendig for at ikke tilfeldige sammenfall mellom morsmålenes implikasjoner skal få gruppene til å se mer like ut enn de faktisk er, altså tilsløre intergruppe-heterogenitetseffekten. Ved å dele inn datamaterialet på denne måten, framhever jeg altså de områdene hvor transfereffekten intergruppe-heterogenitet faktisk vil komme til syne om det foreligger transfer.

Som jeg vil komme tilbake til, er det inndeling i ulike referensielle typer (indefinit, definit og generisk referanse) som er utgangspunktet for å skille ut kontekster for intergruppe-heterogenitet. Innen hver referensielle type har jeg skilt ut kontekster hvor den engelske og den polske gruppen er forventet å ha *ulike* språklige

valg, og kontekster hvor de er forventet å ha *like* språklige valg. Disse vil bli omtalt som henholdsvis «markørkontekst» og «sammenfallskontekst».

Markørkontekster: Språklige kontekster hvor prediksjonene for polsk morsmålssamsvar ikke sammenfaller med ventede språklige valg for den engelske gruppen. I disse kontekstene er det forventet at den polske og den engelske gruppen har ulike formvalg om transfer foreligger, altså intergruppe-heterogenitet.

Sammenfallskontekster: Språklige kontekster hvor prediksjonene for polsk morsmålssamsvar tilfeldigvis sammenfaller med ventede språklige valg for den engelske gruppen. I disse kontekstene er det forventet at den polske og den engelske gruppen har like formvalg, altså *ikke* intergruppe-heterogenitet.

En inndeling i sammenfalls- og markørkontekster forutsetter at jeg også presiserer hvilke språklige valg de *engelske* innlærerne er ventet å gjøre, og dette har jeg gjort med utgangspunkt i morsmålet deres, altså engelsk. Jeg tar dermed for gitt at også de språklige valgene til engelske innlærere er påvirket av morsmålet, noe som i seg selv *ikke* er gjenstand for undersøkelse i denne avhandlingen. Denne antakelsen er ikke uten videre uproblematisk, ettersom vi ikke kan vite om det faktisk er morsmålet som styrer de språklige valgene i den engelske gruppen. Jeg vil komme tilbake til denne problemstillingen i avslutningskapittelet.

8.5.3.4 *Intralingval kontrast: sammenligning med målspråksgruppen*

Operasjonaliseringen av den siste transfereffekten, *intralingval kontrast*, krever litt mer diskusjon ettersom jeg har valgt å gjøre det på en noe annerledes måte enn det som beskrives av Jarvis (2010), og det kan argumenteres for at min løsning ligger nærmere opp til en undersøkelse av *intergruppe-heterogenitet i forhold til målspråket*, enn det intralingval kontrast er ment å være. Jeg vil først si litt om hva Jarvis (2010) ser ut til å legge i denne effekten, og så vil jeg forklare hva min løsning er.

Jarvis (2010: 175-176) omtaler intralingval kontrast som en transfereffekt som undersøkes ved å dele inn i områder hvor morsmål og målspråk står i ulik kontrastiv relasjon til hverandre, og så undersøke om det er signifikante forskjeller i språkbruken til innlærerne avhengig av hvilken type relasjon det er snakk om. For eksempel kan en inndeling i språklige trekk etter om det er en likhetsrelasjon, en ulikhetsrelasjon eller en

nullrelasjon, brukes til dette formålet (Jarvis 2010: 176). Etter det jeg kan se i omtalen av denne effekten og de studiene Jarvis (2010: 176) viser til at har undersøkt den, så er denne effekten først og fremst undersøkt med tanke på *mestring* av et språklig trekk i forhold til andre språklige trekk. I motsetning til de andre tre transfereffektene, ser altså denne effekten først og fremst ut til å være anvendelig i studier som undersøker tilegnelse og mestring, og hvor man måler kvantitative utslag av transfer, slik jeg gjorde i undersøkelse 1. Siden jeg ikke ser på mestring av trekk i undersøkelse 2, ser jeg ikke at denne måten å undersøke intralingval kontrast på er direkte anvendelig til mitt formål. Jeg har derfor forsøkt å operasjonalisere denne effekten på en måte som passer bedre inn i min undersøkelse, og som på en mer tydelig måte kan bidra til å påvise transfer i denne studien.

Min løsning deler også inn i områder hvor morsmål og målspråk er like eller ulike, og den ser også på om det er forskjeller i de polske innlærernes språkbruk avhengig av om morsmål og målspråk er like eller ulike, men uten fokus på mestring eller tilegnelse. Måten jeg har valgt å undersøke intralingval kontrast på, fungerer som en understøtting av effekten *morsmålssamsvar*, og den tar i betraktning at det fins kontekster der prediksjonene om morsmålssamsvar sammenfaller med formvalg som er vanlig i målspråket. Dette er i prinsippet den samme logikken som lå bak inndelingen i sammenfalls- og markørkontekster for sammenligningen med den engelske gruppen. Om hypotesene om polsk morsmålpåvirkning delvis predikerer formvalg som tilfeldigvis sammenfaller med målspråkets typiske formvalg, kan det som ser ut som morsmålssamsvar, alternativt tilskrives tilegnelse av målspråkets distribusjon av former (Jarvis & Pavlenko 2008: 44). Jeg har derfor valgt også for intralingval kontrast å dele opp materialet i sammenfalls- og markørkontekster, altså kontekster hvor prediksjonene om polsk morsmålssamsvar sammenfaller med målspråket og kontekster hvor de ikke gjør det. Det vil i større grad støtte opp under et argument om at det er morsmålet som er kilde til formvalget, om en finner morsmålssamsvar i markørkontekster. Morsmålssamsvar som derimot forekommer i sammenfallskontekster, gir svakere støtte til et argument om transfer, siden tilegnelse av målspråket like gjerne kan forklare formvalget.

I praksis ser jeg altså også i sammenligningen mellom den polske og den norske gruppen på områder hvor gruppene er forventet *heterogene* (ulike) og områder hvor gruppene er forventet *homogene* (like), på samme måte som for transfereffekten intergruppe-heterogenitet mellom den polske og den engelske gruppen. Dette betyr at det er en viss likhet mellom hvordan transfereffektene intergruppe-heterogenitet og intralingval kontrast undersøkes. Undersøkelsen av begge effektene har til felles at de skiller mellom områder hvor den polske gruppen er forventet å være lik og ulik en av de andre gruppene, men mens intergruppe-heterogenitet undersøkes i forhold til den andre innlærergruppen, undersøkes intralingval kontrast i forhold til målspråksgruppen. Formålene er også ulike: *Intergruppe-heterogenitet* viser om den polske gruppen har gruppespesifikke formvalg der det er forventet å være spesifikke formvalg, og utfyller effekten *intragruppe-homogenitet* (grupperepresentative formvalg). *Intralingval kontrast* tydeliggjør derimot om form/kontekst-assosiasjonene i den polske gruppen kan tilskrives morsmålspåvirkning, eller om slike assosiasjoner likegodt kan tilskrives målspråkstilegnelse, og denne effekten utfyller transfereffekten *morsmålsansvar*.

Jarvis (Jarvis 2010: 174) påpeker at en rekke ekstra transfereffekter har blitt foreslått for ham, for eksempel sammenligning mellom språket til innlærere og morsmålsbrukere, men han mener dette går inn under intergruppe-heterogenitet. Det kan derfor argumenteres for at det jeg her kaller intralingval kontrast egentlig er intergruppe-heterogenitet i forhold til målspråket, og at det jeg undersøker er om polsk har spesifikke språklige valg i forhold til målspråksgruppen. I denne fortolkningen er det dermed heller snakk om to typer intergruppe-heterogenitet i min undersøkelse. Den polske gruppen vil i så fall undersøkes for om formvalgene er *spesifikke* på to ulike områder og med to ulike formål: 1) utelukke at det er snakk om form/kontekst-assosiasjoner som er felles for alle innlærere, som språklige universalier og utviklingstrekk, og 2) utelukke at det er formvalg som kan forklares med tilegnelse av målspråket. Jeg har likevel valgt å omtale dette som intralingval kontrast, og vil komme tilbake til dette valget i avslutningskapittelet.

8.5.3.5 *Referansetype som utgangspunkt for sammenligning av gruppene*

Det er altså nødvendig med en inndeling i sammenfallskontekster og markørkontekster både for sammenligningen mellom den polske og den engelske gruppen, og mellom den

polske og den norske. Dette krever en differensiering av materialet, slik at det er mulig å skille fra hverandre områder hvor gruppene er forventet å være like og ulike i sine formvalg.

Et hensiktsmessig grunnlag for inndeling i markør- og sammenfallskontekster i forhold til de to andre gruppene, er *referensielle typer*. I både norsk og engelsk er referanse avgjørende for formvalg: Ved indefinitt referanse har begge språkene typisk ubestemte former, mens man ved definitt referanse typisk har bestemte former. Ved generisk referanse er imidlertid formvalgene mer varierte, selv om både engelsk og norsk typisk har ubestemte former oftere enn bestemte (se kap. 7). Prediksjonene for polsk morsmålssamsvar tilsier derimot at formvalgene er like på tvers av disse referensielle typene, så mens transfereffekten *morsmålssamsvar* i seg selv kan undersøkes i materialet som helhet, vil vurderingen av om et eventuelt morsmålssamsvar er spesifikt for polske innlærere og ikke sammenfallende med målspåket, først komme til syne når materialet deles opp i områder som er relevante for engelske og norske formvalg, altså ulike referensielle typer. I resultatkapitlene for hver av hypotesene vil jeg komme tilbake til mer spesifikke redegjørelser for hva som utgjør sammenfalls- og markørkontekster innen hver av de referensielle typene.

8.5.4 Om inndeling i individuelle profiler

Siden gruppetall kun kan si noe om den generelle fordelingen av former i gruppen som helhet, og ikke tar hensyn til individuell variasjon, har jeg klassifisert alle de individuelle tekstene etter om de har en distribusjon av former som stemmer med prediksjonene for polsk morsmålssamsvar eller ikke, nærmere bestemt i *positive*, *negative* og *kombinerte* profiler. Klassifiseringen av individuelle tekster har jeg gjort både i den polske, engelske og norske gruppen, og tekstene i alle de tre gruppene er klassifisert etter om de stemmer med prediksjonene om *polsk morsmålssamsvar* eller ikke. Dette er for å kunne se om det predikerte formvalget er spesifikt for den polske gruppen, eller om det også fins i de andre gruppene.

En opptelling av ulike typer individuelle profiler gjør det mulig å både undersøke morsmålssamsvar på individnivå og homogenitet innad i den polske gruppen, i tillegg til intergruppe-heterogenitet og intralingval kontrast i forhold til de to andre gruppene. I den grad hypotesene mine testes statistisk, er det på individnivå, hvor jeg som nevnt

bruker en kji kvadrat-test for å undersøke om det er signifikante forskjeller mellom gruppene i antall *positive* individuelle profiler, altså tekster hvor formvalget er i tråd med prediksjonene for polsk morsmålssamsvar.

Det er to steg i klassifiseringen av hver individuelle profil: først en vurdering av hver og en av konteksttypene i profilen, deretter en samlet vurdering av hele profilen.⁴⁹

Steg 1

Vurdering av hver konteksttype isolert sett, og klassifisering av om den har sin predikerte form (= positiv) eller ikke (= negativ).

Steg 2

En helhetlig vurdering av alle konteksttypene innen en individuell profil, og klassifisering etter om profilen som helhet er *positiv*, *negativ* eller *kombinert*.

Steg 1: Klassifisering av konteksttyper:

Som vi så over, predikerer begge hypotesene visse form/kontekst-assosiasjoner; tre ulike assosiasjoner for hypotese 1 og to assosiasjoner for hypotese 2. Klassifiseringen av konteksttyper tar derfor utgangspunkt i om den predikerte formen er brukt eller ikke. Om en konteksttype er forventet å ha bestemt form og har dette, er den positiv, men har den derimot ubestemt form, er den negativ.

Siden de færreste av de individuelle tekstene har et en-til-en-forhold mellom form og kontekst, men derimot variasjon med bruk av flere ulike former i én og samme konteksttype, har jeg satt et krav om at det må være *minst to forekomster* mer av den forventede formen for at konteksttypen skal regnes som positiv. Om det bare er én i overvekt av den forventede formen, en jevn fordeling av ulike former, mindretall av den forventede formen, eller ingen forekomster av den forventede formen, er konteksttypen negativ.

1) Positiv konteksttype:

- a) Kategorisk: Kun den forventede formen
- b) Variasjon: Minst to flere av den forventede formen enn de øvrige formene.

2) Negativ konteksttype:

- a) Kategorisk: Ingen forekomster av den forventede formen
- b) Jevn variasjon: Likt antall av forventede former og øvrige former
- c) Ujevn variasjon: Bare én mer av den forventede formen enn de øvrige formene

⁴⁹ Med «konteksttyper» mener jeg her de tre konteksttypene *aktiverte sluttfelt*, *ikke-aktiverte forfelt* og *aktiverte forfelt/ikke-aktiverte sluttfelt* for hypotese 1, og *subjekt* og *objekt* for hypotese 2.

Steg 2: Klassifisering av individuelle profiler:

Selve klassifiseringen av individuelle profiler baserer seg på hvordan konteksttypene innen en og samme tekst samlet sett ser ut: Er det kun positive konteksttyper, er profilen *positiv*, er det kun negative, er den *negativ*, mens en kombinasjon av både negative og positive konteksttyper gir en *kombinert profil*. Det er altså tre hovedtyper av individuelle profiler:

- 1) **Positiv individuell profil:** Fordelingen av former stemmer med prediksjonene om polsk morsmålssamsvar for **alle** de konteksttypene som forekommer i den individuelle teksten.
- 2) **Negativ individuell profil:** Fordelingen stemmer **ikke** med prediksjonene for polsk morsmålssamsvar for noen av de konteksttypene som forekommer i den individuelle teksten.
- 3) **Kombinert individuell profil:** Fordelingen stemmer med prediksjonene for polsk morsmålssamsvar for enkelte av konteksttypene som forekommer i den individuelle teksten, men ikke for alle.

Siden ikke alle de individuelle profilene inneholder alle konteksttypene for hypotesene, tar klassifiseringen utgangspunkt i *den* eller *de* konteksttypene som faktisk forekommer i teksten, enten det er én, to eller tre. For eksempel er det ikke alle innlærerne som har enkle substantivfraser som er subjekter, dermed er det bare objektskontekster som er representert i denne profilen, og klassifiseringen tar utgangspunkt i om denne ene konteksten har sin predikerte form eller ikke. Har den det, er den positiv, har den ikke det, er den negativ.

Jeg vil komme nærmere inn på de mer spesifikke detaljene rundt klassifisering og analyse av individuelle profiler i løpet av undersøkelsen i kapittel 9 og 10.

8.6 Oppsummering av kapittel 8

Jeg har i dette kapittelet presentert de to hypotesene som er utgangspunktet for undersøkelse 2, og jeg har presentert det datamaterialet som er hentet ut fra ASK til denne undersøkelsen, i tillegg til at jeg har redegjort for klassifisering av dette materialet. Til sist har jeg også redegjort for det som er felles prosedyrer for

undersøkelsen av hypotesene. I denne sammenhengen var det et knippe metodiske grep som ble beskrevet:

1. Prediksjoner om morsmålssamsvar i form av **form/kontekst-assosiasjoner**.
2. Undersøkelse av prediksjonene både «**innen form**» og «**innen kontekst**» for å belyse om forholdet mellom form og kontekst er gjensidig.
3. Undersøkelse av materialet både i **gruppen som helhet** og i **individuelle profiler**.
4. Utskilling av **sammenfallskontekster** og **markørkontekster** for å framheve områder hvor den polske gruppen er ventet å skille seg språklig fra de andre to gruppene, om det foreligger transfer fra polsk.
5. **Differensiering av materialet i referensielle typer** for å kunne skille ut sammenfalls- og markørkontekster.

Resultater og mer spesifikke detaljer knyttet til metoden blir presentert i kapittel 9 (hypotese 1) og kapittel 10 (hypotese 2).

9 UNDERSØKELSE 2: HYPOTESE 1

Hypotese 1 ble beskrevet i kapittel 8.1.2, og er formulert som følger:

De polske innlærerne vil overføre morsmålets bruk av pragmatisk ordstilling som signal om tekstuell identifiserbarhet og en oppfatning om at markering av identifiserbarhet er overflødig når pragmatisk ordstilling brukes. De norske bestemthetsmarkørene vil bli tatt i bruk når identifiserbarhet ikke går fram av ordstillingen.

Ifølge hypotesen vil valg av form hos polske innlærere være styrt av hvorvidt markering av identifiserbarhet oppleves som redundant eller ikke. Redundansoppfatningen er knyttet til at identifiserbarhet går fram av ordstillingen når den følger tema/rema-prinsippet, altså når en tekstuelt aktivert referent står i starten av setningen, eller en ikke-aktivert referent står i slutten av setningen. Markering oppleves derimot ikke som redundant når slike fraser står på motsatt sted i setningen, ettersom identifiserbarheten da ikke går fram av ordstillingen. En overordnet konsekvens av dette er dermed at bestemt form og artikkelfrase kun vil brukes når setningens ordstilling bryter med tema/rema-prinsippet.

Til grunn for hypotese 1 ligger også et *morsmålspåvirket identifiserbarhetsbegrep*, som handler om hvorvidt referansen er aktivert eller ikke gjennom tidligere nevning, altså *tekstuell* identifiserbarhet. Dette er fordi det er denne typen identifiserbarhet som kommer til uttrykk gjennom pragmatisk ordstilling (se 3.3 og 8.1.2). Identifiserbarhet som tekstuell identifiserbarhet er mer begrenset enn den typen identifiserbarhet som ligger til grunn for grammatisk bestemthet i norsk, der referenten kan være identifiserbar gjennom andre forhold enn bare teksten, for eksempel gjennom å være unik eller assosiert med en annen referent. Med mindre noe annet er spesifisert, vil derfor termen «identifiserbarhet» i forbindelse med hypotese 1 vise til *tekstuell identifiserbarhet*.

9.1 Utvalg av materiale og prediksjoner

Undersøkelsen av hypotese 1 tar for seg alle de fem typene enkle substantivfraser som er brukt innen indefinit, definit og generisk referanse, men substantivfrasene er gruppert i tre typer basert på fravær eller tilstedeværelse av bestemthetsendelse eller ubestemt artikkel: 1) ubestemt entall/flertall, 2) bestemt entall/flertall og 3) artikkelfrase.

Det skilles altså ikke mellom entalls- og flertallssubstantiv, og grupperingen er gjort med tanke på at det først og fremst er bruk og fravær av bestemthetsbøyning eller ubestemt artikkel som skal undersøkes.

Alle substantivfrasene i materialet er som omtalt i kapittel 8, klassifisert etter to forhold for denne hypotesen: *setningsposisjon* og *tekstuell aktivering*, som sammen indikerer om referentens identifiserbarhet går fram av ordstillingen eller ikke, og dermed om markering kan oppleves som redundant eller ikke. Som nevnt i 8.4.3, er klassifiseringen av tekstuell aktivering basert på hvorvidt substantivfrasens referent er nevnt i samme eller foregående hovedsetning eller ikke. En referent kan være aktivert gjennom et tidligere referensielt uttrykk med *samme* referanse eller gjennom et tidligere uttrykk som viser til en referent av samme type uten at det er samreferanse (se 8.4.3). Klassifiseringen av setningsposisjon er basert på om substantivfrasen står i setningens *forfelt* eller *slutfelt*, der jeg med «forfelt» for enkelhets skyld viser til både hovedsetningers forfelt og den faste posisjonen til subjekter i leddsetninger eller andre innføyde setninger (se 8.4.2).

Ifølge hypotesen vil substantivfraser som viser til *ikke-aktiverte* referenter og står i *forfeltet*, ha artikkelfrase, mens substantivfraser som viser til *aktiverte* referenter og står i *slutfeltet*, vil ha bestemt form. De øvrige substantivfrasene vil ha ubestemt entall eller flertall, altså de frasene som viser til *aktiverte* referenter og står i *forfeltet* eller de som viser til *ikke-aktiverte* referenter og står i *slutfeltet*.

Jeg vil under vise eksempler på de fire ulike konteksttypene som kan dannes ved å kombinere tekstuell aktivering og setningsposisjon: 1) ikke-aktiverte forfelt, 2) ikke-aktiverte slutfelt, 3) aktiverte forfelt, og 4) aktiverte slutfelt.⁵⁰ De substantivfrasene som illustrerer konteksttypene, er understreket. Eksempelene er hentet fra innlærertekster skrevet både av polske og engelske innlærere, og den aktuelle tekstens *tekst-id* fra ASK står i parentes.

⁵⁰ Disse benevnelsene (aktivert forfelt, ikke-aktivert slutfelt osv.) kan virke misvisende på den måten at det fremstår som om det er setningsposisjonene forfelt og slutfelt som er aktivert/ikke-aktivert, mens det egentlig er referansen til substantivfrasen som er plassert i disse setningsposisjonene, som er aktivert/ ikke-aktivert. Benevnelsene skal derfor forstås slik at de viser til substantivfraser som befinner seg i forfelt/slutfelt og hvis referanse er aktivert/ikke-aktivert.

Ikke-aktiverte forfelt

De understrekede substantivfrasene i setning 9.1-a og 9.1-b (under) er begge plassert i forfeltet til en hovedsetning, og verken museum eller leger er nevnt i samme eller foregående hovedsetning, derfor regnes disse referentene som ikke-aktiverte.

9.1-a Mange er interessert i ekstra utdanning etter de er ferdig med videregående skole. For eksempel kan de lære kunsthistorie eller de kan lære å tegne og male. **Et museum** blir et godt sted for å skape inspirasjon. (s0297)

9.1-b Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikshospital sykehuset. **Legene** sa at operasjon gikk bra. (s0745)

I den siste av de to setningene (9.1-b) ser vi at «legene» ikke har en aktivert referent, men denne substantivfrasen har likevel definit referanse siden de omtalte legene kan regnes som identifiserbare ved assosiasjon til «sykehuset». Dette eksemplet illustrerer dermed hvordan en substantivfrasens referent kan være identifiserbar (definit) om en legger kriteriene for grammatisk bestemthet til grunn, men *ikke* identifiserbar om en kun legger til grunn tekstuell identifiserbarhet, som er utgangspunktet for pragmatisk ordstilling.

Ikke-aktiverte slutfelt

I de to setningene under (9.1-c og 9.1-d) er substantivfrasene plassert i slutfeltet av henholdsvis en hovedsetning og en infinitivkonstruksjon. Heller ikke «ambulanse» og «vakter» er nevnt tidligere, og disse frasenes referenter regnes derfor også som *ikke-aktiverte*.

9.1-c Først og fremst kan mobiltelefonene hjelpe oss fordi vi har bestandig kontakt med andre. For eksempel: Vi kan bestille **ambulanse** hvis det skjer noe farlig. (s0765)

9.1-d Jeg har arbeidet som sykepleier. Mange av mine kolleger trenger ofte å bytte **vakter** eller de vi informere meg at de f. eks ikke kan komme på jobb. (s0763)

Aktiverte forfelt

De neste to setningene (9.1-e og 9.1-f) viser substantivfraser som er klassifisert som aktiverte. Substantivfrasene «katten» og «kvinner» er plassert initialt i henholdsvis en hovedsetning og en leddsetning, plasseringer som jeg i begge tilfeller har valgt å omtale som «forfelt».

9.1-e To år siden opplevde jeg veldig trist for meg ulikker. Det var min katten som var påkjørt foran mitt hus. **Katten** gikk tilbake fra skogen hjem og ble påkjørt på veien. (s0940)

9.1-f Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at **kvinner** fikk stemmerett. (s0654)

Både «katten» og «kvinner» er klassifisert som aktiverte, men på litt ulike grunnlag. Når det gjelder «katten» er den samme referenten nevnt i forrige setning gjennom substantivfrasen «min katten», og referenten «katten» regnes dermed som aktivert gjennom samreferanse. Substantivfrasen «kvinner» har generisk referanse, da den viser generelt til kvinner, og denne regnes som aktivert på grunnlag av at det i setningen før også refereres generisk til kvinner i frasen «kvinnen possision i samfunnet». Se også 8.4.3 for omtale av disse to forekomstene.

Aktiverte slutfelt

De siste to eksemplene (9.1-g og 9.1-h) viser to substantivfraser som er plassert i slutfeltet, den første i en hovedsetning og den andre i en leddsetning.

9.1-g Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

9.1-h Det er vanligvis lett å få barn, mens alle problemer kommer når vi må oppdra **barna**. Først og fremst er det foreldre som bør oppdra barna. (s0725)

Begge disse eksemplene er klassifisert som aktiverte på grunn av at mobiltelefoner og barn er nevnt tidligere, men i motsetning til eksemplene over, hvor det var snakk om samreferanse mellom to substantivfraser, er det i disse to eksemplene bare en referent *av samme type* som er nevnt før. I det øverste eksemplet viser den understrekede «mobiltelefonen» til individuelle mobiltelefoner som folk har, men i setningen før vises det bare generelt til mobiltelefoner, altså generisk referanse. (Se 8.4.3 for nærmere omtale av dette eksemplet.) Det samme gjelder substantivfrasen «barna» i den nederste setningen, som viser til individuelle barn, altså de individuelle barna til foreldrene som nevnes i samme setning. Referenten til «barna» klassifiseres likevel som aktivert siden det har blitt vist generisk til barn i setningen før.

Eksemplene over har fokusert på tekstuell aktivering og setningsposisjon, og bare til en viss grad på referensiell type. Jeg vil komme tilbake til tilsvarende eksempel og omtale av indefinitte, definite og generiske substantivfraser i de underkapitlene som omhandler de ulike referensielle typene.

9.2 Transfereffekter

I 8.5.3 omtalte jeg hvordan de fire transfereffektene undersøkes generelt for de to hypotesene, og jeg vil her si litt mer spesifikt om transfereffekter for hypotese 1.

9.2.1 Morsmålssamsvar

Transfereffekten *morsmålssamsvar* er formulert som prediksjoner om hvilke formvalg den polske gruppen forventes å ha i ulike språklige kontekster om hypotesen stemmer, altså hvilke former og kontekster som vil være assosiert med hverandre. For hypotese 1 er det predikert assosiasjoner mellom:

1. **Ubestemt entall/flertall og**
 - a) aktiverte referenter i forfeltet
 - b) ikke-aktiverte referenter i slutfeltet
2. **Bestemt entall/flertall og**
aktiverte referenter i slutfeltet
3. **Artikkelfrase og**
ikke-aktiverte referenter i forfeltet

Det er altså prediksjoner om *tre* ulike form/kontekst-assosiasjoner for hypotese 1. Jeg har valgt å gi de tre ulike konteksttypene navn etter hvilken form som forventes ifølge hypotesen: Aktiverte forfelt og ikke-aktiverte slutfelt (1a og 1b over) forventes å ha ubestemt entall/flertall, og blir derfor omtalt som **UB-kontekster**. Aktiverte slutfelt forventes å ha bestemt form, og omtales derfor som **BE-kontekster**, og ikke-aktiverte forfelt forventes å ha artikkelfrase, og omtales derfor som **ART-kontekster**. Navnet på konteksttypene viser altså til *hvilken form som er forventet i denne konteksten hos de polske innlærerne om hypotesen stemmer*, og viser *ikke* til hvilke former som er forventet etter målspåketets norm eller hvilken form som faktisk er brukt.

Grunnen til at jeg har valgt å navngi konteksttypene på denne måten, er at presentasjonen av resultatene vil fokusere på om materialet stemmer med prediksjonene eller ikke i hver og en av disse konteksttypene. Da er det hensiktsmessig å velge et navn som bidrar til at leseren lettere husker hva slags prediksjon som gjelder for den aktuelle konteksttypen. Alternativet hadde vært å kalle konteksttypene «aktivert forfelt», «ikke-aktivert slutfelt» osv., men jeg har vurdert det til at slike benevnelser kan gjøre det krevende for leseren å huske hvilken prediksjon som gjelder til enhver tid. Jeg vil

imidlertid enkelte steder omtale konteksttypene på denne måten også, selv om BE-kontekst, UB-kontekst og ART-kontekst vil være de primære benevnelsene.

De form/kontekst-assosiasjonene som undersøkes videre er altså:

- 1) Assosiasjon mellom UB-kontekster og ubestemt entall/flertall
- 2) Assosiasjon mellom BE-kontekster og bestemt entall/flertall
- 3) Assosiasjon mellom ART-kontekster og artikkel frase

Eksempler på disse form/kontekst-assosiasjonene gis under, der de understrekede substantivfrasene har den formen som er predikert for den aktuelle konteksttypen. Alle setningene er hentet fra polske innlærertekster, og alle eksemplene unntatt 9.2-d sammenfaller med eksemplene i 9.1.

UB-kontekst med ubestemt entall/flertall

De to eksemplene under illustrerer en assosiasjon mellom UB-kontekst og ubestemt entall/flertall. Det øverste eksemplet viser en UB-kontekst av typen *aktivert forfelt* med *ubestemt flertall* (9.2-a), mens det nederste illustrerer en UB-kontekst av typen *ikke-aktivert slutfelt* med *ubestemt entall* (9.2-b). Formvalgene er i tråd med prediksjonen om at UB-kontekster skal ha ubestemt entall/flertall siden slike substantivfraser har en plassering i setningen som gjør at markering av tekstuell identifiserbarhet (aktivert/ikke aktivert) oppleves som redundant.

9.2-a Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at **kvinner** fikk stemmerett. (s0654)

9-2-b Først og fremst kan mobiltelefonene hjelpe oss fordi vi har bestandig kontakt med andre. For eksempel: Vi kan bestile **ambulanse** hvis det skjer noe farlig. (s0765)

BE-kontekst med bestemt form

Det neste eksemplet (9.2-c) illustrerer en assosiasjon mellom BE-kontekst og bestemt form av substantivet. Den understrekede substantivfrasen er klassifisert som *aktivert slutfelt*, og formvalget er i tråd med prediksjonen om at slike kontekster skal ha bestemt form siden substantivfrasen er plassert slik at det oppleves som nødvendig å markere at referenten er identifiserbar i den foregående teksten (aktivert).

9.2-c Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

ART-kontekst med artikkelfrase

Det siste eksemplet (9.2-d, under) illustrerer assosiasjonen mellom en ART-kontekst og artikkelfrase. Substantivfrasen «en by» er plassert *preverbalt* i en innføydd setning, og representerer derfor det jeg omtaler som «forfelt». Det har ikke vært nevnt noe om en by i de foregående setningene og derfor er forekomsten «en by» klassifisert som ikke-aktivert. Formvalget er i tråd med prediksjonen om at ART-kontekster (ikke-aktiverte forfelt) skal ha artikkelfrase siden substantivfrasen er plassert slik at det oppleves som nødvendig å markere at referenten *ikke* er identifiserbar i den foregående teksten (ikke-aktivert).

9.2-d Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan **en by** passer for barn. (s0295)

De fire eksemplene over varierer med hensyn til hvilken type referanse det er snakk om, men som nevnt vil eksempler på predikerte form/kontekst-assosiasjoner for henholdsvis indefinit, definit og generisk referanse bli gitt senere i de aktuelle underkapitlene.

I innlærertekstene til de polske innlærerne ventes det at assosiasjoner mellom kontekst og form skal komme til uttrykk både gjennom at den gitte konteksttypen først og fremst har den predikerte formen, og gjennom at en gitt form først og fremst brukes i den predikerte konteksttypen, og ikke i andre kontekster. Graden av gjensidighet oppgis dermed både som tall som viser prosentandelen av formen **innen kontekst**, altså hvor stor andel av en konteksttype som har den forventede formen, og prosentandelen av konteksten **innen form**, altså hvor stor andel av en formtype som brukes i sin forventede konteksttype.

Prosentandel innen kontekst:

Andelen forekomster av en konteksttype (UB, BE eller ART) som har den predikerte formen.

Prosentandel innen form:

Andelen forekomster av en formtype som er brukt i sin predikerte konteksttype (UB, BE eller ART).

Prediksjonene om morsmålssamsvar undersøkes både på gruppenivå gjennom en vurdering av tall som gjelder for gruppen som helhet, og på individnivå, gjennom en vurdering av om individuelle profiler stemmer med prediksjonene.

Hypotesen gjelder på tvers av referensielle typer, altså på samme måte ved både indefinitt, definitt og generisk referanse, ettersom dette skillet ikke markeres i polsk. Morsmålssamsvaret vil derfor være mulig å påvise i hele det samlede materialet, men som jeg var inne på i omtalen av transfereffekter, er det likevel nødvendig å dele opp materialet etter ulike referensielle typer for å kunne vurdere transfereffektene *intergruppe-heterogenitet* og *intralingval kontrast* (se 8.5.3). Transfereffekten *morsmålssamsvar* er dermed forventet å foreligge både i det samlede materialet og innen hver av de referensielle typene om hypotesen stemmer.

9.2.2 Intragruppe-homogenitet

Siden undersøkelsen på gruppenivå ikke tar hensyn til individuell variasjon, vil transfereffekten *intragruppe-homogenitet* undersøkes kun gjennom opptelling av individuelle profiler. Homogenitet vil kunne komme til syne gjennom en høy andel profiler som har de forventede form/kontekst-assosiasjonene i den polske gruppen (positive individuelle profiler). Dette vil kunne belyse om det predikerte mønsteret er representativt for gruppen. Homogenitet undersøkes bare med tanke på det predikerte morsmålssamsvaret, og er derfor kun aktuelt om det faktisk er morsmålssamsvar i det polske materialet. Andre former for representative formvalg i den polske gruppen blir imidlertid påpekt der det er relevant.

9.2.3 Intergruppe-heterogenitet og intralingval kontrast

Som nevnt i gjennomgangen av den generelle metoden i forrige kapittel, vil både intergruppe-heterogenitet i forhold til den engelske gruppen og intralingval kontrast i forhold til den norske gruppen innebære en inndeling i *sammenfallskontekster* og *markørkontekster*, altså kontekster der de predikerte form/kontekst-assosiasjonene for den polske gruppen er forventet å sammenfalle eller *ikke* sammenfalle med formvalgene til den engelske og den norske gruppen.

Tabell 9-1 (neste side) viser de forventede likhetene og forskjellene i formvalg mellom de tre gruppene innen de ulike referansetyperne. Markørkontekster er uthevet.

Tabell 9-1 Oversikt over forventede formvalg i de tre gruppene

	POLSK	ENGELSK	NORSK
Definitt referanse			
BE-kontekster	Bestemte former	Bestemte former	Bestemte former
ART-kontekster	Artikkelfraser	Bestemte former	Bestemte former
UB-kontekster	Ubestemt ent/flt	Bestemte former	Bestemte former
Indefinitt referanse			
BE-kontekster	Bestemte former	Ubestemte former	Ubestemte former
ART-kontekster	Artikkelfraser	Ubestemte former	Ubestemte former
UB-kontekster	Ubestemt ent/flt	Ubestemte former	Ubestemte former
Generisk referanse			
BE-kontekster	Bestemte former	Alle former mulige	
ART-kontekster	Artikkelfraser		
UB-kontekster	Ubestemt ent/flt		

Ved **definitt referanse** er det bestemt form som er det typiske formvalget i både engelsk og norsk. Siden hypotese 1 predikerer at de polske innlærerne bruker bestemte former i BE-kontekster (aktiverte slutfelt), er det i disse kontekstene forventet sammenfall mellom de tre gruppene. For UB- og ART-kontekster er den polske gruppen derimot forventet å være ulik både den engelske og den norske gruppen, ettersom den polske gruppen ikke er forventet å ha bestemte former i disse kontekstene i det hele tatt.

Ved **indefinitt referanse** er det ubestemte former som er det typiske formvalget i engelsk og norsk, som vanligvis er artikkelfrase og ubestemt flertall ved tellelige substantiv, og ubestemt entall ved ikke-tellelige substantiv, men ikke-tellelige substantiv er ikke inkludert i materialet. Den polske gruppen er spesielt forventet å skille seg fra de andre to i BE-kontekster, hvor den polske gruppen er forventet å ha bestemte former, noe som sjeldent brukes i engelsk og norsk ved indefinitt referanse. Innen indefinitt ART- og UB-kontekster er det derimot forventet et visst sammenfall mellom gruppene, men den polske gruppen vil ha artikkelfraser og ubestemt entall/flertall i ulike kontekster (UB og ART), mens de andre to gruppene forventes å bruke disse formtypene om hverandre. I tillegg er språkene engelsk og norsk ulike her, ettersom ubestemt entall sjelden brukes ved tellelige substantiv i engelsk, noe som er mer vanlig i norsk (se 7.4.2). En vurdering forskjellen mellom engelsk og norsk jeg kommer jeg tilbake til senere.

Siden forholdet mellom kontekst og form er mer sammensatt ved **generisk referanse** enn det er ved definitt og indefinitt i både norsk og engelsk, er det

vanskeligere å si noe om forventninger til likt og ulikt formvalg. Generelt vet vi at ubestemt flertall er mye brukt i både engelsk og norsk, og en kan til en viss grad derfor anta at det vil være en del sammenfall i UB-kontekster, men jeg vil komme tilbake til dette i 9.7, da jeg vil basere meg på den faktiske bruken av former i det engelske og norske materialet.

De forventningene som er satt opp over, er basert på beskrivelsen av engelsk og norsk i kapittel 7, i tillegg til prediksjonene om polsk morsmålssamsvar, men jeg vil også vurdere de *faktiske* formvalgene den engelske og norske gruppen har i materialet, underveis i undersøkelsen. For den norske gruppen er det faktiske formvalget forventet å stemme godt med det som er beskrevet over, siden det er snakk om morsmålsbrukere, men for den engelske gruppen er det ikke like sikkert at de faktiske formvalgene i materialet stemmer med de forventningene som er satt opp over, Dette er fordi det ikke er gitt at formvalgene til de engelske innlærerne er påvirket av morsmålet, noe jeg også var inne på i 8.5.3.3.

9.3 Undersøkelse av hypotese 1 i hele materialet

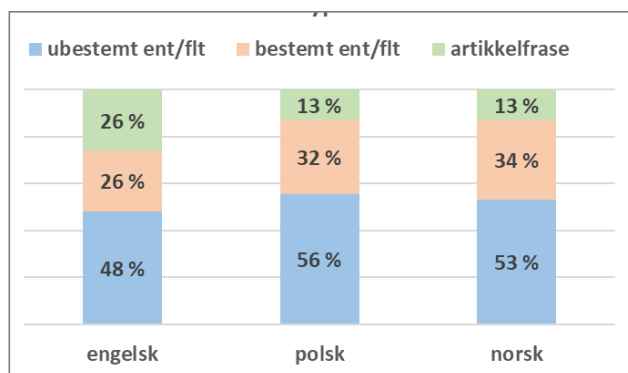
Jeg vil først undersøke transfereffektene for hypotese 1 i hele materialet *samlet*, altså uten å dele opp i referensielle typer. Før jeg ser på selve form/kontekst-assosiasjonene, vil jeg si litt om den generelle forekomsten av former og konteksttyper i datamaterialet.

9.3.1 Deskriptivt for alle gruppene

Det er ulikt antall enkle substantivfraser i de tre gruppene: Den polske gruppen har 317 slike fraser, mens den engelske og den norske gruppen har henholdsvis 285 og 232, altså er det relativt stor forskjell på antallet i de tre gruppene.

Figur 9-1 og tabell 9-2 (øverste neste side) viser andelen av de ulike *formtypene* i de tre gruppene. Alle gruppene har mest ubestemt entall/flertall, men er ellers noe ulik i andelen både av bestemt form og artikkelfrase. I den polske og den norske gruppen er andelen av de tre formtypene veldig like, mens den engelske gruppen skiller seg mer ut, først og fremst med å ha høyere andel artikkelfrase (pol. 13 %; eng. 26 %; nor. 13 %), og tilsvarende lavere andel av ubestemt entall/flertall (pol. 56 %; eng. 48 %, nor. 53 %) og bestemt form (pol. 32 %; eng. 26 %; nor. 34 %).

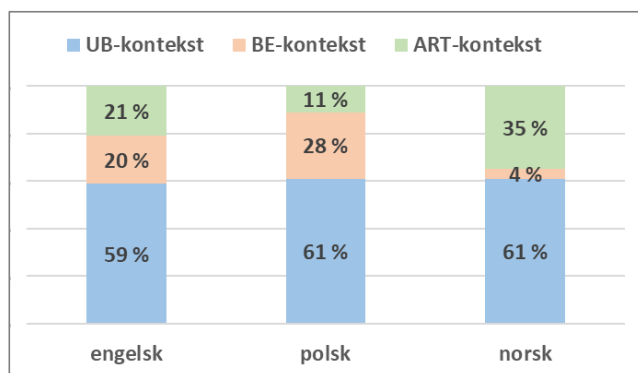
Figur 9-1 Andel av de ulike formtypene i de tre gruppene



Tabell 9-2 Andel av de ulike formtypene i de tre gruppene

Morsmål	Ubestemt ent/flt	Bestemt ent/flt	Artikkelfrase	Totalt
Polsk	56 % (176)	32 % (100)	13 % (41)	317
Engelsk	48 % (137)	26 % (73)	26 % (75)	285
Norsk	53 % (124)	34 % (79)	13 % (29)	232

Figur 9-2 Andel av de ulike konteksttypene i de tre gruppene



Tabell 9-3 Andel av de ulike konteksttypene i de tre gruppene

Morsmål	UB-kontekst	BE-kontekst	ART-kontekst	Totalt
Polsk	61 % (193)	28 % (89)	11 % (35)	317
Engelsk	59 % (168)	20 % (58)	21 % (59)	285
Norsk	61 % (141)	4 % (9)	35 % (82)	232

Figur 9-2 og tabell 9-3 (nederst forrige side) viser andel og antall av de tre *konteksttypene* for hypotese 1, altså UB-, BE- og ART-kontekster. De tre gruppene har en relativt lik andel av UB-kontekster (pol. 61 %; eng. 59 %; nor. 61 %), men skiller seg fra hverandre når det gjelder BE-kontekster og ART-kontekster. Den norske gruppen har betraktelig lavere andel BE-kontekster (pol. 28 %; eng. 20 %, nor. 4 %), og tilsvarende høyere ande ART-kontekster (pol. 11 %; eng. 21 %; nor. 35 %).

Antallet BE-kontekster i det norske materialet er det eneste antallet som er spesielt lavt i det samlede materialet, det er kun snakk om *ni* BE-kontekster (aktiverte slutfelt). Utover dette ligger antallet av de ulike formtypene mellom 29 og 176, og antallet av de ulike konteksttypene mellom 35 og 193.

9.3.2 Morsmålssamsvar i den polske gruppen

I gjennomgangen av det samlede materialet vil jeg først ta for meg prediksjonene om morsmålssamsvar i den polske gruppen ved å se på hver og en av de predikerte form/kontekst-assosiasjonene, og deretter sammenligne med de andre to gruppene. Som jeg gjorde rede for i 9.2.1, predikerer hypotese 1 assosiasjoner mellom:

- 1) UB-kontekster og ubestemt entall/flertall
- 2) BE-kontekster og bestemt entall/flertall
- 3) ART-kontekster og artikkelfrase

Om det foreligger en assosiasjon mellom form og kontekst etter de prediksjonene hypotesen innebærer, vil det komme til uttrykk både gjennom at den gitte formtypen er den mest utbredte innen sin konteksttype, og at det først og fremst er i denne konteksten at formen brukes, ikke i andre kontekster.

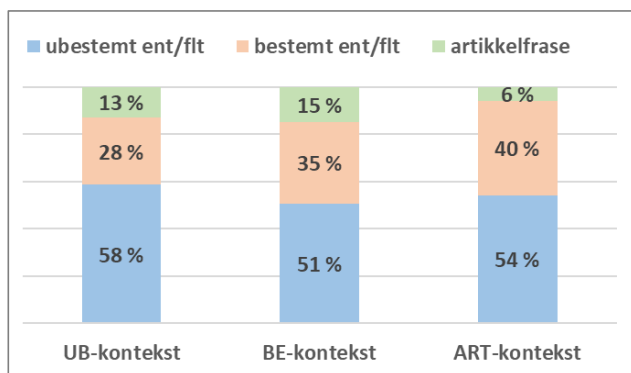
Tabellen på neste side og de påfølgende søylediagrammene viser fordelingen av former både *innen kontekst* og *innen form*, altså hvilke former som brukes innen en konteksttype og hvilke konteksttyper de ulike formene brukes i. I tabellen er UB-konteksten delt opp i de to ulike variantene (aktiverte forfelt og ikke-aktivert slutfelt). I gjennomgangen vil disse to i hovedsak bli behandlet samlet som UB-kontekst, men noen steder vil jeg likevel kommentere forskjeller mellom de to variantene.

Tabell 9-4 Polsk gruppe: Fordeling av former etter konteksttype for hypotese 1

S1=POLSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekst: Form:	Aktivert forfelt	Ikke-aktivert sluttfelt	Aktivert sluttfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	25	87	45	19	176
Andel innen form	14 %	49 %	26 %	11 %	100 %
Andel innen kontekst	32 %	75 %	51 %	54 %	56 %
BESTEMT ENT/FLT	49	6	31	14	100
Andel innen form	49 %	6 %	31 %	14 %	100 %
Andel innen kontekst	64 %	5 %	35 %	40 %	32 %
ARTIKKELFRASE	3	23	13	2	41
Andel innen form	7 %	56 %	32 %	5 %	100 %
Andel innen kontekst	4 %	20 %	15 %	6 %	13 %
Totalt	77(24%)	116 (37%)	89 (28%)	35 (11%)	317

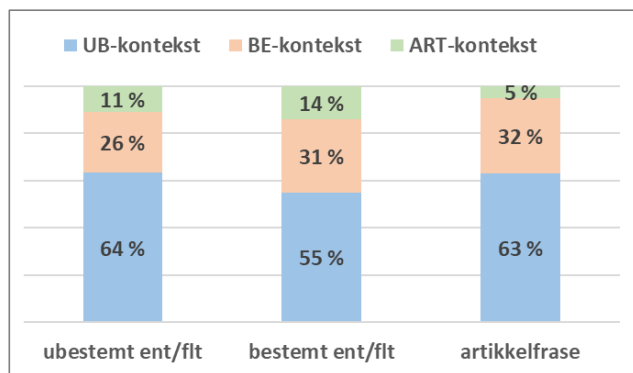
Figur 9-3 (under) viser hvilke former som brukes i de ulike konteksttypene, og det er tydelig at fordelingen av former er relativt lik i alle de tre konteksttypene i det polske materialet: Det er brukt mest ubestemt entall/flertall (51-58 %) både i UB-, BE- og ART-kontekster, nest mest bestemt form (28-40 %), og minst artikkel frase (6-15 %). Dette sier oss allerede at formvalgene til den polske gruppen strider imot prediksjonene, som sier at disse tre ulike konteksttypene fører til *ulike* formvalg.

Figur 9-3 Andel form innen kontekst (polsk gruppe)



Figur 9-4 (under) viser hvilke konteksttyper de ulike formene er brukt i, og her kommer det fram at alle de tre formtypene brukes relativt likt i alle tre konteksttypene: Alle tre formene brukes oftest i UB-kontekster (55-64 %), nest mest i BE-kontekster (26-32 %), og minst i ART-kontekster (5-14 %). Disse observasjonene er heller ikke i tråd med prediksjonene om ulike formvalg i de ulike konteksttypene.

Figur 9-4 Andel kontekst innen form (polsk gruppe)



Jevnheten som blir observert, tyder på at denne typen språklige kontekster ikke er styrende for formvalget i de polske tekstene, så allerede nå ser vi at morsmålsansvaret i den polske gruppen sannsynligvis er lavt, og at den type transfer som hypotese 1 handler om, mest sannsynlig ikke foreligger. Det er likevel litt tidlig å avvise både transfereffekten morsmålsansvar og transfer i seg selv, derfor vil jeg videre se nærmere på hver og en av de tre form/kontekst-assosiasjonene som er predikert for det polske materialet.

9.3.2.1 Assosiasjon mellom UB-kontekster og ubestemt entall/flertall

I UB-kontekster (aktiverte forfelt og ikke-aktiverte sluttfelt) er det ifølge hypotesen forventet at den polske gruppen bruker *ubestemt entall/flertall* ettersom markering av tekstuell identifiserbarhet her antas å oppleves som redundant. Dette gjør at verken bestemt form eller artikkelfrase vil tas i bruk i disse konteksttypene. Det er videre forventet at det er først og fremst i UB-kontekster at ubestemt entall/flertall brukes, ikke i andre konteksttyper.

De fire eksemplene under viser setninger der den understrekede substantivfrasen står i en UB-kontekst, men hvor formvalgene er ulike. De to første eksemplene viser

fraser hvor formvalget er i tråd med prediksjonen for UB-kontekster, mens de to siste viser fraser hvor formvalget *ikke* er i tråd med prediksjonen.

Substantivfrasene «kvinner» og «ambulanse» i eksempel 9.3-a og 9.3-b (under) ble også omtalt i 9.1, og illustrerer henholdsvis et *aktivert forfelt* og et *ikke-aktivert slutfelt*, som begge regnes som UB-kontekster. Formene som er brukt, er henholdsvis ubestemt flertall og ubestemt entall uten artikkel, noe som er i tråd med prediksjonen for UB-kontekster.

9.3-a Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at **kvinner** fikk stemmerett. (s0654)

9.3-b Først og fremst kan mobiltelefonene hjelpe oss fordi vi har bestandig kontakt med andre. For eksempel: Vi kan bestile **ambulanse** hvis det skjer noe farlig. (s0765)

Substantivfrasen «katten» i eksempel 9.3-c (under), er omtalt tidligere, og illustrerer en UB-kontekst av typen *aktivert forfelt*, der formen som er valgt er *bestemt form*, noe som *ikke* er i tråd med prediksjonen for UB-kontekster.

9.3-c To år siden opplevde jeg veldig trist for meg ulikker. Det var min katten som var påkjørt foran mitt hus. **Katten** gikk tilbake fra skogen hjem og ble påkjørt på veien. (s0940)

Substantivfrasen «en påfillingskort» i eksempel 9.3-d (under) står i slutfeltet til en hovedsetning, og påfillingskort har ikke tidligere vært nevnt, derfor regnes denne substantivfrasens referent som ikke-aktivert. Dette er dermed en UB-kontekst av typen *ikke-aktivert slutfelt*, men det er en artikkelfrase, noe som ikke er i tråd med prediksjonen for UB-kontekster.

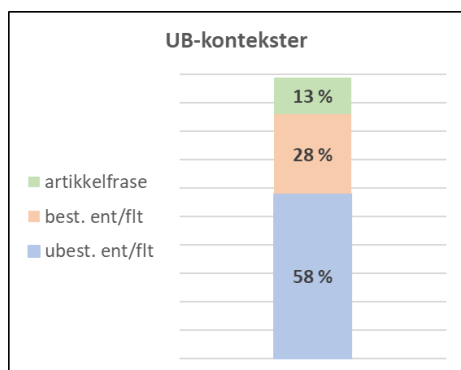
9.3-d Jeg bruker mobiltelefonen ikke ofte, bare av og til. Jeg kjøper **en påfilingkort** månedlig. (s0772)

Søylediagrammene på neste side illustrerer forholdet mellom UB-kontekster og ubestemt entall/flertall i det polske materialet.

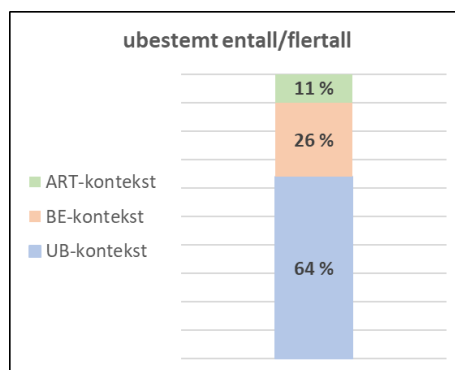
Figur 9-5 (til venstre) viser hvilke former som er brukt i UB-kontekster i de polske innlærertekstene, og det er brukt ubestemt entall/flertall i 58 % av tilfellene. Det er altså ikke noen konsekvent bruk av slike former, men det er likevel en overvekt i forhold til de to andre formtypene (be. 28 %; art. 13 %).

Figur 9-6 (til høyre) viser videre at de fleste frasene med ubestemt entall/flertall er brukt i UB-kontekster (64 %), noe som forteller at denne formtypen ikke utelukkende er brukt i den predikerte konteksttypen, men likevel oftere der enn i de øvrige (BE. 26 %; ART. 11 %).

Figur 9-5 Andel form innen UB-kontekst



Figur 9-6 Andel kontekst innen ub. ent/flt.



Assosiasjonen mellom ubestemt entall/flertall og UB-kontekster er altså til stede når en tar i betraktning både kontekst og form, men mer som en tendens, og på ingen måte kategorisk, ettersom andre former også brukes i UB-kontekster og ubestemt entall/flertall også brukes i andre steder.

Om vi går tilbake til tabell 9-4 (s. 229) og ser nærmere på de to ulike konteksttypene som utgjør UB-kontekster, altså aktiverte forfelt og ikke-aktiverte slutfelt, er det her en *ujevn* fordeling av ubestemt form: Ved *ikke-aktiverte slutfelt* er det 75 % ubestemt entall/flertall, mens ved *aktiverte forfelt* er det bare 32 %. Siden begge disse konteksttypene er forventet å ha ubestemt entall/flertall, stemmer det dårlig med prediksjonene at de to undertypene har ujevn bruk av denne formtypen. I tillegg ser vi at aktiverte forfelt har høyere andel av bestemt form enn ubestemt entall/flertall (ub. 32 %; be. 64 %), så innad i UB-kontekstene er det en fordeling med overvekt av bestemt form i aktiverte kontekster og overvekt av ubestemt i ikke-aktiverte kontekster. I tabellen ser vi også at ubestemt entall/flertall er brukt i *slutfelt* i til sammen 75 % av tilfellene. Det er snakk om 49 % i ikke-aktiverte slutfelt og 26 % i aktiverte slutfelt, altså oftere i slutfelt enn i UB-kontekster (58 %).

For prediksjonen om en assosiasjon mellom UB-kontekster og ubestemt entall/flertall er det altså en tendens til assosiasjon mellom forventet form og kontekst (58 % innen kontekst/64 % innen form), men det kan også se ut til at ubestemt entall/flertall har en sterkere assosiasjon med slutfeltet enn med UB-kontekster, og når en UB-kontekst som står i forfeltet er aktivert, er det derimot bestemt form som dominerer.

9.3.2.2 Assosiasjon mellom BE-kontekster og bestemte former

I BE-kontekster (aktiverte slutfelt) er det ifølge hypotesen forventet at den polske gruppen skal bruke *bestemt form*, ettersom det er snakk om aktiverte slutfelt, en type kontekst hvor det kan oppleves som nødvendig å markere at referenten er tekstuell identifiserbar (aktivert). Det er også ventet at det er i BE-kontekster at bestemt form først og fremst brukes, og ikke i andre konteksttyper.

De tre eksemplene under viser alle setninger der den understrekede substantivfrasen står i en BE-kontekst, men hvor formvalgene er ulike. Substantivfrasen i det første eksemplet har en form som er i tråd med prediksjonen for BE-kontekster, mens substantivfrasene i de to andre eksemplene har former som *ikke* er i tråd med prediksjonen.

Det første eksemplet (9.3-e) har vært omtalt tidligere, og viser en substantivfrase som er klassifisert som *aktivert slutfelt*, altså en BE-kontekst. Substantivet har bestemt form og er derfor i tråd med prediksjonen for denne konteksttypen.

9.3-e Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

I eksempel 9.3-f (under) finner vi substantivet «mobiltelefon» som står i slutfeltet til en innføyd setning. Siden mobiltelefoner også nevnes i setningen før, regnes denne som aktivert. Frasen «mobiltelefon» er dermed et *aktivert slutfelt* (BE-kontekst), men substantivet har ubestemt form og dermed er ikke formvalget i tråd med prediksjonen for BE-kontekster.

9.3-f Så nåværende mobiltelefoner er mye bedre enn jeg hadde ventet. Og sannsynligvis det er bare jeg som har ikke **mobiltelefon** i dag. (s0791)

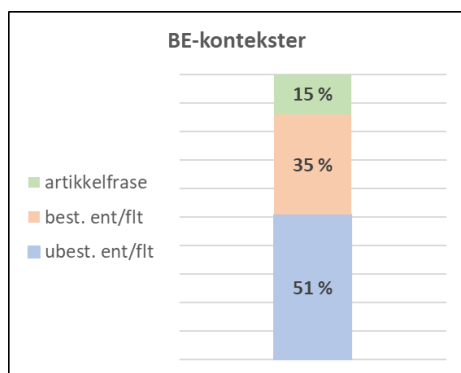
Det samme gjelder «en mobiltelefon» i setning 9-3-g (under), som står i slutfeltet til en infinitivkonstruksjon. Denne referenten er klassifisert som aktivert fordi mobiltelefon er nevnt tidligere gjennom frasen «mobiltelefonen min», selv om de to frasene ikke er samreferente. «En mobiltelefon» er dermed et eksempel på et aktivert slutfelt (BE-kontekst), men det er en artikkel frase, og dermed er ikke formvalget i tråd med prediksjonen for BE-kontekster.

9.3-g Jeg reiser veldig ofte og det er mye lettere for meg hvis jeg kan bruke mobiltelefonen min. Det skjer også mange rare situasjoner i vår liv når det er bra å ha en mobiltelefon med – for eksempel bilulykker. (s0983)

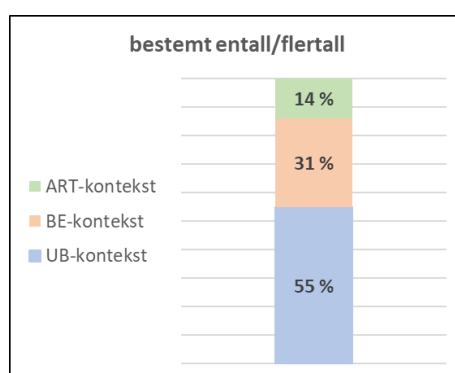
Søylediagrammene under illustrerer forholdet mellom BE-kontekster og bestemt form i det polske materialet.

Figur 9-7 (til venstre) viser hvilke former som er brukt i BE-kontekster, og vi ser at det kun er 35% av BE-kontekstene som har bestemt form, mens hele 51 % har ubestemt entall/flertall og 15 % artikkel frase. Vi ser videre av figur 9-8 (til høyre), som viser hvilke kontekster bestemt form er brukt i, at kun 31 % av frasene med bestemt form brukes i BE-kontekster. Bestemt form brukes derimot oftest i UB-kontekster (55 %). Assosiasjonen mellom bestemt form og BE-kontekster er dermed ganske svak i det polske materialet både innen kontekst og innen form.

Figur 9-7 Andel form innen BE-kontekst



Figur 9-8 Andel kontekst innen best. form



Om vi ser tilbake på tabell 9-4 (s. 229) kommer det fram at bestemt form i størst grad er brukt i *aktiverte forfelt* (49 %), og dernest i *aktiverte slutfelt* (31 %). Det ser dermed ut som aktiveringsgrad, uavhengig av plasseringen i setningen, i større grad sammenfaller

med bruk av bestemt form enn fravær av redundans i aktiverte slutfelt gjør, ettersom det i til sammen 80 % av aktiverte kontekster er brukt bestemt form, mens det bare er brukt i 35 % av BE-kontekstene.

Assosiasjon mellom BE-kontekster og bestemt form er altså svakere enn mellom UB-kontekster og ubestemt entall/flertall, og dermed er også morsmålssamsvaret svakere for denne delprediksjonen (35 % innen kontekst/31 % innen form). Det kan dessuten se ut til at bruk av bestemt form har en sterkere knytning til tekstuell aktiverte substantivfraser uavhengig av plassering i setningen, enn til de forholdene som hypotese 1 predikerer.

9.3.2.3 Assosiasjon mellom ART-kontekster og artikkelfraser

I ART-kontekster (ikke-aktiverte slutfelt) er det ifølge hypotese 1 forventet at den polske gruppen bruker *artikkelfrase* ettersom det er snakk om ikke-aktiverte forfelt, en type kontekst hvor det kan oppleves som nødvendig å markere at referenten *ikke* er tekstuell identifiserbar (ikke-aktivert). Det er dessuten forventet at det er i ART-kontekster at artikkelfrase først og fremst brukes, ikke i andre kontekster.

De tre eksemplene under viser alle setninger der den understrekede substantivfrasen utgjør en ART-kontekst, men hvor formvalgene er ulike. Substantivfrasen i det første eksemplet har en form som er i tråd med prediksjonen for ART-kontekster, mens substantivfrasene i de to andre eksemplene har former som *ikke* er i tråd med prediksjonene.

Substantivfrasen «en by» i eksempel 9.3-h (under) har vært omtalt tidligere i 9.2 og er klassifisert som et *ikke-aktivert slutfelt*, altså en ART-kontekst. Dette er en artikkelfrase og formvalget er derfor i tråd med prediksjonen for ART-kontekster.

9.3-h Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan **en by** passer for barn. (s0295)

De to substantivfrasene «legene» og «lege» i eksempel 9.3-i og 9.3-j står begge i forfeltet til en hovedsetning, og begge er klassifisert som ikke-aktiverte ettersom leger ikke har vært nevnt tidligere. Substantivfrasene har henholdsvis bestemt form og ubestemt entall uten artikkel, og formvalgene er dermed ikke i tråd med prediksjonen for ART-kontekster

9.3-i Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikhospital sykehuset. **Legene** sa at operasjon gikk bra. (s0745)

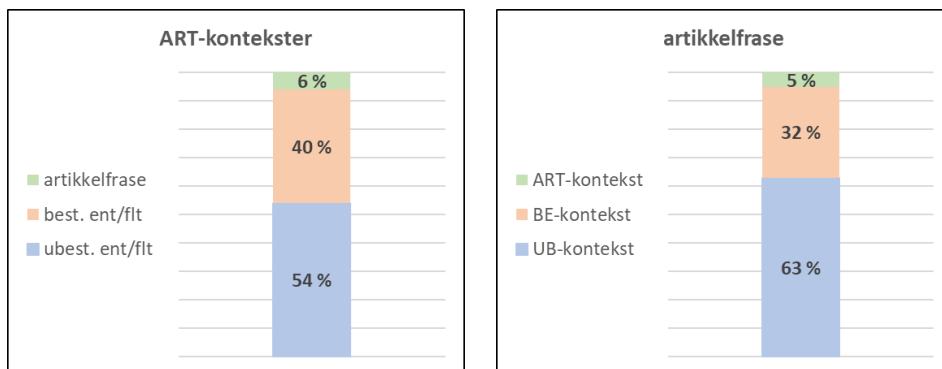
9.3-j Det sjedde noen dager siden i Krakow. Fire andre studenter er på sykehus men **lege** sier at de blir friske. (s0793)

Søylediagrammene under illustrerer forholdet mellom ART-kontekster og artikkelfrase i det polske materialet.

Assosiasjonen *ART-kontekst/artikkelfrase* viser seg å være mye svakere enn det vi så for de to andre delprediksjonene. Figur 9-9 (til venstre) viser hvilke former som er brukt i ART-kontekster, og bare 6 % av ART-kontekstene har artikkelfrase, mens det derimot er 54 % som har ubestemt entall/flertall og 40 % som har bestemt form.

Figur 9-10 (til høyre) viser hvilke kontekster artikkelfrase er brukt i. Bare 5 % av alle artikkelfrasene er brukt i ART-kontekster, og formen er langt oftere brukt i UB-kontekster (63 %) og BE-kontekster (32 %).

Figur 9-9 Andel form innen ART-kontekst **Figur 9-10 Andel kontekst innen art.frase**



Om vi ser tilbake på tabell 9-4 (s. 229) kan vi også se at artikkelfrase er oftest brukt ved *ikke-aktiverte slutfelt* (56 %) og *aktiverte slutfelt* (32 %). Bruk av artikkelfrase ser dermed ut til å i større grad være knyttet til *slutfeltet* uavhengig av tekstuell aktivering, heller enn til ART-kontekster, siden hele 88 % av artikkelfrasene forekommer i slutfelt, mens bare 6 % forekommer i ART-kontekster.

Det ser altså ikke ut til å være noen assosiasjon mellom artikkelfrase og ART-kontekster i det polske materialet (6 % innen kontekst/5 % innen form), men derimot det svakeste morsmålssamsvaret i materialet så langt. Det ser også ut til at bruk

av artikkelfrase i større grad er knyttet til slutfeltet enn til de forholdene hypotese 1 predikerer.

9.3.2.4 *Polsk morsmålssamsvar - foreløpig oppsummering*

I det samlede polske materialet er det altså kun assosiasjonen mellom UB-kontekster og ubestemt entall/flertall som viser tendenser som kan ligne på den type morsmålssamsvar som er predikert av hypotese 1. For prediksjonen om BE-kontekster er assosiasjonen svakere, og aller svakest er det for ART-kontekster, hvor formvalgene i den polske gruppen vanskelig kan sies å være uttrykk for det predikerte morsmålssamsvaret.

Om en ser helhetlig på gruppedataene for den polske gruppen, er det altså ingen sterke indikasjoner på morsmålssamsvar. Selv om delprediksjonen om UB-kontekster viser et visst morsmålssamsvar, gjelder det i liten grad de to andre delprediksjonene, og om hypotesen hadde stemt, ville en forventet et relativt høyt morsmålssamsvar for *samtlig*e tre delprediksjoner. Det er dessuten sterkere indikasjoner på assosiasjoner mellom ubestemte former generelt og slutfeltet, og bestemte former og tekstuelte aktiverte substantivfraser.

Funnene så langt er imidlertid kun basert på data fra gruppen som helhet, uten å ta i betraktning individuell variasjon, og jeg vil komme tilbake til individuelle profiler etter å ha sammenlignet gruppetallene med den engelske og den norske gruppen.

9.3.3 *Sammenligning med engelsk og norsk gruppe*

Jeg vil nå sammenligne funnene for den polske gruppen med den engelske og den norske gruppen med tanke på transfereffekten *intergruppe-heterogenitet* i det samlede materialet. Selv om materialet enda ikke er differensiert etter referensiell type for å tydeliggjøre områder for sammenfall og ulikhet mellom gruppene, kan det uddifferensierte materialet likevel si oss noe om den generelle forskjellen mellom de polske og de engelske innlærerne, og i hvor stor grad formvalgene til den polske gruppen skiller seg fra det en finner i målspråksgruppen, med det forbehold at mangel på differensiering av materialet kan innebære en tilsløring av både forskjeller og likheter mellom gruppene (se. 8.5.3.5).

Tabell 9-5 Polsk gruppe: Fordeling av former etter konteksttype for hypotese 1

S1=POLSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	25	87	45	19	176
Andel innen form	14 %	49 %	26 %	11 %	100 %
Andel innen kontekst	32 %	75 %	51 %	54 %	56 %
BESTEMT ENT/FLT	49	6	31	14	100
Andel innen form	49 %	6 %	31 %	14 %	100 %
Andel innen kontekst	64 %	5 %	35 %	40 %	32 %
ARTIKKELFRASE	3	23	13	2	41
Andel innen form	7 %	56 %	32 %	5 %	100 %
Andel innen kontekst	4 %	20 %	15 %	6 %	13 %
Totalt	77(24%)	116 (37%)	89 (28%)	35 (11%)	317

Tabell 9-6 Engelsk gruppe: Fordeling av former etter konteksttype for hypotese 1

S1=ENGELSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	30	53	30	24	137
Andel innen form	22 %	39 %	22 %	18 %	100 %
Andel innen kontekst	49 %	50 %	52 %	41 %	48 %
BESTEMT ENT/FLT	23	10	8	32	73
Andel innen form	32 % E	14 %	11 %	44 %	100 %
Andel innen kontekst	38 %	9 %	14 %	54 %	26 %
ARTIKKELFRASE	8	44	20	3	75
Andel innen form	11 %	59 %	27 %	4 %	100 %
Andel innen kontekst	13 %	41 %	34 %	5 %	26 %
Totalt	61(21%)	107 (38%)	58 (20%)	59 (21%)	285

Tabell 9-7 Norsk gruppe: Fordeling av former etter konteksttype for hypotese 1

S1= NORSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	total
UBESTEMT ENT/FLT	12	74	3	35	124
Andel innen form	10 %	60 %	2 %	28 %	100 %
Andel innen kontekst	55 %	62 %	33 %	43 %	53 %
BESTEMT ENT/FLT	9	27	6	37	79
Andel innen form	11 %	34 %	8 %	47 %	100 %
Andel innen kontekst	41 %	23 %	67 %	45 %	34 %
ARTIKKELFRASE	1	18	0	10	29
Andel innen form	3 %	62 %	0 %	34 %	100 %
Andel innen kontekst	5 %	15 %	0 %	12 %	13 %
Totalt	22 (9%)	119 (51%)	9 (4%)	82 (35%)	232

9.3.4 Fordeling av former i samlet materiale

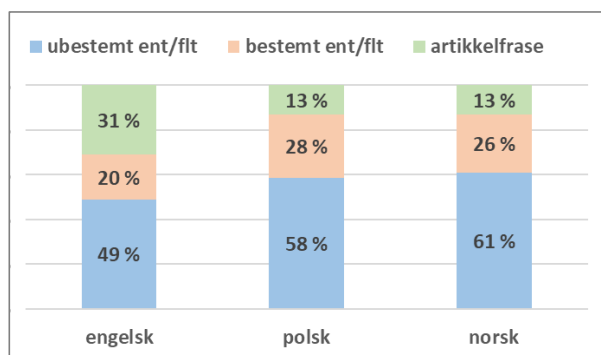
Tabellene på forrige side (9-5 til 9-7) viser fordelingen av former etter konteksttype i de tre gruppene. I tabellene er UB-konteksten delt opp i de to ulike variantene som er predikert å ha ubestemt entall/flertall, altså *aktiverte forfelt* og *ikke-aktivert slutfelt*. I gjennomgangen vil disse to i hovedsak være slått sammen, men noen steder vil jeg likevel kommentere forskjeller innad i UB-konteksten. Tabellene viser også prosentandelen både «innen kontekst» og «innen form», altså hvor ofte en form brukes innen en viss konteksttype i forhold til andre former, og hvor ofte en form er brukt i en viss kontekst i forhold til i andre kontekster (se kap. 8.5.1). Tallene i tabellene vil også bli framstilt gjennom søylediagrammer med søylen for den polske gruppen i midten.

9.3.4.1 UB-kontekster og ubestemt entall/flertall

I undersøkelsen av morsmålsamsvar i den polske gruppen over kom det fram at det for delprediksjonen om UB-kontekster og ubestemt entall/flertall var en viss grad av morsmålsamsvar ettersom 58 % av UB-kontekstene hadde ubestemt entall/flertall, og 64 % av frasene med ubestemt entall/flertall var brukt i UB-kontekster.

Figur 9-11 (under) viser hvilke former som er brukt i UB-kontekster, og vi ser at det er en relativ lik fordeling av former innen UB-kontekster i de tre gruppene. Spesielt gjelder det den polske og den norske gruppen, som har en nesten lik fordeling av former (pol./nor.: ub. 58/61 %; be. 28/26 %; art.13/13 %), mens den engelske gruppen skiller seg ut med å ha en merkbart større andel artikkel frase, og tilsvarende lavere andel av de andre formene (ub. 49 %; be. 20 %; art. 31 %).

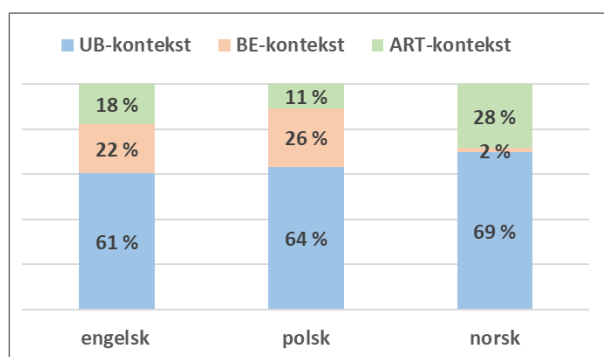
Figur 9-11 Andel av form innen UB-kontekster



Tendensen til assosiasjon mellom UB-kontekster og ubestemt entall/flertall i den polske gruppen er altså i stor grad sammenfallende med det en finner i den norske gruppen, dermed er det lite heterogenitet mellom den polske gruppen og målspråksgruppen på dette punktet. Den engelske og den polske gruppen viser derimot en viss heterogenitet, da den polske har en noe større andel av ubestemt entall/flertall i UB-kontekster (pol. 58 %; eng. 49 %), og den engelske gruppen har en større andel av artikkelfrase (pol. 12 %; eng. 21 %). Med tanke på at den norske og den polske gruppen her er såpass jevne, kan det se ut til at det egentlig er den engelske gruppen som skiller seg ut, og at forskjellen mellom den polske og den engelske gruppen er knyttet til en høyere og gruppespesifikk bruk av artikkelfrase i den engelske gruppen, og ikke til gruppespesifikk bruk av ubestemt entall/flertall i den polske gruppen.

Figur 9-12 (under) viser hvilke kontekster ubestemt entall/flertall er brukt i, og vi ser at de tre gruppene er like med tanke på at de fleste frasene med slik form brukes i UB-kontekster (pol. 64 %; eng. 61 %; nor. 69 %). Det er dermed ingen nevneverdig forskjell mellom den polske gruppen og de to andre gruppene når det kommer til hvilke kontekster disse formene brukes i, og heterogeniteten mellom gruppene må derfor sies å være liten også innen form.

Figur 9-12 Andel av kontekst innen ubestemt ent/flt



Tendensen til assosiasjon mellom UB-kontekster og ubestemt entall/flertall virker med andre ord ikke å være spesifikk for den polske gruppen ettersom en slik tendens fins både i den engelske og den norske gruppen (pol. 58/64 %; eng. 49/61 %; nor. 61/69 %). Dette indikerer at formvalgene er relativt like for de to innlærergruppene, i tillegg til å

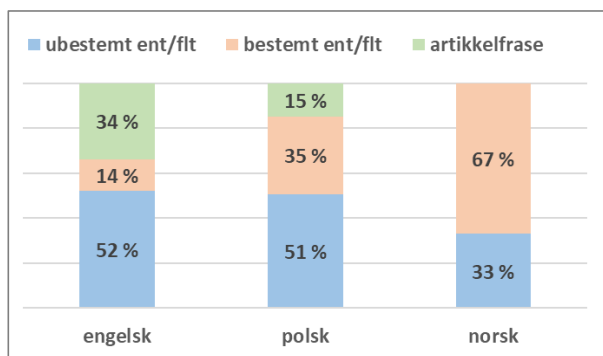
være sammenfallende med norsk, noe som gjør det vanskelig å utelukke både forhold som er universelt for innlærere og målspråkstilleggelse som styrende for formvalgene i den polske gruppen.

9.3.4.2 BE-kontekster og bestemte former

I undersøkelsen av morsmålssamsvar i den polske gruppen over kom det fram at det var en relativt svak kobling mellom BE-kontekster og bestemt form: 35 % av BE-kontekstene hadde bestemt form, mens bare 31 % av substantivfrasene med bestemt form ble brukt i BE-kontekster.

Figur 9-13 (under) viser hvilke former som er brukt i BE-kontekster i de tre gruppene. Om vi sammenligner bruken av bestemt form, ser vi at den polske gruppen skiller seg fra både den norske og den engelske gruppen med å ha mer bruk av bestemt form enn den engelske, men langt mindre enn den norske (pol. 35 %; eng. 14 %; nor. 67 %). I den grad en assosiasjon mellom BE-kontekster og bestemt form er spesifikt for den polske gruppen i forhold til den engelske, er likevel den norske gruppen *mer* i tråd med prediksjonene for den polske gruppen enn det den polske gruppen selv er.

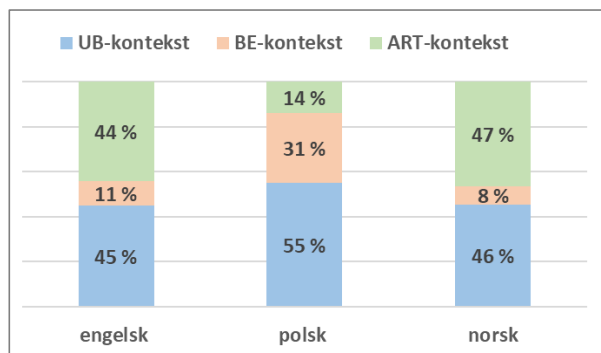
Figur 9-13 Andel av form innen BE-kontekster



Figur 9-14 (neste side) viser hvilke kontekster bestemt form er brukt i, og også her er gruppene forskjellige. Det er den polske gruppen som oftest bruker frasene med bestemt form i BE-kontekster (31 %), mens de andre to gruppene gjør dette mer sjeldent (eng. 11 %; nor. 8 %). Den polske gruppen er altså forskjellig fra de to andre gruppene også innen form, men morsmålssamsvaret er fremdeles lavt, og bestemt form er langt

oftere brukt i UB-kontekster (55 %) hos de polske innlærerne, noe som gjelder de andre to gruppene også (nor. 46 %; eng. 45 %).

Figur 9-14 Andel av kontekst innen bestemt form



Om vi ser på formvalg på andre områder enn bare de predikerte form/kontekst-assosiasjonene, er det den engelske og den polske gruppen som er mest lik, ettersom de begge bruker alle tre formtypene i BE-kontekstene, mens den norske gruppen ikke bruker artikkelfrase her i det hele tatt. Den engelske og den polske gruppen er derimot også her ulike i at den engelske gruppen har høyere andel artikkelfrase (engelsk 34 %, polsk 15 %), og tilsvarende lavere andel bestemt form. Igjen er det funn som tyder på at den engelske gruppen har formvalg som er spesifikke for denne gruppen sammenlignet med både den polske og den norske gruppen.

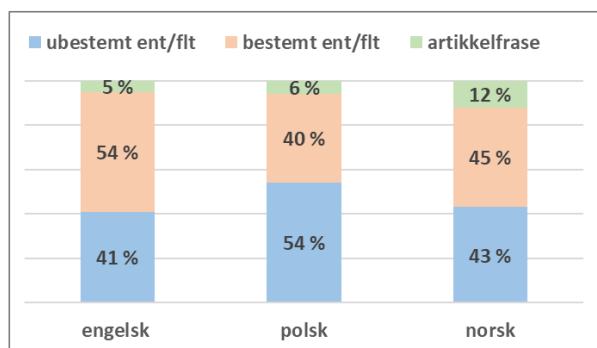
For koblingen mellom BE-kontekster og bestemt form finner vi altså lavt morsmålssamsvar i det polske materialet (35 % innen kontekst/31 % innen form), og dette er ledsaget av en viss heterogenitet i forhold til de andre gruppene (eng. 14 %/11 %, nor. 67 %/8 %). Heterogeniteten mellom gruppene kommer likevel til uttrykk på ulike måter: Forskjellen mellom den norske og den polske gruppen innebærer at den norske gruppen er *mer* i tråd med prediksjonene for den polske gruppen enn den polske gruppen selv er, mens forskjellen mellom de to innlærergruppene ser ut til å være knyttet til høyere andel av artikkelfrase i den engelske gruppen, og tilsvarende lavere andel ubestemt entall/flertall.

9.3.4.3 ART-kontekster og artikkel frase

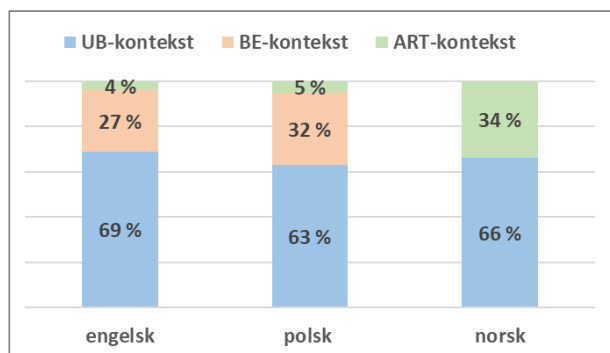
I undersøkelsen av morsmålsamsvar i den polske gruppen over kom det fram at assosiasjonen mellom ART-kontekster og artikkel frase var lav. Det var kun 6 % av ART-kontekstene som hadde artikkel frase, kun 5 % frasene med artikkel sto i ART-kontekster. Det var dermed lavt morsmålsamsvar for denne assosiasjonen.

Som vi ser av figur 9-15 og figur 9-16 (under), er den begrensede bruken av artikkel frase i ART-kontekster felles for de to innlærergruppene både innen kontekst (pol. 6 %; eng. 5 %) og innen form (pol. 5 %; eng. 4 %). Den norske gruppen har derimot en litt sterkere assosiasjon (12 % innen kontekst/34 % innen form). Det er dermed lav intergruppe-heterogenitet mellom den polske og den engelske gruppen, mens det er en viss heterogenitet mellom den polske og den norske. Også her er den norske gruppen noe *mer* i samsvar med prediksjonene for polsk morsmålsamsvar enn den polske gruppen, men assosiasjonen er ikke spesielt sterk i den norske gruppen heller.

Figur 9-15 Andel av form innen ART-kontekster



Figur 9-16 Andel av kontekst innen artikkel fraser



For koblingen mellom artikkelfrase og ART-kontekster er det altså veldig lavt morsmålsamsvar i den polske gruppen (6 % innen kontekst/ 5 % innen form), noe som er ledsaget av lav intergruppe-heterogenitet i forhold til den engelske gruppen (5 % innen kontekst/4 % innen form), og at den norske gruppen er mer i tråd med prediksjonene enn den polske (12 % innen kontekst/34 % innen form). Gruppene er imidlertid like alle tre med tanke på at ingen av dem har noen særlig sterk assosiasjon mellom artikkelfrase og ART-kontekster. I motsetning til det vi så ved de to andre form/kontekst-assosiasjonene, er det ingen spesielle forskjeller mellom de to innlærergruppene når det gjelder forholdet mellom artikkelfrase og ART-kontekster.

9.3.4.4 Oppsummering av funn på gruppenivå

I undersøkelsene av hypotese 1 på gruppenivå virker tendensen å være at den av de tre form/kontekst-assosiasjonene som har høyest samsvar med prediksjonene i den polske gruppen (UB-kontekst og ubestemt entall/flertall), også er det området hvor den polske gruppen er *mest lik* de andre to gruppene. For delprediksjonene om bestemte former og artikkelfraser, har den polske gruppen liten grad av morsmålsamsvar, og her er gruppen først og fremst ulik den norske gruppen, som dessuten har et sterkere samsvar med de polske prediksjonene enn det den polske gruppen selv har.

Om hypotesen stemmer, forventes det at den polske gruppen har morsmålsamsvar for samtlige av de tre delprediksjonene, men som vi ser er det bare for assosiasjonen mellom UB-kontekster og ubestemt entall/flertall at en finner tendenser som er i tråd med prediksjonene. Det at bare én av delprediksjonene har slike tendenser, peker dermed ikke i retning av transfer av den typen som hypotese 1 handler om, noe som også understøttes av at det er ved denne delprediksjonen en finner de minst gruppespesifikke tendensene for den polske gruppen.

Det ble også observert noen interessante tendenser av intergruppe-heterogenitet mellom den engelske og den polske gruppen som ikke så ut til å handle om gruppespesifikke formvalg for den *polske* gruppen, men heller for den engelske: Både i UB-kontekster og BE-kontekster har den engelske gruppen langt mer bruk av artikkelfrase enn både den polske og den norske gruppen. Siden dette kan knyttes til at artikkelfrase gjerne er mer utbredt i engelsk enn i norsk (se 7.4.2.1), kan det være et tegn på transfer hos de engelske innlærerne.

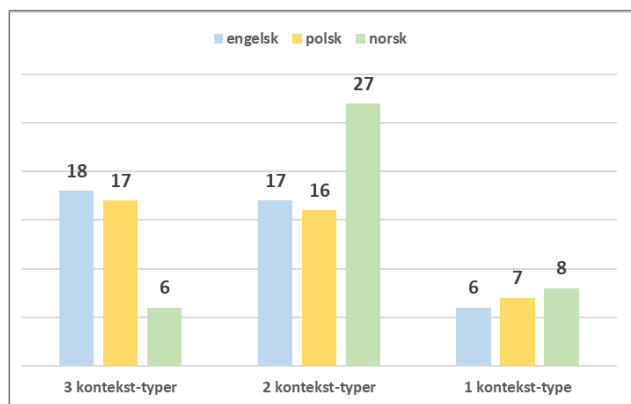
Undersøkelsen så langt har bare sett på tall for gruppene som helhet, og har derfor ikke tatt i betraktning individuell variasjon, så jeg vil nå se på hvordan de individuelle tekstene stemmer med prediksjonene for hypotese 1.

9.3.5 Individuelle profiler

I det polske materialet er det 40 av de 41 innlærertekstene som har enkle substantivfraser, mens dette gjelder alle de 41 tekstene i det engelske og norske materialet. Før jeg ser på forskjeller i formvalg i de individuelle profilene, vil jeg kort beskrive i hvor stor grad de ulike konteksttypene er representert i de enkelte tekstene.

Figur 9-17 (under) viser antallet av individuelle tekster som har alle de tre konteksttypene (UB, BE og ART), eller bare to eller én av typene. I den polske gruppen er det 17 innlærertekster som har alle tre konteksttypene, mens 16 har to typer, og 7 har bare én. Samtlige av de polske innlærertekstene har UB-kontekster, så om de ikke har alle tre typene, er det enten BE- eller ART-kontekster som mangler. Det er ART-konteksten som opptrer hos færrest av de polske innlærerne.

Figur 9-17 Antall tekster med tre, to eller bare én av konteksttypene



Den engelske gruppen er ganske lik den polske, med 18 innlærertekster som har alle tre konteksttyper, 17 som har to og 6 som bare har én av typene. UB-kontekster er også den mest utbredte hos de engelske innlærerne, da den forekommer hos alle unntatt én. Den norske gruppen gir derimot et noe annet bilde. Det er kun 6 tekster i den norske gruppen som har alle tre konteksttypene, mens hele 27 stykker har to. Antallet som har bare én type, er jevnt med de andre to gruppene (8). Tallene for den norske gruppen kan knyttes

til at det er langt færre BE-kontekster (aktiverte slutfelt) i det norske materialet. Til sammen er det bare 9 av 41 i den norske gruppen som har BE-kontekster, til forskjell fra den polske og den engelske gruppen som har henholdsvis 20 og 28 innlærere som har BE-kontekster.

9.3.5.1 *Prosedyre for inndeling i profiler*

Formålet og prosedyre for inndeling i individuelle profiler ble beskrevet i 8.5.4, og for hypotese 1 er det formvalgene i konteksttypene UB, BE og ART som klassifiseres. Første steg i prosedyren er å avgjøre om hver og en av konteksttypene i hver tekst har den predikerte formen eller ikke (*positiv eller negativ*), og konteksttypen regnes som positiv om det er kun den predikerte formen som brukes, eller om det er en overvekt på mer enn to forekomster av den predikerte formen sammenlignet med de andre formene.

Andre steg i prosedyren er å klassifisere hver individuelle profil etter om den samlet sett utgjør en *positiv, negativ eller kombinert profil*. Inndelingen i profiler for hypotese 1 har basert seg på følgende:

Positiv profil: Formvalget stemmer med prediksjonene om polsk morsmålssamsvar for alle de UB-, BE-, og ART-kontekstene som forekommer i den aktuelle teksten.

Negativ profil: Formvalget stemmer *ikke* med prediksjonene for polsk morsmålssamsvar for de UB-, BE-, og ART-kontekstene som forekommer i den aktuelle teksten.

Kombinert profil: Formvalget stemmer med prediksjonene for polsk morsmålssamsvar for én eller to av konteksttypene som forekommer i den aktuelle teksten, men det stemmer ikke med den/de resterende.

For de profilene som bare har én av konteksttypene, tar jeg utgangspunkt i om denne har den predikerte formen eller ikke, men i disse profilene er det ikke mulig å sammenligne formvalg i flere konteksttyper.

I undersøkelsen av de individuelle profilene vil jeg se på om en representativ andel av de polske profilene følger prediksjonene (positive profiler), og om denne andelen er høyere enn i den engelske og den norske gruppen. Merk at også den engelske og den norske gruppen er klassifisert etter de samme kriteriene som den polske gruppen, altså etter om de er i tråd med prediksjonene for *polsk* morsmålssamsvar eller ikke. Dette

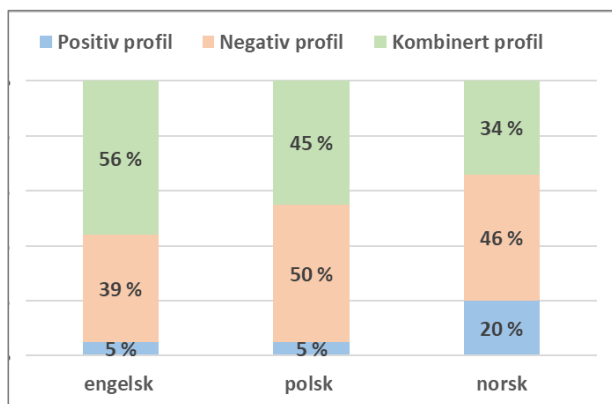
er for å kunne se om eventuelt morsmålssamsvar i de polske profilene er spesifikt for den polske gruppen, eller ikke. For å være et argument for transfer bør andelen av positive profiler i det polske materialet for det første være relativt høy, noe som vil indikere både morsmålssamsvar og intragruppe-homogenitet. For det andre bør andelen positive profiler være signifikant høyere enn i de andre gruppene, noe som indikerer intergruppe-heterogenitet og ulikhet fra målspåket.

Som jeg har vært inne på før (8.5.2) bruker jeg kjikvadrat-test for å teste eventuelle forskjeller mellom gruppene, og jeg regner det som en statistisk signifikant forskjell om p-verdien er lavere enn 0.05 ($p < 0.05$). Siden hypotesen er *retningsstyrt*, altså ikke bare sier at det skal være en forskjell mellom gruppene, men også hvilken retning forskjellen går, er p-verdien fra testen halvert, som i undersøkelse 1.⁵¹

9.3.5.2 Resultat - individuelle profiler

Figur 9-18 og tabell 9-8 (under) viser antall og prosentandel av positive, negative og kombinerte profiler i de tre gruppene.

Figur 9-18 Profiltypene i de tre gruppene



Tabell 9-8 Profiltypene i de tre gruppene

Profiltipe	Engelsk	Polsk	Norsk
Positiv	5 % (2)	5 % (2)	20 % (8)
Negativ	39 % (16)	50 % (20)	46 % (19)
Kombinert	56 % (23)	45 % (18)	34 % (14)
Total	41	40	41

⁵¹ Se 6.1.4 for en forklaring på denne praksisen.

Av de 40 polske innlærertekstene er det kun *to* tekster (5 %) som har en positiv profil, altså den type morsmålssamsvar som hypotese 1 predikerer. Hele 20 stykker (50 %) har derimot en negativ profil, og 18 en kombinert (45 %). Begge de positive profilene har kun UB-kontekster, og mangler dermed BE- og ART-kontekster helt. Til sammenligning har 16 av de 20 negative profilene to eller tre konteksttyper. Siden de positive profilene bare har én av konteksttypene representert, utgjør disse et noe tynnere grunnlag enn de negative profilene, som kan vise formvalg i mer enn én konteksttype, og dermed er mer egnet til å gi et helhetlig inntrykk av form/kontekst-assosiasjoner.

Sammenlignet med de to andre gruppene, er det den norske gruppen som har flest positive profiler, da åtte norske profiler (20 %) stemmer med prediksjonene for polsk morsmålssamsvar. Forskjellen mellom den polske og den norske gruppen er signifikant om en tester antall positive profiler versus negative og kombinerte samlet ($\chi^2 = 3.9404$, $p < 0.05$). Forskjellen går imidlertid i feil retning av det predikerte siden det er den norske gruppen som har flest positive profiler, ikke den polske.

I den engelske gruppen er det også bare to innlærere (5 %) som har en positiv profil, altså likt med den polske, men det er den polske gruppen som har flest negative profiler (pol. 20; eng. 16; nor. 19), mens kombinerte profiler er mest vanlig i det engelske materialet (pol. 18; eng. 23; nor. 14). Den polske gruppen er altså mer lik den engelske enn den norske gruppen, og i den grad de to innlærergруппene er ulike, er det fordi den polske har flere *negative* profiler, men denne forskjellen mellom den engelske og den polske gruppen er ikke signifikant ($\chi^2 = 0.9878$, $p > 0.05$).

De individuelle profilene for det samlede materialet viser altså at polsk har veldig få positive profiler, og flest negative profiler av de tre gruppene. At så få innlærere har en positiv profil, viser at det er lavt *morsmålssamsvar* i gruppen, men også at det er lav *intragruppe-homogenitet*, altså at det det predikerte mønsteret ikke er særlig representativt for de polske innlærerne. Det er dermed ikke belegg for verken morsmålssamsvar eller intragruppe-homogenitet i de polske individuelle profilene. Det er heller ikke særlig *intergruppe-heterogenitet*, siden den polske og den engelske gruppen ikke er signifikant forskjellige. *Intralingval kontrast* lar seg ikke undersøke her, men den norske gruppen har altså signifikant flere positive profiler enn den polske

gruppen. Dette stemmer med det som kom fram i gruppetallene, nemlig at den norske gruppen på flere områder er mer i tråd med prediksjonene enn den polske gruppen selv er.

9.3.6 Oppsummering av funn for hypotese 1 i det samlede materialet

Tabell 9-9 (under) gir en oversikt over hovedfunnene fra undersøkelsen av hele det samlede materialet på gruppe- og individnivå. På gruppenivå har jeg i tabellen kun oppgitt prosentandelen for de tre form/kontekst-assosiasjonene som er predikert for den polske gruppen, ettersom det er disse som er hovedfokus for undersøkelsen. På individnivå har jeg slått sammen de negative og kombinerte profilene, siden begge disse i prinsippet er negative (ikke-positive) med hensyn til prediksjonene.

Tabell 9-9 Oversikt over hovedfunn for hypotese 1 i det samlede materialet

GRUPPENIVÅ	Polsk	Engelsk	Norsk
UB-kontekst/ubestemt ent/flt			
- Innen kontekst	58 %	49 %	61 %
- Innen form	64 %	61 %	69 %
BE-kontekst/bestemt ent/flt			
- Innen kontekst	35 %	14 %	67 %
- Innen form	31 %	11 %	8 %
ART-kontekst/artikkel frase			
- Innen kontekst	6 %	5 %	12 %
- Innen form	5 %	4 %	34 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	5 %	5 %	20 %
- Negative/kombinerte	95 %	95 %	80 %

På gruppenivå så vi at den polske gruppen hadde ulik grad av morsmålsamsvar for de tre form/kontekst-prediksjonene. Det var høyest samsvar for assosiasjonen *UB-kontekst/ubestemt entall/flertall*, lavere for *BE-kontekst/bestemt form* og aller lavest, nesten fraværende, for *ART-kontekst/artikkel frase*.

Sammenligningen med de to andre gruppene viste at assosiasjonen med sterkest tendens til morsmålsamsvar (*UB-kontekst/ubestemt entall/flertall*) var til stede også hos de to andre gruppene, mens den polske gruppen først og fremst skilte seg ut for assosiasjonene *BE-kontekst/bestemt form* og *ART-kontekst/artikkel frase*. Der det var

tendenser til morsmålssamsvar, var dette med andre ord ikke spesifikt for den polske gruppen, og der det var tendenser til gruppespesifikke formvalg i den polske gruppen, var det ikke understøttet av morsmålsamsvar. Konklusjonen er i alle tilfeller at argumentene for transfer på gruppenivå er svake for alle de tre predikerte form/kontekst-assosiasjonene, enten på grunn av lavt morsmålssamsvar eller på grunn av lav intergruppe-heterogenitet og stor grad av målspåklighet.

Den type transfer som er formulert gjennom hypotese 1, innebærer dessuten at det fins tegn på transfer for *samtlig*e av de predikerte form/kontekst-assosiasjonene, men her har vi altså sett det motsatte: Ingen av de predikerte assosiasjonene så ut til å kunne knyttes tydelig til transfer.

Funnene på gruppenivå ble understøttet av funnene i de individuelle profilene, hvor den polske gruppen hadde kun to positive profiler. Det var dessuten ikke signifikante forskjeller i antall positive profiler mellom den polske og den engelske gruppen. Den norske gruppen hadde imidlertid signifikant *flere* positive profiler enn den polske. De individuelle profilene viste altså lite morsmålssamsvar, lite intragruppe-homogenitet, lite intergruppe-heterogenitet, at den norske gruppen var mer i tråd med prediksjonene for polsk morsmålssamsvar enn den polske gruppen selv, noe som også var tilfelle på gruppenivå.

Konklusjonen etter å ha sett på det samlede materialet på både gruppe- og individnivå, er altså at det ikke foreligger transfer fra polsk av den typen som hypotese 1 handler om. Dette var det allerede tydelige tegn på tidlig i undersøkelsen av morsmålssamsvar, hvor det kom fram at den polske gruppen hadde en veldig lik fordeling av former i de tre konteksttypene, og at de ulike formtypene var brukt jevnt i alle konteksttypene. Prediksjonene for hypotese 1 tilsier imidlertid *ulike* formvalg for hver av konteksttypene.

Transfereffektene intralingval kontrast og intergruppe-heterogenitet vil imidlertid best belyses når materialet er delt opp i ulike referensielle typer, og dette vil jeg se nærmere på i de neste delene av undersøkelsen. Før det vil jeg kort kommentere noen andre relevante observasjoner fra gjennomgangen av det samlede materialet.

Alternative form/kontekst-assosiasjoner i materialet

Prediksjonene for hypotese 1 baserer seg på en kombinasjon av substantivfrasereferentenes identifiserbarhet (aktivering) og setningsposisjon, men jeg har også underveis kommentert alternative form/kontekst-assosiasjoner knyttet til setningsposisjon og aktivering uavhengig av hverandre. Jeg fant at både ubestemt entall/flertall og artikkelfrase i større grad var assosiert med *slutfeltet* enn de var assosiert med UB- og ART-kontekster (ub. 75 %; art. 88 % i slutfelt). I tillegg fant jeg at bestemt form i større grad var assosiert med *tekstuelte aktive* referenter enn med BE-kontekster (be. 80 % aktive). Sammenlignet med den engelske og den norske gruppen er det både likheter og forskjeller på dette området. Den engelske gruppen (se tabell 9-6, side 238) har en tilsvarende, men noe svakere tendens til at ubestemt entall/flertall og artikkelfrase brukes oftest i slutfeltet (ub. 61 %; art. 85%), og bestemt form i størst grad er assosiert med *forfeltet* (76 %). Den norske gruppen (se Tabell 9-7, side 238) har derimot en helt annen profil, nemlig en sterk tendens til å bruke samtlige formtyper mest ved *ikke-aktive* substantivfraser, uavhengig av setningsposisjon (ub. 88 %; be. 81 %; art. 96 %). Dette henger imidlertid mest sannsynlig sammen med at det er veldig få substantivfraser med aktivert referent i det norske materialet; av 232 substantivfraser er kun 9 % aktive forfelt, og 4 % aktive slutfelt.

Jeg vil ta med meg disse observasjonene videre til det differensierte materialet for å se om ulike type referanse kan belyse disse alternative form/kontekst-assosiasjonene.

Transfer hos de engelske innlærerne?

Som jeg har vært inne på flere steder, er det funn i det samlede materialet som peker i retning av transfer hos de engelske innlærerne. I UB- og BE-kontekster ble det funnet en høyere andel artikkelfrase hos den engelske gruppen enn hos begge de andre gruppene, og det kom også fram i de deskriptive tallene innledningsvis at den engelske gruppen samlet sett også bruker mest artikkelfrase av gruppene (eng. 26 %; pol./nor. 13 %). Materialet for hypotese 1 inneholder bare tellelige substantiv, og satt i sammenheng med at man i engelsk språk sjeldnere bruker ubestemt entall ved tellelige substantiv enn man gjør i norsk, og oftere artikkelfrase (7.4.2.1), er dette observasjoner som kan peke i retning av både *morsmålssamsvar* og *intergruppe-heterogenitet* knyttet

til bruk av artikkel frase i den engelske gruppen. Om de engelske innlærerne er påvirket av morsmålet, er det sannsynlig at det slår ut nettopp i form av mer utstrakt bruk av artikkel frase, og tilsvarende mindre bruk av ubestemt entall, som igjen slår ut som lavere andel av det som her omtales som ubestemt entall/flertall.

Tendensen til overrepresentasjon av artikkel frase var ikke til stede i ART-kontekster, hvor den polske og den engelske gruppen begge hadde lavere andel av artikkel frase enn den norske, og det er uklart hvorfor en eventuell transfer hos den engelske gruppen ikke skulle slå ut i disse kontekstene, altså ikke-aktiverte forfelt, men det ligger utenfor denne undersøkelsen å gå nærmere inn i forhold som har med transfer i den engelske gruppen å gjøre.

9.4 Materiale differensiert etter ulike referensielle typer

Videre i undersøkelsen av hypotese 1 vil jeg se på substantivfraser med indefinit, definit og generisk referanse hver for seg. Som nevnt tidligere, har oppdelingen av materialet i referensielle typer både som formål å tydeliggjøre kontekster for intergruppe-heterogenitet mellom den polske og den engelske gruppen, og å tydeliggjøre kontekster for morsmålssamsvar som ikke sammenfaller med målspråket (intralingval kontrast).

Innen hver referensielle type vil jeg undersøke transfereffektene på samme måte som jeg så langt har gjort for det samlede materialet, men framover vil også det jeg har kalt *sammenfallskontekster* og *markørkontekster* vurderes. Disse ble forklart i 8.5.3.3, men forklaringen kan gjentas her:

Markørkontekster: Språklige kontekster hvor formvalgene til den polske gruppen er forventet å være *ulik* det en finner i den norske og/eller engelske gruppen om transfer foreligger.

Sammenfallskontekster: Språklige kontekster hvor formvalgene til den polske gruppen er forventet å være *lik* det en finner hos den norske og/eller engelske gruppen, selv om transfer foreligger.

Jeg vil komme tilbake til hva som utgjør markør- og sammenfallskontekster for indefinit, definit og generisk referanse i de respektive underkapitlene.

9.5 Indefinit referanse

I neste del av undersøkelsen vil jeg undersøke transfereffekter i substantivfraser med *indefinit referanse*. Prediksjonene om de polske innlærernes formvalg i UB-, BE- og ART-kontekster er de samme som tidligere, og under er eksempler på indefinite substantivfraser som står i hver av de tre konteksttypene og har den predikerte formen.

Indefinit referanse foreligger når en substantivfrase viser til én eller flere individuelle referenter som ikke vurderes som identifiserbare for tilhøreren. Indefinit referanse står i kontrast til *definit referanse*, hvor én eller flere individuelle referenter regnes som identifiserbare for tilhøreren gjennom anafori, situasjonen, kunnskap om verden eller ved assosiasjon. Indefinit referanse står også i kontrast til *generisk referanse*, som viser til klasser av referenter heller enn individuelle referenter. Se 8.4.4 for mer om klassifiseringen av referensielle typer.

Under er eksempler på de ulike konteksttypene for hypotese 1 med indefinite substantivfraser.

Indefinit UB-kontekst

I eksempel 9.5-a (under) ser vi substantivfrasen «kalender», som er vurdert som indefinit av følgende grunner: 1) Den viser til en individuell referent, ikke til en klasse av referenter, og 2) det er ikke snakk om en kalender som er identifiserbar for mottageren av ytringen innen noen av de områdene som gjør referenter definite, og som ble nevnt i forrige avsnitt.

9.5-a Nå har vi nesten alt i denne små telefon. Vi har **kalender** som kan minne oss om viktige saker. (s0767)

Substantivfrasen «kalender» står i hovedsetningens slutfelt, og referenten er ikke nevnt før, verken gjennom et samreferent uttrykk eller gjennom et uttrykk som viser til samme type referent uten å være samreferent. Den er derfor klassifisert som *ikke-aktivert slutfelt*, som er en type UB-kontekst. Substantivfrasen har ubestemt form uten artikkel, og formvalget er derfor i tråd med prediksjonen for UB-kontekster.

Indefinit BE-kontekst

Det neste eksemplet (9.5-b) viser substantivfrasen «mobilen», som er klassifisert som indefinit fordi det er snakk om en individuell, ikke-spesifikk mobil(telefon) som ikke

er identifiserbar, og det vises med denne frasen heller ikke generisk til mobiler som klasse.

9.5-b Jeg synes at i den moderne verden, hvor alt skjer veldig fort og mennesker trenger å ha masse opplysninger, det er viktig og nødvendig å ha en mobiltelefon. Jeg synes også at før vi kjøper **mobilen** for vår barn vi bør tenke litt på både positive og negative sider av den og hvordan påvirker den til helsen våre. (s0983)

Substantivfrasen «mobilen» står i slutfeltet til en leddsetning, og siden samme type referent er nevnt i forrige setning («en mobiltelefon»), regnes substantivfrasens referent som aktivert, altså er dette et eksempel på et aktivert slutfelt (BE-kontekst). Siden substantivet har bestemt form, er formvalget her i tråd med prediksjonene for BE-kontekster.⁵²

De to ulike substantivfrasene som viser til mobiltelefoner i eksemplet over, viser ikke til samme mobiltelefon, egentlig viser ingen av dem til en spesifikk telefon, men substantivfrasen «en mobiltelefon» regnes likevel som aktiverende for «mobilen» i neste setning fordi uttrykket viser til samme *type* av referent. Eksemplet over illustrerer dermed hvordan substantivfraser kan være indefinitte, men samtidig aktiverte, noe som kan virke som et paradoks ettersom indefinitt referanse innebærer en ikke-identifiserbar referent. Om aktivert i denne undersøkelsen bare hadde handlet *samreferanse* mellom det aktuelle uttrykket og et tidligere nevnt uttrykk, hadde egenskapene indefinitt og aktivert vært gjensidig utelukkende: Om en referent er nevnt i forrige setning, vil den nemlig være definitt ved påfølgende nevning. Som jeg var inne på i 8.4.3, er aktivering i denne undersøkelsen både knyttet til samreferanse og referanse til en referent av samme type. På grunn av dette er det fullt mulig for en indefinitt referent å være aktivert. Indefinitte substantivfraser som har aktivert sin referanse, vil dermed alltid være aktivert gjennom et uttrykk som viser til samme type av referent, og aldri gjennom samreferanse, siden det siste vil innebære definitt referanse.

Indefinitt ART-kontekst

I det siste eksemplet (9.5-c under) ser vi den indefinitte substantivfrasen «en by». Denne frasen viser ikke til byer generelt, men til en individuell, ikke-spesifikk by, og den er

⁵² Jeg har her sett på «mobil» og «mobiltelefon» som refererende til samme type referent *med samme substantiv* siden det ene ordet kan ses på som en kortform av det andre. Dette er et unntak fra kravet om at det skal være samme substantiv, som ble nevnt i 8.4.3.

derfor ikke generisk. Byen er ikke identifiserbar for mottageren gjennom de domene som etablerer definit referanse, og er derfor heller ikke definit.

9.5-c Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan **en by** passer for barn.(s.0295)

Frasen «en by» er plassert initialt i en innføydd setning i den forstand at den står preverbalt, altså det jeg har valgt å kalle «forfelt». Den har i tillegg en ikke-aktivert referent siden det ikke er nevnt noen by i de foregående setningene. Dette er derfor et eksempel på et ikke-aktivert forfelt (ART-kontekst), og siden det er snakk om en artikkelfrase, er formvalget i tråd med prediksjonen for denne konteksttypen.

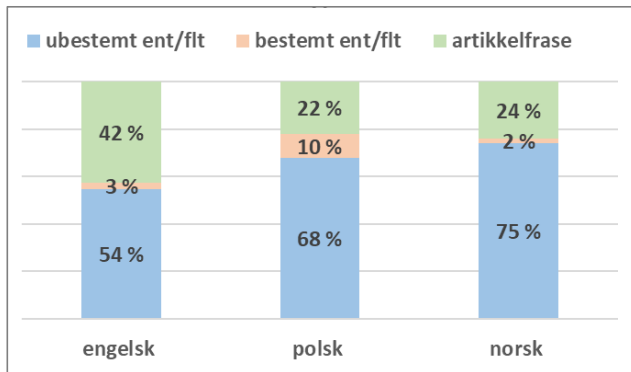
Jeg vil komme tilbake til flere eksempler i gjennomgangen av resultatene. Før selve analysen av dataene vil jeg gi en deskriptiv oversikt over det indefinite materialet i de tre gruppene, og gjøre rede for hva som er sammenfallskontekster og markørkontekster ved indefinit referanse. Etter dette tar jeg for meg hver og en av de predikerte form/kontekst-assosiasjonene på gruppe- og individnivå.

9.5.1 Deskriptivt for alle gruppene

Det er til sammen 180 indefinite, enkle substantivfraser i det polske materialet for hypotese 1, noe som er relativt jevnt med det engelske (172), mens det norske har en god del færre slike fraser (110). Denne forskjellen avspeiler mest sannsynlig antallet individuelle tekster som har enkle substantivfraser med indefinit referanse i de ulike gruppene, da både den polske og den engelske gruppen har 39 tekster som har dette, mens det bare er snakk om 33 tekster i den norske gruppen.

Figur 9-19 og tabell 9-10 (øverste neste side) viser andelen av de ulike formtypene i det indefinite materialet. Det er ubestemt entall/flertall og artikkelfrase som dominerer, og sammenlagt utgjør disse over 90 % av substantivfrasene for alle tre gruppene (pol. 90 %; eng. 97 %; nor. 98 %). Den polske gruppen har likevel en noe lavere andel av ubestemt form, og noe mer bestemt form enn de andre to gruppene (pol. 10 %; eng. 3 %; nor. 2 %). Siden dette er indefinit referanse, er det heller ikke ventet at den engelske og den norske gruppen skal bruke bestemt form, og det er kun snakk om seks slike tilfeller i det engelske materialet, og to i det norske. Utover dette er det stor likhet mellom den polske og den norske gruppen, mens den engelske skiller seg noe ut

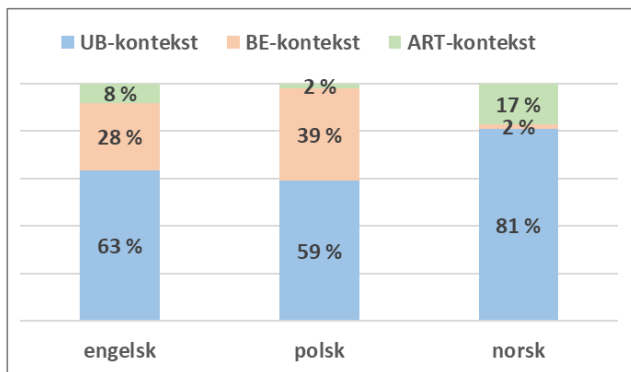
Figur 9-19 Andel former i de tre gruppene (indefinit referanse)



Tabell 9-10 Andel av de ulike formtypene i de tre gruppene (indefinit referanse)

Morsmål	Ubestemt ent/flt	Bestemt ent/flt	Artikkelfrase	Totalt
Polsk	68 % (122)	10 % (18)	22 % (40)	180
Engelsk	54 % (93)	3 % (6)	42 % (73)	172
Norsk	75 % (82)	2 % (2)	24 % (26)	110

Figur 9-20 Andel konteksttyper i de tre gruppene (indefinit referanse)



Tabell 9-11 Andel av de ulike konteksttypene i de tre gruppene (indefinit referanse)

Morsmål	UB-kontekst	BE-kontekst	ART-kontekst	Totalt
Polsk	59 % (106)	39 % (70)	2 % (4)	180
Engelsk	63 % (109)	28 % (49)	8 % (14)	172
Norsk	81 % (89)	2 % (2)	17 % (19)	110

ved å ha en noe høyere andel av artikkelfrase enn det de to andre gruppene har (pol. 22 %; eng. 42 %; nor. 24 %), og tilsvarende lavere andel av ubestemt entall/flertall (pol. 68 %; eng. 54 %; nor. 75 %).

Figur 9-20 og tabell 9-11 (nederst forrige side) viser andelen av de ulike kontekststypene for hypotese 1, og her er den engelske og den polske gruppen relativt like i å ha rundt 60 % UB-kontekster. De har også en lav andel ART-kontekster (pol. 2 %; eng. 8 %), og antallet er spesielt lavt i det polske gruppen, som bare har fire slike kontekster. Den polske gruppen har dessuten en større andel BE-kontekster (pol. 39 %; eng. 28 %). Den norske gruppen skiller seg merkbart ut fra begge de andre gruppene ved å ha langt flere UB-kontekster (81 %) og ART-kontekster (17 %), i tillegg til kun *to* BE-kontekster (2 %).

9.5.2 Sammenfalls- og markørkontekster for indefinitt referanse

I 9.2 viste jeg hvordan prediksjonene om morsmålssamsvar for den polske gruppen forventes å *sammenfalle med* eller *skille seg fra* den engelske og norske gruppens formvalg innen de ulike referansetyperne. I språkene norsk og engelsk er det typisk ubestemte former (ubestemt entall, ubestemt flertall og artikkelfrase) som brukes ved indefinitt referanse, og dette er forventet å reflekteres i formvalgene til den engelske og den norske gruppen. Hos den polske gruppen er det derimot ifølge prediksjonene forventet ulike formvalg avhengig av om det er UB-, BE- eller ART-kontekster det er snakk om. Eksempler på dette så vi over i eksempel 9.5a - 9.5-c, hvor substantivfrasenes form varierte selv om alle frasene var indefinitte.

Innen indefinitt referanse er det BE-kontekstene som er markørkontekster, altså det området der den polske gruppen ventes å skille seg fra de andre to gruppene om hypotesen stemmer. Den polske gruppen forventes å ha bestemt form i denne konteksttypen, mens de to andre gruppene er forventet å ha ubestemt form.

For UB- og ART-kontekster er kontrastene mellom de tre gruppene mer sammensatt. I UB-kontekster forventes den polske gruppen å være lik de andre to gruppene i å bruke ubestemt entall/flertall, men ulik i at den polske gruppen er predikert til å bruke *kun* dette, og ikke artikkelfrase. For ART-kontekster er det motsatt: Den polske gruppen er forventet å være lik de andre gruppene i å bruke artikkelfrase, men ulik i å *kun* bruke dette.

Forskjellen og likheten mellom den polske gruppen og de to andre gruppene er også forventet å være noe ulik etter om det er den norske eller den engelske gruppen som sammenlignes med. Forskjellen mellom engelsk og norsk er her først og fremst knyttet til at materialet kun inneholder tellelige substantiv, og det at man i engelsk først og fremst bruker artikkelfrase eller flertallsform ved tellelige substantiv. I norsk er det derimot vanlig å bruke ubestemt entall også ved tellelige substantiv. Denne forskjellen mellom engelsk og norsk gjør at den norske og den engelske gruppen kan ha noen forskjeller i fordelingen av artikkelfrase versus ubestemt entall/flertall, der den engelske gruppen kan antas å ha større andel artikkelfrase enn den norske. Konsekvensen av dette er at prediksjonene for den polske gruppen er mindre sammenfallende med den engelske gruppen i UB-kontekster, men mer sammenfallende i ART-kontekster, mens det er motsatt med den norske gruppen, altså *mer* sammenfall i UB-kontekster, og mindre i ART-kontekster.

Tabellen på neste side viser hvordan formvalgene forventes å være i de tre gruppene. Den mest entydige markørkonteksten er uthevet (BE), men det er altså også noen kontraster for de ulike ubestemte formene. «Ubestemte former» viser i tabellen generelt til ubestemte former, enten ubestemt entall, flertall eller artikkelfrase.

Tabell 9-12 Sammenfalls- og markørkontekster for indefinitt referanse

Indefinitt referanse	Polsk	Engelsk	Norsk
UB-kontekst	ubestemt ent/flt	ubestemte former	ubestemte former
BE-kontekst	bestemte former	ubestemte former	ubestemte former
ART-kontekst	artikkelfrase	ubestemte former	ubestemte former

9.5.3 Samlet oversikt for indefinitt referanse

Jeg vil nå ta for meg hver og en av de tre form/kontekst-prediksjonene i det indefinitte materialet. Tabellene på neste side gir, på samme måte som tidligere, full oversikt over former og kontekster i de tre gruppene både innen kontekst og innen form. UB-kontekstene er også her delt opp i de to ulike variantene (aktiverte forfelt og ikke-aktiverte slutfelt). Tallene vil bli visualisert ved hjelp av søylediagram i gjennomgangen, med polsk i midten av hver klynge med søyler.

Tabell 9-13 Polsk gruppe: Fordeling av former etter kontekststype (indefinit).

S1=POLSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster:	Aktivert	Ikke-aktivert	Aktivert	Ikke-aktivert	
Former:	forfelt	slutfelt	slutfelt	forfelt	
UBESTEMT ENT/FLT	2	77	41	2	122
Andel innen form	2 %	63 %	34 %	2 %	100 %
Andel innen kontekst	50 %	75 %	59 %	50 %	68 %
BESTEMT ENT/FLT	0	2	16	0	18
Andel innen form		11 %	89 %		100 %
Andel innen kontekst		2 %	23 %		10 %
ARTIKKELFRASE	2	23	13	2	40
Andel innen form	5 %	58 %	33 %	5 %	100 %
Andel innen kontekst	50 %	23 %	19 %	50 %	22 %
TOTALT	4 (2 %)	102 (57 %)	70 (39 %)	4 (2 %)	180

Tabell 9-14 Engelsk gruppe: Fordeling av former etter kontekststype (indefinit)

S1=ENGELSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster:	Aktivert	Ikke-aktivert	Aktivert	Ikke-aktivert	
Former:	forfelt	slutfelt	slutfelt	forfelt	
UBESTEMT ENT/FLT	6	50	27	10	93
Andel innen form	6 %	54 %	29 %	11 %	100 %
Andel innen kontekst	46 %	52 %	55 %	71 %	54 %
BESTEMT ENT/FLT	0	2	2	2	6
Andel innen form		33 %	33 %	33 %	100 %
Andel innen kontekst		2 %	4 %	14 %	3 %
ARTIKKELFRASE	7	44	20	2	73
Andel innen form	10 %	60 %	27 %	3 %	100 %
Andel innen kontekst	54 %	46 %	41 %	14 %	42 %
TOTALT	13 (8 %)	96 (56 %)	49 (28 %)	14 (8 %)	172

Tabell 9-15 Norsk gruppe: Fordeling av former etter kontekststype ved (indefinit)

S1= NORSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster:	Aktivert	Ikke-aktivert	Aktivert	Ikke-aktivert	
Former:	forfelt	slutfelt	slutfelt	forfelt	
UBESTEMT ENT/FLT	2	67	2	11	82
Andel innen form	2 %	82 %	2 %	13 %	100 %
Andel innen kontekst	67 %	78 %	100 %	58 %	75 %
BESTEMT ENT/FLT	0	2	0	0	2
Andel innen form	0 %	100 %	0 %	0 %	100 %
Andel innen kontekst	0 %	2 %	0 %	0 %	2 %
ARTIKKELFRASE	1	17	0	8	26
Andel innen form	4 %	65 %	0 %	31 %	100 %
Andel innen kontekst	33 %	20 %	0 %	42 %	24 %
TOTALT	3 (3 %)	86 (78 %)	2 (2 %)	19 (17%)	110

9.5.4 Assosiasjon mellom UB-kontekster og ubestemt entall/flertall

I UB-kontekster (aktiverte forfelt og ikke-aktiverte slutfelt) er det forventet at den polske gruppen skal bruke ubestemt entall/flertall ettersom markering av tekstuell identifiserbarhet antas å oppleves som redundant. Dette gjør at verken bestemt form eller artikkelfrase vil tas i bruk. Det er også forventet at det er først og fremst i UB-kontekster at ubestemt entall/flertall brukes, ikke i andre konteksttyper.

De følgende eksemplene viser UB-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for UB-kontekster, deretter to eksempler på formvalg som *ikke* er i tråd med prediksjonen.

Eksempel på en UB-kontekst med predikert formvalg ble omtalt allerede innledningsvis i 9.5, men er gjentatt under (9.5-d). Substantivfrasen «kalender» er her klassifisert som ikke-aktivert slutfelt og har ubestemt flertallsform, som er i tråd med prediksjonen.

9.5-d Nå har vi nesten alt i denne små telefon. Vi har **kalender** som kan minne oss om viktige saker. (s0767)

De to eksemplene under viser UB-kontekster hvor formvalget ikke stemmer med prediksjonen. Eksempel 9.5-e viser et ikke-aktivert slutfelt som har bestemt form, mens 9.5-f viser tilsvarende kontekst med artikkelfrase. Siden UB-kontekster er predikert å ha ubestemt entall/flertall, er ingen av disse formvalgene i tråd med prediksjonen.

9.5-e Det er mange butikker, banker, restauranter, busser osv. Vi har også funnet ut at det er viktig å ha **pengene** for å kjøpe møbler, hvite varer og noe annet at man kunne trenge. (s0309)

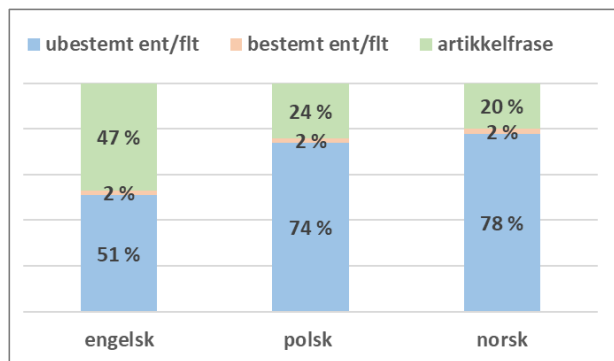
9.5-f Jeg bruker mobiltelefonen ikke ofte, bare av og til. Jeg kjøper **en påfilingskort** månedlig. (s0772)

Søylediagrammene på neste side illustrerer forholdet mellom UB-kontekster og ubestemt entall/flertall i de tre gruppene. Jeg vil først se på den polske gruppen.

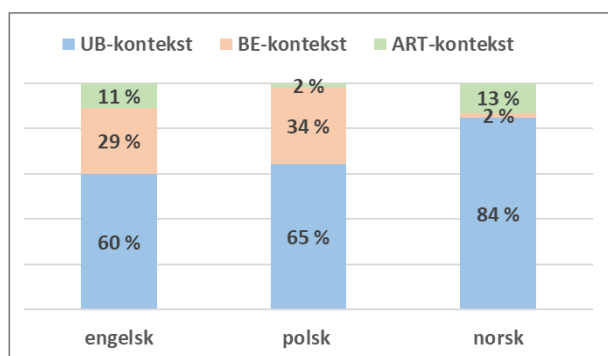
Figur 9-21 (øverst neste side) viser hvilke former som brukes i UB-kontekster i det indefinite materialet, og det er som ventet ubestemt entall/flertall som er mest brukt av de polske innlærerne (74 %). Figur 9-22 (nederst neste side) viser hvilke kontekster denne formtypen brukes mest i, og vi ser at 65 % av alle fraser med ubestemt entall/flertall er brukt i nettopp UB-kontekster, altså langt oftere enn i BE- og

ART-kontekster. Det er med andre ord et relativt høyt morsmålsamsvar for assosiasjonen mellom UB-kontekster og ubestemt entall/flertall ved indefinitt referanse, slik tendensen også var i det samlede materialet.

Figur 9-21 Andel av form innen UB-kontekst



Figur 9-22 Andel av kontekst innen ubest. ent/flt



Sammenlignet med den norske og den engelske gruppen er det ventet et *delvis* sammenfall for formvalget i UB-kontekster, da alle tre gruppene forventes å bruke ubestemt entall/flertall, mens polsk i motsetning til de andre to ikke er forventet å bruke artikkelfrase. Forskjellene ventes å være størst i forhold til den engelske gruppen, siden engelsk har større begrensninger på bruk av ubestemt entall ved tellelige substantiv enn norsk.

I figur 9-21 (øverst over) ser vi at den norske og den polske gruppen generelt sett like i formvalgene innen UB-konteksten, og andelen ubestemt entall/flertall er veldig jevn i de to gruppene (pol. 74 %; nor. 78 %). I figur 9-22 (nederst forrige side) ser vi at den norske gruppen bruker ubestemt entall/flertall først og fremst i UB-kontekster

framfor andre kontekster, men i enda større grad enn den polske gruppen gjør (pol. 65 %; nor. 84 %). I begge gruppene er det også veldig lite bruk av bestemt form, kun 2 %.

Tendensen til en assosiasjon mellom kontekst og form som ble funnet i det polske materialet, er altså i stor grad sammenfallende med det som fins i det norske materialet, noe som også er ventet. Ser vi imidlertid på andelen av *artikkelfrase* i UB-kontekster er også denne relativt jevn i begge gruppene (pol. 24 %; nor. 20 %), noe som *ikke* er forventet om hypotesen stemmer. Polsk ser dermed ut til å være mer i tråd med målspråket enn med prediksjonene om morsmålssamsvar, som jo utelukker artikkelfrase i UB-kontekster.

Forskjellen mellom den engelske og polske gruppen er noe større. Vi ser i figur 9-21 (øverst forrige side) at det hos den engelske gruppen er *lavere* andel av ubestemt entall/flertall i UB-kontekster enn hos den polske (pol. 74 %; eng. 51 %), men tilsvarende *mer* artikkelfrase hos den engelske gruppen (pol. 24 %; eng. 47 %). I figur 9-22 (nederst forrige side) ser vi at forskjellen ikke like stor når det gjelder hvilke konteksttyper frasene med ubestemt entall/flertall brukes i, da begge gruppene oftest bruker slike i UB-kontekster (pol. 65 %; eng. 60 %).

Den polske og den engelske gruppen er dermed like i å ha tendenser til assosiasjon mellom UB-kontekst og ubestemt entall/flertall både innen form og kontekst, men assosiasjonen er sterkere i den polske gruppen, og de to gruppene har en viss heterogenitet knyttet til mer bruk av artikkelfrase hos de engelske innlærerne. Med tanke på at den norske og den polske gruppen er veldig jevne i andelen av artikkelfrase i UB-kontekster, mens den engelske gruppen har en høyere andel (pol. 24 %; eng. 37 %; nor. 20 %), kan det se ut som det er den engelske gruppen som har de mest gruppespesifikke formvalgene her. Dette er med andre ord et lignende mønster som det som kom fram tidligere i undersøkelsen, og som jeg mente kunne være tegn på transfer hos de engelske innlærerne.

Gjennomgangen så langt viser altså at alle de tre gruppene har en tendens til assosiasjon mellom UB-kontekster og ubestemt entall/flertall, men denne er sterkest i det norske materialet (74 % innen kontekst/84 % innen form), nest sterkest i det polske (74 % innen kontekst/65 % innen form), og svakest i det engelske materialet (51 % innen kontekst/ 60 % innen form). Tendensen til morsmålssamsvar i den polske gruppen

er dermed ledsaget av stor grad *homogenitet* i forhold til den norske gruppen, både der det er forventet sammenfall (bruk av ubestemt entall/flertall), og der det skal være forskjeller, om hypotesen stemmer (bruk av artikkelfrase). Mellom den polske og den engelske gruppen er det derimot større grad av *heterogenitet*, noe som var ventet.

9.5.5 Assosiasjon mellom BE-kontekster og bestemt form

I BE-kontekster (aktiverte slutfelt) er det forventet at den polske gruppen skal bruke bestemt form, og dette er også den mest tydelige markørkonteksten ved indefinit referanse, altså det området hvor den polske gruppen er forventet å skille seg mest fra de andre to gruppene om hypotesen stemmer.

Eksempelene under viser BE-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for BE-kontekster, deretter to eksempler på formvalg som *ikke* er i tråd med prediksjonen.

Det første eksemplet (9.5-g) ble også omtalt innledningsvis i 9.5, og viser en BE-kontekst med bestemt form, som er det predikerte formvalget.

9.5-g Jeg synes at i den moderne verden, hvor alt skjer veldig fort og mennesker trenger å ha masse opplysninger, det er viktig og nødvendig å ha en mobiltelefon. Jeg synes også at før vi kjøper **mobilen** for vår barn vi bør tenke litt på både positive og negative sider av den og hvordan påvirker den til helsen våre. (s0983)

De neste to eksemplene (9.5-h og 9.5-i) viser også aktiverte slutfelt, altså BE-kontekster. I begge eksemplene er mobiltelefoner nevnt i foregående setning og derfor regnes referentene for både «mobiltelefon» i 9.5-h og «en mobiltelefon» i 9.5-i som aktiverte. Ingen av disse to frasene viser generelt til mobiltelefoner, de viser til individuelle referenter som ikke er identifiserbare, og referansen er derfor indefinit. Siden de to frasene har henholdsvis ubestemt entall og artikkelfrase, stemmer ikke formvalget med prediksjonen for BE-kontekster.

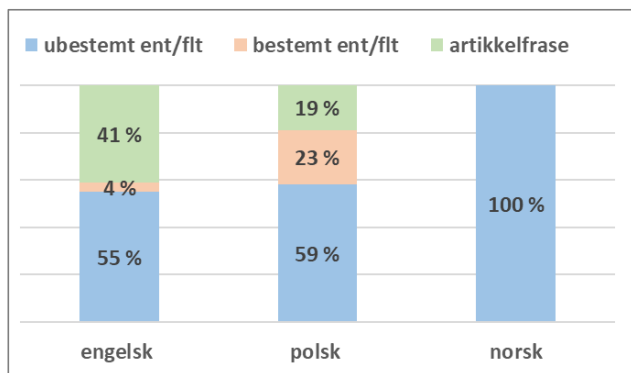
9.5-h Så nåværende mobiltelefoner er mye bedre enn jeg hadde ventet. Og sannsynligvis det er bare jeg som har ikke **mobillefeton** i dag. (s0791)

9.5-i Jeg reiser veldig ofte og det er mye lettere for meg hvis jeg kan bruke mobiltelefonen min. Det skjer også mange rare situasjoner i vår liv når det er bra å ha **en mobiltelefon** med – for eksempel bilulykker. (s0983)

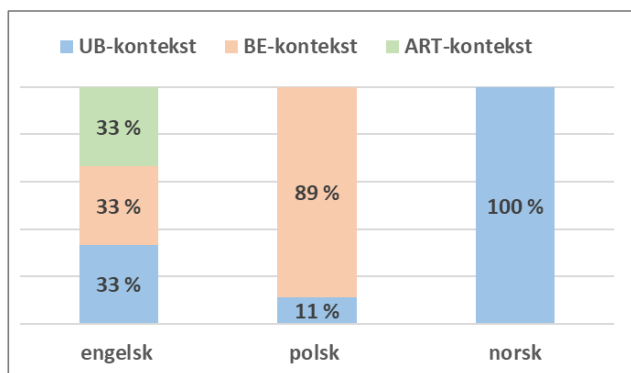
Søylediagrammene under illustrerer forholdet mellom BE-kontekster og bestemt form i de tre gruppene. Jeg vil først se på den polske gruppen.

Figur 9-23 (øverst) viser hvilke former som blir brukt i BE-kontekster av de tre gruppene, og vi ser at bare 23 % av BE-kontekstene har bestemt form i det polske materialet. Samtidig ser vi av figur 9-24 (nederst), som viser hvilke kontekster bestemt form brukes i, at hele 89 % av frasene med bestemt form står i en BE-kontekst hos den polske gruppen. Assosiasjonen mellom form og kontekst for den polske gruppen er altså ganske lav med tanke på at BE-kontekster sjelden har denne formen, men høy med tanke på at fraser med bestemt form som regel brukes i BE-kontekster og sjelden i andre konteksttyper.

Figur 9-23 Andel av form innen BE-kontekst



Figur 9-24 Andel av kontekst innen bestemt form



BE-konteksten er den konteksten der det er ventet størst forskjell mellom den polske og den norske gruppen, ettersom den norske gruppen først og fremst ventes å bruke ubestemt form. Som vi ser av søylediagrammene på forrige side, har den norske gruppen *ingen* bruk av bestemt form i BE-kontekster, men derimot utelukkende ubestemt entall/flertall (figur 9-23, øverst). I tillegg er samtlige fraser med bestemt form brukt i UB-kontekster i den norske gruppen (figur 9-24, nederst). Den norske gruppen skiller seg altså fra den polske ved å helt mangle en assosiasjon mellom bestemt form og BE-kontekster. Det er uventet at den norske gruppen i det hele tatt bruker bestemt form ved indefinit referanse, men det er kun snakk om *to* forekomster. Det er også kun to tilfeller av BE-kontekster i det norske materialet, så sammenligningsgrunnlaget er tynt, men om en ser bort fra dette, er altså forskjellene merkbare mellom den norske og den polske gruppen både innen form og kontekst. Selv om den polske gruppen oppviser et noe ujevnt og lavt morsmålssamsvar for assosiasjonen *BE-kontekst/bestemt form*, så er gruppen altså ulik den norske gruppen, noe som vil si at den bruken av bestemt form som vi finner i det polske materialet, ikke er målspråklik.

Innen BE-kontekster er det også forventet å finne størst forskjell mellom den polske og den engelske gruppen, og vi ser av de samme søylediagrammene på forrige side at den polske gruppen bruker bestemt form oftere enn den engelske i BE-kontekster (pol. 23 %; eng. 4 %), og at den polske gruppen i større grad forbeholder denne formen til disse kontekstene (pol. 89 %; eng. 33 %). Innlærergruppene er altså relativt heterogene, men de er like i at begge grupper har en god del ubestemt entall/flertall i BE-kontekster, som jo også var det kategoriske formvalget for den norske gruppen (pol 59 %; eng. 55 %; nor. 100 %). En tydelig forskjell er likevel at den engelske gruppen også i BE-kontekster har langt høyere andel av artikkel frase enn de to andre gruppene (pol. 19 %; eng. 41 %; nor. 0 %).

Antydning til morsmålssamsvar for assosiasjonen *BE-kontekst/bestemt form* i det polske materialet er altså utfylt av forskjeller fra begge de andre gruppene, ettersom den polske gruppen har en noe sterkere assosiasjon mellom form og kontekst (23 % innen kontekst / 89 % innen form) enn de to andre gruppene (eng. 4/33 %; nor. 0/0 %). Morsmålssamsvaret var sterkere innen form enn innen kontekst, noe som betyr at selv

om formen ikke så ofte brukes i BE-kontekster i forhold til andre former, så er det likevel først og fremst i BE-kontekster bestemt form brukes.

9.5.6 Assosiasjon mellom ART-kontekster og artikkelfrase

I ART-kontekster (ikke-aktiverte forfelt) er det forventet at den polske gruppen skal bruke artikkelfrase. Eksempelene under viser eksempler på ART-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for ART-kontekster, deretter et eksempel på formvalg som *ikke* er i tråd med prediksjonen. I de polske innlærertekstene fins det ingen ART-kontekster med bestemt form, derfor mangler eksempel på dette.

Eksemplet under (9.5-j) viser en ART-kontekst med artikkelfrase, altså det predikerte formvalget. Eksemplet ble også omtalt som indefinitt innledningsvis i 9.5.

9.5-j Samtidig er det lett å finne mange fine steder å gå på tur, både på fjellene og havet. En annen ting som er viktig for meg i det siste er hvordan **en by** passer for barn.(s.0295)

Eksemplet under (9.5-k) viser en ART-kontekst med ubestemt form uten artikkel. Den understrekede substantivfrasen «konsekvens» står i forfeltet til en hovedsetning, og konsekvenser har ikke vært nevnt tidligere, altså er referenten ikke-aktivert. Formvalget stemmer likevel ikke med prediksjonen for slike kontekster, da dette ikke er en artikkelfrase. Det kan være vanskelig å si om dette skal regnes som en entallsform eller en flertallsform på grunn av avviket i endelsen, men det avgjørende her er uansett at det ikke er en artikkelfrase.

9.5-k De bør snakke mye med dem, spesiell hva er godt og hva er dårlig. **Konsekvens** er også viktig. (s0163)

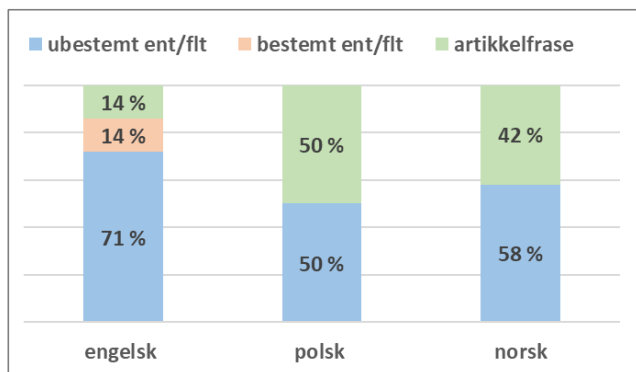
Søylediagrammene på neste side illustrerer forholdet mellom ART-kontekster og artikkelfrase i de tre gruppene. Jeg vil først se på den polske gruppen.

Figur 9-25 (øverst neste side) viser hvilke former som er brukt i ART-kontekster i de tre gruppene. Prediksjonen for den polske gruppen om bruk av artikkelfrase innfris i 50 % av ART-kontekstene, noe som er jevnt med andelen av ubestemt entall/flertall som er brukt.

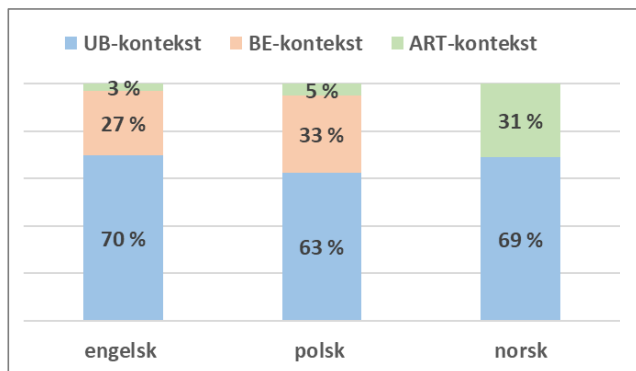
Figur 9-26 (nederst neste side) viser hvilke kontekster artikkelfrase er brukt i, og her ser vi at av alle artikkelfraser, er det bare 6 % som er brukt i ART-kontekster. Det

er imidlertid kun *fire* ART-kontekster i det polske materialet, så grunnlaget er veldig tynt, men det at bare *to* av de 40 artikkelfrasene i polske tekster er brukt i ART-kontekster, viser tydelig at denne formen ikke i særlig grad blir brukt sin predikerte kontekst.

Figur 9-25 Andel av form innen ART-kontekst



Figur 9-26 Andel av kontekst innen artikkelfraser



ART-konteksten er i likhet med UB-konteksten bare delvis en sammenfallskontekst for sammenligningen med de andre to gruppene. Prediksjonene om polsk morsmålsamsvar er sammenfallende på den måten at alle tre gruppene forventes å bruke artikkel frase i denne konteksten, men den polske gruppen ventes å skille seg ut ved å *ikke* bruke ubestemt entall/flertall. Likhetene ventes å være *størst* i forhold til den engelske gruppen, siden engelsk oftere har artikkel frase ved tellelige substantiv enn norsk.

Når man sammenligner med den norske gruppen i figur 9-25 (øverst forrige side), viser det seg at begge gruppene har en god del artikkelfraser i ART-kontekster, men den polske har noe mer den norske (pol. 50 %; nor. 42 %). I figur 9-26 (nederst forrige side) kommer det derimot fram at den norske gruppen bruker flere av sine artikkelfraser i ART-kontekster enn den polske gjør (pol. 5 %; nor. 31 %).

Den polske og den norske gruppen er altså relativt homogene innen kontekst, men mer heterogene innen form, og homogeniteten finner vi der morsmålssamsvaret er høyest, mens heterogeniteten er knyttet til at den norske gruppen er mer i tråd med prediksjonene enn den polske gruppen er.

De to gruppene er også like i at det brukes en relativt lik andel av ubestemt entall/flertall (pol. 50 %; nor. 58 %), riktignok noe mer i den norske gruppen, men ubestemt entall/flertall er ikke i tråd med prediksjonene for ART-kontekster hos den polske gruppen. Det samlede inntrykket av forskjeller og likheter mellom den polske og den norske gruppen tyder også her på at formvalgene til de polske innlærerne er mer i tråd med målspråket enn med prediksjonene om morsmålssamsvar.

Om man sammenligner den polske og den engelske gruppen i de samme søylediagrammene, ser vi at disse to gruppene er mer heterogene i den forstand at den polske bruker langt mer artikkelfrase enn den engelske i ART-kontekster (pol. 50 %; eng. 14 %). Når det kommer til hvilke kontekster artikkelfraser brukes i, er gruppene derimot relativt like, da begge sjeldent bruker artikkelfrasene sine i ART-kontekster (pol. 5 %; eng. 3 %). De to innlærergruppene er med andre ord forskjellige innen kontekst, men mer like innen form.

Faktisk er alle de tre gruppene like innen form her, med tanke på at artikkelfraser brukes først og fremst i UB-kontekster (>60 % i alle gruppene), dernest i BE-kontekster (≈30 % i alle gruppene), og i veldig liten grad i ART-kontekster (< 5% i alle gruppene). Innen kontekst, altså i hvilke former som brukes i ART-kontekster, er det imidlertid en forskjell mellom de tre gruppene, men det er den engelske gruppen som skiller seg mest ut, med å ha en langt *lavere* andel artikkelfrase enn de to andre gruppene (pol. 50 %; eng. 14 %; nor. 42 %). Som vi også har sett tidligere, er det den engelske gruppen som ser ut til å ha de mest gruppespesifikke formvalgene, i form av mer bruk av artikkelfrase

enn det en finner i de andre gruppene, men i dette tilfellet er det altså *mindre* bruk av artikkelfrase som er tendensen hos de engelske innlærerne.

9.5.7 Oppsummering av funn på gruppenivå

Før jeg ser på de individuelle profilene, vil jeg gi en oppsummering av funnene på gruppenivå.

Morsmålssamsvar

Morsmålssamsvaret i det polske materialet ser noe ujevnt ut når både bruk innen kontekst og bruk innen form er tatt i betraktning. Det er først og fremst assosiasjonen mellom UB-kontekst og ubestemt entall/flertall som skiller seg ut med å ha formvalg som til en viss grad stemmer med prediksjonene for polsk morsmålssamsvar (74 % innen kontekst/65 % innen form), mens de andre to kontekstene har en mer variert tendens innen form og innen kontekst: For assosiasjonen *BE-kontekst/bestemt form* ser vi at det er lavt samsvar innen kontekst, men høyt innen form, da de aller fleste fraser med bestemt form brukes nettopp i BE-kontekster. For assosiasjonen *ART-kontekst/artikkelfrase* så vi at artikkelfrase brukes sjeldent i ART-kontekster (5 %), men at 50 % av ART-kontekstene faktisk har denne formen. Her er det imidlertid kun snakk om to av fire forekomster.

Intralingval kontrast

Vurderingen av intralingval kontrast i forhold til den norske gruppen har hatt som mål å belyse om tendenser til morsmålssamsvar i den polske gruppen er spesifikke for denne gruppen i forhold til den norske gruppen, noe som eventuelt kan utelukke målspåkstilignelse som forklaring på formvalgene. Her var det spesielt markørkontekster som var sentrale, fordi prediksjonene for disse konteksttypene ikke sammenfaller med det typiske formvalget i norsk.

Markørkonteksten ved indefinitt referanse er BE-kontekster, da dette er kontekster hvor morsmålet og målspåket har ulike implikasjoner for formvalg, og det er også her den polske og den norske gruppen er mest forskjellige: Den norske gruppen hadde ingen kobling mellom BE-kontekster og bestemt form i det hele tatt, mens den polske gruppen brukte 89 % av frasene med bestemt form i BE-kontekster. At den polske gruppen skiller seg fra den norske, viser at tendensene i det polske materialet

ikke kan tilskrives tilegnelse av målspåket, og da står argumentet for transfer sterkere. Likevel var det kun 23 % av BE-kontekstene som hadde bestemt form i den polske gruppen, og ubestemt entall/flertall var langt oftere brukt (74 %). Morsmålssamsvaret er dermed ganske lavt innen kontekst, og det ser ikke ut som bestemt form er spesielt sentralt for å markere det bruddet på tema/rema-prinsippet som finner sted i BE-kontekster (aktiverte slutfelt).

Der morsmålssamsvaret i den polske gruppen var høyest, altså assosiasjonen mellom UB-kontekster og ubestemt entall/flertall, var det derimot stort sammenfall mellom den polske og den norske gruppen. Begge grupper hadde ustrakt bruk av ubestemt entall/flertall, men dette var forventet ettersom dette til dels er en sammenfallskontekst.

Når det gjelder assosiasjonen mellom ART-kontekstene og artikkelfrase, er den polske og den norske gruppen også like: Begge har en andel på nesten 50 % artikkelfraser i disse kontekstene, men den polske gruppen bruker artikkelfrase oftere i BE-kontekster og ART-kontekster enn den norske gruppen.

Undersøkelsen av intralingval kontrast peker altså ikke i retning av morsmålpåvirkning i indefinite UB-kontekster og ART-kontekster hos den polske gruppen, men det fins tendenser i BE-kontekster, hvor det er stor grad av heterogenitet i forhold til målspåksgruppen: Til tross for å ha et ujevnt morsmålssamsvar, har den polske gruppen en sterkere kobling mellom BE-kontekst og bestemt form enn den norske gruppen har. Det taler imidlertid mot hypotesen at den polske gruppen bruker mer ubestemt form der det er forventet bestemt, og at det er like vanlig å bruke artikkelfrase som bestemt form i BE-kontekstene.

Intergruppe-heterogenitet

I sammenligningen med den engelske gruppen fant jeg intergruppe-heterogenitet mellom engelsk og polsk gruppe innen alle de tre form/kontekst-assosiasjonene. Forskjellene var knyttet både til en sterkere assosiasjon mellom BE-kontekst og bestemt form i det polske materialet, men også til fordelingen av artikkelfrase versus ubestemt entall/flertall i UB- og BE-kontekster.

I markørkonteksten for indefinitt referanse (BE) var assosiasjonen mellom BE-kontekst og bestemt form generelt sterkere i det polske materialet enn i det engelske,

og dette tyder på språklige valg som er spesifikke for den polske innlærergruppen. Til tross for at det er snakk om indefinitt referanse, hvor det i engelsk ikke brukes bestemt form, brukte den engelske gruppen en god del av dette, men ikke like mye som den polske.

I UB-kontekster var det en forskjell mellom de to gruppene knyttet til at den engelske gruppen hadde langt mer bruk av artikkelfrase enn den polske, og selv om det er ventet en del sammenfall mellom de to gruppene her, er det nettopp når det kommer til artikkelfraser at prediksjonene for den polske og den engelske gruppen ikke har sammenfall ved indefinitt referanse. De polske prediksjonene tilsier nemlig at artikkelfrase ikke brukes her.

I ART-kontekster er også gruppene ulike, men det er den polske gruppen som har mest artikkelfrase i disse kontekstene. Formvalgene i den polske gruppen er altså delvis spesifikke i forhold til den engelske.

Selv om jeg bare har foretatt en skjønsmessig vurdering av forskjeller og likheter, er det flere områder hvor det er merkbare forskjeller mellom gruppene, derfor kan det se ut til at det innen indefinitt referanse er en god del intergruppe-heterogenitet.

Målspråkstilegnelse?

I de deskriptive tallene for indefinitte substantivfraser innledningsvis (9.5.1) så vi at den polske gruppen hadde generelt utstrakt bruk av ubestemt entall/flertall og artikkelfrase, og at til sammen 90 % av alle frasene hadde ubestemt form (ub. 68 %; art. 22 %), noe som er relativt likt den norske gruppen, som hadde 98 % ubestemt form (ub. 75 %; art. 24 %). Her ligger altså den polske gruppen nært opp til den norske, og dermed også nært opp til målspraksregelen om ubestemt form ved indefinitt referanse. Høy andel av ubestemt form hos de polske innlærerne viste seg også å gjelde i alle de tre konteksttypene (UB. 98 %; BE. 78 %; ART. 100 %), noe som kan tyde på at de polske innlærerne først og fremst velger form i tråd med målspråket. Likevel ser vi at det spesielt er markørkonteksten (BE) som skiller seg mest ut her med å ha en god del lavere andel av ubestemt form, og høyere andel bestemt form, slik også hypotesen tilsier. Selv om det ikke er overvekt av fraser med bestemt form i BE-kontekster, så er det altså nettopp i denne markørkonteksten for indefinitt referanse at en finner en forskjell både fra den norske og den engelske gruppen. Supplert med at 89 % av frasene med bestemt

form er brukt BE-kontekster, så er det her et interessant mønster som kan være en indikasjon på transfer.

Morsmålpåvirkning fra engelsk?

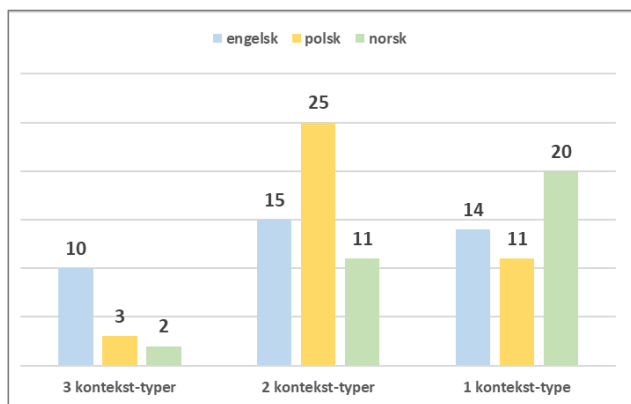
I likhet med i det samlede materialet var det funn på gruppenivå som peker i retning av transfer i den engelske gruppen. Dette vi jeg komme tilbake til etter å ha sett på de individuelle profilene for indefinitt referanse.

9.5.8 Individuelle profiler

Jeg vil nå se på de individuelle profilene for å se om disse kan belyse hvordan hver enkelt innlærertekst står i forhold til prediksjonene.

Det er flere individuelle tekster som har indefinitte enkle substantivfraser i de to innlærergruppene enn det er i den norske gruppen (pol/eng. 39; nor. 33). Figur 9-27 (under) viser hvor mange av konteksttypene for hypotese 1 de ulike individuelle tekstene inneholder. I den polske gruppen er det bare to tekster som har både UB-, BE- og ART-kontekster, og det samme gjelder den norske gruppen, mens det er snakk om hele 10 tekster i den engelske gruppen.

Figur 9-27 Antall tekster med tre, to eller én konteksttype (indefinitt referanse)



Det er den polske gruppen som har flest tekster med *to* av konteksttypene (pol. 25; eng. 15; nor. 11), og her er det en forskjell mellom de to innlærergruppene og den norske gruppen når det gjelder *hvilke* kontekster som er representert. De fleste i den engelske og polske gruppen har UB- og BE-kontekster, mens alle i den norske gruppen har UB- og ART-kontekster. Dette stemmer godt med det vi så på gruppenivå, hvor det var

veldig få BE-kontekster i det norske materialet, og at spesielt den polske gruppen hadde få ART-kontekster.

Det er den norske gruppen som har flest tekster med bare én av konteksttypene (pol. 11; eng.14; nor. 20), og det gjelder for alle unntatt én tekst i hver gruppe at det kun er UB-kontekster som er representert.

Hvilke konteksttyper som er representert i de individuelle profilene, er viktig for analysen av transfereffekter på individnivå, ettersom det er avgjørende at spesielt polske innlærere har BE-kontekster, som er markørkontekster ved indefinitt referanse.

For undersøkelsen av individuelle profiler har jeg talt opp antall av de ulike profiltypene (positiv, negativ eller kombinert), og jeg har i tillegg vurdert hver enkelt polske profil med tanke på sammenfallskontekster og markørkontekster. Spesielt de positive profilene har blitt nærmere undersøkt for om de inneholder både markør- og sammenfallskontekster. Siden det kun er markørkontekster som kan si om formvalgene i de polske tekstene er forskjellig fra det en finner i de engelske og norske tekstene, og dermed si noe om gruppespesifikke formvalg og utelukke målspråket som kilde til formvalg, vil positive profiler være sterkere om de inneholder markørkontekster, mens positive profiler som kun har sammenfallskontekster ikke like godt støtter argumentet for transfer.

Som vi så over, er det veldig få polske innlærertekster som har alle tre konteksttypene, men de aller fleste har to av dem, og da er det stort sett UB-kontekster og BE-kontekster som er representert. Det er 28 av 39 polske innlærertekster som har både sammenfalls- og markørkontekster. De fleste som bare har én konteksttype, har bare sammenfallskontekster (UB), så for disse tekstene vil en positiv profil ikke kunne utelukke målspråkstilegnelse som kilde til formvalget.

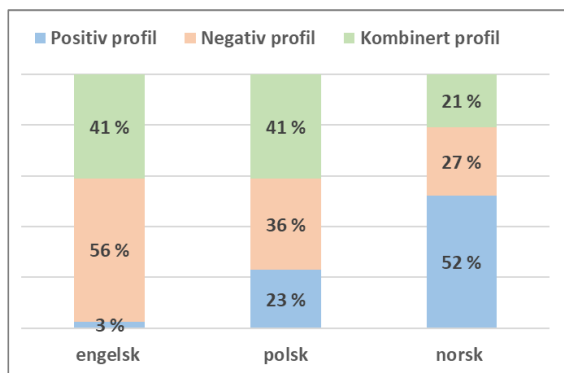
Som tidligere sammenligner jeg antallet positive profiler i den polske gruppen med antallet i de andre gruppene, og der det er forskjeller, bruker jeg kjiqvadrat-test for å si om forskjellen er signifikant.

9.5.8.1 *Resultat - Individuelle profiler*

Figur 9-28 og tabell 9-16 (neste side) viser antall og andel av de ulike profiltypene i de tre gruppene. Antall profiler som har indefinitte substantivfraser, er ulikt i de tre

gruppene (pol. 39; eng. 39; nor. 33). Jeg vil derfor ta utgangspunkt i prosentandeler i sammenligningen av gruppene.

Figur 9-28 Profiltyper i de tre gruppene (indefinit referanse)



Tabell 9-16 Profiltyper i de tre gruppene (indefinit referanse)

Profiltipe	Engelsk	polsk	Norsk
Positiv	3 % (1)	23 % (9)	52 % (17)
Negativ	56 % (22)	36 % (14)	27 % (9)
Kombinert	41 % (16)	41 % (16)	21 % (7)
Total	39	39	33

Det er 23 % av de polske innlærerne (9 stk.) som har en positiv profil, og dermed det predikerte morsmållsamsvaret, mens 36 % har en negativ profil, og 41 % en kombinert. Av de ni innlærerne som har positive profiler har hele åtte *kun* sammenfallskontekster (UB og/eller ART). Bare én av de positive profilene har en markørkontekst i tillegg. For de fleste polske positive profilene vil dermed morsmållsamsvaret være målspåkløst, og det er ikke mulig å undersøke om formvalget for øvrig skiller seg fra målspåkløst siden markørkontekster mangler.

Om man ser nærmere på kombinerte og negative profiler, kan det også kaste lys over hvordan de positive profilene kan fortolkes. Kombinerte profiler er profiler som er i tråd med prediksjonene for noen konteksttyper, men ikke for andre. Hele 15 av de 16 polske kombinerte profilene er positive i sammenfallskontekster, men negative i markørkontekster. Det er også slik at 11 av 14 negative profiler har markørkontekster, men samtlige markørkontekster har en annen form enn predikert. Av de totalt 28 polske innlærerne som har markørkontekster, er det dermed bare 2 som først og fremst bruker

bestemt form der. Dette er vesentlig for hvordan fraværet av markørkontekster i 8 av de 9 positive profilene kan fortolkes. På den ene siden kan man si at fravær av markørkontekst i seg selv ikke *utelukker* morsmålssamsvar, men at det bare ikke er godt nok grunnlag for å konkludere. På den andre siden ser vi at de innlærerne som faktisk har disse markørkontekstene, stort sett *ikke* har det predikerte formvalget.

Konklusjonen er altså at kun én av de polske innlærerne har en tydelig positiv profil som stemmer med prediksjonene, ettersom formvalgene er i tråd med prediksjonene også i en markørkontekst. I tillegg fins det ytterligere 8 positive profiler som ikke har BE-kontekster i det hele tatt, og disse utelukker dermed ikke målspråkstilegnelse. Transfereffekten morsmålssamsvar står dermed ikke veldig sterkt i de polske individuelle profilene.

Transfereffekten *intragruppe-homogenitet*, altså hvor representativt formvalget er for den polske gruppen, ble ikke undersøkt på gruppenivå ettersom det krever informasjon om individuelle formvalg. De individuelle profilene viser imidlertid at homogeniteten i den polske gruppen er lav: Bare 23 % har de predikerte formvalgene, mens de resterende ikke har dette, altså er det bare et mindretall av gruppen som distribuerer formene på denne måten. En langt større andel polske innlærertekster hadde negativ eller kombinert profil, så i den grad gruppen er homogen, så er ikke homogeniteten først og fremst knyttet til det predikerte morsmålssamsvaret. Ettersom disse andre profiltypene ikke er knyttet til ett bestemt distribusjonsmønster, kan man ikke si om disse innebærer noe som kan kalles representative formvalg uten å gå nærmere inn i profilene, men det ligger utenfor denne undersøkelsen.

Sammenlignet med den norske gruppen har den polske gruppen *lavere* andel positive profiler (pol. 23 %; nor. 53 %), og denne forskjellen mellom de to gruppene er signifikant om en sammenligner dem for positive profiler versus negativ/kombinert samlet ($\chi^2=4.4937$, $p < 0.05$). Det vil si at i tillegg til at nesten alle de polske positive profilene hadde formvalg som var sammenfallende med målspråket, så predikerer også hypotesen formvalget til 50 % av den norske gruppen. Samtlige positive profiler i det norske materialet har kun sammenfallskontekster, altså er dette forklaringen på det høye antallet positive profiler. Det er kun to norske profiler som har markørkontekst, og disse er negative, som ventet. Forskjellene mellom den polske og den norske gruppen er altså

knyttet til hvilke konteksttyper som er representert i tekstene i disse to gruppene, spesielt det at det er veldig få BE-kontekster (aktiverte slutfelt) i det norske materialet. Hadde antallet av disse vært høyere i det norske materialet, hadde det nok vært flere negative profiler der også.

Sammenligningen med den norske gruppen viser altså signifikante forskjeller, men de positive profilene i de to gruppene har til felles at positive profiler først og fremst er et resultat av mangel på markørkontekster, med unntak av hos én polsk innlærer.

Det er også forskjeller mellom den polske og den engelske gruppen når det gjelder andel av positive profiler. Den engelske gruppen har kun én eneste positiv profil (3 %), og hele 56 % har en negativ profil, mot 36 % i den polske gruppen. Forskjellen mellom gruppene er signifikant om en sammenligner antall positive profiler med antall negative/kombinerte samlet ($\chi^2=8.5726$, $p < 0.01$). Den ene engelske positive profilen har kun én sammenfallskontekst, og i motsetning til i den norske gruppen er det langt flere individuelle profiler med markørkontekster i den engelske gruppen, da hele 25 engelske innlærere har BE-kontekster (64 %). Dette gjør at forskjellen mellom den engelske og den polske gruppen er tydeligere knyttet til de faktiske formvalgene, og ikke bare til hvilke kontekster som er representert i profilene.

Det kan dermed se ut som at profilmønstrene til en viss grad er spesifikke for polske innlærere i forhold til engelske innlærere, men den store graden av sammenfall med målspåket og den norske gruppen, svekker mistanken om at dette er knyttet til det predikerte morsmålsamsvaret. Inndelingene i profiler er ikke designet for å framheve morsmålspåvirkning hos de engelske innlærerne, men det kan være at intergruppetheterogeniteten vi ser på individnivå mellom den engelske og den polske gruppen, har sammenheng med tendensene til overrepresentasjon av artikkelfrase som vi så i det engelske materialet på gruppenivå.

Totalinntrykket fra undersøkelsen av de individuelle profilene er at polsk som innlærergruppe skiller seg fra den andre innlærergruppen i og med at det er signifikant flere polske innlærertekster som følger prediksjonene enn engelske. Det er likevel bare cirka en fjerdedel av de polske innlærerne som velger form i tråd med prediksjonene, og som vi så hadde bare én av disse markørkontekster, mens de øvrige bare hadde sammenfallskontekster. I tillegg var det signifikant flere i den norske gruppen som

hadde en positiv profil. Dette sier oss at hypotesen for polsk i stor grad også favner formvalg som gjøres i målspråket, og at mangelen på BE-kontekster i de positive profilene gjør det vanskelig å konkludere med at dette er morsmålssamsvar eller transfer. Det at de aller fleste BE-kontekstene i den polske gruppen forekommer i negative profiler, altså ikke har de predikerte formvalgene, taler også imot transfer.

9.5.9 Oppsummering av funn for hypotese 1 ved indefinitt referanse

Tabell 9-17 (under) gir en oversikt over funnene på gruppe- og individnivå. På gruppenivå oppgir tabellen kun prosentandelene for de form/kontekst-assosiasjonene som er predikert for den polske gruppen, og på individnivå er negative og kombinerte profiler slått sammen.

Tabell 9-17 Oversikt over hovedfunn for hypotese 1 ved indefinitt referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
UB-kontekst/ubestemt ent/flt			
- Innen kontekst	74 %	51 %	78 %
- Innen form	65 %	60 %	84 %
BE-kontekst/bestemt ent/flt			
- Innen kontekst	23 %	4 %	0 %
- Innen form	89 %	33 %	0 %
ART-kontekst/artikkel frase			
- Innen kontekst	50 %	14 %	42 %
- Innen form	5 %	3 %	31 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	26 %	3 %	52 %
- Negative/kombinerte	74 %	97 %	48 %

Morsmålssamsvar

I likhet med det vi fant i undersøkelsen av det samlede materialet, er det i det indefinitte materialet ulik grad av morsmålssamsvar for de tre predikerte form/kontekst-assosiasjonene. Høyest samsvar finner vi for assosiasjonen *UB-kontekst/ubestemt entall/flertall*, mens det er lavere og mer ujevnt for assosiasjonen *BE-kontekst/bestemt form*, og enda lavere for *ART-kontekst/artikkel frase*. Selv om antydningen til morsmålssamsvar er mer jevnt innen UB-kontekster, på den måten at det kommer til syne både innen kontekst og innen form, er det en interessant tendens også innen

BE-kontekster: Selv om det stort sett er andre former enn bestemte som brukes i BE-kontekster, så er hele 89 % av frasene med bestemt form faktisk brukt i BE-kontekster (89 %).

Innen indefinitt referanse tok jeg også i betraktning sammenfalls- og markørkontekster på gruppenivå, og for indefinitt referanse var det først og fremst BE-kontekster som var markørkontekst. Dermed ser vi at det er høyest morsmålssamsvar i en sammenfallskontekst (UB), men også noen tendenser til morsmålssamsvar i markørkonteksten (BE).

Intralingval kontrast og intergruppe-heterogenitet

Innen det indefinitte materialet har den polske gruppen vist seg å være mest lik den norske gruppen, og noe mer ulik den engelske gruppen.

Den norske og den polske gruppen har utvist størst forskjell i markørkonteksten (BE), da det ikke er noen assosiasjon mellom BE-kontekster og bestemt form i det hele tatt hos den norske gruppen. Gruppene er eller relativt like for de andre to delprediksjonene. Dette betyr at de største forskjellene mellom den polske og den norske gruppen kom til syne i nettopp markørkonteksten, mens de var mer like i sammenfallskontekster. Den polske gruppen er imidlertid *generelt* lik den norske i formvalget i UB- og ART-kontekster, ettersom det her ikke bare er andelen av den predikerte formen som ligner, men også de øvrige formene: Både bruk av artikkelfrase i UB-kontekster og ubestemt entall/flertall i ART-kontekster, er relativt likt i de to gruppene, selv om disse formvalgene strider imot prediksjonene for den polske gruppen.

Når det gjelder transfereffekten *intergruppe-heterogenitet* mellom de to innlærergruppene, oppviste den polske og den engelske gruppen ulikheter for alle de tre delprediksjonene, noe som har vist at den polske gruppen har noen gruppe-spesifikke form/kontekst-assosiasjoner. Alle de tre predikerte assosiasjonene er sterkere i den polske enn i den engelske gruppen, men for UB- og BE-kontekster har forskjellene først og fremst vært knyttet til høyere andel av artikkelfrase i den engelske gruppen. I ART-kontekster er derimot bruken *lavere* i den engelske gruppen. Dette viser at det er en del intergruppe-heterogenitet for alle de tre predikerte assosiasjonene, men tar en i betraktning at den polske gruppen var veldig lik den norske, er vi igjen inne på det vi så i det samlede materialet, nemlig at det er den engelske gruppen som skiller seg mest ut.

Utover dette er det BE-kontekster som er det området der den polske og den engelske gruppen er *mest* ulike, slik det også var mellom den polske og den norske.

På gruppenivå kan vi dermed konkludere med at det for assosiasjonene *UB-kontekst/ubestemt entall/flertall* og *ART-kontekster/artikkel frase* ikke er noe tydelig argument for transfer. Dette kommer fram gjennom at der det er tendenser til morsmålssamsvar hos de polske innlærerne, er det også *liten* forskjell fra målspråkgruppen, og der det *er* forskjell, er det lavt morsmålssamsvar. De forskjellene som forekommer i forhold til den engelske gruppen, kan se ut til å best fortolkes som at det er den engelske gruppen som har gruppespesifikke formvalg, ikke den polske.

For den siste assosiasjonen (*BE-kontekster/bestemt form*) er det derimot sterkere tegn til transfer: Det foreligger høyere morsmålssamsvar enn i de to andre gruppene, og generelle forskjeller fra de andre gruppene når det gjelder bruk av bestemt form i BE-kontekster. Morsmålssamsvaret var imidlertid lavt innen kontekst, men nesten alle bestemt form brukes i nettopp BE-konteksten. Sett i sammenheng med at dette er den konteksten den polske gruppen skiller seg mest ut i, og at det i tillegg er markørkonteksten, er dette et interessant delfunn.

Hypotesen tilsier uansett at morsmålssamsvar for den polske gruppen skal foreligge for *samtlig*e av de tre delprediksjonene, og som vi har sett er morsmålssamsvaret varierende, og spesielt lavt for assosiasjonen *ART-kontekst/artikkel frase*. Samlet sett er det altså ikke noe godt argument for transfer, men det fins altså noen interessante tendenser med tanke på bruk av bestemt form i BE-kontekster.

Individuelle profiler

Når det gjelder de individuelle profilene, så vi at det var 23 % av de polske innlærerne som hadde en positiv profil, men at mange av disse kun hadde sammenfallskontekster, noe som gjorde det vanskelig å skille morsmålssamsvar fra målspråklikhet.

Antallet polske positive profiler var signifikant *færre* enn i den norske gruppen, men signifikant *flere* enn i den engelske gruppen, men det høye antallet i den norske gruppen så ut til å handle om et høyt antall individuelle profiler som bare besto av sammenfallskontekster, og dermed var positive av den grunn. Det samme gjaldt de aller fleste av de polske profilene som var positive.

De kvantitativt signifikante forskjellene mellom den polske gruppen og de andre to gruppene ble dermed svekket av de mer kvalitative tendensene for hvilke konteksttyper som er representert i de ulike profilene. Den polske og den norske gruppen hadde en god del positive profiler, men de fleste av disse hadde kun sammenfallskontekster, mens det i den engelske gruppen var stort belegg for markørkontekster og få positive profiler.

De individuelle profilene viste ikke bare at morsmålssamsvaret er svakt, men også at de predikerte assosiasjonene mellom form og kontekst ikke er representative for den polske gruppen, altså er transfereffekten *intragruppe-homogenitet* også lav.

Oppsummering: transfereffekter

Innen delmaterialet med indefinitt referanse har jeg nå undersøkt både gruppene som helhet og individuelle profiler med tanke på de fire transfereffektene *morsmålssamsvar*, *intragruppe-homogenitet*, *intergruppe-heterogenitet* og *intralingval kontrast*. Verken på gruppenivå eller individnivå fant jeg noe sterkt argument for transfer. Der morsmålssamsvaret var høyt, var det lite heterogenitet i forhold til de andre gruppene, og de positive profilene i det polske materialet viste seg å stort sett bare ha sammenfallskontekster og lite markørkontekster, noe som gjør at morsmålssamsvaret i disse profilene i stor grad er målpråkslikt.

I den deskriptive delen for indefinitt referanse (9.5.1) så vi at alle de tre gruppene hadde utstrakt bruk av ubestemte former (ubestemt entall, flertall og artikkelfrase) totalt sett, selv om den polske gruppen hadde noe mindre andel enn de andre to (pol. 90 %; eng. 97 %; nor. 98 %). Denne tendensen er ikke uventet i den norske og den engelske gruppen, ved indefinitt referanse er jo ubestemte former det typiske formvalget i både norsk og engelsk, men det at dette også gjelder den polske gruppen, tyder på at de polske innlærerne generelt ligger nært opp til målpråket i sine formvalg. Tatt i betraktning at det er svake argument for transfer hos den polske gruppen, spesielt at det er sterkt sammenfall med målpråksgruppen, så kan det se ut som de polske innlærernes formvalg er målpråksbaserte, heller enn morsmålsbaserte, ved indefinitt referanse. Jeg vil komme tilbake til dette når jeg har sett på om tilsvarende mønster også forekommer ved definitt og generisk referanse.

Indefinit referanse og aktivering og posisjon isolert sett

I undersøkelsen av det samlede materialet så jeg i tillegg på forholdet mellom de ulike formtypene og setningsposisjon og aktiveringsgrad uavhengig av hverandre, og fant at den polske gruppen brukte ubestemt form i slutfeltet, og bestemt form ved aktiverte referenter, oftere enn i de predikerte kontekstene. Denne tendensen gjaldt stort sett den engelske gruppen også, mens den norske gruppen brukte alle former først og fremst ved ikke-aktiverte referenter.

Om vi ser tilbake på hvordan dette ser ut i det indefinite materialet, i tabellene på side 259, ser vi at fordelingen er annen i alle de tre gruppene. I den polske gruppen (tabell 9-13) er *alle* de tre formtypene først og fremst brukt i slutfeltet (ub. 97 %; be. 100 %; art. 91 %). Det er imidlertid først og fremst slutfelt som er representert i dette materialet, og det er kun 4 % av substantivfrasene som står i et forfelt.

En finner den samme tendensen i den engelske gruppen (tabell 9-14, s. 259), hvor også alle formtypene brukes mest i slutfeltet uavhengig av aktivering (ub. 83 %; be. 66 %; art. 87 %), men her er andelen forfelt høyere enn i det polske materialet (16 %), og tendensen er dermed noe svakere hos den engelske gruppen enn hos den polske.

Den norske gruppen viser i likhet med i det samlede materialet en annen profil enn innlærergruppene (tabell 9-15, s. 259). Her er ubestemt entall/flertall og artikkelfrase først og fremst brukt når referenten er *ikke-aktivert* (ub. 84 %; art. 96 %), mens bestemt form er mest brukt ved *ikke-aktiverte slutfelt*, men det er bare snakk om *to* fraser med bestemt form. Også her virker det som tallene avspeiler den generelle forekomsten av ulike konteksttyper: I det norske materialet er det først og fremst ikke-aktiverte kontekster som er representert (95 %), og det er bare *fem* aktiverte kontekster (4,5 %).

Det kan dermed virke som disse mønstrene først og fremst er en avspeiling av hvilke konteksttyper som er til stede i tekstene i de ulike gruppene, heller enn at formene er spesielt assosiert med en viss posisjon eller aktiveringsgrad. At nesten alle formene er brukt i samme type kontekst *innen* hver av gruppene, peker også i den retningen. Forskjellene mellom innlærergruppene på den ene siden og norsk på den andre, er i så fall knyttet til at de indefinite referentene i det norske materialet som regel er

ikke-aktiverte, mens substantivfraser med indefinitte referenter som regel er plassert i slutfeltet av innlærergruppene. Dette er en interessant forskjell mellom innlærere og morsmålsbrukere, men det ligger utenfor denne undersøkelsen å gå nærmere inn på dette.

Transfer hos de engelske innlærerne

Som jeg også var inne på i undersøkelsen av det samlede materialet, er det funn innen det indefinitte materialet som kan tyde på at det er transfer hos de engelske innlærerne. I både UB- og BE-kontekster hadde den engelske gruppen høyere andel av artikkelfrase enn begge de andre gruppene, mens andelen var *lavere* i ART-kontekster. Vi så også i de deskriptive tallene for indefinitt referanse at andelen av artikkelfrase i den engelske gruppen var høyere totalt enn for de andre to gruppene (eng. 42 %; pol. 22 %; nor. 24 %). Det er spesielt innen indefinitt referanse at forskjellen i bruk av artikkelfrase versus ubestemt entall ved tellelige substantiv i norsk og engelsk kommer til uttrykk, så dette kan være et tegn på at engelske innlærere er påvirket av morsmålet i formvalget for tellelige substantiv.

Som i det samlede materialet, var det observasjoner som tyder på både morsmålsamsvar, intergruppe-heterogenitet og forskjeller fra målspråket når det gjaldt de engelske innlærernes bruk av artikkelfrase i deler av det indefinitte materialet. Den tendensen vi så i det samlede materialet, er altså mest sannsynlig et utslag av det som kom fram innen indefinitt referanse, siden det virker mer sannsynlig at den engelske gruppen vil ha overrepresentasjon av artikkelfrase ved indefinitt referanse enn ved definitt og generisk referanse. Dette fordi det er indefinitt referanse som er kjerneområdet for bruk av ubestemt form i engelsk, og det er spesielt her forskjellen mellom engelsk og norsk i bruk av artikkelfrase ligger.

Konklusjon

Den endelige konklusjonen for det indefinitte materialet er altså at det ikke foreligger argument for transfer fra polsk av den typen som hypotese 1 predikerer, verken på gruppenivå eller på individnivå. Det er tendenser til gruppespesifikke formvalg i den polske gruppen knyttet til den ene delprediksjonen om assosiasjon mellom

BE-kontekst/bestemt form, men isolert sett er ikke dette tilstrekkelig for å understøtte hypotese 1.

9.6 Definit referanse

I neste del av undersøkelsen vil jeg undersøke transfereffekter for de definte substantivfrasene. Definit referanse foreligger når en substantivfrase viser til én eller flere individuelle referenter som vurderes som identifiserbar for mottageren av ytringen, enten gjennom *anafori, situasjon, kunnskap om verden* eller ved *assosiasjon til en annen referent* (se 8.4.4). Definit referanse står i kontrast til indefinit referanse, hvor én eller flere individuelle referenter *ikke* regnes som identifiserbare for mottageren innen et av domene som er nevnt over. Definit referanse står også i kontrast til *generisk referanse*, som viser generelt til hele klasser av referenter heller enn individuelle referenter. Prediksjonene om de polske innlærernes formvalg i UB-, BE- og ART-kontekster er de samme som tidligere, og jeg vil under vise noen eksempler på de ulike konteksttypene med substantivfraser som har definit referanse.

Definit UB-kontekst

Eksempel 9.6-a (under) viser den definte substantivfrasen «telefoner» som er plassert i forfeltet til en hovedsetning, og som er klassifisert som aktivert gjennom at mobiltelefoner er nevnt i forrige setning. Dette gjør dette til en UB-kontekst av typen *aktivert forfelt*. Siden det er brukt ubestemt flertallsform, er formvalget i tråd med prediksjonen for UB-kontekster.

9.6-a Desverre vi kan høre signaler fra mobiltelefon på skoler, på kinoer, på teater og andre steder hvor telefoner skulle være på stille. Jeg tror at det er også uhoflig vis vi er på fest og **telefoner** ligger på bordet foran oss. (s0791)

Referansen til frasen «telefoner» er regnet som definit fordi det er snakk om individuelle telefoner, ikke hele klassen av telefoner, og fordi disse telefonene er identifiserbare i den forstand at det er vanlig at folk («vi») eier telefoner. Denne typen etablering av definit referanse er derfor knyttet til en persons personlige sfære, og referenten til frasen «telefoner» er dermed identifiserbar som «telefonene sine» (se 8.4.4 for omtale av denne typen definit referanse). Det er med andre ord ikke fordi telefonene

er nevnt i forrige setning at denne frasen regnes som definitt, ettersom det ikke er noe anaforisk forhold mellom de to telefonreferansene i eksemplet.

Definit BE-kontekst

Det neste eksemplet (9.6-b) viser den definte substantivfrasen «mobiltelefonen» som er plassert i slutfeltet til en hovedsetning, og som er klassifisert som aktivert fordi mobiltelefoner er nevnt i foregående setning. Dette er altså en BE-kontekst (aktivert slutfelt), og siden det er brukt bestemt form, er formvalget i tråd med prediksjonen.

9.6-b Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

«Mobiltelefonen» regnes som definitt på samme grunnlag som i forrige eksempel (personlig sfære). Frasen «mobiltelefoner» i den foregående setningen er derimot *generisk*. Den understrekede substantivfrasen er altså heller ikke her definitt i kraft av å være nevnt før, og referansen er ikke aktivert på grunn av samreferanse, men fordi en referent av samme *type* har vært nevnt før.

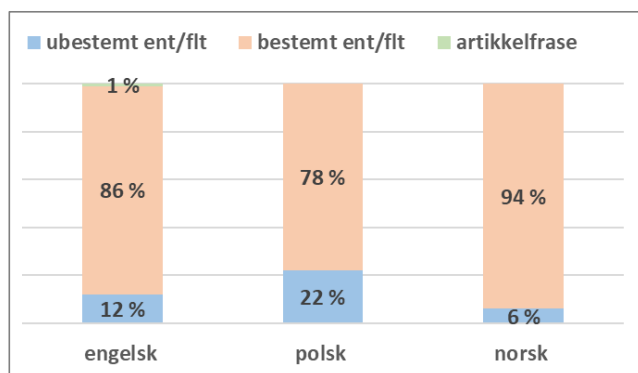
Definit ART-kontekst

Det fins ingen ART-kontekster (ikke-aktiverte forfelt) med artikkelfrase i det definte materialet hos noen av gruppene, derfor har jeg ikke noe eksempel på en slik kontekst med sin predikerte form. Eksemplet under (9.6-c) viser i stedet en ART-kontekst med ubestemt form uten artikkel.

9.6-c Det sjette noen dager siden i Krakow. Fire andre studenter er på sykehus men **lege** sier at de blir friske. (s0793)

Substantivfrasen «lege» er her plassert i forfeltet til en hovedsetning, og referenten regnes som ikke-aktivert siden det ikke nevnes noen lege i samme eller foregående setning. Frasen «lege» regnes som definitt i kraft av å være assosiert med en annen referent, nemlig «sykehus» i setningen før. Siden alle sykehus har leger, blir «lege» indirekte introdusert og identifiserbar. Dette er dermed et eksempel på identifiserbarhet gjennom assosiasjon.

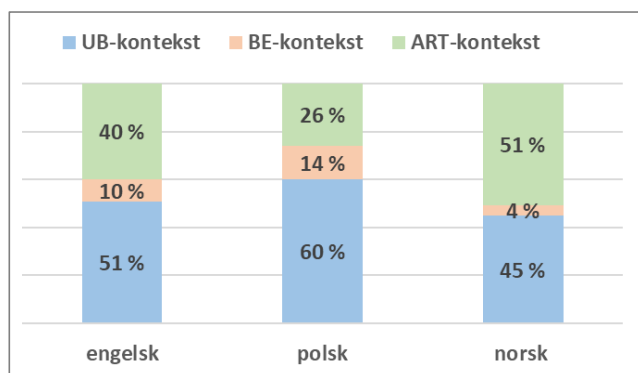
Figur 9-29 Andel av de ulike formtypene i de tre gruppene (definit referanse)



Tabell 9-18 Andel av de ulike formtypene i de tre gruppene (definit referanse)

Morsmål	Ubestemt ent/flt	Bestemt ent/flt	Artikkelfrase	Totalt
Polsk	22 % (19)	78 % (66)	-	85
Engelsk	12 % (9)	86 % (63)	1 % (1)	73
Norsk	6 % (4)	94 % (65)	-	69

Figur 9-30 Andel av de ulike kontekststypene i de tre gruppene (definit referanse)



Tabell 9-19 Andel av de ulike kontekststypene i de tre gruppene (definit referanse)

Morsmål	UB-kontekst	BE-kontekst	ART-kontekst	Totalt
Polsk	60 % (51)	14 % (12)	26 % (22)	85
Engelsk	51 % (37)	10 % (7)	40 % (29)	73
Norsk	45 % (31)	4 % (3)	51 % (35)	69

Jeg vil komme tilbake til flere eksempler. Gjennomgangen av det definte materialet vil være som tidligere i kapittelet, altså først form/kontekst-assosiasjoner på gruppenivå og deretter til individuelle profiler. Før gjennomgangen av resultat vil jeg se på deskriptive tall for materialet, og hva som regnes som sammenfalls- og markørkontekst ved defintt referanse.

9.6.1 Deskriptivt for alle gruppene

Det definte materialet er en god del mindre enn det indefinte, og det er den polske gruppen som har høyest antall enkle substantivfraser med defintt referanse (pol. 85; eng. 73; nor. 69). Til tross for at det polske materialet slikt sett er størst, er det likevel denne gruppen som har færrest individuelle tekster med defintt referanse, kun 25 stykker, mens den engelske og norske gruppen har henholdsvis 34 og 33 tekster. Søylediagrammene og tabellene på neste side viser andelen av former og konteksttyper i de tre gruppene.

Figur 9-29 og tabell 9-18 (øverst forrige side) viser andelen av de ulike formtypene, og vi ser at det er bestemt form som dominerer i alle de tre gruppene, selv om det er aller mest av denne formen i det norske materialet, og en del mindre i det polske (pol. 78 %; eng. 86 %; nor. 94 %). Tilsvarende har alle gruppene en liten andel ubestemt entall/flertall, men aller mest i den polske gruppen (pol. 22 %; eng. 2 %; nor. 6 %). Den norske gruppen har bare fire fraser med ubestemt form, noe som ikke er overraskende siden det er snakk om defintt referanse. Det er dessuten kun én artikkel frase i hele det definte materialet, og denne er i en engelsk innlærtekst.

Figur 9-30 og tabell 9-19 (nederst forrige side) viser andelen av de ulike konteksttypene. Den norske gruppen skiller seg ut med å ha bare *tre* BE-kontekster, som er en del færre enn de andre to gruppene (pol. 14 %; eng. 10 %; nor. 4 %), en tendens som også var til stede ved indefintt referanse. Det er også en merkbar forskjell i andelen av ART-kontekster, som er høyere i den engelske og norske gruppen enn i den polske (pol. 26 %; eng. 40 %; nor. 51 %).

9.6.2 Sammenfalls- og markørkontekster for defintt referanse

I 9.2.3 viste jeg hvordan prediksjonene for den polske gruppen er ventet å sammenfalle

med eller være ulik engelske og norske formvalg innen definitt referanse. Der det polske materialet er forventet å fordele formene etter konteksttypene UB, BE og ART, er den engelske og norske gruppen forventet å ha bestemt form i *alle* de tre konteksttypene, siden det er snakk om definitt referanse. Det er dermed BE-kontekster som er sammenfallskontekster ved definitt referanse, siden alle tre gruppene er forventet å bruke bestemt form. UB- og ART-kontekster vil derimot være markørkontekster, ettersom den polske gruppen er forventet å bruke ubestemt form, mens de to andre gruppene er forventet å bruke bestemt form. Eksempler på dette ble omtalt tidligere som eksempel 9.6-a og 9.6-b, og vi så at de to substantivfrasenes form var ulik selv om begge var defintitte. Disse eksemplene viste bare variasjonen mellom UB-kontekster og BE-kontekster, siden det ikke fins noen eksempler på ART-kontekster med artikkelfrase i materialet.

Tabell 9-20 (under) viser forventede formvalg for de tre gruppene, med markørkontekstene uthevet.

Tabell 9-20 Sammenfalls- og markørkontekster for definitt referanse

Definitt referanse	Polsk	Engelsk	Norsk
UB-kontekst	Ubestemt ent/flt	Bestemte former	Bestemte former
BE-kontekst	Bestemte former	Bestemte former	Bestemte former
ART-kontekst	Artikkelfrase	Bestemte former	Bestemte former

Med dette som utgangspunkt er det ventet at intergruppe-heterogenitet mellom den polske og den engelske gruppen først og fremst vil slå ut i UB- og ART-kontekster, siden de to morsmålene impliserer ulike formvalg i disse kontekstene, og det er også disse kontekstene prediksjonene om polsk morsmålssamsvar ikke sammenfaller med målspåket. Individuelle positive profiler i den polske gruppen som har UB- og ART-kontekster, står altså sterkere som tegn på morsmålssamsvar, ettersom disse i større grad kan isolere morsmålet som kilde til formvalget.

9.6.3 Samlet oversikt for definitt referanse

Tabellene på neste side viser statistikk for former og kontekster i de tre gruppene ved definitt referanse, og disse tallene vil bli visualisert forløpende med søylediagram.

Tabell 9-21 Polsk gruppe: Fordeling av former etter kontekststype (definit)

SI= POLSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster: Former:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	7	2	1	9	19
Andel innen form	37 %	11 %	5 %	47 %	100 %
Andel innen kontekst	15 %	40 %	8 %	41 %	22 %
BESTEMT ENT/FLT	39	3	11	13	66
Andel innen form	59 %	5 %	17 %	20 %	100 %
Andel innen kontekst	85 %	60 %	92 %	59 %	78 %
ARTIKKELFRASE	0	0	0	0	0
Andel innen form					
Andel innen kontekst					
TOTALT	46 (54%)	5 (6%)	12 (14%)	22 (26%)	85

Tabell 9-22 Engelsk gruppe: Fordeling av former etter kontekststype (definit)

SI= ENGELSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster: Former:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	5	2	1	1	9
Andel innen form	56 %	22 %	11 %	11 %	100 %
Andel innen kontekst	19 %	20 %	14 %	3 %	12 %
BESTEMT ENT/FLT	21	8	6	28	63
Andel innen form	33 %	13 %	10 %	44 %	100 %
Andel innen kontekst	78 %	80 %	86 %	97 %	86 %
ARTIKKELFRASE	1	0	0	0	1
Andel innen form	100 %				100 %
Andel innen kontekst					1 %
TOTALT	27 (37%)	10 (14%)	7 (20%)	29 (40%)	73

Tabell 9-23 Norsk gruppe: Fordeling av former etter kontekststype (definit)

SI= NORSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Kontekster: Former:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	0	0	0	4	4
Andel innen form	0 %	0 %	0 %	100 %	100 %
Andel innen kontekst	0 %	0 %	0 %	11 %	6 %
BESTEMT ENT/FLT	9	22	3	31	65
Andel innen form	14 %	34 %	5 %	48 %	100 %
Andel innen kontekst	100 %	100 %	100 %	89 %	94 %
ARTIKKELFRASE	0	0	0	0	0
Andel innen form					
Andel innen kontekst					
TOTALT	9 (13%)	22 (32%)	3 (4%)	35 (51%)	69

9.6.4 Assosiasjon mellom UB-kontekster og ubestemt entall/flertall

I UB-kontekster (aktiverte forfelt og ikke-aktiverte slutfelt) er det forventet at den polske gruppen bruker ubestemt entall/flertall, og at denne typen form først og fremst brukes i denne konteksttypen. Dette er fordi markering av tekstuell identifiserbarhet (aktivert/ikke-aktivert) antas å oppleves som redundant, noe som gjør at verken bestemt form eller artikkelfrase vil bli brukt i disse konteksttypene.

Eksemplene under viser eksempler på UB-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for UB-kontekster, deretter et eksempel på formvalg som *ikke* er i tråd med prediksjonen. Det fins ingen UB-kontekster med artikkelfrase, så dette kan ikke gis eksempler på.

Eksemplet på en UB-kontekst med predikert formvalg ble omtalt som definitivt innledningsvis i 9.6, men er gjentatt under (9.6-d). Frasen «telefoner» er klassifisert som aktivert forfelt, og har ubestemt flertallsform, som er i tråd med prediksjonen for UB-kontekster.

9.6-d Desverre vi kan høre signaler fra mobiltelefon på skoler, på kinoer, på teater og andre steder hvor telefoner skulle være på stille. Jeg tror at det er også uhøflig vis vi er på fest og **telefoner** ligger på bordet foran oss. (s0791)

Det neste eksemplet (9.6-e) viser en UB-kontekst der formvalget ikke stemmer med prediksjonen fordi det er brukt bestemt form.

9.6-e To år siden opplevde jeg veldig trist for meg ulikker. Det var min katten som var påkjørt foran mitt hus. **Katten** gikk tilbake fra skogen hjem og ble påkjørt på veien. (s0940)

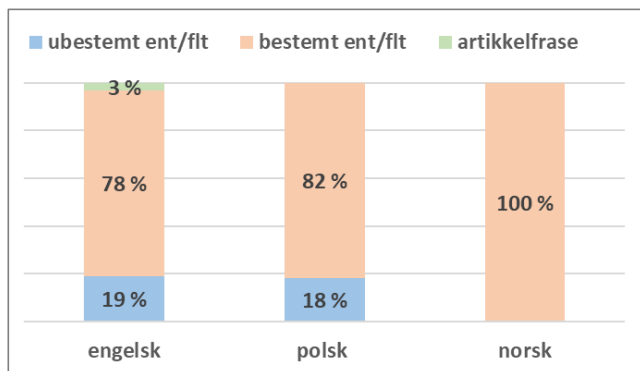
Her er frasen «katten» plassert i forfeltet til en hovedsetning, og har samme referanse som «min katten» i foregående setning, altså er det snakk om et aktivert forfelt, og det er også definitivt referanse gjennom anafori.

Søylediagrammene på neste side illustrerer forholdet mellom UB-kontekster og ubestemt entall/flertall for de tre gruppene. Jeg vil først se på den polske gruppen.

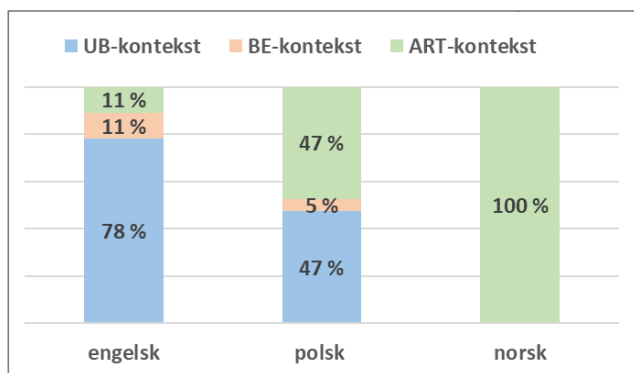
Figur 9-31 (øverst neste side) viser hvilke former som er brukt i definitive UB-kontekster. Det er kun en relativt liten andel av UB-kontekstene til den polske gruppen som har ubestemt entall/flertall (18 %), mens det er brukt hele 82 % bestemt form. Figur 9-32 (nederst neste side) viser hvilke kontekster fraser med ubestemt

entall/flertall er brukt i, og det er her en jevn fordeling mellom bruk i UB- og ART-kontekster (UB. 47 %; ART. 47%). Det må derfor sies å være liten grad av morsmållsamsvar for denne form/kontekst-assosiasjonen.

Figur 9-31 Andel av form innen UB-kontekst



Figur 9-32 Andel av kontekst innen ubest. ent/flt



Om vi ser på den norske gruppen i de samme søylediagrammene over, ser vi at det er en helt konsekvent bruk av *bestemt form* i UB-kontekster (figur 9-31), og at samtlige fraser med ubestemt entall/flertall er brukt i ART-kontekster (figur 9-32). Her er det imidlertid snakk om kun fire forekomster. UB-konteksten er en av markørkontekstene ved definitt referanse, og det er altså en forskjell mellom den norske og den polske gruppen: Den norske gruppen har *ingen* assosiasjon mellom denne formen og konteksttypen i det hele tatt, mens den polske gruppen har en svak assosiasjon (18 % innen kontekst/47 % innen

form). Dette viser at den polske gruppen har formvalg som ikke ligner det en finner i målspråket, noe som stemmer bra med at dette er en markørkontekst, men det er likevel ikke noen sterk tendens til morsmålssamsvar.

Også den engelske og polske gruppen er forventet å ha ulike formvalg i UB-kontekster. Om man sammenligner disse to gruppene i de samme søylediagrammene på forrige side, ser man at de er ganske like når det gjelder hvilke former som brukes i UB-kontekster. Begge gruppene har nemlig rundt 20 % ubestemt entall/flertall og 80 % bestemt form (figur 9-31, øverst). I figur 9-32 (nederst) er det derimot tydelige forskjeller mellom gruppene. Den engelske gruppen har en noe *sterkere* tendens enn den polske til å bruke fraser med ubestemt entall/flertall i UB-kontekster (pol. 47%; eng. 78 %). I tillegg bruker den polske gruppen oftere fraser med ubestemt entall/flertall i ART-kontekster enn det den engelske gjør (pol. 47 %; eng. 11 %). De to innlærergruppene er altså homogene i formvalget innen UB-kontekster, men heterogene når det kommer til hvor slike former brukes. Heterogeniteten er dessuten preget av at formvalgene i den engelske gruppen er *mer* i tråd med prediksjonene for den polske gruppen, enn den polske gruppen selv er.

Det samlede bildet er derfor at det ikke er sterke tegn på transfer for assosiasjonen mellom UB-kontekster og ubestemt entall/flertall ved definitt referanse. Den polske gruppen har lite bruk av ubestemt entall/flertall i UB-kontekster, og i den grad disse formene brukes, brukes de like mye i de to innlærergruppene. Den polske gruppen bruker først og fremst *bestemt form* i denne konteksten, noe som er felles for alle tre gruppene, selv om den norske er mer kategorisk her enn de to innlærergruppene. Denne kvantitative forskjellen fra den norske målspråksgruppen er imidlertid ikke spesifikk for de polske innlærerne.

9.6.5 Assosiasjon mellom BE-kontekster og bestemt form

I BE-kontekster (aktiverte sluttfelt) er det forventet at den polske gruppen bruker bestemt form, og det er også forventet at bestemt form først og fremst brukes i BE-kontekster, ikke i andre konteksttyper. Dette er fordi substantivfrasene har en slik plassering at det oppleves som nødvendig å markere at referenten er identifiserbar (aktivert).

Eksemplene under viser eksempler på BE-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for BE-kontekster, deretter et eksempel på formvalg som *ikke* er i tråd med prediksjonene.

Eksempel på en BE-kontekst med predikert formvalg ble omtalt allerede innledningsvis i 9.6, men er gjentatt under (9.6-f). Det er her snakk om den aktiverte referenten «mobiltelefonen», som er plassert i slutfeltet. Bruk av bestemt form stemmer med prediksjonen for BE-kontekster.

9.6-f Mobiltelefoner ble veldig populært i de siste årene. Nesten alle mennesker har **mobiltelefonen** i lomma eller i veska. (s0791)

Det fins ingen eksempler på BE-kontekster som har artikkelfrase i materialet, men det fins ett eksempel på bruk av ubestemt form i BE-kontekster (9.6-g):

9.6-g Det er viktig føler jeg at man må forstå andre og hva de tenker og sier og respektere **tradisjoner** så man leve et komfortabel og stabil liv. (s0299)

Formvalget i dette siste eksemplet stemmer altså ikke med prediksjonen for denne konteksttypen.

Søylediagrammene på neste side illustrerer forholdet mellom BE-kontekster og bestemt form for de tre gruppene. Jeg vil først se på den polske gruppen.

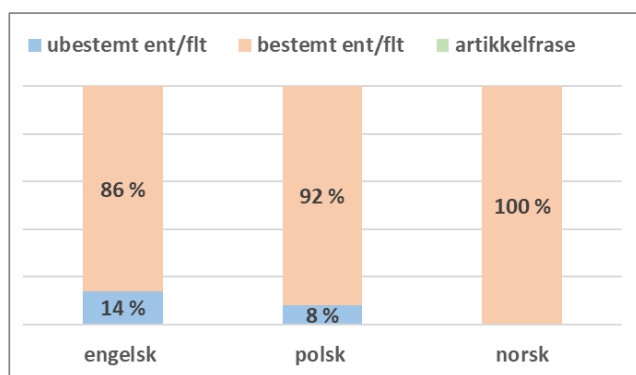
Figur 9-33 (øverst neste side) viser hvilke former som brukes i BE-kontekster i de tre gruppene. I den polske gruppen har 11 av 12 BE-kontekster bestemt form (92 %), men av alle frasene med bestemt form er det bare 17 % som er brukt i BE-kontekster (figur 9-34, nederst neste side). Morsmålssamsvaret er altså veldig høyt innen kontekst, da de fleste BE-kontekstene faktisk har bestemt form, men lavere innen form, siden formen først og fremst brukes i UB- og ART-kontekster.

Om vi ser på den norske gruppen i de samme søylediagrammene, ser vi at denne gruppen utelukkende har brukt bestemt form i BE-kontekster, som forventet i norsk ved definitiv referanse (figur 9-33). Av alle frasene med bestemt form er det likevel bare 5 % som brukes i BE-kontekster (figur 9-34). Dette kommer nok av det lave antallet BE-kontekster i det norske materialet (3 stk). I BE-kontekster er det forventet sammenfall mellom den polske og den norske gruppen, noe det også er *innen kontekst*, hvor begge grupper først og fremst bruker bestemt form (nor. 100 %; pol. 92 %), men

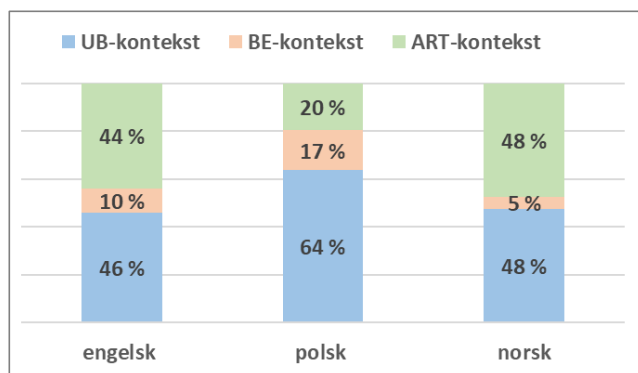
det er forskjeller i hvor stor del av frasene med bestemt form som brukes i denne konteksten (pol. 17 %; nor. 5 %).

Om vi sammenligner den polske gruppen med den engelske i de samme søylediagrammene, ser vi at den polske gruppen også er veldig lik den engelske når det gjelder å ha størst andel fraser med bestemt form i BE-kontekster (eng. 86 %; pol. 92 %). Når det gjelder hvilke kontekster bestemt form brukes i, gjør den samme typen forskjell som mellom den polske og den norske gruppen seg gjeldende, nemlig at den polske gruppen oftere bruker fraser med bestemt form i BE-kontekster (pol. 17 %; eng. 10 %).

Figur 9-33 Andel av form innen BE-kontekst



Figur 9-34 Andel av kontekst innen bestemt form



Det altså høyt morsmålsamsvar i den polske gruppen *innen kontekst* (92%), men lavere innen form (17%), og alle tre gruppene har utstrakt bruk av bestemt form, altså er den polske gruppen relativt lik begge de andre gruppene. Dette viser at tendensene til morsmålsamsvar i den polske gruppen ikke er gruppespesifikke verken i forhold til den

norske eller den engelske gruppen. En slik likhet mellom gruppene er imidlertid ikke uventet siden BE-konteksten er sammenfallskontekst ved definit referanse.

9.6.6 Assosiasjon mellom ART-kontekster og artikkelfrase

I ART-kontekster (ikke-aktiverte forfelt) er det forventet at den polske gruppen bruker artikkelfrase, og det er også forventet at artikkelfrase først og fremst brukes i ART-kontekster, ikke i andre konteksttyper. Dette er fordi substantivfrasene har en slik plassering at det oppleves som nødvendig å markere at referenten *ikke* er identifiserbar (ikke-aktivert).

Som nevnt innledningsvis i 9.6, fins det *ingen* ART-kontekster med artikkelfrase i materialet, derfor har jeg ingen eksempler på kontekster som innfrir prediksjonen for denne konteksttypen. Eksempelene under viser i stedet to ART-kontekster der formvalget bryter med prediksjonen fordi det er brukt bestemt form (9.6-h) og ubestemt form uten artikkel (9-6-i). Begge eksemplene er omtalt tidligere i henholdsvis 9.1 og 9.6.

9.6-h Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikshospital sykehuset. **Legene** sa at operasjon gikk bra. (s0745)

9.6-i Det sjedde noen dager siden i Krakow. Fire andre studenter er på sykehus men **lege** sier at de blir friske. (s0793)

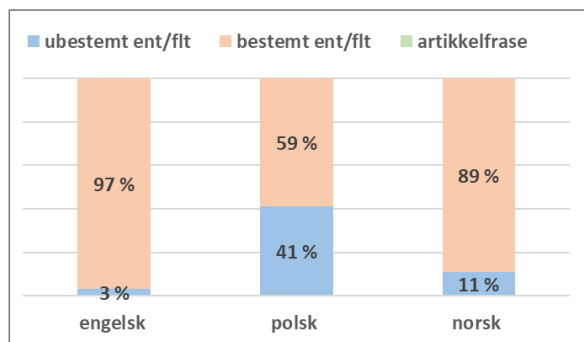
I begge disse eksemplene er de understrekede substantivfrasene plassert i forfeltet, og siden det ikke er nevnt noe om leger tidligere, regnes referansen som ikke-aktivert. Dette er dermed eksempler på ART-kontekster. Både «legene» og «lege» regnes som definte substantivfraser fordi referentene er identifiserbare gjennom å være assosiert med sykehus, som *er* nevnt i foregående setning i begge eksemplene.

Søylediagrammene på neste side illustrerer forholdet mellom ART-kontekster og artikkelfrase for de tre gruppene. Jeg vil først se på den polske gruppen.

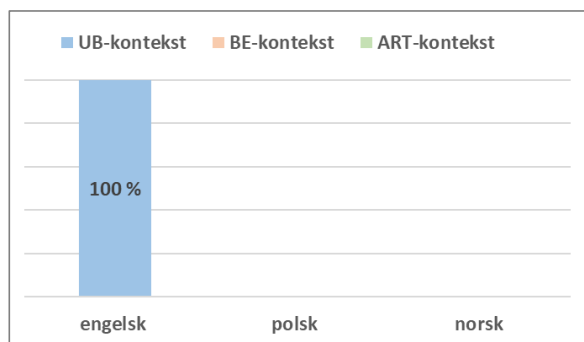
Figur 9-35 og 9-36 (neste side) viser henholdsvis hvilke former som er brukt i ART-kontekster og hvilke kontekster artikkelfrase er brukt i. Det fins *ingen* artikkelfraser hos den polske gruppen, og dermed fins det heller ingen ART-kontekster med artikkelfrase. Fraværet av artikkelfraser gjelder også den norske gruppen, og den engelske gruppen har kun én eneste slik frase.

Dette vil si at det for delprediksjonen om *ART-kontekst/artikkelfrase* mangler både morsmålsamsvar, intergruppe-heterogenitet og forskjell fra målspråksgruppen. ART-kontekster er markørkontekst ved definitt referanse, og vi ser at til tross for at den polske gruppen er ventet å være ulik de andre gruppene her om hypotesen stemmer, så er likheten mellom gruppene stor.

Figur 9-35 Andel av form innen ART-kontekst



Figur 9-36 Andel kontekst innen artikkelfraser



Selv om den polske gruppen er lik de andre to gruppene med hensyn til å ikke bruke artikkelfrase i ART-kontekster, er det noen forskjeller mellom gruppene når det gjelder hvilke former som *er* brukt i ART-kontekster. Mens den norske og den engelske gruppen stort sett bruker bestemt form her, har den polske gruppen en merkbart lavere andel av dette (pol. 59 %; eng. 97 %; nor. 89 %). Av de tre konteksttypene for hypotese 1 er det i ART-kontekster at den polske gruppen har *lavest* andel av bestemt form (UB 82 %; BE. 92 %; ART. 59 %), som er en interessant observasjon jeg vil komme tilbake til under, med tanke på om det kan være noe med ART-kontekster (ikke-aktiverte forfelt)

som gjør at de polske innlærerne sjeldnere bruker bestemt form, slik det er korrekt å gjøre i målspråket.

9.6.7 Oppsummering av funn på gruppenivå

Jeg har så langt sett på de tre predikerte form/kontekst-assosiasjonene på gruppenivå, og ved definnitt referanse er det UB- og ART-kontekster som er markørkontekster, mens BE-kontekster er sammenfallskontekster.

For assosiasjonen *UB-kontekst/ubestemt entall/flertall*, hvor den polske gruppen var forventet å være ulik de to andre gruppene, var morsmålssamsvaret lavt og sammenfallet med de to andre gruppene tydelig, altså er det verken tegn til morsmålssamsvar eller relevant intergruppe-heterogenitet og forskjell fra målspråket for denne delprediksjonen.

For assosiasjonen *BE-kontekst/bestemt form* brukte den polske gruppen ofte bestemt form i BE-kontekster, men dette er sammenfallskontekst for definnitt referanse, og det var også store likheter mellom gruppene. At bare 17 % av alle frasene med bestemt form ble brukt i BE-kontekster av de polske innlærerne, taler også mot en særlig assosiasjon mellom form og kontekst.

For assosiasjonen *ART-kontekst/artikkelfrase* var det forventet forskjeller mellom gruppene, og i tillegg til at det ikke var morsmålssamsvar i den polske gruppen, var de tre gruppene helt like i å ikke bruke artikkelfrase i ART-kontekster i det hele tatt. I den grad den polske gruppen var forskjellig fra de to andre, var det knyttet til mindre bruk av bestemt form her.

De transfereffektene som er undersøkt så langt, er altså svake for alle de tre delprediksjonene om form/kontekst-assosiasjon. Transfereffektene er spesielt svake i de to markørkontekstene (UB og ART), hvor det var ventet at forskjellen mellom gruppene skulle være tydeligst om hypotesen stemte, så det er lite på gruppenivå som peker i retning av transfer.

Det som heller ser ut til å være tendensen i den polske gruppen, er utstrakt bruk av bestemt form (UB. 82 %; BE. 92 %; ART. 59 %). Dette er i stor grad likt de to andre gruppene, som også stort sett bruker bestemt form. Den engelske gruppen bruker mellom 78 % og 97 % bestemt form i de ulike konteksttypene, og den norske mellom

89 % og 100 %. Den polske gruppen ligger altså noe under de to andre gruppene, spesielt i ART-kontekster, som jeg var inne på over.

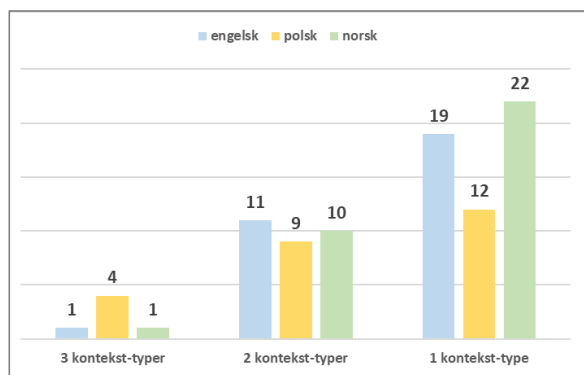
Jeg vil nå se på de individuelle profilene for de tre gruppene ved definitt referanse.

9.6.8 Individuelle profiler

I den polske gruppen er det bare 25 innlærertekster som har enkle substantivfraser med definitt referanse, mens det i den norske og engelske gruppen er henholdsvis 34 og 33, altså er det en del færre individuelle profiler i den polske gruppen. Figur 9-37 (under) viser antallet av de individuelle profilene som har alle tre, eller bare to eller én av konteksttypene.

Det er veldig få tekster som har alle tre konteksttypene, men noe mer i den polske enn i de andre to gruppene (pol. 4; eng./nor. 1). Utover dette er det et relativt jevnt antall som har to konteksttyper i de tre gruppene (pol. 9; eng. 11; nor.10), og det er færre i den polske gruppen som bare har én av konteksttypene, mens det gjelder godt over halvparten i den norske og den engelske gruppen (pol. 12; eng. 19; nor. 22).

Figur 9-37 Antall tekster med tre, to eller én konteksttype (definitt referanse)



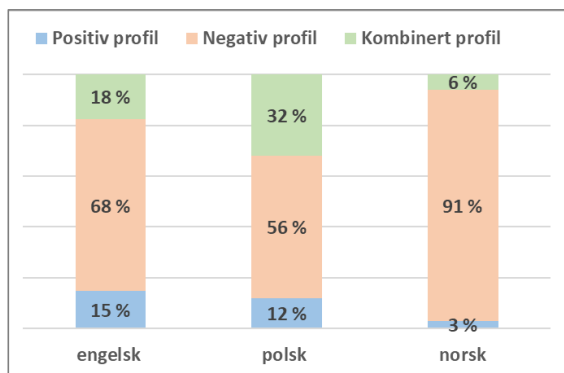
Av de profilene som bare har to konteksttyper, er det i stor grad snakk om UB- og ART-kontekster for alle gruppene. Det er også disse to konteksttypene som dominerer hos de som har bare én av typene. Siden det er UB- og ART-kontekster som er markørkontekster ved definitt referanse, er det positivt for undersøkelsen at det først og fremst er disse som dominerer. I den polske gruppen er det kun to tekster som *ikke* har markørkontekster, mens det er snakk om fire engelske og én norsk.

9.6.8.1 Resultat - Individuelle profiler

Figur 9-38 og tabell 9-24 (under) viser antall og andel av individuelle profiltyper i de tre gruppene. Innen definitiv referanse er det langt færre positive profiler i den polske gruppen enn det var innen indefinit referanse. Det er kun *to* av profilene som er positive (12 %), mens 56 % er negative, og 32 % er kombinerte.

Andelen av positive profiler er også lav i de to andre gruppene (eng. 15 %; nor. 3 %), men disse har en god del flere negative profiler: Den engelske gruppen har 68 % negative, mens den norske gruppen har hele 92 % negative profiler. Forskjellen i antall negative profiler versus de øvrige profilene i den polske og norske gruppen er signifikant ($\chi^2=9.4663$, $p < .01$), mens forskjellen mellom den polske og den engelske *ikke* er signifikant ($\chi^2=0.8358$, $p > .05$). Det er altså signifikante forskjeller mellom den polske og den norske gruppen, mens forskjellen mellom de to innlærergруппene ikke er tilstrekkelig stor. Det lave antallet positive profiler i den polske gruppen vitner uansett ikke om sterkt morsmålssamsvar på individnivå, og homogeniteten innad i den polske gruppen er dermed heller ikke høy.

Figur 9-38 Profiltyper i de tre gruppene (definitiv referanse)



Tabell 9-24 Profiltyper i de tre gruppene (definitiv referanse)

Profiltipe	Engelsk	polsk	Norsk
Positiv	15 % (5)	12 % (3)	3 % (1)
Negativ	68 % (23)	56 % (14)	91 % (30)
Kombinert	18 % (6)	32 % (8)	6 % (2)
Total	34	25	33

Heller ikke en nærmere undersøkelse av sammenfallskontekster og markørkontekster i de polske positive profilene gir særlig støtte til hypotesen. Ved definitiv referanse er UB-kontekster og ART-kontekster markørkontekster, mens det er BE-kontekster som er sammenfallskontekster. For å bedre kunne støtte oppunder at morsmålssamsvaret har opphav i morsmålet, er det derfor styrkende for hypotesen om de positive profilene er i tråd med prediksjonene innen UB- og ART-kontekster. Av de tre polske positive profilene er det bare én som har en positiv markørkontekst (UB), mens de andre to kun har sammenfallskontekster (BE), noe som vil si at kun én av de tre positive profilene, kan vise til morsmålssamsvar som ikke også er målspråkligt. Dette er også det eneste tilfellet av en positiv markørkontekst i hele den polske gruppen. Til sammenligning har den ene norske positive profilen kun en sammenfallskontekst, mens 2 av de 5 engelske positive profilene er positive i en markørkontekst.

Av de 8 kombinerte profilene i den polske gruppen er samtlige positive kontekster sammenfallskontekster, mens de negative kontekstene er markørkontekster. Dette gjelder også i de to andre gruppene, med unntak av én engelsk profil som er positiv i en markørkontekst.

I likhet med for indefinitiv referanse er det altså ikke veldig sterk antydning til morsmålssamsvar på individnivå ved definitiv referanse, og morsmålssamsvaret er enda svakere ved definitiv referanse, siden kun tre profiler er positive, men to av dem uten å kunne utelukke målspråkstilegnelse. Selv om det er en viss heterogenitet mellom gruppene, er denne kun signifikant mellom den norske og den polske gruppen, og morsmålssamsvaret i den polske gruppen er uansett for svakt til å si at formvalgene er i tråd med prediksjonene for hypotese 1.

9.6.9 Oppsummering av funn for hypotese 1 ved definitiv referanse

Tabell 9-25 (neste side) gir en oversikt over hovedfunnene for hypotese 1 ved definitiv referanse. Som ved indefinitiv referanse oppgir tabellen bare tall for de form/kontekst-assosiasjonene som er predikert av hypotesen, og på individnivå er negative og kombinerte profiler slått sammen.

Tabell 9-25 Oversikt over hovedfunn for hypotese 1 ved definit referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
UB-kontekst/ubestemt ent/flt			
- Innen kontekst	18 %	19 %	0 %
- Innen form	47 %	78 %	0 %
BE-kontekst/bestemt ent/flt			
- Innen kontekst	92 %	86 %	100 %
- Innen form	17 %	10 %	5 %
ART-kontekst/artikkelfrase			
- Innen kontekst	0 %	0 %	0 %
- Innen form	-	0 %	-
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	12 %	15 %	3 %
- Negative/kombinerte	88 %	85 %	97 %

Morsmålsamsvar og intralingval kontrast

Når både bruken innen kontekst og innen form ble tatt i betraktning på gruppenivå, var det lite morsmålssamsvar i den polske gruppen. Den delprediksjonen som hadde mest morsmålssamsvar, var assosiasjonen *BE-kontekster/bestemt form*, hvor 92 % av alle kontekstene hadde bestemt form, men hvor likevel bare 17 % av alle bestemt form var brukt i BE-kontekster. For de to andre delprediksjonene var morsmålssamsvaret enda svakere, spesielt for assosiasjonen *ART-kontekst/artikkelfrase*, siden artikkelfrase ikke ble brukt av den polske gruppen i det hele tatt. Tatt i betraktning at BE-konteksten er en sammenfallskontekst, står den utstrakte bruken av bestemt form i denne konteksten svakere som morsmålssamsvar, da det like gjerne kan forklares som målspåkstilleggelse. Vi så også at den polske gruppen var relativt lik den norske gruppen i BE-kontekster med tanke på å først og fremst bruke bestemt form (pol. 92 %; nor. 100 %).

Funnene på gruppenivå var også reflektert i de individuelle profilene, da bare tre av de polske innlærertekstene var positive profiler, og av disse tre hadde to bare sammenfallskontekster. Også de kombinerte profilene var positive bare i sammenfallskontekster. Sammenlignet med den norske gruppen var antall positive profiler ganske jevn, men den polske gruppen hadde signifikant færre *negative* profiler enn den norske.

Vurderingen av morsmålsamsvar og intralingval kontrast i forhold til målspråket og den norske gruppen, viste altså at det ikke er noe sterkt morsmålsamsvar for hypotese 1 verken på gruppenivå eller individnivå. Det lave antallet positive profiler viste også at de predikerte form/kontekst-assosiasjonene ikke er representative for den polske gruppen ved definnitt referanse.

Sammenligningen med den engelske gruppen viste at gruppene var relativt like med hensyn til assosiasjonen mellom UB-kontekster og ubestemt/entall. Begge gruppene hadde en andel på ca. 20 % ubestemt entall/flertall, men brukte ellers bestemt form. Lite bruk av ubestemt entall/flertall i disse kontekstene, var altså ikke noe som var spesifikt for den polske innlærergruppen, og gruppene var også like i å bruke bestemt form mindre enn den norske gruppen, som *kun* brukte bestemt form.

Innen BE-kontekster var det derimot stor likhet mellom den engelske og polske gruppen i form av ustrakt bruk av bestemt form, noe som også stemmer med at dette er en sammenfallskontekst. Innen ART-kontekster var gruppene like i å så å si ikke bruke artikkelfrase i det hele tatt, dvs. ingen artikkelfraser for den polske gruppen og bare én for den engelske.

Gjennomgangen har altså vist at den engelske og polske gruppen er mest like når det kommer til assosiasjoner mellom BE-kontekst og bestemt form, noe som er ventet siden dette er sammenfallskontekst ved definnitt referanse. Vurderingen av transfereffekten intergruppe-heterogenitet hviler derfor på at det finnes forskjeller for de andre delprediksjonene, noe det ikke ser ut til å gjøre.

Disse funnene bekreftes videre av de individuelle profilene, hvor den engelske gruppen ikke så ut til å skille seg merkbart fra den polske. Det var kun en litt større, og ikke signifikant, andel av positive profiler i den engelske gruppen (pol. 12 %; eng. 15 %).

Konklusjon om transfer fra polsk ved definnitt referanse

Konklusjonen for det definnitte materialet er altså at det ikke foreligger argument for transfer av den typen hypotese 1 predikerer, verken på gruppenivå eller på individnivå. Tvert imot er det flere ting som tyder på at formvalgene til den polske gruppen i større grad sammenfaller med målspråket enn med transferprediksjonene. Dette vil jeg si litt om før jeg undersøker hypotese 1 i materialet av generiske substantivfraser.

Målspråkstilegnelse?

I de deskriptive tallene for definitt referanse innledningsvis (0) så vi at den polske gruppen generelt hadde utstrakt bruk av bestemt form, da til sammen 78 % av substantivfrasene hadde bestemt form, mens resten hadde ubestemt entall/flertall. Dette er ikke fullt så mye som den norske gruppen, som hadde 100 % bestemt form, men likevel er det nært opp til målspråksregelen om bestemt form ved definitt referanse.

Denne omfattende bruken av bestemt form gjelder mest i BE- og UB-kontekster, og mindre i ART-kontekster (UB. 82 %; BE. 92 %; ART. 59 %), så forskjellen mellom den norske og den polske gruppen er størst i ART-kontekster. Manglende bestemt form er med andre ord mer utbredt i ART-kontekster (ikke-aktiverte forfelt), og som vi så var den polske gruppen heller ikke i samsvar med prediksjonene her ettersom ingen artikkelfraser ble brukt, så bak dette mønsteret ligger det sannsynligvis andre forhold enn både målspråkstilegnelse eller den type transfer som hypotese 1 handler om. ART-kontekster tilsvarer ikke-aktiverte forfelt, og siden det er snakk om definitt referanse så innebærer dette definte referenter som ikke er identifiserbare gjennom teksten, men som da må være *situasjonelt* identifiserbare. Slike situasjonelt aktiverte substantivfraser blir av Ekiert (2010: 140) framhevet nettopp som et område hvor bruken av bestemt artikkel er både variabel og faller senere på plass hos polske innlærere av engelsk, og observasjonene som er beskrevet over, kan tyde på noe lignende.

Setningsposisjon og aktiveringsgrad hver for seg

Tidligere i undersøkelsen har jeg også sett på hvordan valg av form står i forhold til setningsposisjon og aktiveringsgrad uavhengig av hverandre. For indefinitt referanse så vi at både den polske og den engelske gruppen i stor grad brukte alle formene først og fremst i slutfeltet, mens den norske gruppen først og fremst brukte dem i ikke-aktiverte kontekster. Jeg knyttet dette mønsteret til den generelt høye andelen av slutfelt i det polske og engelske materialet, og ikke-aktiverte substantiv i det norske, og ikke til at dette representerte noen spesifikk form/kontekst-assosiasjon. I det definte materialet er bildet noe annerledes (se tabellene side 288). For det første er det så å si ingen artikkelfraser i noen av gruppene, så disse er ikke med i vurderingen under.

I den polske gruppen er både ubestemt entall/flertall og bestemt form brukt mest i *forfeltet* (ub. 84 %; be. 76 %), altså motsatt av ved indefinitt referanse, hvor de stort

sett var brukt i slutfeltet. Også her er nok tallet et resultat av at det er nettopp forfelt det er mest av ved definitte referanse: 80 % av de definitive substantivfrasene står i forfeltet.

I den engelske gruppen er det veldig få fraser med ubestemt entall/flertall (9), og disse har ikke noen distribusjon som skiller seg nevneverdig ut. For bestemt form er derimot den engelske gruppen lik den polske i å bruke formen oftest i forfeltet (77 %), men det er også her forfelt som utgjør størstedelen av materialet, da 77 % av de definitive substantivfrasene er i forfeltet.

Den norske gruppen har få fraser med ubestemt entall/flertall (4), og alle disse er brukt i ikke-aktiverte forfelt. Fraser med bestemt form er derimot først og fremst brukt ved ikke-aktiverte referenter, uavhengig av setningsposisjon (82 %), noe som ser ut til å avspeile at størstedelen av det norske definitive materialet består av ikke-aktiverte forekomster (83 %).

Som ved indefinitt referanse ser altså bruksmønstrene ved definitte referanse også ut til å være en avspeiling av hvilke kontekster som er representert i tekstene, heller enn spesifikke form/kontekst-assosiasjoner: I de to innlærergруппene er definitive substantivfraser først og fremst *forfelt*, mens de i den norske gruppen først og fremst er *ikke-aktiverte*.

For den norske gruppen ser det ut til å være generell overvekt av ikke-aktiverte substantivfraser i hele materialet for hypotese 1, uavhengig av om det er definitte eller indefinitt referanse. For de to innlærergруппene ser det derimot ut til å fordele seg forskjellig innen de to referensielle typene, med indefinitt substantiv i slutfeltet og definitive i forfeltet.

Jeg vil nå gå over til å se på generisk referanse.

9.7 Generisk referanse

Den siste av de referensielle typene som skal undersøkes, er generisk referanse. Generisk referanse står i kontrast til indefinitt referanse og definitte referanse ved at det ikke refereres til et eller flere individuelle referenter, men heller til en hel klasse av referenter av samme type (se 3.2.1/8.4.4). Prediksjonene om de polske innlærernes formvalg i UB-, BE- og ART-kontekster ved generisk referanse er de samme som ved indefinitt og definitte referanse. Under er noen eksempler på de ulike konteksttypene for hypotese 1 med generiske substantivfraser.

Generisk UB-kontekst

De understrekede substantivfrasene i begge eksemplene under (9.7-a og 9.7-b) har generisk referanse. Både «mennesker» og «kvinner» viser generelt til mennesker og kvinner som klasse, ikke til individuelle mennesker og kvinner.

9.7-a Jeg har truffet en gutt på jobben min. Han har hatet verden, han har hatet mennesker, han ville ikke snakke med noen. (s0794)

9.7-b Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at kvinner fikk stemmerett. (s0654)

Frasen «mennesker» i 9.7-a er plassert i slutfeltet til en hovedsetning, og siden mennesker ikke har vært nevnt tidligere, regnes referansen som ikke-aktivert. Frasen «kvinner» i 9.7-b er derimot plassert i forfeltet til en leddsetning, og referansen er her aktivert gjennom å være nevnt i foregående setning. Disse to frasene illustrerer dermed hver sin type UB-kontekst, og som vi ser er det ubestemt flertallsform som er brukt i begge tilfellene, derfor stemmer formvalgene med prediksjonen for denne konteksttypen.

Generisk BE-kontekst

Det neste eksemplet (9-7-c) viser en generisk substantivfrase i en BE-kontekst (aktivert slutfelt). Den understrekede substantivfrasen «barna» refererer generelt til barn som klasse og er en generell påstand om alle barn, ikke om ett eller flere individuelle barn.

9.7-c Det er en stor forskjell mellom barna fra gamle dager og våre dager. Forskjellen ligger i oppførelsen av barna. Det er ikke lett å oppdra barna, det er våre vanskeligste oppgave i livet vårt. (s0078)

Når bestemt form blir brukt her, kan det skape inntrykk av at referansen er definnitt og at det er snakk om noen spesifikke barn, noe som kan forsterkes av at det også i setningene forut har vært to tilfeller av *ikke-generisk referanse* til barn, altså referanse til gruppene av individuelle barn før i tiden og nå. Den understrekede forekomsten av «barna» oppfatter jeg derimot som et utsagn om barn generelt, om alle barn, hvor en naturlig rekonstruksjon til norsk ville vært «Det er ikke lett å oppdra barn.». Denne regnes derfor som generisk, og den er videre klassifisert som et aktivert slutfelt, da referansen er aktivert gjennom tidligere nevning av barn.

Siden dette er en BE-kontekst, stemmer det med prediksjonen å bruke bestemt form, men vi ser at bestemt form vil kunne være misvisende som formvalg på norsk. Dette illustrerer godt en markørkontekst, der det predikerte formvalget for den polske gruppen ikke sammenfaller med det typiske formvalget i norsk.

ART-kontekst

Det siste eksemplet (9.7-d) viser en generisk substantivfrase i en ART-kontekst (ikke-aktivert forfelt). Frasen «et museum» viser her til museum generelt. Det er en påstand om at museum kan være inspirerende, og denne påstanden gjelder for alle museum.

9.7-d Mange er interessert i ekstra utdanning etter de er ferdig med videregående skole. For eksempel kan de lære kunsthistorie eller de kan lære å tegne og male. **Et museum** blir et godt sted for å skape inspirasjon. (s0297)

Frasen «et museum» er videre plassert i forfeltet på en hovedsetning, og referansen er ikke aktivert siden det ikke har vært nevnt noe om museum tidligere. Dette er dermed en ART-kontekst, og siden det er brukt artikkelfrase, stemmer formvalget med prediksjonene for denne konteksttypen.

Det kan likevel diskuteres hvorvidt frasen i eksemplet over representerer *generisk* eller *ikke-spesifikk indefinit* referanse, altså om det vises til museum generelt eller bare til et hvilket som helst individuelt museum. Forskjellen mellom generisk og indefinit referanse kommer tydeligere fram i setninger som «Et museum er en samling av historiske gjenstander» (generisk) og «Jeg ønsker å gå på et museum i dag» (indefinit). I setningen som definerer museum, må referansen nødvendigvis være generisk, mens det i setningen som uttrykker et ønske om å gå på museum, heller vises til et individuelt museum selv om det ikke er bestemt hvilket. Setning 9.7-d oppfatter jeg som mer lik den definerende setningen, derfor oppfatter jeg «et museum» som generisk heller enn indefinit.

Jeg vil komme tilbake til flere eksempler senere. Gjennomgangen av transfereffekter i det generiske materialet vil være på samme måte som tidligere i kapittelet, altså først form/kontekst-assosiasjoner på gruppenivå og deretter individuelle

profiler. Før dette vil jeg se på deskriptive tall for materialet og hva som regnes som sammenfalls- og markørkontekst ved generisk referanse.

9.7.1 Deskriptivt for alle gruppene

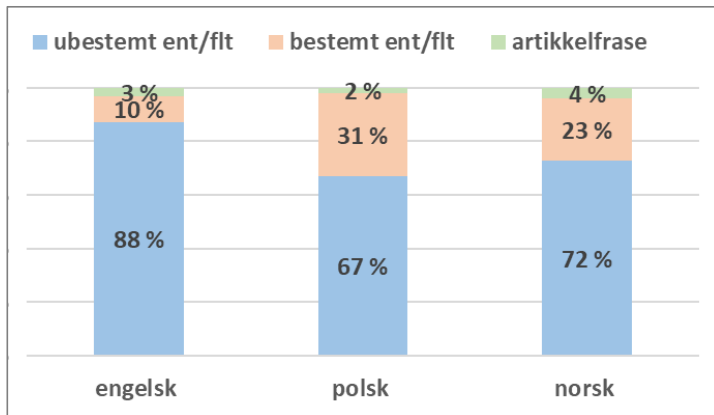
Det generiske materialet er det minst omfattende av de tre referensielle delmaterialene, med kun 52 enkle substantivfraser i de polske tekstene, 53 i de norske, og 40 i de engelske. Den engelske gruppen har altså færrest slike fraser, noe som mest sannsynlig avspeiler at det totalt sett er færre engelske innlærertekster som har substantivfraser med generisk referanse, bare 21 tekster, i motsetning til henholdsvis 25 og 26 i den polske og norske gruppen.

Figur 9-39 og tabell 9-26 (øverst neste side) viser andelen av de ulike formtypene ved generisk referanse. Alle former er tatt i bruk i alle tre gruppene, men det er snakk om veldig få artikkelfraser, og det er i fordelingen av ubestemt entall/flertall og bestemt form vi finner hovedforskjellene mellom gruppene: Den polske gruppen har en del mer bestemt form enn de andre to gruppene, spesielt mer enn den engelske (pol. 31 %; eng. 10 %, nor. 23 %). Tilsvarende har den engelske gruppen mer ubestemt entall/flertall (pol. 67 %; eng. 88 %; nor. 72 %). Når det gjelder artikkelfraser har alle tre gruppene veldig få slike, det er bare snakk om én artikkelfrase for den polske og engelske gruppen, og *tre* for den norske.

Figur 9-40 og tabell 9-27 (nederst neste side) viser andelen av konteksttyper i de ulike gruppene. Alle tre gruppene har størst andel UB-kontekster (pol. 69 %; eng. 55 %; nor. 40 %), men andelen er aller størst i den polske gruppen, som dessuten har en merkbart lavere andel av ART-kontekster enn den norske og engelske gruppen (pol. 17 %; eng. 40 %; nor. 53 %). Alle tre gruppene har relativt få BE-kontekster, men det er noe mer i den polske gruppen (pol. 13 %; eng. 5 %; nor. 8 %).

Det generiske materialet har altså få artikkelfraser i alle de tre gruppene (pol. 1; eng. 1; nor. 3), og det er også relativt få BE-kontekster (pol. 7; eng. 2; nor. 4). I tillegg

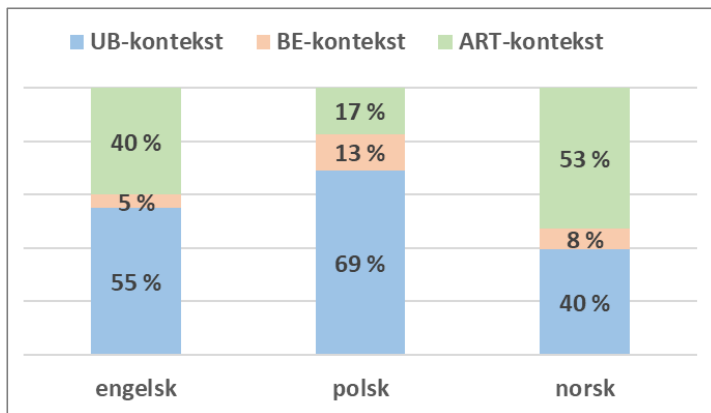
Figur 9-39 Andel av de ulike formtypene i de tre gruppene (generisk referanse)



Tabell 9-26 Andel av de ulike formtypene i de tre gruppene (generisk referanse)

Morsmål	Ubestemt ent/flt	Bestemt ent/flt	Artikkelfrase	Totalt
Polsk	67 % (35)	31 % (16)	2 % (1)	52
Engelsk	88 % (35)	10 % (4)	3 % (1)	40
Norsk	72 % (38)	23 % (12)	4 % (3)	53

Figur 9-40 Andel av de ulike kontekststypene i de tre gruppene (generisk referanse)



Tabell 9-27 Andel av de ulike kontekststypene i de tre gruppene (generisk referanse)

Morsmål	UB-kontekst	BE-kontekst	ART-kontekst	Totalt
Polsk	69 % (36)	13 % (7)	17 % (9)	52
Engelsk	55 % (22)	5 % (2)	40 % (16)	40
Norsk	40 % (21)	8 % (4)	53 % (28)	53

er det veldig få fraser med bestemt form i det engelske materialet (4). Dette gjør at det er et noe tynt grunnlag for undersøkelsen av form/kontekst-assosiasjonene for BE- og ART-kontekster i det generiske materialet.

9.7.2 Sammenfalls- og markørkontekster ved generisk referanse

I 9.2.3 beskrev jeg områder for homogenitet og heterogenitet mellom den polske og de andre to gruppene. Det ble da nevnt at mens reglene om ubestemt form ved indefinitt referanse og bestemt form ved definitt referanse er relativt fast i både norsk og engelsk, er distribusjonen av former mer varierende ved generisk referanse, i tillegg til at det er større forskjeller mellom språkene norsk og engelsk. Fordelingen av former ved generisk referanse er til en viss grad styrt av generiske undertyper (se 3.2.1.1, 7.2.4.2 og 7.4.2.2), og materialet mitt er ikke kodet for disse undertypene, noe som ikke gjør det mulig å dele opp denne delen av materialet på et så finmalt nivå som kunne ha vært hensiktsmessig. Dette gjør det også vanskeligere å beskrive sammenfalls- og markørkontekster nøyaktig. Jeg vil på grunn av dette ta utgangspunkt i de deskriptive tallene for generisk referanse som nettopp ble presentert i 9.7.1.

Figur 9-39 (øverst forrige side) viser at både den engelske og den norske gruppen bruker alle tre formtypene ved generisk referanse, og at det er lite bruk av artikkelfrase. Hovedforskjellen er at den norske gruppen bruker mer bestemt form enn den engelske (nor. 23 %; eng. 10 %), og tilsvarende mindre ubestemt form (nor. 72 %; eng. 88 %). Om man antar at dette er en fordeling som vil være til stede jevnt både i UB-, BE- og ART-kontekster hos den engelske og den norske gruppen, kan man forvente at begge gruppene først og fremst vil ha ubestemt entall/flertall i alle disse tre kontekstene, men at den norske gruppen vil kunne ha noe mer bestemt form enn den engelske. Det vil dessuten være få artikkelfraser.

Siden den polske gruppen er predikert til å velge form etter om det er UB-, BE- eller ART-kontekst, og den engelske forventes å stort sett bruke ubestemt entall/flertall i alle disse konteksttypene, vil heterogeniteten mellom den polske og den engelske gruppen først og fremst komme til syne i BE-kontekster og ART-kontekster, hvor den polske gruppen ventes å bruke henholdsvis bestemt form og artikkelfrase, mens den engelske ventes å stort sett bruke ubestemt entall/flertall. I UB-kontekster ventes begge gruppene å bruke ubestemt entall/flertall, noe som fører til sammenfall.

Mellom den polske og den norske gruppen vil stort sett det samme gjelde, men det at man i norsk bruker mer bestemt form enn i engelsk, gjør at heterogeniteten ikke ventes å være like stor mellom den norske og den polske gruppen i BE-kontekster, og at sammenfallet ikke nødvendigvis er så stort i UB-kontekster. I ART-kontekster er de derimot ventet å være ulike siden den polske gruppen ventes å bruke artikkelfrase.

Summen av dette er at det først og fremst er UB-kontekster som er sammenfallskontekster ved generisk referanse, mens BE- og ART-kontekster er markørkontekster både i forhold til den norske og engelske gruppen. Tabell 9-28 (under) viser den ventede fordelingen av former i de tre gruppene, med markørkontekster uthevet. I motsetning til ved indefinit og definit referanse er ikke formene som er oppført for den engelske og norske gruppen ventet å være konsekvent brukt, bare oftere.

Tabell 9-28 Sammenfalls- og markørkontekster for generisk referanse

Generisk referanse	Polsk	Engelsk	Norsk
UB-kontekst	Ubestemt ent/flt	Ubestemte former	Ubestemte former
BE-kontekst	Bestemte former	Ubestemte former	Ubestemte former
ART-kontekst	Artikkelfrase	Ubestemte former	Ubestemte former

Ifølge dette vil intergruppe-heterogenitet mellom den engelske og den polske gruppen først og fremst komme til syne i BE- og ART-kontekster, og det er også disse to konteksttypene som best avdekker morsmålsansvar som ikke er målspråkligt. Forskjellen er ventet å være størst ved den engelske gruppen. Forskjellen er ventet å være størst i forhold til den engelske gruppen i BE-kontekster.

Som vi så i den deskriptive gjennomgangen av det generiske materialet, er det nettopp artikkelfraser og BE-kontekster det er færrest av i materialet, noe som gjør at det er tynneste grunnlag for undersøkelsen nettopp i markørkontekstene for generisk referanse, noe som ikke er positivt for undersøkelsen.

Det er imidlertid kun mulig å bruke denne inndelingen til undersøkelsene på gruppenivå, og ikke til å undersøke om individuelle profiler sammenfaller eller ikke sammenfaller med målspråket. Dette er fordi alle formene potensielt sett er mulig å bruke i norsk, derfor vil alle formvalg som gjøres i den polske gruppen, uavhengig av hvilken kontekst det er, også sammenfalle med målspråkets mulige formvalg. Dette skyldes at jeg ikke har delt inn materialet i finere undertyper som kunne sagt noe om hvilken form som brukes i norsk i en spesifikk undertype.

Tabell 9-29 Polsk gruppe: Fordeling av former etter konteksttype (generisk)

S1=POLSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	16	8	3	8	35
Andel innen form	46 %	23 %	9 %	23 %	100 %
Andel innen kontekst	59 %	89 %	43 %	89 %	67 %
BESTEMT ENT/FLT	10	1	4	1	16
Andel innen form	63 %	6 %	25 %	6 %	100 %
Andel innen kontekst	37 %	11 %	57 %	11 %	31 %
ARTIKKELFRASE	1	0	0	0	1
Andel innen form	100 %	0 %	0 %	0 %	100 %
Andel innen kontekst	4 %	0 %	0 %	0 %	2 %
TOTALT	27 (52 %)	9 (17 %)	7 (13 %)	9 (17 %)	52

Tabell 9-30 Engelsk gruppe: Fordelinger av former etter konteksttype (generisk)

S1=ENGELSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	19	1	2	13	35
Andel innen form	54 %	3 %	6 %	37 %	100 %
Andel innen kontekst	90 %	100 %	100 %	81 %	88 %
BESTEMT ENT/FLT	2	0	0	2	4
Andel innen form	50 %			50 %	100 %
Andel innen kontekst	10 %			13 %	10 %
ARTIKKELFRASE	0	0	0	1	1
Andel innen form				100 %	100 %
Andel innen kontekst				6 %	3 %
TOTALT	21 (53 %)	1 (3 %)	2 (5 %)	16 (40 %)	40

Tabell 9-31 Norsk gruppe: Fordelinger av former etter konteksttype (generisk)

S1= NORSK	UB-kontekster		BE-kont.	ART-kont.	Totalt
Form: \ Kontekst:	Aktivert forfelt	Ikke-aktivert slutfelt	Aktivert slutfelt	Ikke-aktivert forfelt	
UBESTEMT ENT/FLT	10	7	1	20	38
Andel innen form	26 %	18 %	3 %	53 %	100 %
Andel innen kontekst	100 %	64 %	25 %	71 %	72 %
BESTEMT ENT/FLT	0	3	3	6	12
Andel innen form	0 %	25 %	25 %	50 %	100 %
Andel innen kontekst	0 %	27 %	75 %	21 %	23 %
ARTIKKELFRASE	0	1	0	2	3
Andel innen form	0 %	33 %	0 %	67 %	100 %
Andel innen kontekst	0 %	9 %	0 %	7 %	4 %
TOTALT	10 (19 %)	11 (21 %)	4 (8 %)	28 (53 %)	53

9.7.3 Samlet oversikt for generisk referanse

Tabellene på forrige side viser forholdet mellom form og konteksttyper ved generisk referanse i de tre gruppene, og vil bli fortløpende visualisert med søylediagram.

9.7.4 Assosiasjon mellom UB-kontekster og ubestemte former

I UB-kontekster (aktiverte forfelt og ikke-aktiverte sluttfelt) er det forventet at den polske gruppen bruker ubestemt entall/flertall, og at denne typen form først og fremst brukes i denne konteksttypen. Eksemplene under viser eksempler på UB-kontekster med ulike formvalg, først to eksempler der formvalget er i tråd med prediksjonen for UB-kontekster, deretter to eksempler på formvalg som *ikke* er i tråd med prediksjonene.

Eksempel på generiske substantivfraser i UB-kontekster som har predikert formvalg ble omtalt allerede innledningsvis i 9.7, men er gjentatt under (9.7-e og 9.7-f). Begge disse har ubestemt flertallsform, som stemmer med prediksjonen for denne konteksttypen.

9.7-e Jeg har truffet en gutt på jobben min. Han har hatet verden, han har hatet **mennesker**, han ville ikke snakke med noen. (s0794)

9.7-f Det som forandret seg meste er kvinnen possision i samfunnet. Det er ikke så veldig lenge siden at **kvinner** fikk stemmerett. (s0654)

De to eksemplene under (9.7-g og 9.7-h) viser UB-kontekster der det er brukt henholdsvis artikkelfrase og bestemt form, og som derfor bryter med prediksjonen. I det første eksemplet ser vi «en venn» som har generisk referanse fordi det er en beskrivelse av hva en venn er. Dette er et eksempel på den undertypen av generisk referanse som kalles *taksonomisk*, hvor det refereres til egenskaper som er iboende eller definerende for en klasse (se 3.2.1.1). Som vi ser er frasen «en venn» plassert i forfeltet til en leddsetning, og referansen er aktivert gjennom tidligere nevning av venner, altså er det et eksempel på en UB-kontekst av typen *aktivert forfelt*.

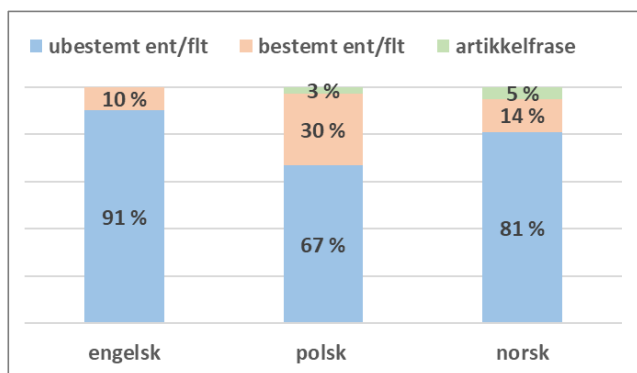
9.7-g Noen kaller «venner» alle personer, som de kjenner. Men det er ikke riktig, fordi **en venn** er det en person, som man kan stole på, kan regne med, kan få hjelp og forståelse. (s0593)

9.7-h Det var bare meg fordi jeg gav ham hjertet mitt. Han elsket **fjellet**, det var spennende å snakke med han om det. (s0794)

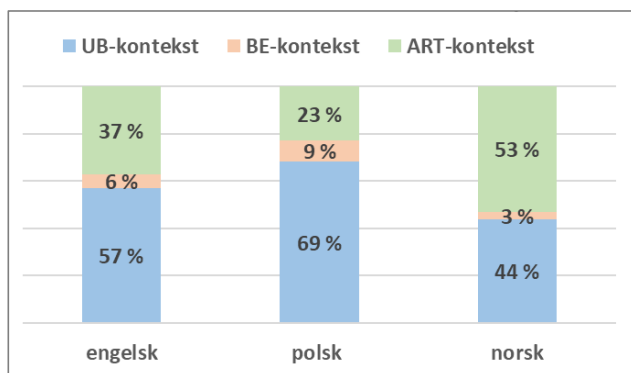
I det andre eksemplet over ser vi frasen «fjellet», som er generisk fordi det her vises til fjell generelt, ikke et eller flere individuelle fjell. Frasen er plassert i sluttfeltet til en hovedsetning, referenten er ikke nevnt før, og dermed er denne forekomsten klassifisert som ikke-aktivert. Eksemplet viser altså en UB-kontekst av typen *ikke-aktivert sluttfelt*.

Søylediagrammene under illustrerer forholdet mellom UB-kontekster og ubestemt entall/flertall for de tre gruppene.

Figur 9-41 Andel av form innen UB-kontekst



Figur 9-42 Andel av kontekst innen ubest. ent/flt



Figur 9-41 (øverst) viser hvilke former som er brukt i UB-kontekster, mens figur 9-42 (nederst) viser hvilke kontekster ubestemt entall/flertall er brukt i for de tre gruppene.

For den polske gruppen er assosiasjonen mellom UB-kontekster og ubestemt entall/flertall relativt sterk: 67 % av UB-kontekstene har den predikerte formen, og 69 %

av alle substantivfraser med ubestemt entall/flertall brukes i nettopp UB-kontekster. Dette er imidlertid en sammenfallskontekst, og som forventet fins det tilsvarende utstrakt bruk av ubestemt entall/flertall både i det norske og det engelske materialet, men den engelske gruppen har klart mest bruk av disse formene i UB-kontekster (innen kontekst/innen form eng. 91 % / 57 %; nor. 81 % / 44 %).

Det fins altså en relativt sterk assosiasjon mellom UB-kontekst og ubestemt entall/flertall for alle tre gruppene, og den polske gruppen er dermed relativt homogen med de andre gruppene, men først og fremst innen kontekst. I den grad de er ulike, handler det om at den engelske og norske gruppen bruker *mer* ubestemt entall/flertall, og at den polske bruker mer bestemt form. Tendensen til morsmålsvarsvar for denne delprediksjonen utfylles dermed ikke med intergruppe-heterogenitet eller forskjeller fra målspråksgruppen.

9.7.5 Assosiasjon mellom BE-kontekster og bestemt form

I BE-kontekster (aktiverte sluttfelt) er det forventet at den polske gruppen bruker bestemt form, og det er også forventet at bestemt form først og fremst brukes i BE-kontekster, ikke i andre konteksttyper. Eksemplene under viser eksempler på BE-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen for BE-kontekster, deretter et eksempel på formvalg som *ikke* er i tråd med prediksjonene.

Eksempel på en generisk BE-kontekst med predikert formvalg ble omtalt allerede innledningsvis i 9.7, men er gjentatt under (9.7-i). Den understrekede frasen «barna» har her bestemt form, som stemmer med prediksjonen for denne konteksttypen.

9.7-i Det er en stor forskjell mellom barna fra gamle dager og våre dager. Forskjellen ligger i oppførelsen av barna. Det er ikke lett å oppdra **barna**, det er våre vanskeligste oppgave i livet vårt. (s0078)

Det finnes ikke eksempler på bruk av artikkelfrase i BE-kontekster i det generiske materialet, men eksempel 9.7-j (under) viser en BE-kontekst med ubestemt form uten artikkel. Substantivfrasen «barn» er her generisk ettersom den viser til barn generelt, ikke til individuelle barn.

9.7-j Foreldre kan aldri si at vi har ikke tid til dem [barna] eller svare på. Å oppdra **barn** betyr å lære dem. (s0725, min innsetting)

«Barn» er her plassert i slutfeltet på en infinitivkonstruksjon, og er klassifisert som aktivert ettersom det har vært referanse til barn i setningen forut, da gjennom pronomenet «dem» som har «barn» som antesedent i en tidligere setning.

Søylediagrammene på neste side illustrerer forholdet mellom BE-kontekster og bestemt form for de tre gruppene. Jeg vil først se på den polske gruppen.

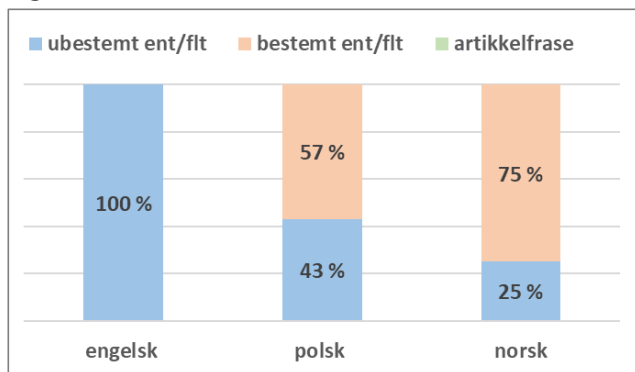
Figur 9-43 (øverst neste side) viser hvilke former som er brukt i BE-kontekster, og i de polske tekstene er det en svak overvekt av bestemt form (57 %), som prediksjonen tilsier, men det er også en god del ubestemt entall/flertall (43 %). Figur 9-44 (nederst neste side) viser hvilke kontekster bestemt form er brukt i, og frasene med bestemt form i det polske materialet er først og fremst brukt i UB-kontekster, mens kun 25 % er brukt i BE-kontekster. Det er derfor svakere morsmålssamsvar for prediksjonen om BE-kontekster enn vi så for prediksjonen om UB-kontekster.

Om vi sammenligner den polske og den norske gruppen i de samme søylediagrammene, ser vi i figur 9-43 (øverst) at den norske gruppen er relativt lik den polske i å først og fremst bruke bestemt form i BE-kontekster. Tendensen er imidlertid sterkere i den norske gruppen, hvor hele 75 % av alle BE-kontekster har bestemt form (pol. 57 %). I figur 9-44 (nederst) ser vi at gruppene derimot er helt like i at 25 % av alle fraser med bestemt form brukes i BE-kontekster. Forskjellene mellom den polske og den norske gruppen er altså først og fremst knyttet til at den norske gruppen har formvalg som er mer i tråd med prediksjonene om polsk morsmålssamsvar enn det den polske gruppen selv har.

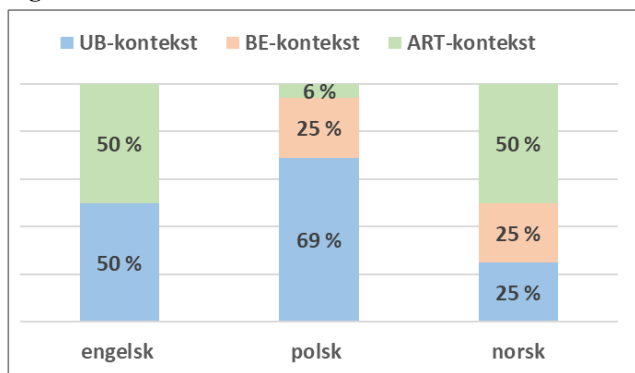
Når man sammenligner den polske og den engelske gruppen i søylediagrammene på neste side, viser det seg at det er forskjeller mellom de to innlærergruppene. Den engelske gruppen bruker ikke bestemt form i denne konteksten i det hele tatt, og de to eneste BE-kontekstene i de engelske tekstene har ubestemt entall/flertall.

Siden BE-konteksten er markørkontekst, er det ventet at den polske gruppen er ulik begge de andre gruppene, men forskjeller finnes altså først og fremst i forhold til den engelske gruppen. Med tanke på at bestemt form er mer utbredt i norsk enn i engelsk ved generisk referanse, er ikke dette så overraskende.

Figur 9-43 Andel av form innen BE-kontekst



Figur 9-44 Andel av kontekst innen bestemt form



Det ser forøvrig ut som bestemt form er brukt langt oftere i BE-kontekster enn i de andre konteksttypene i det norske materialet (UB. 14 %; BE. 75 %; ART. 21 %). Dette viser at antakelsen om at fordelingen av ubestemt og bestemt form skulle være lik innen de ulike konteksttypene, ikke stemmer. At den norske gruppen viser seg å spesielt bruke bestemt form i BE-kontekster (aktiverte slutfelt), kan kanskje ha noe å gjøre med at slutfelt typisk er objekter, og at det er mer akseptert med bestemt form i norsk når det er generisk referanse i objektsposisjon (Hagen 1998: 61). Den polske gruppen har noe av den samme tendensen til å oftere bruke bestemt form i BE-kontekster (UB. 30 %; BE. 57 %; ART. 11 %). Dette viser at selv om det er en god del bruk av bestemt form i BE-kontekster i den polske gruppen, som er i tråd med prediksjonene, så er denne ujevne fordelingen av fraser med bestemt form også til en viss grad målspråkslik. Det er imidlertid et forbehold her at det er ganske få BE-kontekster i alle de tre gruppene

(pol. 7; eng. 2; nor. 4), så sammenligningsgrunnlaget er ganske tynt for denne konteksttypen.

Helhetsvurderingen for assosiasjonen *BE-kontekst/bestemt form* er at den er sterkere i den polske gruppen enn i den engelske gruppen, men noe svakere i den polske gruppen enn i den norske. Det er få BE-kontekster, men det kan se ut til at det er den engelske gruppen som har den mest gruppespesifikke tendensen på dette området, i form av mindre bruk av bestemt form enn de andre gruppene, og dette kan kanskje knyttes til at man i engelsk sjeldnere bruker bestemt form enn i norsk ved generisk referanse, altså kan det være snakk om transfer.

9.7.6 Assosiasjon mellom ART-kontekster og artikkelfrase

I ART-kontekster (ikke-aktiverte forfelt) er det forventet at den polske gruppen bruker artikkelfrase, og det er også forventet at artikkelfrase først og fremst brukes i ART-kontekster, ikke i andre konteksttyper.

Eksempelene under viser eksempler på ART-kontekster med ulike formvalg, først et eksempel der formvalget er i tråd med prediksjonen, deretter to eksempler der formvalget *ikke* er det.

Eksempel på en generisk substantivfrase i en ART-kontekst, som har det predikerte formvalget, ble omtalt allerede innledningsvis i 9.7, men er gjentatt under (9.7-k). Den understrekede frasen «et museum» er en artikkelfrase, og formvalget stemmer dermed med prediksjonen for denne konteksttypen.

9.7-k Mange er interessert i ekstra utdanning etter de er ferdig med videregående skole. For eksempel kan de lære kunsthistorie eller de kan lære å tegne og male. **Et museum** blir et godt sted for å skape inspirasjon. (s0297)

De to neste eksemplene (9.7-l og 9.7-m) viser bruk av bestemt form og ubestemt form uten artikkel i ART-kontekster. Substantivfrasen «forskerne» i det første eksemplet viser generelt til forskere og er derfor generisk, og vi ser at denne frasen er plassert i forfeltet til en hovedsetning og at forskere ikke er nevnt tidligere.

9.7-l Jeg må skrive om at mobiltelefoner kunne bli farlig for helsa. **Forskerne** tror at mobiltelefoner kan bli årsaken til den alvorlige sykdommen kreft. (s0770)

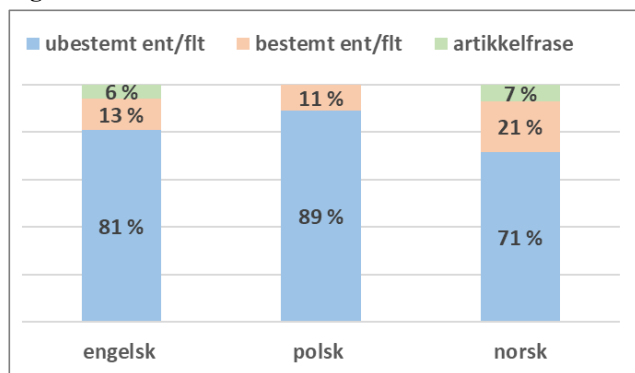
9.7-m **Mobiltelefoner** ble veldig populært i de siste årene.

I det siste eksemplet har «mobiltelefoner» generisk referanse fordi det viser generelt til mobiltelefoner, og frasen er plassert i forfeltet på en hovedsetning. Dette er den første setningen i teksten den er hentet fra, så referansen er ikke aktivert. De understrekede substantivfrasene i de to setningene over, har ikke artikkelfrase, så de innfrir ikke prediksjonen for ART-kontekster.

Søylediagrammene under illustrerer forholdet mellom ART-kontekster og artikkelfrase for de tre gruppene. Det er veldig få artikkelfraser i det generiske materialet (pol. 1; eng. 1, nor. 3), noe som påvirker muligheten for å undersøke assosiasjonen mellom artikkelfrase og ART-kontekster. Det fins imidlertid en god del ART-kontekster, så det lar seg gjøre å undersøke denne assosiasjonen innen kontekst, altså se på hvilke former som er brukt i ART-kontekster.

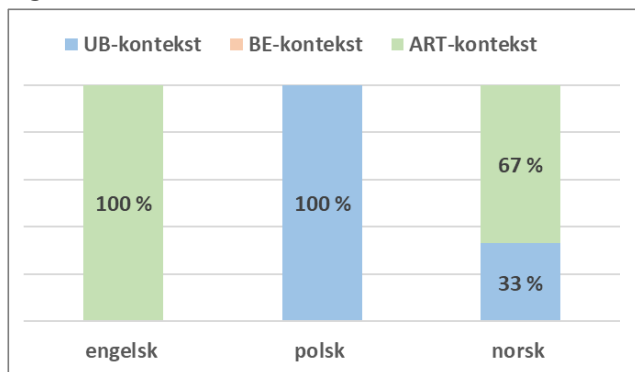
Figur 9-45 (under) viser hvilke former som er brukt i ART-kontekster, og det er ingen artikkelfraser i ART-kontekster i det polske materialet. Det er derimot brukt en stor andel ubestemt entall/flertall (89 %). Den norske og den engelske gruppen har en minimal andel artikkelfraser i ART-kontekster (eng. 6 %; nor. 7 %), men det er altså snakk om kun 1 og 3 stykker.

Figur 9-45 Andel av form innen ART-kontekst



Figur 9-46 (neste side) viser hvilke kontekster artikkelfrase er brukt i. Disse tallene er av mindre verdi siden det er så få artikkelfraser, men vi ser at den ene generiske artikkelfrasen i det polske materialet (tilsvarende 100 %) er brukt i en UB-kontekst.

Figur 9-46 Andel av kontekst innen artikkelfraser



For assosiasjonen *ART-kontekst/artikkelfrase* er det altså *ikke* morsmålssamsvar i den polske gruppen, og vi ser at både den engelske og den norske gruppen har en noe sterkere assosiasjon mellom disse to enn det polske materialet har, med forbehold knyttet til at materialet er tynt.

9.7.7 Oppsummering av funn på gruppenivå

Materialet på gruppenivå bærer noe preg av å ha få artikkelfraser og få BE-kontekster, noe som gjør at det er svakt grunnlag for å undersøke markørkontekstene for generiske referanse (ART og BE). Det som kommer fram av gjennomgangen er uansett at det er få tegn til transfer innen generisk referanse.

For assosiasjonen *UB-kontekst/ubestemt entall/flertall* (sammenfallskontekst), så vi høyt morsmålssamsvar, men lite forskjell fra de andre to gruppene, og at den engelske gruppen var noe mer i tråd med prediksjonene for polsk morsmålssamsvar enn den polske gruppen selv var. For assosiasjonen *BE-kontekst/bestemt form* (markørkontekst) var det et noe ujevnt morsmålssamsvar, som viste seg å være ulikt den engelske gruppen, men veldig likt den norske gruppen, og den norske gruppen var dessuten mer i tråd med prediksjonene for polsk morsmålssamsvar enn den polske gruppen selv var. For assosiasjonen *ART-kontekst/artikkelfrase* (markørkontekst), ble det ikke funnet noe morsmålssamsvar og det var stor likhet mellom gruppene.

I likhet med både indefinit og definit referanse, er det altså ved generisk referanse sterkest tendens til morsmålssamsvar i sammenfallskontekster, mens morsmålssamsvaret er lavere eller fraværende i markørkontekster, i dette tilfellet

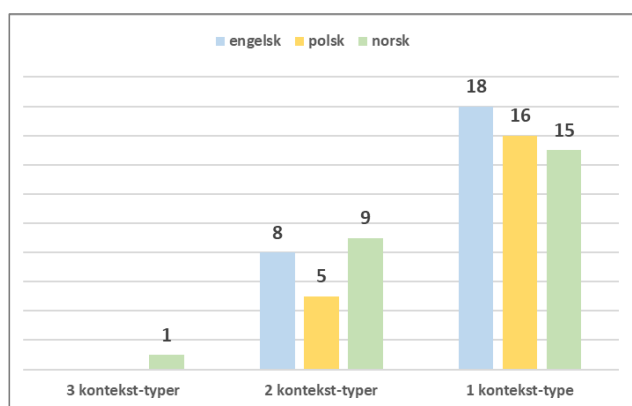
BE-kontekster. Det er derfor lite på gruppenivå som taler for transfer hos den polske gruppen. Det ble imidlertid observert noen forskjeller mellom gruppene i form av større andel ubestemt entall/flertall i UB- og BE-kontekster hos engelske innlærere. Dette kan potensielt knyttes til transfer i den engelske gruppen, ettersom dette både ga utslag som intergruppe-heterogenitet, og kan kobles til at ubestemt form er det mest vanlige i engelsk ved generisk referanse.

9.7.8 Individuelle profiler

Det er 26 polske innlærertekster som har enkle substantivfraser med generisk referanse, mens det er snakk om 25 i den norske og 21 i den engelske gruppen. Figur 9-47 (under) viser antall av individuelle profiler som har tre, to eller bare én av konteksttypene.

Det er bare én tekst som både har UB-, BE- og ART-kontekster, og denne er fra den norske gruppen. De aller fleste tekstene har kun én av konteksttypene (pol. 16; eng. 18; nor. 15), og det er da UB- og ART-kontekster som dominerer. Bare én tekst har kun BE-kontekster, og denne er polsk. Hos de som har to konteksttyper, er det også UB-kontekster, i kombinasjon med enten BE- eller ART-kontekster, som er mest vanlig i alle tre gruppene. Dette er i tråd med det vi så på gruppenivå, nemlig at UB-konteksten dominerer i det generiske materialet, men de aller fleste tekstene har likevel minst én forekomst av enten BE- eller ART-kontekster (pol; 18; eng. 14; nor. 21).

Figur 9-47 Antall tekster med tre, to eller én konteksttype (generisk referanse)



Jeg var tidligere inne på at det ikke lar seg gjøre å undersøke sammenfallskontekster og markørkontekster på individnivå for generisk referanse, siden alle former teoretisk sett

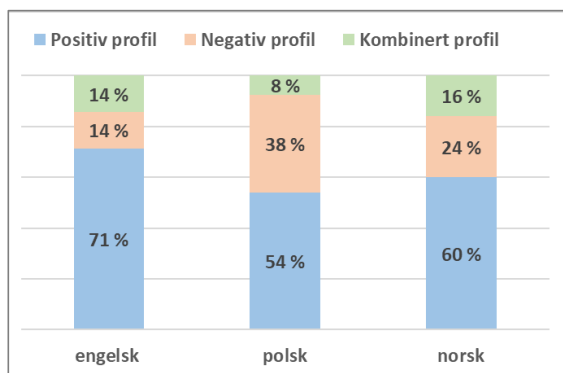
er mulig å bruke i målspråket. Det hadde derfor krevd en finere inndeling i generiske undertyper for å vite hvilke områder de polske prediksjonene sammenfaller og ikke sammenfaller med målspråket eller engelsk. Dette gjør at jeg på individnivå kun har sett på de rent kvantitative forskjellene mellom gruppene, og ikke på mer kvalitative forskjeller knyttet til sammenfalls- og markørkontekster, slik jeg har gjort tidligere.

9.7.8.1 Resultat - Individuelle profiler

Tabell 9-32 og figur 9-48 (under) viser antall og prosentandel av de ulike profiltypene ved generisk referanse i de tre gruppene.

Den polske gruppen har 14 positive profiler (54 %), og dette er den så langt høyeste andelen av positive polske profiler i undersøkelsen. Det kan dermed se ut som om det først og fremst er innen generiske referanse at prediksjonene slår positivt ut for hypotese 1, men ettersom det er umulig å skille ut markørkontekster på individnivå ved generisk referanse, er det også vanskelig å validere disse individuelle profilene med tanke på om morsmålsansvaret sammenfaller med målspråket eller ikke.

Figur 9-48 Profiltyper i de tre gruppene (generisk referanse)



Tabell 9-32 Profiltyper i de tre gruppene (generisk referanse)

Profiltipe	Engelsk	polsk	norsk
Positiv	71 % (15)	54 % (14)	60 % (15)
Negativ	14 % (3)	38 % (10)	24 % (6)
Kombinert	14 % (3)	8 % (2)	16 % (4)
Total	21	26	25

I tillegg viser det seg at selv om det er en god del positive profiler i den polske gruppen, så er dette også tilfellet for de to andre gruppene. I den norske gruppen er det 60 %

positive profiler (15 stk.), mens det i den engelske er hele 71 % (15 stk.). Vi så også på gruppenivå at både den engelske og den norske gruppen noen ganger hadde formvalg som var mer i tråd med prediksjonene for den polske gruppen, så at disse to gruppene også har flere positive profiler, stemmer godt med funnene på gruppenivå. Forskjellene i antall positive profiler er imidlertid ikke signifikante for antall positive profiler versus negative/kombinerte (pol./nor.: $\chi^2 = 0.1968$; $p = 0.66$; pol./eng: $\chi^2 = 0.5198$; $p = 0.22$). Det er dessuten den polske gruppen som har flest negative profiler av gruppene (pol. 38 %; eng. 14 %; no. 24 %).

De individuelle profilene viser altså at det ikke er signifikante forskjeller mellom gruppene når det kommer til positive profiler, men polsk er likevel den gruppen som har *færrest* positive og *flest* negative profiler, så selv om det er mange innlærere som velger form i tråd med prediksjonene, så fremstår dette som lite spesifikt for polske innlærere.

9.7.9 Oppsummering av funn for hypotese 1 ved generisk referanse

Tabell 9-33 (under) gir en oversikt over funnene for hver av de predikerte form/kontekst-assosiasjonene ved generisk referanse, og andel positive versus negative/kombinerte profiler i de tre gruppene.

Tabell 9-33 Oversikt over hovedfunn for hypotese 1 ved generisk referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
UB-kontekst/ubestemt ent/flt			
- Innen kontekst	67 %	91 %	81 %
- Innen form	69 %	57 %	44 %
BE-kontekst/bestemt ent/flt			
- Innen kontekst	57 %	0 %	75 %
- Innen form	25 %	0 %	25 %
ART-kontekst/artikkelfrase			
- Innen kontekst	0 %	6 %	7 %
- Innen form	0 %	100 %	67 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	54 %	71 %	60 %
- Negative/kombinerte	46 %	29 %	40 %

På gruppenivå så vi at det var ulik grad av *morsmålssamsvar* for de ulike assosiasjonene. Det var høyest for *UB-kontekst/ubestemt entall/flertall*, nest høyest for

BE-kontekst/bestemt form, og morsmålsvars manglet nesten helt for *ART-kontekst/artikkel frase*. Dette vil si at morsmålsvars var høyest i sammenfallskonteksten (UB), og lavere i de to markørkontekstene (BE og ART). Videre var over halvparten av de polske individuelle profilene *positive* (54 %), noe som gjør de predikerte form/kontekst-assosiasjonene representative for over halvparten av innlærerne, altså en større *intragruppe-homogenitet* enn vi har sett ved de andre referensielle typene. Siden det ikke har vært mulig å undersøke sammenfalls- og markørkontekster på individnivå, er det imidlertid ukjent i hvor stor grad formvalgene for disse profilene sammenfaller med målspråket.

Sammenligning med den norske gruppen viste små forskjeller mellom gruppene på gruppenivå, og i BE-kontekster har den norske gruppen også større andel av bestemt form enn det den polske gruppen har, til tross for at det er ventet at forskjellen skulle gå den andre veien, med mest bestemt form i den polske gruppen. Sammenligningen viste altså relativt stor likhet med målspråksgruppen for alle de tre predikerte form/kontekst-assosiasjonene. Den norske gruppen hadde også noen flere positive profiler enn den polske (pol. 54 %; nor. 60 %), men denne forskjellen viste seg å ikke være signifikant, altså er gruppene relativt homogene også på individnivå.

Sammenligningen med den engelske gruppen viste større forskjeller på gruppenivå, og her er forskjellen størst for assosiasjonen *BE-kontekst/bestemt form*, hvor den polske gruppen har en sterkere assosiasjon, mens det ikke er noen assosiasjon i den engelske gruppen i det hele tatt. Dette stemmer godt med at dette er en markørkontekst. Når det gjelder individuelle profiler har den engelske gruppen en god del flere positive profiler enn den polske (pol. 54 %; eng. 71 %), men denne forskjellen var altså ikke signifikant.

Det samlede inntrykket for generisk referanse er at det er lite som peker i retning av transfer. Til tross for en god del positive profiler, har det kommet fram en generell likhet med målspråksgruppen som heller kan peke i retning av at det er tilegnelse av målspråket som står bak de mønstrene en ser i den polske gruppen.

Forskjellene mellom den polske og den engelske gruppen vitner til en viss grad om intergruppe-heterogenitet, men tatt i betraktning at den norske og den polske gruppen er såpass like, så kan det igjen se ut til at det er den engelske gruppen som har

de mest gruppespesifikke tendensene. I alle tre konteksttypene hadde den engelske gruppen utstrakt bruk av ubestemt entall/flertall, og mer enn de andre to gruppene i både UB- og BE-kontekster, noe jeg tidligere nevnte kan være tegn på både intergruppetheterogenitet, morsmålsamsvar og forskjeller fra målspråket i det engelske materialet, og dermed kanskje transfer. Morsmålsamsvaret ble her knyttet til at man i engelsk sjeldnere bruker bestemt form ved generisk referanse, noe de engelske innlærerne i dette materialet gjorde.

Valg av form etter setningsposisjon og aktiveringsgrad

Jeg har for de andre referensielle typene sett på hvordan formtypene fordeler seg over setningsposisjon og aktiveringsgrad uavhengig av hverandre, og jeg vil også gjøre dette for generisk referanse. I det generiske materialet var det veldig lite bruk av artikkel frase, så disse er ikke tatt med i vurderingen. Tabellene på side 310 viser de relevante tallene for denne vurderingen.

De tre gruppene er mer like ved generisk referanse enn det vi har sett tidligere. Alle tre gruppene bruker ubestemt entall/flertall oftest i forfeltet (pol. 69 %; eng. 91 %, nor. 79 %), mens det er ulikheter i hvor frasene med *bestemt* form brukes: De polske innlærerne bruker bestemt form oftest ved *aktiverte* substantivfraser uavhengig av plassering i setningen (88 %), de engelske i forfeltet (100 %), men den norske gruppen bruker bestemt form aller mest ved ikke-aktiverte substantivfraser (75 %).

I det indefinitte og definitte materialet så det ut til at mønstrene bare var en indirekte effekt av at visse kontekster var mer utbredte enn andre, og ved generisk referanse ser vi noe lignende, men heller enn at det bare er ett av trekkene som dominerer, for eksempel forfelt eller aktivert, er det kombinasjoner av disse: I den polske gruppen er 52 % av alle de generiske substantivfrasene *aktiverte forfelt*, noe som påvirker sterkt til at formtypene først og fremst brukes i enten aktiverte eller forfelt. I den engelske gruppen er 93 % av substantivfrasene i forfeltet, og i den norske gruppen er 53 % *ikke-aktiverte forfelt*, som påvirker til at formene brukes enten ved ikke-aktiverte substantiv eller i *forfelt*. Tallene er med andre ord også her delvis påvirket av hvilke konteksttyper som er representert og dominerer i de ulike gruppene, og ikke til noen spesifikk form/kontekst-assosiasjon.

Andre mønster

I den deskriptive beskrivelsen av det generiske materialet i 9.7.1 så vi at alle tre gruppene hadde størst andel ubestemt entall/flertall, men at det var forskjeller mellom gruppene i fordelingen av ubestemt versus bestemt form. For det første skilte den polske gruppen seg fra begge de andre gruppene med å ha større andel bestemt form totalt (pol. 31 %; eng. 10 %; nor. 23 %), og for det andre skilte den engelske gruppen seg ut med å ha større andel ubestemt entall/flertall (pol. 67 %; eng. 88 %; nor. 72 %). Ingen av de to innlærergruppene er dermed helt lik målspråksgruppen på disse punktene.

Den lavere andelen av bestemt form i den engelske gruppen har jeg vært inne på at kan knyttes til morsmålspåvirkning som har sitt opphav i at bestemt form er mindre utbredt i engelsk enn i norsk ved generisk referanse. Hva den høyere andelen av bestemt form i det polske materialet kan tilskrives, er mer uklart, men det kan handle om hvilke generiske undertyper som er representert i materialet, og at det polske materialet bare tilfeldigvis har flere forekomster som krever eller tillater bestemt form enn den norske gruppen har. Eventuelt kan det også skyldes transfer som ligger utenfor det hypotesen handler om.

Konklusjon om transfer ved generisk referanse

Gjennomgangen av det generiske materialet viser at det heller ikke her er funn i de polske innlærertekstene som indikerer transfer av den typen hypotese 1 tar for seg. Under vil jeg oppsummere og sammenligne funnene for den polske gruppen innen hver av de referensielle typene.

9.8 Polsk morsmålsassosiasjon på tvers av referensiell type

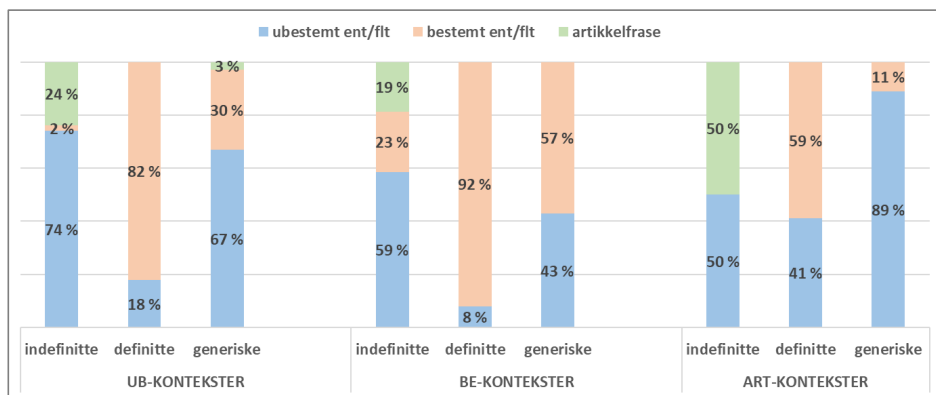
Selv om det har vært lite som tyder på transfer hos de polske innlærerne i gjennomgangen av materialet, vil jeg her kort sammenligne form/kontekst-assosiasjonene innen de ulike referensielle typene. Det meste av det som tas fram her, har allerede kommet fram indirekte tidligere i undersøkelsen, så dette er mer en oppsummering som retter fokus mot bare den polske gruppen og hvordan formvalgene kommer til uttrykk innen de referensielle typene. Det er en generell forventning om at det for den polske gruppen skal være likheter *mellom* de ulike referensielle typene, men

forskjeller *innad* i hver referansetype, avhengig av hvilken konteksttype det er snakk om.

Søylediagrammene på neste side setter opp de ulike referensielle typene ved siden av hverandre både innen kontekst og innen form. Begge diagrammene viser at det ikke er noen nevneverdig *likhet* mellom det indefinitte, definitte og generiske materialet, men tvert imot ulike profiler for hver av dem.

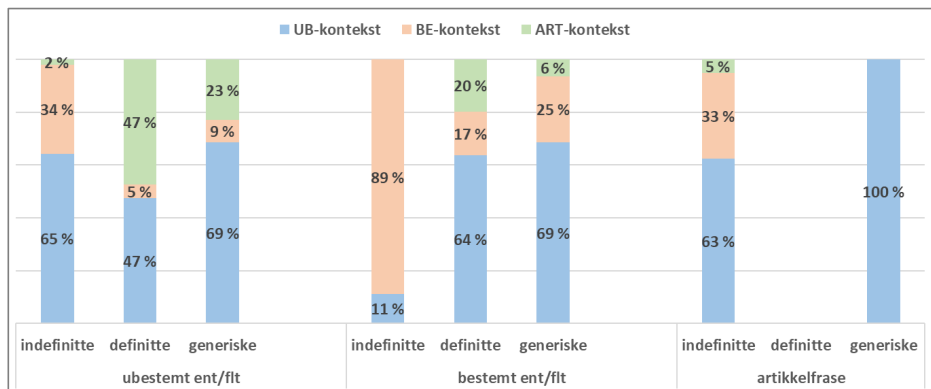
Figur 9-49 (under) hvilke former som er brukt i de ulike konteksttypene. I UB-kontekster varierer bruken av ubestemt entall/flertall etter referensiell type, og bruken er spesielt lav ved definit referanse (ind. 74 %; def. 18 %; gen. 67 %). I BE-kontekster varierer bruken av bestemt form, og den er spesielt lav ved indefinit referanse (ind. 23 %; def. 92 %; gen. 57 %). I ART-kontekster kommer forskjellen først og fremst til uttrykk gjennom at artikkelfrase kun brukes ved indefinit referanse, og ikke ved definit og generisk. Hver referansetype har dermed sin egen fordeling av former, selv om hypotesen tilsier at det skal være likt.

Figur 9-49 Andel av form innen kontekst etter referensiell type i den polske gruppen



Figur 9-50 (neste side) viser hvilke kontekster de ulike formene brukes i, og det er forskjeller her også: Ubestemt entall/flertall brukes mest i UB-kontekster ved generisk referanse, og noe mindre ved de andre to referansetyper (ind. 65 %; def. 47 %; gen. 69 %). Bestemt form brukes oftest i BE-kontekster ved indefinit referanse, og en god del mindre ved de andre to referansetyper (ind. 89 %; def. 17 %; gen. 25 %). Artikkelfrase er først og fremst brukt ved indefinit referanse, og brukes ikke ved definit og generisk referanse (ind. 50 %; def. 0 %; gen. 0 %).

Figur 9-50 Andel av kontekst innen form etter referensiell type i den polske gruppen



Et raskt blikk på fordelingen av former *innad* i hver referansetype, viser at det er store likheter mellom de ulike konteksttypene: Ved indefinitt referanse er det generelt høy andel av ubestemt entall/flertall i alle de tre konteksttypene (UB. 74 %; BE. 59 %; ART. 50 %), og det samme gjelder andel av bestemt form ved definitte referanse (UB. 82 %; BE. 92 %; ART. 59 %). Også i ART-kontekster er det gjennomgående lav andel av artikkel frase (UB. 3 %, BE. 0 %; ART. 0 %).

Når den polske gruppen sammenlignes med seg selv innen referensielle typer, ser vi altså at det ikke er den forventede likheten på tvers av referensiell type. Det er tvert imot veldig ulikt mellom referansetyperne, og derimot ganske likt innad i hver referansetype, der det er ventet variasjon etter om det er snakk om en UB-, BE- eller ART-kontekst.

9.9 Oppsummering hypotese 1

I dette kapitlet har jeg undersøkt den første hypotesen om morsmålpåvirkning fra polsk, som tok utgangspunkt i ideen om at bestemt form og ubestemt artikkel brukes for å markere substantivfraser som har en setningsplassering som byter med tema/rema-prinsippet. Bestemt form og ubestemt artikkel antas derimot å *ikke* brukes når substantivfrasen er plassert i tråd med tema/rema-prinsippet. Hypotesen hadde tre delprediksjoner om morsmålsamsvar i form av spesifikke form/kontekst-assosiasjoner:

1) Assosiasjon mellom UB-kontekster og ubestemt entall/flertall

Ubestemt entall/flertall brukes ved *aktiverte forfelt* eller *ikke-aktiverte slutfelt*

2) Assosiasjon mellom BE-kontekster og bestemte former

Bestemte former brukes ved *aktiverte slutfelt*

3) Assosiasjon mellom ART-kontekster og artikkelfraser

Artikkelfraser brukes ved *ikke-aktiverte forfelt*

Hypotesen impliserte i tillegg at *alle* disse assosiasjonene skal være til stede i det polske materialet, og at de skal være det uavhengig av om det er snakk om indefinitt, definitt eller generisk referanse.

Hypotesen har blitt undersøkt både i gruppen som helhet og i individuelle profiler, og jeg har forsøkt å påvise alle de fire transfereffektene *morsmålssamsvar*, *intragruppe-homogenitet*, *intergruppe-heterogenitet* og *intralingval kontrast*. Tre av transfereffektene ble undersøkt både på gruppenivå og i individuelle profiler, mens *intragruppe-homogenitet* kun ble undersøkt gjennom individuelle profiler.

Mens transfereffekten *morsmålssamsvar* i prinsippet kunne undersøkes i det samlede materialet ettersom den ikke var ventet å variere med referensiell type, var det nødvendig å dele opp materialet etter referensiell type for å undersøke effektene *intergruppe-heterogenitet* og *intralingval kontrast*. Dette var for å kunne avdekke områder der prediksjonene for den polske gruppen var sammenfallende og ikke sammenfallende med det formvalget en kunne forvente i den engelske og den norske gruppen.

Den overordnede konklusjonen fra undersøkelsen var at det *ikke* foreligger transfer fra polsk av den typen hypotese 1 handler om. Denne konklusjonen baserer seg på funn både i det samlede materialet og i sammenligningen mellom polsk og de to andre gruppene i det differensierte materialet. Jeg vil nå oppsummere de viktigste tendensene og argumentene for at det ikke foreligger transfer hos de polske innlærerne.

9.9.1 Morsmålssamsvar og intergruppe-heterogenitet i samlet materiale

Eventuell morsmålpåvirkning hos de polske innlærerne skulle ifølge hypotesen komme til syne uten at materialet var delt opp i referensielle typer, altså i det jeg har omtalt som *det samlede materialet*, ettersom innlærernes oppfatning av redundans ble antatt å gjelde uavhengig av referansetype. Dette innebærer at jeg allerede i denne første delen av

undersøkelsen kunne anta at det ikke forelå transfer av den typen hypotese 1 handler om. For det første fant jeg ikke morsmålssamsvar for noen av de tre predikerte form/kontekst-assosiasjonene i det samlede materialet. For assosiasjonen *UB-kontekst/ubestemt entall/flertall* var det tendenser som var i tråd med prediksjonene, mens assosiasjonen var svakere for *BE-kontekst/bestemt form* og nesten fraværende for *ART-kontekst/artikkelfrase*.

Vi så også at antall positive profiler i det samlede materialet var lavt (2 stykker), og at det var stor variasjon i antall positive profiler innen de ulike referensielle typene: størst andel ved generisk referanse og lavest andel ved definitt.

Disse observasjonene i det samlede polske materialet er i seg selv nok til å utelukke transfereffektene *morsmålssamsvar* og *intragruppe-homogenitet*, altså at de predikerte form/kontekst-assosiasjonene er til stede i innlærerspråket og samtidig representative for den polske gruppen. Det var likevel noen tendenser som først kunne utelukkes fra å være transfer etter en sammenligning med den engelske og den norske gruppen. Tendensen til assosiasjon mellom UB-kontekst og ubestemt entall/flertall var relativt sterk i det polske materialet, men sammenligningen med den norske og den engelske gruppen viste at en slik assosiasjon sto sterkt også hos disse, dermed var det lav *intergruppe-heterogenitet* og stor grad av sammenfall med målspråket. Intergruppe-heterogeniteten sto sterkere for BE- og ART-prediksjonene, men her var det som sagt lavt morsmålssamsvar. Det var altså ulike argument for manglende transfer for de tre delprediksjonene, enten lavt morsmålssamsvar eller at de språklige valgene ikke var spesifikke for den polske gruppen, men helhetsbildet er at det i det samlede materialet ikke var mulig å finne transfer av den typen hypotese 1 handler om.

9.9.2 Morsmålssamsvar og heterogenitet i det differensierte materialet

Et sentralt forbehold ved det samlede materialet var at det ikke skilte ut kontekster hvor de tre gruppene kunne forventes å være like i formvalget fordi prediksjonene for polsk morsmålssamsvar tilfeldigvis sammenfaller med det som er de vanlige formvalgene i språkene norsk og engelsk, noe som potensielt kunne tilsløre faktiske forskjeller mellom gruppene. Gjennomgangen av de tre ulike referensielle typene hver for seg hadde dermed som mål å skille ut disse *sammenfallskontekstene*, og eventuelt avdekke om mangel på eller små forskjeller mellom gruppene kunne skyldes en slik tilsløring av

faktiske forskjeller. Morsmålssamsvar og forskjeller mellom gruppene i *markørkontekster* ble regnet som viktige for å kunne argumentere for transfer, ettersom det var i disse kontekstene at forskjellene mellom polsk og de andre to gruppene var ventet å komme til syne om hypotesen stemte.

Mens dette i *praksis* var en undersøkelse av *intergruppe-heterogenitet* både i forhold til den engelske og den norske gruppen, var det ulike formål med å sammenligne med den norske og den engelske gruppen: Intergruppe-heterogenitet med den engelske gruppen i markørkontekster kunne vise at de to innlærergruppene hadde ulike formvalg, og at mønstrene i den polske gruppen dermed ikke var universelle for innlærere uavhengig av morsmål. Intergruppe-heterogenitet med den norske gruppen i markørkontekster kunne derimot vise at formvalgene til den polske gruppen ikke alternativt kunne forklares som målspråkstilegnelse, og denne sammenligningen knyttet jeg til transfereffekten *intralingval kontrast* (se 8.5.3.4). Både gruppetall og individuelle profiler i den polske gruppen som viste morsmålssamsvar og skilte seg fra de andre gruppene i markørkontekster, var altså regnet som sterkere tegn på transfer. I sammenfallskontekster var det derimot ikke ventet forskjeller mellom gruppene, og for disse språklige kontekstene var det dermed ikke noe krav om intergruppe-heterogenitet.

Jeg vil nå oppsummere hva undersøkelsen av sammenfalls- og markørkontekster sammen med morsmålssamsvar fikk fram fra det differensierte materialet. Alle de tre konteksttypene for hypotese 1 (UB/BE/ART) vekslet på å være sammenfallskontekst og/eller markørkontekst innen de ulike referansetyperne. Jeg vil i oppsummeringen ta for meg hver og en av de tre delprediksjonene og se på hvordan forskjellene og likhetene med de andre gruppene var, avhengig av om det var snakk om markør- eller sammenfallskontekst.

Prediksjon 1: UB-kontekst og ubestemt ent/flt: ikke støtte i markørkontekster

Assosiasjonen mellom *UB-kontekster* og *ubestemt entall/flertall* er den delprediksjonen som har hatt sterkest morsmålsamsvar gjennom hele undersøkelsen av hypotese 1, med gjennomgående over 50 % andel både innen form og kontekst både i det samlede materialet og innen de ulike referensielle typene, bortsett fra ved definitivt referanse (markørkontekst), der formen bare ble brukt i 20 % av kontekstene. Dette vil si at morsmålssamsvaret er høyest når UB-konteksten er sammenfallskontekst, men lavere

når den er markørkontekst, noe som er negativt for et argument om transfer ettersom den delprediksjonen som har lavest morsmålssamsvar, er den eneste hvor prediksjonen for polsk ikke sammenfaller med det en kan vente i målspåket.

Det høye morsmålssamsvaret i sammenfallskontekster er også generelt ledsaget av *likhet* med de andre gruppene, noe som er ventet siden det er en sammenfallskontekst. Til tross for lavt morsmålssamsvar når UB er markørkontekst (ved definit referanse), er forskjellen faktisk større her både i forhold til den norske og den engelske gruppen: Polsk har lavt morsmålssamsvar, men likevel *sterkere* enn det en finner i den norske gruppen, men igjen *svakere* enn i den engelske gruppen. At den polske gruppen faktisk er forskjellig fra de andre gruppene når UB er markørkontekst, er positivt for hypotesen, men lavt morsmålssamsvar gjør at denne gruppespesifikke tendensen ikke er tilstrekkelig koblet til hva det er i morsmålet som motiverer formvalgene. Det kan likevel være at det er en type transfer her som ikke fanges opp av hypotesen.

Vurderingen av sammenfalls- og markørkontekster for prediksjonen om assosiasjon mellom UB-kontekster og ubestemt entall/flertall viste altså at høyt morsmålssamsvar først og fremst opptrer når UB er sammenfallkontekst, og da ledsaget av stor likhet med de andre gruppene, mens markørkonteksten hadde lavt morsmålssamsvar, lavere enn den andre innlærergruppen, men likevel høyere enn målspåkegruppen. Som i det samlede materialet er det altså ulike argument mot transfer for hver av assosiasjonene, men det helhetlige bildet av morsmålssamsvar og gruppeforskjeller viser at argumentet for transfer er svakt for delprediksjonen om en assosiasjon mellom UB-kontekster og ubestemt entall/flertall. Tabell 9-34 gir en oversikt over transfereffekter for denne delprediksjonen.

Tabell 9-34 Transfereffekter for assosiasjonen UB-kontekst og ubestemt entall/flertall

	Indefinit referanse	Definit referanse	Generisk referanse
UB-kontekster =	Sammenfallskontekst	Markørkontekst	Sammenfallskontekst
Morsmålssamsvar	Ja	Nei	Ja
Ulik engelsk gruppe	Nei	Ja	Nei
Ulik norsk gruppe	Nei	Ja	Nei

Prediksjon 2: BE-kontekst og bestemte former – tegn til transfer?

Assosiasjonen mellom *BE-kontekster* og *bestemt form* hadde ikke spesielt høyt morsmålssamsvar i det samlede materialet, gjennomgående rundt 30 % innen både form

og kontekst. Dette viste seg likevel å være den assosiasjonen som har mest variert morsmålsamsvar innen de ulike referensielle typene, i tillegg til å være preget av større distanse mellom andel innen form og andel innen kontekst enn de andre delprediksjonene. Tendensen har vært at om andelen er høy innen form, så er den lav innen kontekst, eller motsatt, altså er ikke forholdet mellom form og kontekst særlig *gjensidig*. Ved definitiv referanse var morsmålsamsvaret høyt på den måten at de fleste BE-kontekstene hadde bestemt form, mens det var høyt innen indefinitiv og generisk referanse på den måten at de aller fleste tilfellene av bestemt form ble brukt i BE-kontekster.

Ved indefinitiv referanse er BE-kontekster markørkontekst, og den polske gruppen viste seg også å være forskjellig og å ha en sterkere form/kontekst-assosiasjon enn begge de andre gruppene. Når BE-konteksten var sammenfallskontekst (definitiv), var derimot likheten mellom gruppene større, som ventet. Generisk referanse er også markørkontekst for denne prediksjonen, og her var det forskjeller mellom alle gruppene: Den polske gruppen hadde sterkere samsvar enn den engelske, mens den norske hadde enda sterkere enn den polske. Mønstrer er altså at likheten mellom gruppene opptrer i sammenfallskonteksten, mens det er forskjeller i markørkontekstene, noe som er positivt for argumentet om transfer. Funnene for transfereffekter er oppsummert i tabell 9-35.

Tabell 9-35 Transfereffekter for assosiasjonen BE-kontekst og bestemte form

	Indefinitiv referanse	Definitiv referanse	Generisk referanse
BE-kontekster =	Markørkontekst	Sammenfallskontekst	Markørkontekst
Morsmålsamsvar	Kun innen form	Kun innen kontekst	Kun innen kontekst
Ulik engelsk gruppe	Ja	Nei	Ja
Ulik norsk gruppe	Ja	Nei	Ja

Sammenligningen av sammenfalls- og markørkontekster for assosiasjonen *BE-kontekst/bestemt form* viste dermed at det var et noe ujevnt morsmålsamsvar innen alle de tre referensielle typene, og at morsmålsamsvaret sto sterkest innen indefinitiv og definitiv referanse, men noe lavere innen generisk. Det var også tendenser til forskjeller mellom gruppene, minst ved indefinitiv referanse (sammenfallskontekst) og mest ved definitiv referanse (markørkontekst), noe som taler positivt for et argument for transfer.

For delprediksjonen om assosiasjon mellom BE-kontekst og bestemt form var det altså ikke sterke tegn til transfer, men det ble observert ett interessant mønster innen

indefinit referanse, hvor de fleste bestemt form faktisk ble brukt i BE-konteksten, slik prediksjonen sier (89 %), en tendens som ikke ble funnet i de to andre gruppene. Det som er interessant, er at dette faktisk skjer i *aktiverte slutfelt*, hvor plasseringen til substantivfrasen nettopp bryter med tema/rema-prinsippet. Imidlertid inneholder det indefinite materialet flere slike kontekster, og bare 23 % av disse hadde bestemt form. Bestemt form brukes dermed neppe for å markere brudd på tema/rema-prinsippet, men når formene først brukes, begrenses det altså i stor grad til slike kontekster. Dette kan potensielt handle om individuell variasjon, der enkelte av de polske innlærerne bruker formene på denne måten, mens andre ikke gjør det. De individuelle profilene viser imidlertid at dette ikke er tilfellet, da det er kun to av innlærerne som bare bruker eller har overvekt av bestemt form i BE-kontekster. I alle de andre individuelle profilene er forekomstene av bestemt form i BE-kontekster i variasjon med en av de ubestemte formene, som vi så at det også på individnivå er tydelig variasjon i formvalget.

Prediksjon 3: ART-kontekst/artikkel frase – ikke transfer

Assosiasjonen mellom *ART-kontekst* og *artikkel frase* er den delprediksjonen som har hatt lavest morsmålssamsvar både i det samlede materialet og innen hver av de referensielle typene, i de fleste tilfellene 0-5 % både innen kontekst og form. Assosiasjonen var noe sterkere ved indefinit referanse, hvor 50 % av alle ART-kontekster hadde artikkel frase, men det var her bare fire ART-kontekster totalt i det polske materialet. Innen definit og generisk referanse var det nesten ingen artikkel fraser, men det var en god del ART-kontekster, så det lave antallet artikkel fraser er i seg selv en indikasjon på lavt morsmålssamsvar. Hadde hypotesen stemt, skulle det generelt vært flere artikkel fraser, og de ville vært brukt i ART-kontekster, noe de ikke er.

Det er også generelt stor likhet mellom den polske gruppen og de andre to gruppene for assosiasjonen *ART-kontekst/artikkel frase*, noe som gjør formvalgene lite spesifikke for polske innlærere, og i stor grad sammenfallende med målspåket. At gruppene også er like når ART-konteksten er markørkontekst (definit referanse), tyder på at den polske gruppen generelt er mer lik de to andre gruppene enn den er i tråd med prediksjonene om morsmålssamsvar. Vurderingen av transfereffekter for denne assosiasjonen er oppsummert i tabell 9-36 på neste side.

Tabell 9-36 Transfereffekter for assosiasjonen ART-kontekst og artikkelfrase

	Indefinit referanse	Definit referanse	Generisk referanse
ART-kontekster =	Sammenfallskontekst	Markørkontekst	Markørkontekst
Morsmålssamsvar	Nei	Nei	Nei
Ulik engelsk gruppe	Nei	Nei	Nei
Ulik norsk gruppe	Nei	Nei	Nei

Med tanke på prediksjonene for assosiasjonen *ART-kontekst/artikkelfrase* er det altså en kombinasjon av lavt morsmålssamsvar og manglende forskjell fra de to andre gruppene som taler mot at transfer foreligger. De polske innlærerne har ikke den predikerte assosiasjonen mellom form og kontekst, og er ikke spesielt ulik de andre to gruppene, heller ikke når ART-konteksten er markørkontekst. Dette viser altså lavt morsmålssamsvar, lav intergruppe-heterogenitet, lav intragruppe-homogenitet, og stort grad av sammenfall med målspråket.

Individuelle profiler

Undersøkelsen har også gått inn i individuelle profiler for å finne ut hvor mange polske profiler som har formvalg som er i tråd med prediksjonene. Undersøkelsene av profiler viste for det første at det generelt sett var få positive profiler i det polske materialet, for det andre at andelen positive profiler i den polske gruppen stort sett ikke var signifikant forskjellig fra i den engelske gruppen, og for det tredje at den norske gruppen i flere tilfeller hadde signifikant *flere* positive profiler enn den polske gruppen.

Mens det i det samlede materialet og det definte materialet ble funnet veldig få positive profiler i den polske gruppen (2 % i begge), var det noe mer i det indefinte (23 %) og enda noen flere i det generiske (54 %). Ved indefint referanse viste det seg imidlertid at alle de positive profilene unntatt én hadde bare sammenfallskontekster, noe som gjorde at profilene også stemte med formvalget i målspråket. Ved generisk referanse lot sammenfallskontekster seg ikke undersøke, men både den engelske og den norske gruppen hadde her *flere* positive profiler enn den polske, selv om forskjellene ikke var signifikante.

Få positive profiler viser at det både er lite morsmålssamsvar og lite intragruppe-homogenitet, altså at de predikerte formvalgene ikke var representative for den polske gruppen. Ved generisk referanse var predikerte formvalg representative for over halvparten av gruppen, men enda sterkere representert i de to andre gruppene, noe

som gjør det vanskelig utelukke universelle tendenser eller målspråkstilegnelse. I alle delmaterialene bortsett fra generisk referanse, hadde den norske signifikant flere positive profiler enn den polske, og ved indefinitt referanse gjaldt dette også den engelske gruppen. For den norske gruppen var de positive profilene et resultat av stor mengde sammenfallskontekster.

Tabell 9-37 gir en oversikt over transfereffekter i individuelle profiler, og parentesene angir når den engelske eller den norske gruppen hadde signifikant *flere* positive profiler enn den polske (nor.>pol. og (eng.> pol.).

Tabell 9-37 Transfereffekter for individuelle profiler

Transfereffekt:	Samlet materiale	Indefinitt referanse	Definitt referanse	Generisk referanse
Morsmålssamsvar	Nei	Nei	Nei	Nei
Intragruppe-homogenitet	Nei	Nei	Nei	Nei
Ulik engelsk gruppe	Nei	Ja* (eng.> pol.)	Nei	Nei
Ulik norsk gruppe	Ja* (nor.> pol.)	Ja* (nor.> pol.)	Ja* (nor.> pol.)	Nei

De individuelle profilene viste altså at det var svakt morsmålssamsvar og homogenitet i den polske gruppen, men også at resultatene kunne vært misvisende om ikke sammenfalls- og markørkontekster hadde blitt studert. Det at prediksjonene for hypotesene har såpass sammenfall med formvalgene i norsk og engelsk er altså en viktig faktor å ta hensyn til i undersøkelser som dette.

Konklusjon om manglende transfer hos de polske innlærerne

Denne gjennomgangen av funn for hypotese 1 har vist at når man tar i betraktning alle de fire transfereffektene *morsmålssamsvar*, *intragruppe-homogenitet*, *intergruppe-heterogenitet* og *intralingval kontrast*, så vokser det ikke fram noe sterkt argument for transfer av den typen hypotese 1 tar for seg. Det fins riktignok noen tendenser til gruppespesifikke formvalg i det polske materialet, som delvis kan kobles til prediksjonene til hypotesen, men disse utgjør bare en liten del av det helhetlige bildet, og hypotesen og prediksjonene, slik de er formulert, krever mer gjennomgående funn både i hele materialet og innen hver av de referensielle typene.

9.9.3 Sekundære funn i materialet

Selv om tegnene på transfer hos de polske innlærerne har vært få, er det en del andre observasjoner som har gått igjen i løpet av undersøkelsen. For det første at den polske gruppen i stor grad velger form i tråd med det som er typisk i norsk målspråk, for det andre forholdet mellom form og setningsposisjon og tekstuell aktivering uavhengig av hverandre, og for det tredje tegn på transfer hos de *engelske* innlærerne.

Målspråkslike formvalg i den polske gruppen

Det har gjennom undersøkelsen flere ganger kommet fram at den polske gruppen i stor grad har like formvalg som den norske gruppen, noe som betyr at de polske innlærerne ligger nært opp til målspråket i valg av form. De polske innlærerne ser derimot ikke ut til å bruke formene *helt* målspråkligt, men de ser ut til å være *mer* i tråd med målspråksnormen enn med prediksjonene om morsmålsamsvar.

I det samlede materialet så vi at alle de tre gruppene hadde en relativt lik fordeling av de ulike formtypene innen hver av konteksttypene: Innen alle konteksttypene (UB, BE og ART) var det størst andel ubestemt entall/flertall, nest størst andel bestemt form og lavest andel av artikkelfrase. De ulike formtypene var tilsvarende brukt oftest i UB-kontekster, nest mest i BE-kontekster og minst i ART-kontekster. Denne jevnheten både innen form og kontekst, tyder på at konteksttypene for hypotese 1 er *lite* styrende for formvalgene i den polske gruppen, men også at formvalgene i den polske gruppen er relativt like det en finner målspråket.

Differensieringen i ulike referensielle typer viste videre at det i det polske materialet, i likhet med de norske, var *ulike* mønster i formvalgene etter hvilken type referanse det var snakk om. Det var størst andel av ubestemte former (entall, flertall og artikkelfrase) ved indefinitt referanse (90 %), mest bestemt form ved definitt referanse (78 %), og en blanding av bestemt form og ubestemt entall/flertall ved generisk referanse (ub. 67 %; be. 31 %; art. 2 %). Dette er med andre ord et distribusjonsmønster som stemmer godt med hvordan formene brukes i målspråket, og som også er relativt lik distribusjonen som ble funnet hos den norske gruppen. Den polske gruppen hadde likevel ofte en *kvantitativ* forskjell fra målspråksgruppen: en litt lavere andel ubestemt form ved indefinitt referanse (pol. 90 %; nor. 99 %) og en noe lavere andel bestemt form

ved definitt referanse (pol. 78 %; nor. 94 %). I tillegg var andelen bestemt form noe høyere i den polske gruppen ved generisk referanse (pol. 31 %; nor. 23 %).

Både innen indefinitt og definitt referanse har det vært mulig å identifisere én konteksttype der andelen av den typiske målspåkformen har vært spesielt lav i den polske gruppen i forhold til målspåket. Det er snakk om: 1) indefinitte BE-kontekster, og 2) definte ART-kontekster. Jeg vil nå se litt nærmere på disse tendensene.

Ved indefinitt referanse er det typiske formvalget i norsk ubestemt form, og det er spesielt i BE-kontekster at andelen av ubestemt form er lav i den polske gruppen sammenlignet med den norske, mens de øvrige kontekstene ligger nært opp til 100 % (UB. 98 %; BE. 78 %; ART. 100 %). BE-kontekster er *aktiverte slutfelt*, altså substantivreferenter som har vært nevnt før, men som står i slutfeltet. I denne konteksten skal det ifølge hypotesen brukes *bestemt* form, men som det gikk fram av undersøkelsen, stemte ikke dette; bare 23 % av disse kontekstene hadde bestemt form. Likevel viste det seg at hele 89 % av alle indefinitte fraser med bestemt form nettopp sto i BE-kontekster. At dette er både den konteksten hvor den polske gruppen avviker *mest* fra målspåkslik bruk, og hvor nesten alle indefinitte fraser med bestemt form brukes, er interessant, og kanskje en annen side av spørsmålet om transfer.

Ved definitt referanse, som typisk har bestemt form i målspåket, er det spesielt i ART-kontekster at andelen av bestemt form er spesielt lav i den polske gruppen (UB. 82 %; BE. 92 %; ART. 59 %). ART-kontekster er *ikke-aktiverte forfelt*, og ifølge hypotesen er det ventet artikkelfrase her, men dette er det lite av. At det er lavere andel av bestemt form i forhold til målspåksgruppen i akkurat denne konteksten, knyttet jeg i oppsummeringen av definitt referanse til at disse substantivfrasenes referenter ikke er aktiverte, altså ikke nevnt før. Jeg var inne på at Ekiert (2010: 140) har funnet at substantivfraser som ikke er tekstuelte aktiverte har mer variabel bruk av bestemt artikkel hos engelskinnlærere enn tekstuelte aktiverte. I mitt eget materiale viser det seg faktisk også at det generelt er mindre bruk av bestemt form ved ikke-aktiverte enn ved aktiverte fraser akkurat ved definitt referanse (aktiverte 86 %; ikke aktiverte 60 %).

Dette var et blick på de stedene hvor den polske gruppen avviker mest i forhold til målspåksgruppen med tanke på målspåkets typiske formvalg rent kvantitativt, men det er her et forbehold knyttet til at enkeltforekomstene i undersøkelsen bare er vurdert

etter om de har en form som stemmer med prediksjonene eller ikke, og ikke etter om formen er korrekt eller ikke i målspråket.

Formvalg i forhold til aktivering og setningsposisjon isolert sett

En annen ting som ble undersøkt ved siden av transfereffektene var forholdet mellom formvalg og aktivering eller setningsposisjon, uavhengig av hverandre, noe som ble omtalt som «alternative form/kontekst-assosiasjoner».

Tendensene som ble påpekt var at den polske gruppen i det samlede materialet hadde en viss assosiasjon mellom *ubestemt form* og *sluttfeltet*, og mellom *bestemt form* og *aktiverte* substantivfraser. Dette viste seg imidlertid å mest sannsynlig være en avspeiling av hvilke typer kontekster som var representert ved definit og indefinit referanse: Ved indefinit referanse sto nemlig 96 % av substantivfrasene i sluttfeltet, mens ved definit referanse sto 80 % av frasene i forfeltet. Den engelske gruppen hadde hovedsakelig de samme tendensene som den polske, mens den norske gruppen hadde en overrepresentasjon av substantivfraser med ikke-aktiverte referenter sammenlignet med innlærergruppene.

Konklusjonen av å se på setningsposisjon og aktivering hver for seg, var dermed at mønster som tilsynelatende tilsa at visse former ble brukt i visse kontekster, heller så ut til å være en avspeiling av hvilke kontekster som er representert i tekstene innen hver av de referensielle typene, og ikke spesifikke form/kontekst-assosiasjoner: I de to innlærergruppene er indefinitte substantivfraser først og fremst *sluttfelt* og definte substantivfraser først og fremst *forfelt*, mens det i den norske gruppen ser det ut til å være generell overvekt av ikke-aktiverte substantivfraser i hele materialet for hypotese 1, uavhengig av referansetype.

Transfer i det engelske materialet

For sammenligningen mellom den polske og den engelske gruppen har jeg hatt som utgangspunkt at også de engelske innlærerne er påvirket av morsmålet i sine formvalg, men dette har i seg selv ikke vært et empirisk mål med undersøkelsen, ettersom den engelske gruppen kun fungerer som kontrollgruppe. Det har likevel dukket opp enkelte tendenser underveis, som jeg har antydnet kan være tegn på transfer hos de engelske innlærerne. Disse tendensene har for det første vært knyttet til 1) høyere andel av

artikkelfrase ved indefinitt referanse, og 2) høyere andel av ubestemt entall/flertall ved generisk referanse. Begge disse tendensene har vært ledsaget av gruppeforskjeller både i forhold til den polske og den norske gruppen, altså er det funn som peker både i retning morsmålsamsvar med engelsk, og formvalg som er spesifikke for de engelske innlærerne.

I det *samlede* materialet så vi at det var høyere andel artikkelfrase generelt i det engelske materialet i forhold til begge de to andre gruppene (eng. 26 %; pol. 13 %; nor. 13 %), noe som også var tilfelle i det indefinitte materialet (eng. 42 %; pol. 22 %; nor. 24 %). Fraværet av artikkelfraser har derimot vært påfallende i alle gruppene ved definitt og generisk referanse, og her skilte den engelske gruppen seg ikke ut. Forskjellen i bruk av artikkelfraser har altså vært knyttet til indefinitt referanse, som jo også er det området hvor artikkelfraser er mest brukt i både norsk og engelsk. Siden tendensene i det samlede materialet nok bare er en refleksjon av forskjellene ved indefinitt referanse, er det bare indefinitt referanse som omtales videre.

I undersøkelsen av indefinitt referanse så vi at tendensen til større andel artikkelfrase i det engelske materialet var til stede både i UB-kontekster og BE-kontekster, men *ikke* i ART-kontekster. I UB-kontekster hadde den engelske gruppen 47 % artikkelfraser, mens den polske og den norske gruppen hadde henholdsvis 24 % og 20 %. I BE-kontekster hadde den engelske gruppen 34 % artikkelfraser, mens den polske hadde 19 % og den norske ingen. Disse forskjellene er spesielt interessante i forhold til den norske gruppen, som hadde lavere andel artikkelfrase og høyere andel ubestemt entall/flertall, noe som viser at kontrasten mellom den engelske gruppen er knyttet til nettopp en av hovedforskjellene mellom norsk og engelsk, nemlig at man i norsk kan bruke både ubestemt *entall* og *flertall* i tillegg til artikkelfraser ved tellelige substantiv, mens man i engelsk først og fremst må bruke *ubestemt flertall* og artikkelfraser. Undersøkelsens design har ikke framhevet forskjellene mellom ubestemt entall og flertall, men derimot slått de to formene sammen, derfor kan man ikke se om forskjellene her knytter seg til de to alternative entallsformene, men tallene tyder allikevel på at det er det som er tilfelle. Høy andel av artikkelfrase, reduserer andelen av ubestemt entall, noe som igjen påvirker andelen av formtypen ubestemt entall/flertall i den engelske gruppen.

Det var ikke noen tendens til overrepresentasjon av artikkelfraser i ART-kontekster, der var det faktisk den engelske gruppen som hadde *lavest* andel av denne formen (eng. 14 %; pol. 50 %; nor. 42 %), og det er uklart hva dette kan skyldes. Én forklaring kan være at artikkelfraser generelt sett ikke er spesielt utbredt ved ikke-aktiverte forfelt i engelsk. En annen forklaring kan være at det i akkurat ART-kontekster først og fremst er *flertallsreferenter* som er representert i de engelske tekstene, noe som gjør at problemstillingen med artikkelfrase og ubestemt entall ikke gir like mye utslag i disse kontekstene. Dette kan man imidlertid heller ikke se slik undersøkelsen er lagt opp, siden entall og flertall er slått sammen.

Den andre tendensen vi så i det engelske materialet, var i andelen av ubestemt entall/flertall ved generisk referanse, som var høyere i den engelske gruppen enn i de andre to gruppene (eng. 88 %; pol. 67 %; nor. 72 %). Denne tendensen var til stede både i UB-kontekster (eng. 91 %; pol. 67 %; nor. 81 %) og i BE-kontekster (eng. 100 %; pol. 43 %; nor. 25 %), mens det innen ART-kontekster var en lavere andel i den engelske gruppen enn i den polske, men høyere enn den norske gruppen (eng. 81 %; pol. 89 %; nor. 71 %).

Som jeg var inne på i kapittel 7, er ubestemt flertall den mest utbredte formen ved generisk referanse i engelsk, og bestemt form brukes sjeldnere i engelsk enn i norsk, så forskjellen mellom den engelske og den norske gruppen kan knyttes til denne kontrasten mellom de to språkene. At tendensen til overrepresentasjon av artikkelfrase kan knyttes til forhold i morsmålet, og i tillegg til at det er funnet forskjeller både i forhold til en annen innlærergruppe og målspråksgruppen, tyder på at dette kan utgjøre transfer hos de engelske innlærerne.

Det som videre er interessant, er at disse to tendensene til transfer i det engelske materialet er funnet ved indefinit og generisk referanse, men ikke ved definit. Dette er interessant fordi det først og fremst er innen indefinit og generisk referanse at språkene engelsk og norsk skiller seg fra hverandre i bruken av bestemthetsmarkører, mens de innen definit referanse er mer like i å typisk bruke bestemt form. Dette mønsteret av likhet og forskjell mellom den engelske og norske gruppen, kan derfor knyttes til transfereffekten *intralingval kontrast*, altså at de engelske innlærerne har transfer på områdene hvor morsmålet er ulikt målspråket, men ikke der de to språkene er like.

Alternativt kan man si at transferen er *mer observerbar* på de områdene hvor morsmål og målspråk er ulike, mens sammenfallet er tydeligst der språkene er like. På grunn av undersøkelsens design er det ikke mulig å undersøke om bruken av artikkelfraser hos de engelske innlærerne har ført til grammatiske feil (negativ transfer). Mest sannsynlig er det kun snakk om *overrepresentasjon* av artikkelfraser, siden artikkelfraser også er fullt mulig ved entalls substantiv i norsk og valget av entallsform først og fremst er knyttet til en ikke-obligatorisk nyansering av innholdet i fokus på type eller eksemplar (se 7.2.4). Den intralingvale kontrasten viser altså ikke nødvendigvis positiv versus negativ transfer, men muligens positiv transfer innen definit referanse, og transfer i form av *overrepresentasjon* innen indefinit og generisk referanse.

9.9.4 Hypotese 1: Avslutning

Jeg har nå gått igjennom materialet for hypotese 1 uten at jeg har funnet særlige tegn på den typen transfer som hypotesen handler om. Jeg har imidlertid funnet tegn på transfer i den engelske gruppen, og at den polske gruppen i stor grad har formvalg som ligner de norske. Jeg vil komme tilbake til en videre diskusjon av de negative funnene og gjennomføringen av undersøkelsen i kapittel 11, men før dette skal resultatene for hypotese 2 presenteres.

10 UNDERSØKELSE 2: HYPOTESE 2

Hypotese 2 ble presentert i 8.1.3, og er formulert slik:

De polske innlærerne vil overføre kasus fra morsmålet, og norsk substantivmorfologi vil bli tatt i bruk for å uttrykke kasusdistinksjoner.

Ifølge hypotesen vil valg av form hos de polske innlærerne være styrt av om substantivfrasen står i en kontekst for et umarkert eller markert kasustrekk. En overordnet konsekvens av dette er at språklige kontekster for nominativ i polsk, altså subjekter, vil ha ubestemt form, mens språklige kontekster for ikke-nominativ (øvrige kasus) vil ha bestemt form.

10.1 Utvalg av materiale og prediksjoner

Undersøkelsen av hypotese 2 tar for seg *fire* av de fem enkle substantivfrasetyper som materialet for undersøkelse 2 består av. Disse er igjen gruppert i to typer basert på om de har bestemt bøyningssendelse eller ikke: 1) ubestemt entall/flertall, 2) bestemt entall/flertall. Artikkelfraser er utelatt. De to formtypene vil bli omtalt som henholdsvis «ubestemte former» og «bestemte former», og det blir dermed ikke skilt mellom entalls- og flertallssubstantiv, siden valg av entall eller flertall ikke ventes å være påvirket av kasus.

Artikkelfraser er ikke inkludert materialet for hypotese 2 ettersom hypotesen handler om substantivmorfologi og derfor ikke har bruk av ubestemt artikkel som en del av sine prediksjoner. Ikke-tellelige substantiv er tatt med ettersom disse også har kasusbøyning i polsk og det dermed antas at denne typen morsmålpåvirkning også gjelder for disse substantivfrasene. Både ekskludering av artikkelfraser og inkludering av ikke-tellelige substantiv skiller utvalget av substantivfraser fra det utvalget som ble gjort for hypotese 1. Materialet for hypotese 2 inkluderer dessuten kun *subjekter* og *objekter*, av grunner som er redegjort for i 8.1.3.1.

Prediksjonene for hypotesen i et materiale som bare består av subjekter og objekter, er at ubestemt form vil brukes ved subjekter, og bestemt form ved objekter. Dette vil bli nærmere omtalt i forbindelse med transfereffekten morsmålsamsvar lenger ned. Jeg vil under vise eksempler på substantivfraser som er subjekter og objekter. Jeg vil også kommentere plasseringen til subjektene i setningen, selv om det ikke er direkte

relevant for hypotesen. Dette for å vise at alle subjekter er inkludert i materialet uavhengig av plassering.

De to første eksemplene (10.1-a/b) viser begge subjekter som har bestemt form. Subjektet «læreren» i 10.1-a står i en hovedsetning med invertert ordstilling, altså der subjektet er plassert bak det finite verbet fordi setningen innledes av adverbialt «i Polen». Subjektet «kongen» i 10.1-b står i en leddsetning.

10.1-a I Polen har **læreren** stor autoritet. (s0652)

10.1-b I går hørte jeg på TV-en at **kongen** ble operert og at han ble lagt på en intensivavdeling. (s0756)

Frasene «biler» og «mennesker» i det neste eksemplet er subjekter for hver sin hovedsetning, men disse har ubestemt form. Den første hovedsetningen har ikke invertert ordstilling til tross for at den innledes av adverbialt «ofte», dermed står også subjektet «biler» *før* det finite verbet, noe som ikke er korrekt i norsk.

10.1-c Ofte **biler** krasjer fordi det er for trangt i boligområde og **mennesker** kjører for fort. (s0940)

Formvalget ved disse to subjektene stemmer med prediksjonen for de polske innlærerne fordi det er viser en assosiasjon mellom subjekt og ubestemt form.

De neste eksemplene viser objekter. Eksempel 10.1-d viser objektet «mennesker», som er objekt i en hovedsetning og har ubestemt form, mens eksempel 10.1-e viser leddsetningsobjektet «bilene», som har bestemt form. Av disse er det kun «bilene» som har en form som er i tråd med prediksjonen for objekter hos den polske gruppen.

10.1-d Han har hatet verden, han har hatet **mennesker**, han ville ikke snakke med noen. (s0794)

10.1-e Jeg synes at det er umulig å redusere biltrafikken i boligområdene fordi i dag nesten alle sammen eier **bilene**. (s0940)

Eksemplene over varierer med hensyn til hvilken type referanse det er snakk om; jeg vil komme tilbake til tilsvarende eksempler for indefinit, definit og generisk referanse i de aktuelle underkapitlene.

10.2 Transfereffekter

I 8.5.3 presenterte jeg hvordan de fire transfereffektene blir undersøkt generelt, og jeg vil her gjøre rede for de spesifikke detaljene knyttet til transfereffekter for hypotese 2.

Morsmålssamsvar

Transfereffekten *morsmålssamsvar* er også for hypotese 2 formulert som prediksjoner om hvilke form/kontekst-assosiasjoner den polske gruppen forventes å ha om hypotesen stemmer.⁵³ Prediksjonene for hypotese 2 ble nærmere gjort rede for i 8.1.3.1, og for et materiale som består bare av subjekter og objekter ble det predikert assosiasjoner mellom:

- 1) Ubestemte former og subjekter
- 2) Bestemte former og objekter

Det er altså forventet at det er en assosiasjon mellom subjekter og ubestemt form på den ene siden, og objekter og bestemt form på den andre. Eksempler på substantivfraser som stemmer med disse predikerte assosiasjonene, så vi på forrige side i eksempel 10.1-c for prediksjon 1, og 10.1-e for prediksjon 2.

De to assosiasjonene skal komme til uttrykk både gjennom at formen er den mest brukte i sin predikerte kontekst, og at denne formen først og fremst brukes i bare denne konteksten. Resultatene oppgis dermed også for hypotese 2 både som tall som viser prosentandelen av formen «innen kontekst» og prosentandelen av konteksten «innen form». Prediksjonene om morsmålssamsvar undersøkes både på gruppenivå gjennom undersøkelse og vurdering av tall som gjelder for gruppen som helhet, og på individnivå, gjennom opptelling og vurdering av individuelle profiler.

Ettersom kasusbøyning av substantiv i polsk skjer både ved definit, indefinit og generisk referanse, vil prediksjonene om polsk morsmålssamsvar være uavhengig av referensiell type. Transfereffekten morsmålssamsvar er derfor ventet å manifestere seg i hele det samlede materialet for hypotese 2, men som tidligere vil jeg også se på de

⁵³ Ordvalget «kontekst» kan diskuteres for hypotese 2: Subjekter og objekter er regnet som syntaktiske *funksjoner*, og det kunne derfor vært naturlig å kalle dette for «form/funksjons-assosiasjoner» i stedet, men for å gjøre delundersøkelsene i avhandlingen mer like har jeg valgt å bruke «kontekst» også for hypotese 2.

referensielle typene hver for seg for å synliggjøre transfereffektene *intergruppe-heterogenitet* og *intralingval kontrast*.

10.2.1 Intragruppe-homogenitet

Transfereffekten *intragruppe-homogenitet* vil kun undersøkes gjennom opptelling av individuelle profiler, siden statistikken på gruppenivå ikke tar hensyn til individuell variasjon. Homogenitet i den polske gruppen vil komme til syne gjennom en høy andel individuelle profiler som er positiv med hensyn til prediksjonene, noe som indikerer at et slike språklige valg er representative for gruppen.

10.2.2 Intergruppe-heterogenitet og intralingval kontrast

Som for hypotese 1 har jeg delt inn undersøkelsen av *intergruppe-heterogenitet* og *intralingval kontrast* etter referensiell type for å kunne framheve hvilke kontekster den polske gruppen er forventet å være lik og ulik de to andre gruppene (sammenfallskontekster og markørkontekster). Inndelingen i sammenfallskontekster og markørkontekster mellom den engelske og den polske gruppen baserer seg på en antakelse om at den engelske gruppen velger form i tråd med morsmålet sitt.

Tabell 10-1 (under) viser de forventede likhetene og forskjellene mellom de tre gruppene. De uthevede radene er markørkontekstene, der den polske gruppen er forventet å ha andre formvalg enn den engelske og den norske gruppen om hypotesen stemmer, mens de øvrige er sammenfallskontekster, hvor eventuelle likheter mellom gruppene ikke vil være i strid med hypotesen.

Tabell 10-1 Oversikt over forventede formvalg i de tre gruppene

	POLSK	ENGELSK	NORSK
Definit referanse			
Subjekter	Ubestemt form	Bestemt form	Bestemt form
Objekter	Bestemt form	Bestemt form	Bestemt form
Indefinit referanse			
Subjekter	Ubestemt form	Ubestemt form	Ubestemt form
Objekter	Bestemt form	Ubestemt form	Ubestemt form
Generisk referanse			
Subjekter	Ubestemt form	Ubestemt eller bestemt form	
Objekter	Bestemt form		

Mens den polske gruppen forventes å ha assosiasjoner mellom *ubestemt form* og *subjekt*, og mellom *bestemt form* og *objekt* innen alle referensielle typer, forventes den norske og engelske gruppen å ha variasjon i formvalget avhengig av om det er indefinitt, definitt eller generisk referanse.

Ved **indefinitt** referanse er det typiske formvalget i norsk og engelsk *ubestemt form*, så ved indefinitt referanse er det *subjekt* som er sammenfallskontekst, ettersom alle tre gruppene forventes å bruke ubestemt form. *Objekt* er derimot markørkontekst, siden den polske gruppen ventes å bruke bestemt form, mens de to andre gruppene forventes å bruke ubestemt.

Ved **definitt referanse** er det motsatt. Her er det typiske formvalget i norsk og engelsk bestemte former, noe som medfører at *objekt* er sammenfallskontekst, ettersom alle tre gruppene forventes å bruke bestemt form. *Subjekt* er markørkontekst siden den polske gruppen vil ha ubestemt form, mens de to andre gruppene forventes å ha bestemt form.

Som for hypotese 1 er det ikke like enkelt å presisere det ventede formvalget ved **generisk referanse** for kontrollgruppene (se 9.2.3). Under undersøkelsen av hypotese 1 kom det fram at både den engelske og den norske gruppen hadde mest ubestemt entall/flertall, men den norske mer bestemte former enn den engelske. Siden materialet for hypotese 2 baserer seg på et noe annet utvalg av substantivfraser, vil jeg komme tilbake til hvordan formvalgene for disse to gruppene er, i delkapittelet om generisk referanse (10.7.2).

10.3 Undersøkelse av hypotese 2 i hele materialet

Første del av undersøkelsen av hypotese 2 ser på prediksjonene og transfereffektene i *det samlede materialet*, altså uten å dele inn i referensielle typer. Før jeg presenterer selve resultatet av undersøkelsen, vil jeg vise noen deskriptive data for de tre gruppene hva angår andel av former og konteksttyper generelt.

10.3.1 Deskriptivt for alle gruppene

I det samlede materialet er det en del forskjell mellom de tre gruppene når det gjelder

det totale antallet enkle substantivfraser: Den polske gruppen har 372 enkle substantivfraser, den norske 305 og den engelske 262.

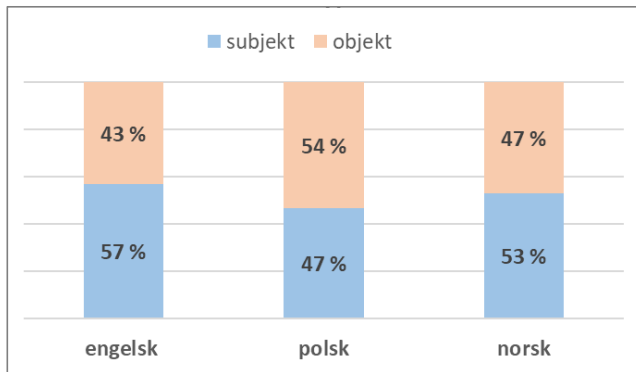
For hypotese 2 er det to formtyper (bestemt/ubestemt) og to konteksttyper (subjekt/objekt) som undersøkes. Til tross for forskjellen i totalt antall enkle substantivfraser er den innbyrdes fordelingen av subjekter/objekter og ubestemt/bestemt veldig jevn i de tre gruppene, som vi kan se av figurene og tabellene på neste side.

Figur 10-1 og tabell 10-2 (øverste neste side) viser andelen av subjekter og objekter, og både den norske og den polske gruppen har en ganske lik fordeling med cirka halvparten av hver (subj/obj: pol. 47/53 %; nor. 53/47 %), mens den engelske har en noe større andel av subjekter enn objekter (subj. 57 %; obj. 43 %). Denne forskjellen *kan* skyldes at artikkelfraser er utelatt fra materialet for hypotese 2, og henge sammen med den morsmålspåvirkningen som ble observert hos de engelske innlærerne i undersøkelsen av hypotese 1, nemlig overrepresentasjon av artikkelfraser ved indefinitt referanse (se 9.9.3.). Som vi så i forrige kapittel, var det ved indefinitt referanse stor overvekt av substantivfrasene som var plassert i slutfeltet, en setningsposisjon som er sterkt assosiert med objekter. Når artikkelfraser utelates fra materialet for hypotese 2, kan andelen av objekter derfor potensielt gå ekstra mye ned i det engelske materialet. Dette har imidlertid ikke noen problematiske konsekvenser for den videre undersøkelsen.

Figur 10-2 og tabell 10-3 (nederste neste side) viser fordelingen av bestemt og ubestemt form i gruppene. Andelen former er *helt* lik i de to innlærergruppene, som begge har en overvekt av ubestemt form (ub. 66 %; be. 34 %), og også den norske gruppen ligger nært opp til dette, men har en *litt* større andel av bestemt form (ub. 63 %; be. 37 %).

Den deskriptive gjennomgangen viser altså at det er godt belegg i alle gruppene for både formtyper og konteksttyper for hypotese 2, og at ingen av gruppene skiller seg spesielt ut med å ha en betydelig høyere eller lavere andel enn de andre.

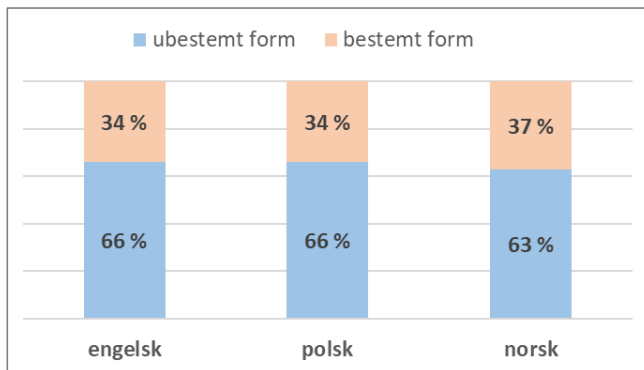
Figur 10-1 Andel av subjekt og objekt i de tre gruppene



Tabell 10-2 Andel av subjekt og objekt i de tre gruppene

Morsmål	Subjekt	Objekt	Totalt
Polsk	47 % (173)	54 % (199)	372
Norsk	53 % (161)	47 % (144)	305
Engelsk	57 % (150)	43 % (112)	262

Figur 10-2 Andel ubestemt og bestemt form i de tre gruppene



Tabell 10-3 Andel av ubestemt og bestemt form i de tre gruppene

Morsmål	Ubestemt	Bestemt	Totalt
Polsk	66 % (246)	34 % (126)	372
Norsk	63 % (192)	37 % (113)	305
Engelsk	66 % (174)	34 % (88)	262

Tabell 10-4 Polsk gruppe: Fordeling av former etter konteksttype

S1= POLSK			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	83	163	246
Andel innen form	34 %	66 %	100 %
Andel innen kontekst	48 %	82 %	66 %
BESTEMT ENT/FLT	90	36	126
Andel innen form	71 %	29 %	100 %
Andel innen kontekst	52 %	18 %	34 %
Totalt	173 (47 %)	199 (54 %)	372

Tabell 10-5 Engelsk gruppe. Fordeling av former etter konteksttype

S1= NORSK			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	89	103	192
Andel innen form	46 %	54 %	100 %
Andel innen kontekst	55 %	72 %	63 %
BESTEMT ENT/FLT	72	41	113
Andel innen form	64 %	36 %	100 %
Andel innen kontekst	45 %	29 %	37 %
Totalt	161 (53 %)	144 (47 %)	305

Tabell 10-6 Norsk gruppe: Fordeling av former etter konteksttype

S1= ENGELSK			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	82	92	174
Andel innen form	47 %	53 %	100 %
Andel innen kontekst	55 %	82 %	66 %
BESTEMT ENT/FLT	68	20	88
Andel innen form	77 %	23 %	100 %
Andel innen kontekst	45 %	18 %	34 %
Totalt	150 (57 %)	112 (43 %)	262

10.3.2 Fordeling av former i samlet materiale

Tabellene på forrige side viser fordelingen av former både *innen kontekst* og *innen form* for gruppene. Tallene vil som tidligere visualiseres underveis gjennom søylediagram.

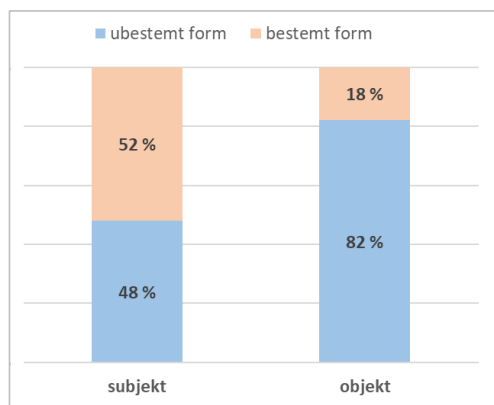
10.3.3 Morsmålssamsvar i den polske gruppen

Morsmålssamsvar i den polske gruppen ventes å komme til uttrykk gjennom for det første en assosiasjon mellom *subjekt* og *ubestemt form*, og for det andre en assosiasjon mellom *objekt* og *bestemt form*. Jeg vil først se på morsmålssamsvaret i den polske gruppen, deretter vil jeg sammenligne med den norske og den engelske gruppen.

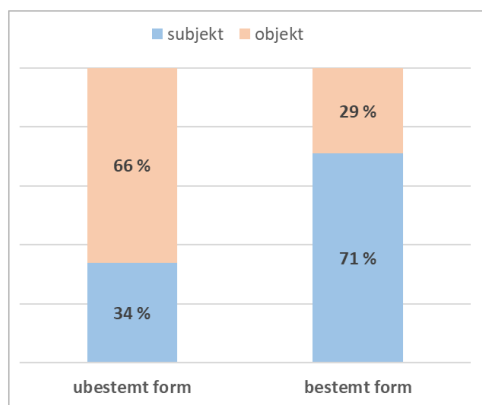
På samme måte som for hypotese 1, vil jeg se på styrken av assosiasjonen både *innen kontekst* og *innen form*. Subjekter forventes dermed å ha ubestemt form, ikke bestemt form, og ubestemt form ventes å bli brukt subjekter, ikke objekter.

Figur 10-3 (under til venstre) viser hvilke former som er brukt ved subjekter og objekter i det polske materialet, altså det jeg kaller andel av form innen kontekst, mens figur 10-4 (under til høyre) viser om formene er brukt ved subjekter eller objekter, altså det jeg kaller andel av kontekst innen form.

Figur 10-3 Andel form innen kontekst



Figur 10-4 Andel kontekst innen form



Den første delen av prediksjonen om morsmålssamsvar for den polske gruppen handler om assosiasjonen *subjekt/ubestemt*. Prediksjonen tilsier en overvekt av ubestemt form ved subjekter, men som vi ser av figur 10-3 (venstre) er det en ganske lik fordeling av ubestemt og bestemt form ved subjekter (ub. 48 %; be. 52 %). Figur 10-4 (høyre) viser dessuten at bare 34 % av alle fraser med ubestemt form er subjekter, og langt oftere

objekter (66 %). Morsmålssamsvaret er med andre ord ikke sterkt for den første delprediksjonen.

Assosiasjonen *objekt/bestemt* er visualisert i de samme to søylediagrammene på forrige side, og denne ser heller ikke ut til å være sterk for den polske gruppen. Figur 10-3 (til venstre) viser at bare 18 % av objektene har bestemt form, mens figur 10-4 (til høyre) viser at kun 29 % av alle frasene med bestemt form brukes som objekt.

I det polske materialet er det med andre ord ikke belegg for å si at det fins morsmålssamsvar i form av assosiasjoner mellom *subjekt/ubestemt* eller *objekt/bestemt*. Tvert imot kan det se ut til at forholdene er motsatt av det predikerte. Spesielt er det en langt tydeligere assosiasjon mellom ubestemt form og objekter (innen kontekst 82 %; innen form 66 %), og en tendens til assosiasjon mellom bestemt form og subjekt (innen kontekst 52 %; innen form 71 %).

10.3.4 Sammenligning med norsk og engelsk gruppe

Jeg vil nå sammenligne den polske gruppen med de to andre gruppene med hensyn til de to delprediksjonene om assosiasjon mellom 1) *subjekt og ubestemt form*, og 2) *objekt og bestemt form*. De to øverste søylediagrammene på neste side illustrerer den første av disse assosiasjonene, mens de to nederste illustrerer den andre.

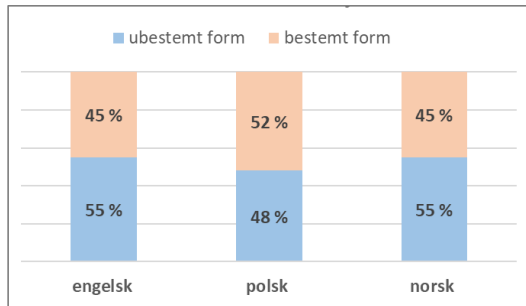
Figur 10-5 og 10-6 (øverste neste side) illustrerer forholdet mellom subjekter og ubestemt form. Det kommer fram en ganske påfallende likhet mellom de tre gruppene i figur 10-5, der alle har en ganske jevn fordeling av ubestemt og bestemt form innen subjekter, selv om den polske gruppen har en noe lavere andel ubestemt enn de andre to gruppene (pol. 48 %; eng/nor. 55 %). Videre viser figur 10-6 at frasene med ubestemt form brukes noe sjeldnere ved subjekter av den polske gruppen (pol. 34 %; eng. 53 % nor. 54 %). Jeg var inne på over at morsmålssamsvaret for delprediksjonen om subjekt og ubestemt form ikke var spesielt sterkt, og her ser vi i tillegg at den polske gruppens tendenser er relativt like det en finner hos de to andre gruppene. I den grad det er forskjeller mellom gruppene, er det knyttet til en *svakere* assosiasjon mellom subjekt og ubestemt form i den polske gruppen.

Figur 10-7 og 10-8 (nederste neste side) viser forholdet mellom objekter og bestemt form. I figur 10-7 ser vi at gruppene er like med hensyn til at de alle bruker mest *ubestemt* form ved objekter, men den norske gruppen bruker bestemt form noe oftere

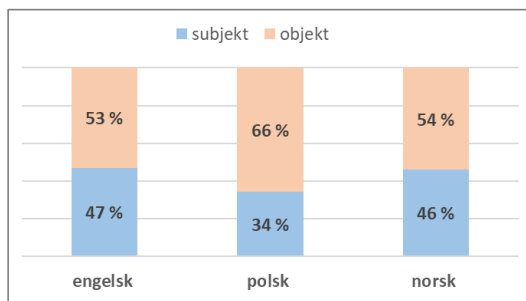
Delprediksjon 1:

subjekt og ubestemt form

Figur 10-5 Andel av form innen subjekter i gruppene



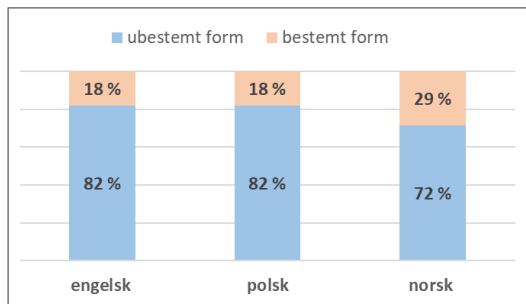
Figur 10-6 Andel av kontekst innen ubestemt form i gruppene



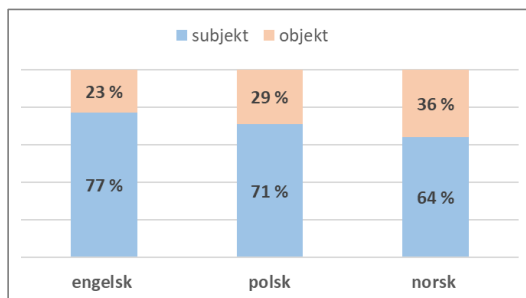
Delprediksjon 2:

objekt og bestemt form

Figur 10-7 Andel av form innen objekter i gruppene



Figur 10-8 Andel av kontekst innen bestemt form i gruppene



i slike tilfeller enn de to innlærergruppene (pol./eng. 18 %; nor. 29 %). I figur 10-8 (nederst) ser vi at den polske gruppen bruker frasene med bestemt form noe sjeldnere ved objekter enn de andre gruppene gjør (pol. 23 %; eng. 29 % nor. 36 %). Som nevnt over, er morsmålsvarsvar for prediksjonen om objekt og ubestemt form lavt, og vi ser også stor grad av homogenitet i de tre gruppernes formvalg. De mindre forskjellene mellom gruppene er knyttet til *svakere* assosiasjon mellom objekt og bestemt form i den polske gruppen.

10.3.5 Oppsummering av funn på gruppenivå

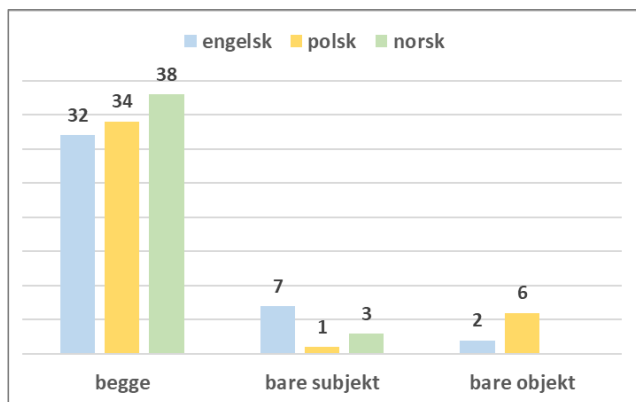
Jeg har så langt sett på morsmålsvarsvar i den polske gruppen og sammenlignet med den engelske og den norske gruppen. Konklusjonen på gruppenivå er at den polske gruppen har lavt morsmålsvarsvar, og at den ikke skiller seg nevneverdig fra den engelske og norske gruppen, bortsett fra små forskjeller på områder der den engelske og den norske gruppen faktisk viste seg å være mer i tråd med prediksjonene for polsk morsmålsvarsvar enn den polske gruppen selv var. I tillegg fant jeg relativt store likheter mellom gruppene med hensyn til å ha *motsatt* tendens av det som er predikert for den polske gruppen, altså at de heller har en assosiasjon mellom *objekt/ubestemt form* og *subjekt/bestemt form*. Her er assosiasjonen sterkere ved objekter enn subjekter.

At det er lite forskjell mellom den polske og den engelske gruppen, tyder på at dette ikke er språklige valg som er spesifikke for polske innlærere, altså er det lav intergruppe-heterogenitet, og likheten med det norske materialet tyder på at de polske innlærerne har formvalg som ligger nært opp til målspråkets. Disse forholdene blir imidlertid bedre belyst når materialet blir delt opp i ulike referensielle typer, ettersom dette bedre synliggjør de språklige kontekstene der det er forventet å være forskjeller mellom gruppene. Dette kommer jeg tilbake til etter å ha sett på individuelle profiler.

10.3.6 Individuelle profiler

I de tre alle gruppene har samtlige av tekstene én eller flere enkle substantivfraser som er subjekter eller objekter. Figur 10-9 (neste side) viser hvor mange tekster fra hver gruppe som har både subjekter og objekter, bare subjekter eller bare objekter.

Figur 10-9: Antall tekster som har subjekt og objekt



De aller fleste tekstene har både subjekter og objekter (pol. 38; eng. 32; nor. 34), men det er også noen tekster som bare har én av delene: Seks polske innlærertekster har bare objekter, syv engelske har bare subjekter, og det samme gjelder tre i den norske gruppen.

10.3.6.1 *Prosedyre for inndeling i profiler*

Som jeg gjorde rede for i 8.5.4, innebærer klassifisering av individuelle profiler først en klassifisering av hver av konteksttypene med tanke på om formvalget stemmer med prediksjonene om polsk morsmålssamsvar eller ikke (*positive/negative*), og deretter en klassifisering av hele profilen etter om den samlet sett er en *positiv*, en *negativ* eller en *kombinert* profil. Positive konteksttyper har enten *kun* den predikerte formen, eller variasjon med et flertall på mer enn to forekomster av den predikerte formen. Inndelingen i profiler for hypotese 2 er basert på følgende tre profiltyper:

Positiv profil: Formvalget stemmer med prediksjonene om polsk morsmålssamsvar både ved subjekter og objekter.

Negativ profil: Formvalget stemmer *ikke* med prediksjonene for polsk morsmålssamsvar verken ved subjekter eller objekter.

Kombinert profil: Formvalget stemmer med prediksjonene for polsk morsmålssamsvar ved enten subjekter eller objekter, men ikke ved begge.

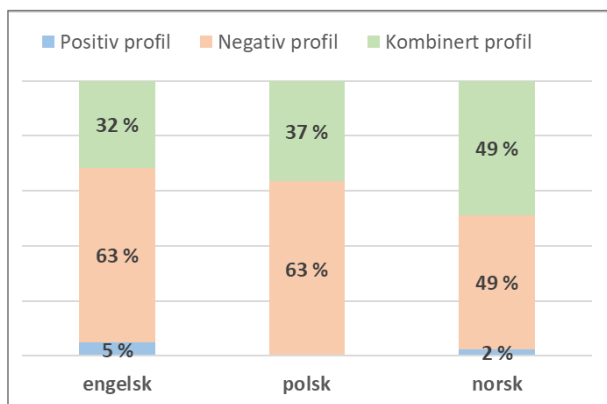
For de tekstene som bare har én av konteksttypene, altså enten bare subjekter eller bare objekter, vurderer jeg profilen etter hvilken form som er brukt der, selv om hypotesen ikke kan undersøkes i begge konteksttyper i slike tekster.

Undersøkelsen av individuelle profiler innebærer både en vurdering av andelen av positive profiler i den polske gruppen, og en sammenligning med de to andre gruppene. Som ved hypotese 1 er også den norske og den engelske gruppen klassifisert etter samme kriterier, altså etter om de er i tråd med morsmålssamsvaret for *polisk* eller ikke. Dette er for å kunne se om et eventuelt morsmålssamsvar i de polske profilene representerer noe som er spesifikt for den polske gruppen, eller ikke.

10.3.6.2 Resultat - individuelle profiler

Figur 10-10 og tabell 10-7 (under) viser antall og andel av de tre ulike profiltypene i hver av gruppene. Det er *ingen* positive profiler i den polske gruppen, men derimot 26 negative (63 %) og 15 kombinerte (37 %), noe som stemmer godt med de negative funnene på gruppenivå.

Figur 10-10 Profiltyper i de tre gruppene



Tabell 10-7 Profiltyper i de tre gruppene

Profiltipe	Engelsk	polisk	Norsk
Positiv	5 % (2)	-	2 % (1)
Negativ	63 % (26)	63 % (26)	49 % (20)
Kombinert	32 % (13)	37 % (15)	49 % (20)
Total	41	41	41

En nærmere undersøkelse av de polske profilene viser at av de 13 *kombinerte* profilene stemmer 11 med prediksjonen om *subjekt/ubestemt*, men ikke med prediksjonen om *objekt/bestemt*. Omtrent halvparten av disse profilene har kun ubestemt form, så her vurderes bare den ene delprediksjonen. De fleste av de 26 *negative* profilene har både subjekter og objekter, mens tre av dem har kun ubestemt form og objekt. Det at både subjekter og objekter er til stede i 23 av de negative profilene, gjør at de er entydig negative, ettersom prediksjonene blir testet i begge konteksttyper. Når dette blir sett i sammenheng med mangelen på positive profiler og funnene på gruppenivå, er det dermed veldig lite som tyder på at tekstene i den polske gruppen følger det predikerte morsmålssamsvaret.

Sammenlignet med de to andre gruppene er det stor likhet i fordelingen av profiler. For det første er antallet positive profiler nesten like fraværende i den engelske og den norske gruppen som i den polske (pol. 0 % eng. 5 %; nor. 2 %). Forskjeller mellom gruppene fins først og fremst i fordelingen av negative og kombinerte profiler, hvor den norske gruppen har noen færre negative profiler enn de to andre gruppene (pol./eng. 63 %; nor. 49 %), men derimot flere kombinerte (pol. 37 %; eng. 32 %; nor. 49 %). Spesielt de to innlærergruppene er veldig like, altså er det også på individnivå lite intergruppe-heterogenitet mellom polske og engelske innlærere.

Forskjellen mellom de to innlærergruppene og den norske gruppen kan være et utslag av at det er færre innlærertekster enn norske tekster som har både subjekter og objekter (pol. 34; eng. 32; nor. 38). Siden en kombinert profil forutsetter tilstedeværelse av både objekter og subjekter, er sjansen større for slike profiler i den norske gruppen.

Undersøkelsen av individuelle profiler peker altså i mer eller mindre samme retning som observasjonene på gruppenivå: lavt morsmålssamsvar i den polske gruppen, og forskjell fra både den engelske og den norske gruppen.

10.3.7 Oppsummering av funn for hypotese 2 i det samlede materialet

Tabell 10-8 (neste side) gir en oversikt over hovedfunnene fra undersøkelsen av det samlede materialet på gruppe- og individnivå. På gruppenivå har jeg i tabellen kun oppgitt prosentandel for de form/kontekst-assosiasjonene som er predikert for den polske gruppen. På individnivå har jeg slått sammen de negative og kombinerte

profilene, siden begge disse i prinsippet er negative med hensyn til prediksjonene for hypotesen.

Tabell 10-8: Oversikt over hovedfunn for hypotese 2 i det samlede materialet

GRUPPENIVÅ	Polsk	Engelsk	Norsk
Subjekt/ubestemt form			
- Innen kontekst	48 %	55 %	55 %
- Innen form	34 %	47 %	46 %
Objekt/bestemt form			
- Innen kontekst	18 %	18 %	29 %
- Innen form	29 %	23 %	36 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	0	5 %	2 %
- Negative/kombinerte	100 %	95 %	98 %

Det har så langt kommet fram en tydelig mangel på morsmålssamsvar i den polske gruppen. På gruppenivå var det kun svake assosiasjoner mellom både *subjekt og ubestemt form*, og *objekt og bestemt form*. På individnivå var det ingen positive polske profiler, men derimot kun ikke-positive. Mangel på positive profiler i den polske gruppen innebærer i tillegg mangel på *intragruppe-homogenitet*, altså kan de predikerte form/kontekst-assosiasjonene ikke regnes som representative i den polske gruppen.

Undersøkelsen av den polske gruppen isolert sett har altså ikke vist noen tegn til morsmålpåvirkning av den typen som hypotese 2 tar for seg, og heller ikke sammenligningen med de to andre gruppene bidrar til å styrke argumentet for transfer, ettersom de tre gruppene viser seg å være påfallende like både på gruppenivå og i fordelingen av profiltyper.

Målspråkslikhet

Til tross for lavt morsmålssamsvar og tilsvarende lav intragruppe-homogenitet for de predikerte form/kontekst-assosiasjonene, så var det andre forhold som virket framtreddende og representativt for den polske gruppen, nemlig en motsatt kobling mellom form og kontekst enn det som er predikert: Det viste seg å være en relativt sterk assosiasjon mellom ubestemt form og objekter (82 % innen kontekst/ 66 % innen form), og til en viss grad også mellom bestemt form og subjekter (52 % innen kontekst /

71 % innen form). De individuelle profilene var ikke klassifisert for å si direkte om denne typen assosiasjoner er representative for den polske gruppen, men et høyt antall negative profiler, som i praksis innebærer at formvalgene er motsatt av det predikerte, tyder på at disse motsatte form/kontekst-assosiasjonene kan være representative for den polske gruppen. Vi så imidlertid at denne tendensen også var til stede i den norske og den engelske gruppen, noe som tyder på at heller ikke denne typen formvalg kan knyttes til transfer, ettersom det ikke er spesifikt for polske innlærere eller ulikt det en finner i målspråket.

Undersøkelse av forskjeller mellom gruppene er bedre lagt til rette for når materialet er delt opp i referensielle typer, så jeg vil komme tilbake til funn og observasjoner fra det samlede materialet etter å ha sett på de tre referensielle delmaterialene.

10.4 Differensiert materiale – ulike referensielle typer

Jeg vil nå se på materialet delt inn etter ulike referensielle typer, noe som innebærer muligheten til å undersøke transfereffektene *intergruppe-heterogenitet* og *intralingval kontrast* på en bedre måte. Inndeling i referensielle typer bidrar til en inndeling i *sammenfallskontekster* og *markørkontekster*, noe som vil tydeliggjøre områdene for intergruppe-heterogenitet mellom de to innlærergruppene, og vise hvor prediksjonene om polsk morsmålssamsvar sammenfaller og ikke sammenfaller med målspråket.

10.5 Indefinitt referanse

Jeg vil i dette delkapittelet undersøke de fire transfereffektene i substantivfraser med indefinitt referanse. Som jeg har nevnt tidligere, skiller indefinitt referanse seg fra definitt og generisk referanse ved å vise til én eller flere individuelle referenter som ikke er identifiserbare for mottageren. Eksemplene under viser subjekter og objekter med indefinitt referanse. Det fins ingen indefinitte subjekter med bestemt form i materialet, derfor er det ingen eksempler på dette.

Det første eksemplet (10.5-a) viser to indefinitte substantivfraser som er subjekter. Dette eksemplet ble også omtalt innledningsvis i 10.1, men er gjentatt her. Begge disse frasene er vurdert som indefinitte fordi de viser til individuelle biler og mennesker som ikke er identifiserbare for mottageren av ytringen.

10.5-a Ofte **biler** krasjer fordi det er for trangt i boligområde og **mennesker** kjører for fort. (s0940)

Det neste eksemplet (10.5-b) viser et objekt med indefinitt referanse. Frasen «mobiltelefon» viser også her til en individuell referent (mobiltelefon), som ikke er identifiserbar for mottageren av ytringen, og det vises ikke generelt til mobiltelefoner som klasse, og det er derfor indefinitt referanse.

10.5-b Det var vanskelig valg til meg, fordi jeg hadde ikke så mye penger, men jeg måtte kjøpe **mobiltelefon**. (s0772)

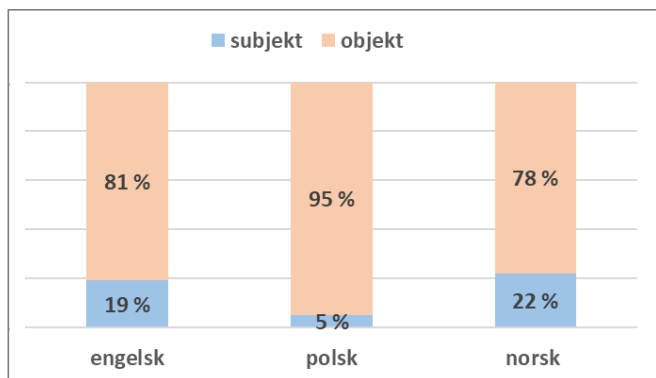
Substantivfrasene i begge eksemplene over kan regnes som svakt referensielle (se 8.4.4). Frasene «biler» og «mennesker» i 10.5-a har ikke nødvendigvis referanse til spesifikke mennesker og biler, men til hvilke som helst referenter som passer til beskrivelsen, altså er det *ikke-spesifikk* referanse. Frasen «mobiltelefon» i det andre eksemplet kan tolkes som den har fokus på type heller enn eksemplar, en annen type svak referanse. Det er nemlig snakk om en spesifikk mobiltelefon som skribenten har kjøpt, men ved å velge naken entallsform kan det i norsk tolkes som at det er mer fokus på handlingen å kjøpe mobiltelefon (type) enn det er på telefonen i seg selv, altså eksemplaret (se 7.2.4). Disse betraktningene om svak referanse i eksemplene over er det likevel knyttet forbehold med siden de er basert på en fortolkning basert på formen som er brukt, noe jeg har ønsket å unngå. Dessuten er eksemplene hentet fra innlærertekster, hvor formene kan brukes på andre måter enn det en ville gjort i norsk. Dette var jeg også inne på i 8.4.4.

Jeg vil lenger ned komme med flere eksempler på indefinitte substantivfraser, da med tanke på om formvalget er i tråd med prediksjonene for hypotese 2 eller ikke.

10.5.1 Deskriptivt for alle gruppene

Før selve gjennomgangen av funn i det indefinitte materialet vil jeg **se** på noen deskriptive tall og gjøre rede for hva som er sammenfalls- og markørkontekster i denne delen av materialet. Søylediagrammene og tabellene på neste side beskriver andelen av konteksttyper (subjekt og objekt) og formtyper (bestemt og ubestemt) i de tre gruppene.

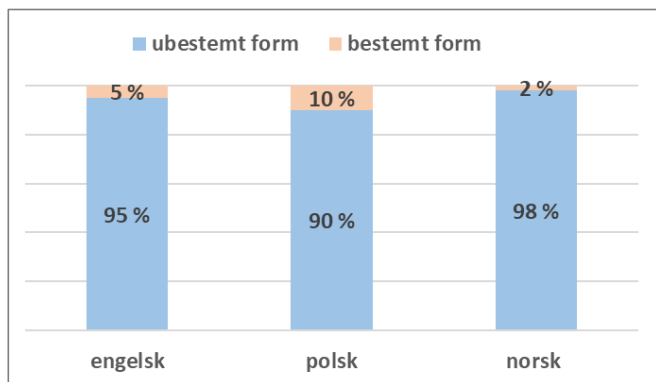
Figur 10-11 Andel av subjekt og objekt i de tre gruppene (indefinitt referanse)



Tabell 10-9 Andel av subjekt og objekt i de tre gruppene (indefinitt referanse)

Morsmål	Subjekt	Objekt	Totalt
Polsk	5 % (9)	95 % (167)	176
Norsk	22 % (27)	78 % (98)	125
Engelsk	19 % (21)	81 % (92)	113

Figur 10-12 Andel av ubestemt og bestemt form i de tre gruppene (indefinitt referanse)



Tabell 10-10 Andel av ubestemt og bestemt form i de tre gruppene (indefinitt referanse)

Morsmål	Ubestemt	Bestemt	Totalt
Polsk	90 % (158)	10 % (18)	176
Norsk	98 % (123)	2 % (2)	125
Engelsk	95 % (107)	5 % (6)	113

De tre gruppene har ulikt antall indefinitte, enkle substantivfraser. Antallet er en god del høyere i den polske gruppen enn i de to andre gruppene (pol. 176; eng. 113; nor. 125), men antallet individuelle tekster som har indefinitte fraser, er bare så vidt høyere for den polske gruppen (pol. 39; eng./nor. 37).

Figur 10-11 og tabell 10-9 (øverst forrige side) viser andelen av subjekter og objekter ved indefinitt referanse. Det er flest objekter i alle de tre gruppene, men den polske har en noe større andel enn de andre to gruppene (pol. 95 %; eng. 81 %; nor. 78 %). Andelen subjekter er tilsvarende lavere i den polske gruppen, som kun har *ni* subjekter, mens de andre gruppene begge har over 20 (pol. 5 %; eng. 19 %; nor. 22 %).

Figur 10-12 og tabell 10-10 (nederst forrige side) viser andelen av bestemte og ubestemte substantivfraser. Alle de tre gruppene har veldig høy andel ubestemt form, men noe lavere i den polske gruppen (pol. 90 %; eng. 95 %; nor. 98 %). Den norske og den engelske gruppen har videre veldig få substantivfraser med bestemt form i forhold til den polske (pol. 18 stk.; eng. 6 stk.; nor. 2 stk.), noe som er i tråd med at bestemt form generelt sett ikke brukes ved indefinitt referanse i språkene engelsk og norsk.

10.5.2 Sammenfalls- og markørkontekster for indefinitt referanse

I 10.2.2 presenterte jeg hva som er sammenfalls- og markørkontekster innen de ulike referensielle typene. Ved indefinitt referanse er det typiske formvalget i norsk og engelsk ubestemt form, og dette er samme formen som er predikert for *subjekter* hos polske innlærere, dermed er subjekter sammenfallskontekst ved indefinitt referanse. Ved objekter forventes derimot de polske innlærerne å bruke bestemt form, som vil si at objekter er markørkontekster. Tabell 10-11 (under) gir en oversikt over forventede formvalg i de tre gruppene med markørkontekstene uthevet.

Tabell 10-11 Sammenfalls- og markørkontekster for indefinitt referanse

Indefinitt referanse	POLSK	ENGELSK	NORSK
Subjekter	Ubestemt form	Ubestemt form	Ubestemt form
Objekter	Bestemt form	Ubestemt form	Ubestemt form

Med utgangspunkt i dette er det ventet at en eventuell intergruppe-heterogenitet mellom den polske og den engelske gruppen kommer til syne ved objekter, mens de to

innlærergruppene derimot forventes å ha like formvalg ved subjekter. Det samme gjelder for sammenligningen med den norske gruppen. Det er dermed morsmålssamsvar og forskjeller i formvalg ved objekter som gir sterkest støtte til et argument om transfer, siden dette vil vise at formvalgene både er gruppespesifikke og at de ikke er målspråkslike.

Som vi så i de deskriptive tallene over i 10.5.1, har alle gruppene, spesielt den polske, høy andel objekter ved indefinitt referanse (pol. 95 %; eng. 78 %; nor. 81 %). Siden det er objekter som er markørkontekster ved indefinitt referanse, er dette positivt for undersøkelsen, og den tilsvarende lave andelen av subjekter utgjør ikke et problem.

10.5.3 Samlet oversikt for indefinitt referanse

Tabellene på neste side viser fordelingen av former som er brukt for subjekter og objekter i de tre gruppene. Tallene blir også visualisert gjennom søylediagrammene som er samlet på påfølgende side.

Prediksjonene for hypotese 2 er at subjekter vil ha ubestemt form, mens objekter vil ha bestemt form. Eksemplene under viser subjekter og objekter med form som enten er i tråd med prediksjonene, eller bryter med dem.

Det første eksemplet, som også ble omtalt innledningsvis i 10.5, viser to *subjekter* som har ubestemt form og som derfor innfrir prediksjonen for hypotese 2. I det polske materialet forekommer det ingen indefinitte subjekter med bestemt form.

10.5-c Ofte **biler** krasjer fordi det er for trangt i boligområde og **mennesker** kjører for fort.

De to neste eksemplene (10.5-d/e) viser indefinitte *objekter* med henholdsvis bestemt og ubestemt form, der formvalget i 10.5-d stemmer med prediksjonen, mens formvalget i 10.5-e ikke gjør det.

10.5-d Jeg synes at det er umulig å redusere biltrafikken i boligområdene fordi i dag nesten alle sammen eier **bilene**. (s0940)

10.5-e Alle har **venner**, som er veldig gode eller gode. (s0593)

Både «bilene» og «venner» i disse to eksemplene er indefinitte fordi det er snakk om individuelle biler og venner som folk har, men som ikke er identifiserbare. Det vises heller ikke generelt til biler og venner som klasse.

Tabell 10-12 Polsk gruppe: Fordeling av former etter konteksttype (indefinit)

S1= POLSK (indefinit referanse)			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	9	149	158
Andel innen form	6 %	94 %	100 %
Andel innen kontekst	100 %	89 %	90 %
BESTEMT ENT/FLT	0	18	18
Andel innen form		100 %	100 %
Andel innen kontekst		11 %	10 %
Totalt	9 (5 %)	167 (95 %)	176

Tabell 10-13 Engelsk gruppe: Fordeling av former etter konteksttype (indefinit)

S1= ENGELSK (indefinit referanse)			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	19	88	107
Andel innen form	18 %	82 %	100 %
Andel innen kontekst	91 %	96 %	95 %
BESTEMT ENT/FLT	2	4	6
Andel innen form	33 %	67 %	100 %
Andel innen kontekst	10 %	4 %	5 %
Totalt	21 (19 %)	92 (81 %)	113

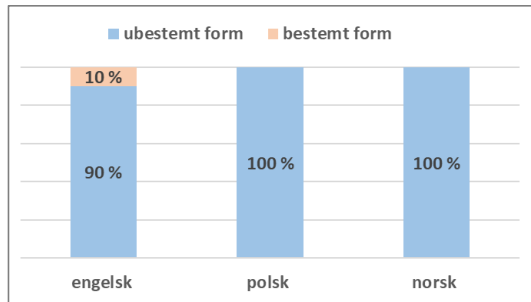
Tabell 10-14 Norsk gruppe: Fordeling av former etter konteksttype (indefinit)

S1= NORSK (indefinit referanse)			
Form: \ Kontekst:	Subjekter	Objekter	Totalt
UBESTEMT ENT/FLT	27	96	123
Andel innen form	22 %	78 %	100 %
Andel innen kontekst	100 %	98 %	98 %
BESTEMT ENT/FLT	0	2	2
Andel innen form	0 %	100 %	100 %
Andel innen kontekst	0 %	2 %	2 %
Totalt	27 (22 %)	98 (78 %)	125

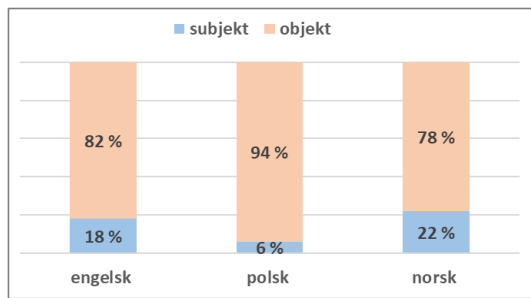
Delprediksjon 1:

subjekt og ubestemt form

Figur 10-13 Andel av form innen subjekter i gruppene



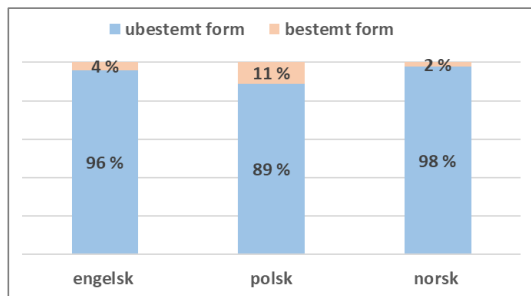
Figur 10-14 Andel av kontekst innen ubest. former i gruppene



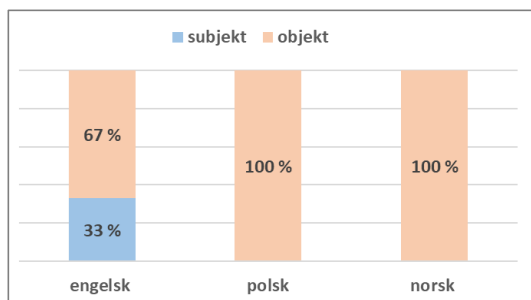
Delprediksjon 2:

objekt og bestemt form

Figur 10-15 Andel av form innen objekter i gruppene



Figur 10-16 Andel av kontekst innen bestemt form i gruppene



10.5.4 Form/kontekst-assosiasjoner på gruppenivå

De to øverste søylediagrammene på forrige side illustrerer forholdet mellom subjekter og ubestemt form. Figur 10-13 viser hvilke former som brukes ved subjekter, og vi ser at samtlige subjekter i det polske materialet har ubestemt form, noe som er helt i tråd med prediksjonene om morsmålssamsvar. Samtidig ser vi av figur 10-14, som viser hvilke kontekster fraser med ubestemt form brukes i, at bare 6 % er subjekter. For assosiasjonen *subjekt/ubestemt* er det altså veldig sterkt morsmålssamsvar innen kontekst, mens det er veldig svakt innen form siden ubestemt form stort sett brukes som objekter, ikke subjekter. Det totale antallet fraser med bestemt form er imidlertid noe lavt, bare 9 stykker.

De to nederste figurene på forrige side illustrerer forholdet mellom objekter og bestemt form, og her ser vi den motsatte tendensen: Kun 11 % av objektene har bestemt form (figur 10-15), mens samtlige fraser med bestemt form nettopp er objekter (figur 10-16). For assosiasjonen *objekt/bestemt form* er altså morsmålssamsvaret lavt innen kontekst ettersom få objekter har den predikerte formen (11 %), men høyt innen form fordi alle fraser med bestemt form er objekter, som predikert.

Dette vil si at det er et noe ujevnt morsmålssamsvar for begge delprediksjonene: For *subjekt/ubestemt* er det 100 % morsmålssamsvar innen kontekst, men veldig lavt innen form, mens det er motsatt for *objekt/bestemt form*. Det kommer også fram at det heller er snakk om en generelt utstrakt bruk av ubestemt form både ved subjekter og objekter (subj. 100 %; obj. 89 %). I den grad bestemt form brukes av de polske innlærerne, er det likevel kun ved objekter, og aldri ved subjekter, som er en interessant observasjon.

Den polske gruppens formvalg er forventet å sammenfalle med de to andre gruppene ved subjekter, mens forskjellene ventes å komme til syne ved objekter. Om vi sammenligner den polske og den norske gruppen i de samme søylediagrammene, ser vi av figur 10-13 og 10-14 (øverst) at gruppene er relativt like både når det gjelder utelukkende bruk av ubestemt form ved subjekter, og med tanke på at begge gruppene bare bruker en liten del av frasene med ubestemt form som subjekt, dog oftere i den norske gruppen enn i den polske (pol. 6 %; nor. 22 %). I figur 10-15 og 10-16 (nederst) ser vi at den norske gruppen også er lik den polske gruppen i at de fleste objekter har

ubestemt form (pol. 89 %; nor. 98 %), og i at samtlige fraser med bestemt form brukes ved *objekter*, ikke subjekter. Det er imidlertid kun to fraser med bestemt form i det norske materialet. Likheten mellom den polske og den norske gruppen er dermed stor, og tendensene til morsmålssamsvar som ble funnet i den polske gruppen, sammenfaller dermed med det en finner i den norske gruppen på alle punkter, noe som tyder på at formvalgene til den polske gruppen er mer i tråd med målspråket enn med det predikerte morsmålssamsvaret.

Forskjellene mellom den polske og den engelske gruppen er også ganske små. Når det gjelder assosiasjonen *subjekt/ubestemt form* (figur 10-13/10-14), har begge gruppene en høy andel ubestemt form ved subjekter, men det er noe mindre i den engelske gruppen (pol. 100 %; eng. 91 %). Begge gruppene bruker også bare en liten del av frasene med ubestemt form ved subjekter (pol. 6 %; eng. 18 %). Gruppene er også like for assosiasjonen *objekt/bestemt form* (figur 10-15/10-16): For det første er det kun et fåtall av objektene som har bestemt form i begge gruppene (pol. 11 %; eng. 4 %). For det andre brukes fraser med bestemt form stort sett bare som objekter i begge gruppene, men her er det likevel en kvantitativ forskjell mellom dem, der den engelske gruppen noe oftere bruker bestemt form ved subjekt (pol. 0 %, eng. 33 %). Det er med andre ord få forskjeller mellom de to innlærergruppene, og dermed liten grad av intergruppeheterogenitet.

Sammenligningen av den polske gruppen med de andre to gruppene viser altså at det er stor grad av likhet mellom dem i sammenfallskontekster (subjekter), hvor likhetene er forventet, men også i markørkontekster (objekter), hvor det er forventet at den polske gruppen skiller seg ut. Gruppene er dessuten generelt like i å ha utstrakt bruk av ubestemt form både ved subjekter og objekter, noe som er typisk for språkene engelsk og norsk ved indefinitt referanse, men som ikke er ventet i den polske gruppen om hypotesen stemmer. Den polske gruppen brukte riktignok samtlige av frasene med bestemt form som objekter, noe som er delvis i tråd med prediksjonen, men denne tendensen ser ut til å gjelde de to andre gruppene også, så heller ikke dette er spesifikt for den polske gruppen, og det er den engelske gruppen som skiller seg mest ut på dette punktet (pol. 100 %; eng. 67 %; nor. 100 %).

10.5.5 Oppsummering av funn på gruppenivå

Det sammenlagte bildet på gruppenivå er at det er lite som tyder på transfer innen indefinitt referanse, men mye som tyder på at den polske gruppen ligger nært opp til målspråkets formvalg, altså utstrakt bruk av ubestemt form. Det er et ujevnt morsmålssamsvar, og de store likhetene mellom de tre gruppene forteller oss for det første at det ikke er noe spesifikt med de polske innlærernes formvalg, ettersom engelske innlærere har samme type formvalg, og for det andre at formvalget i stor grad er målspråkslikt: Formvalg som er preget av ustrakt bruk av ubestemt form, og liten forskjell mellom subjekter og objekter. Likheten med de andre gruppene var dessuten også til stede i markørkontekster (objekter), der forskjellene er forventet å være størst. Som vi så tidligere i kapitlet, var det objekter det var mest av i dette materialet, så materialet er godt egnet for å undersøke forskjellene mellom gruppene nettopp i markørkontekster.

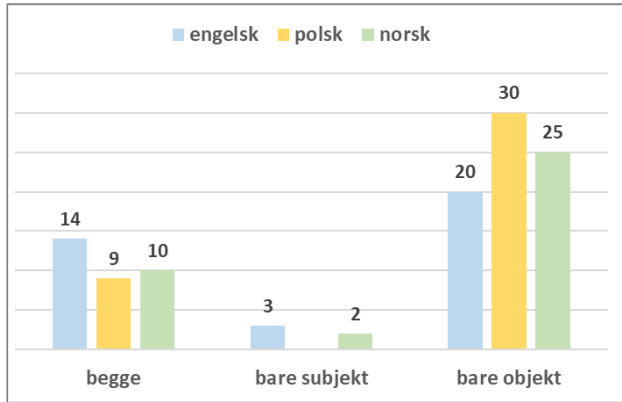
10.5.6 Individuelle profiler

Selv om det på gruppenivå er lite som tyder på transfer, vil jeg nå se på de individuelle profilene for å se om det likevel fins individuelle innlærertekster som følger de predikerte formvalgene. Gjennomgangen av de individuelle profilene vil både innebære en sammenligning av andelen av positive, negative og kombinerte profiler i de tre gruppene, og en vurdering av profilene med tanke på sammenfalls- og markørkontekster. Det er spesielt positive profiler det er interessant å studere markørkontekster i, siden en positiv profil med markørkontekster tydeligere utgjør en forskjell fra forventede formvalg for den engelske og den norske gruppen.

De aller fleste tekstene i materialet har enkle substantivfraser med indefinitt referanse (pol. 39; eng. 37; nor. 37), men ikke alle tekstene har *både* subjekter og objekter. Figur 10-17 (neste side) viser fordelingen av individuelle profiler etter om de har både subjekt og objekt, eller bare en av delene. Bare cirka en fjerdedel av tekstene i hver gruppe har både subjekter og objekter (pol. 9; eng. 10; nor. 14), og bare et fåtall har bare subjekter, mens den aller største delen av hver gruppe har bare *objekter* (pol. 30; eng. 25; nor. 20). Dette gjør at det bare er mulig å sammenligne formvalgene for både subjekt og objekt i en liten del av profilene, mens de øvrige bare har én konteksttype å vurdere. Det positive er at de fleste av disse profilene har *kun objekter*,

som er markørkontekst ved indefinitt referanse, og dermed den mest avgjørende for å undersøke forskjeller mellom gruppene.

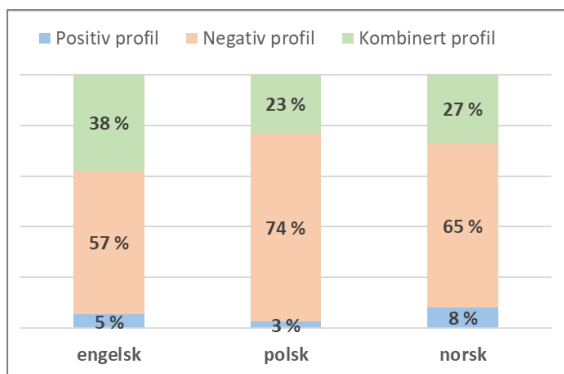
Figur 10-17 Antall tekster som har subjekt og objekt (indefinitt referanse)



10.5.6.1 Individuelle profiler sammenlignet med norsk og engelske gruppe

Figur 10-18 og tabell 10-15 (under) viser fordelingen av ulike typer profiler i de tre gruppene.

Figur 10-18 Profiltyper i de tre gruppene (indefinitt referanse)



Tabell 10-15 Profiltyper i de tre gruppene (indefinitt referanse)

Profiltipe	engelsk	Polsk	Norsk
Positiv	5 % (2)	3 % (1)	8 % (3)
Negativ	57 % (21)	74 % (29)	65 % (24)
Kombinert	38 % (14)	23 % (9)	27 % (10)
Total	37	39	37

Innen indefinitt referanse er det bare én eneste positiv profil i den polske gruppen. Dette er en profil som består av *kun* bestemt form, og *kun* objekter, altså det man kan kalle en «sterk» positiv profil, siden den er positiv i en markørkontekst. Utover dette er det snakk om hele 29 negative profiler (74 %), og ni kombinerte profiler (23 %). Alle de ni kombinerte profilene er negative for assosiasjonen mellom objekt og bestemt form (markørkontekst), men positive for assosiasjonen mellom subjekt og ubestemt form (sammenfallskontekst).

Av de 29 negative profilene har 22 *kun* ubestemt form og *kun* objekter. At disse profilene bare har ubestemt form, framhever tendensen som kom fram i gruppematerialet, nemlig at det er utstrakt bruk av ubestemt form og veldig lite bruk av bestemt form, noe som i seg selv taler imot at formene brukes til å skille mellom subjekter og objekter av de polske innlærerne. Morsmålssamsvaret er med andre ord lavt i den polske gruppen også på individnivå, og det er ikke representativt for gruppen å ha de predikerte form/kontekst-assosiasjonene, derfor har jeg ikke kunnet påvise verken morsmålssamsvar eller intragruppe-homogenitet for den polske gruppen på individnivå.

Det er heller ikke særlig forskjell mellom den polske gruppen og de andre gruppene når det gjelder individuelle profiler. Både den engelske og den norske gruppen har større andel positive profiler enn den polske gruppen, men det er snakk om relativt små forskjeller (pol. 3 %; eng. 5 %; nor. 8 %). Begge disse gruppene har også flere kombinerte profiler enn den polske gruppen (pol. 23 %; eng. 38 %; no. 27 %). De positive og kombinerte profilene i den engelske og norske gruppen, er imidlertid et logisk utfall av at det stort sett er ubestemt form som brukes. Når kun ubestemt form brukes, vil subjektssammenheng være positive og objektsammenheng negative. Om profilen bare har subjekter, blir dette nødvendigvis en positiv profil, og om profilen har både subjekter og objekter, blir profilen nødvendigvis kombinert, og dette gjelder samtlige engelske og norske kombinerte profiler.

Negative profiler er det den polske gruppen som har mest av (pol. 74 %; eng. 57 %; nor. 65 %). Dette er imidlertid *ikke* et signifikant høyere antall negative profiler enn i verken den norske gruppen ($\chi^2 = 2.6138$; $p = 0.11$), eller den engelske gruppen ($\chi^2 = 0.8109$; $p = 0.37$).

De individuelle profilene peker altså i samme retning som gruppetallene, med lite morsmålssamsvar og stor grad av homogenitet mellom gruppene. Det manglende morsmålssamsvaret gir også lite homogenitet *innad* i den polske gruppen, og undersøkelsen av markørkontekster viser at de innlærerne som har slike kontekster, stort sett ikke bruker den predikerte formen, bortsett fra én positiv profil som er i tråd med prediksjonene både i sammenfalls- og markørkontekster.

10.5.7 Oppsummering av funn for hypotese 2 ved indefinitt referanse

Tabell 10-16 (under) gir en oversikt over hovedfunnene ved indefinitt referanse både på gruppenivå og individnivå. Prosentandelene på gruppenivå viser kun de predikerte form/kontekst-assosiasjonene, og på individnivå skilles det kun mellom positive profiler og negative/kombinerte profiler samlet.

Tabell 10-16 Oversikt over hovedfunn for hypotese 2 ved indefinitt referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
Subjekt/ubestemt form			
- Innen kontekst	100 %	91 %	100 %
- Innen form	6 %	18 %	22 %
Objekt/bestemt form			
- Innen kontekst	11%	4 %	2 %
- Innen form	100 %	67 %	100 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	3 %	5 %	8 %
- Negative/kombinerte	97 %	95 %	92 %

I undersøkelsen av hypotese 2 ved indefinitt referanse er det ikke funnet argument for transfer verken på gruppenivå eller individnivå. Det er lite morsmålssamsvar, noe som kommer til uttrykk både gjennom lave gruppetall og få positive profiler, og det er tilsvarende lav intragruppe-homogenitet (representativitet) for de gjeldende prediksjonene. Når man ser dette i sammenheng med at det var små forskjeller mellom den polske og den norske gruppen også i markørkontekstene for indefinitt referanse, viser dette at de polske innlærernes formvalg også her er mer i tråd med målspåket enn med prediksjonene om polsk morsmålssamsvar. Det var dessuten lav intergruppe-heterogenitet i forhold til den engelske gruppen både på gruppe- og individnivå, som

viser at formvalgene til den polske gruppen ikke var spesifikke for denne innlærergruppen.

Det er altså stor likhet mellom funnene i det samlede materialet og i delmaterialet av indefinitte substantivfraser, og muligheten til å skille mellom sammenfallskontekster og markørkontekster har tydeliggjort at det ikke er forskjeller mellom den polske gruppen og de to andre gruppene der forskjeller er forventet om hypotesen stemmer. Mens det samlede materialet potensielt kunne ha tilslørt forskjeller mellom gruppene, viser altså det indefinitte materialet tydelig at det er store likheter mellom gruppene til tross for at polsk, engelsk og norsk impliserer ulike formvalg i visse språklige kontekster.

Målspråklighet

I det indefinitte materialet ser det heller ut til at det er ubestemt form som dominerer uavhengig av syntaktisk funksjon i alle tre gruppene, også den polske, noe som også kom fram i de deskriptive tallene innledningsvis (10.5.1). Når det skilles mellom subjekt og objekt, er andelen av ubestemt form i den polske gruppen 100 % ved subjekter og 89 % ved objekter, mens det er henholdsvis 100 % og 98 % for den norske gruppen, og 91 % og 96 % for den engelske.

I det samlede materialet så vi at det var en assosiasjon mellom objekt og ubestemt form, og til dels mellom subjekt og bestemt form, i alle tre gruppene. Det indefinitte materialet skiller seg fra dette siden ubestemt form er den mest vanlige formen *både* ved subjekter og objekter, noe som stemmer godt med typiske formvalg både i norsk og engelsk ved indefinitt referanse.

10.6 Definit referanse

I den neste delen av undersøkelsen vil jeg studere transfereffektene i substantivfraser med definit referanse. Som jeg har vært inne på tidligere, skiller definit referanse seg fra indefinitt og generisk referanse ved å vise til én eller flere individuelle referenter som er identifiserbare for tilhøreren enten gjennom anafori, situasjon, kunnskap om verden eller gjennom assosiasjon til en annen referent (se 8.4.4.). Eksemplene under viser subjekter og objekter med definit referanse.

Det første eksemplet (10.6-a) viser hvordan definit referanse etableres gjennom *anaphori*, altså at referenten til frasen «sofaen» er identifiserbar gjennom å ha vært nevnt tidligere med frasen «en sofa», og det samme gjelder referenten til «mannen».

10.6-a Jeg vet ikke hvorfor, med som i storbritannia vannligvis er det to programledere, en mann og en kvinne som sitter på en sofa og pratter. Om kvelden er det fremdeles to programledere, imidlertid **sofaen** er byttet på et bord, og **mannen** har slips. (s0328)

Eksempel 10.6-b (under) viser et subjekt i en leddsetning. Subjektet «kongen» er definit gjennom å være en del av den større situasjonen eller konteksten rundt ytringen, i dette tilfellet landet Norge: Det fins bare én konge i Norge, og denne er antatt kjent for både sender og mottager. (Se 3.1 for omtale av denne typen situasjonell identifiserbarhet.)

10.6-b I går hørte jeg på TV-en at **kongen** ble operert og at han ble lagt på en intensivavdeling. (s0756)

Det neste eksemplet viser definit referanse som etableres gjennom *assosiasjon*, der referenten «vinduene» er identifiserbar fordi det allerede er snakk om et nabolag og hjemmet til skribenten i setningene før. Vinduer er assosiert med hjem siden alle hjem normalt har vinduer.

10.6-c Det er også slitsomt med å vaske **vinduene** ofte fordi de er så mekete på grunn av trafikken. (s0921)

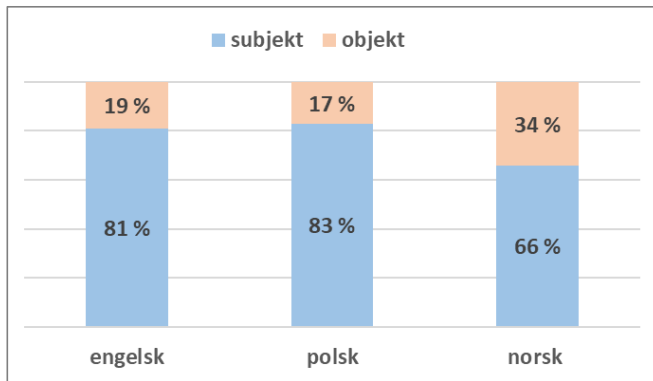
Jeg vil lenger ned komme med flere eksempler på definite fraser, da med tanke på om formvalget er i tråd med prediksjonene for hypotese 2 eller ikke.

Etter en deskriptiv oversikt over definite substantivfraser i de tre gruppene, vil jeg se på hva som er sammenfalls- og markørkontekster for definit referanse, og deretter selve resultatene.

10.6.1 Deskriptivt for alle gruppene

Søylediagrammene og tabellene på neste side viser andelen av konteksttyper (subjekt og objekt) og formtyper (bestemt og ubestemt) i de tre gruppene. Det definite materialet er noe mindre enn det indefinite, og de tre gruppene har alle rundt hundre definite substantivfraser hver, men den polske gruppen har noen flere enn de to andre gruppene (pol. 114; eng. 94; nor. 105). Til tross for at den polske gruppen har flest slike fraser og

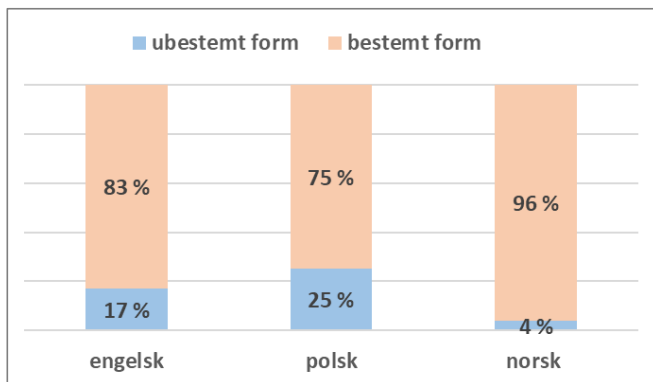
Figur 10-19 Andel av subjekt og objekt i de tre gruppene (definitt referanse)



Tabell 10-17 Andel av subjekt og objekt i de tre gruppene (definitt referanse)

Morsmål	Subjekt	Objekt	Totalt
Polsk	83 % (95)	17 % (19)	114
Norsk	66 % (69)	34 % (36)	105
Engelsk	81 % (76)	19 % (18)	94

Figur 10-20 Andel av ubestemt og bestemt form i de tre gruppene (definitt referanse)



Tabell 10-18 Andel av ubestemt og bestemt form i de tre gruppene (definitt referanse)

Morsmål	Ubestemt		Bestemt	Totalt
Polsk	25 % (29)		75 % (85)	114
Norsk	4 % (4)		96 % (101)	105
Engelsk	17 % (16)		83 % (78)	94

den engelske færrest, har den polske gruppen disse fordelt på færrest tekster (32), og den engelske på flest (38). Den norske gruppen har 36 tekster som har enkle substantivfraser med definitiv referanse.

Figur 10-19 og tabell 10-17 (øverst forrige side) viser andelen av subjekter og objekter. Det er subjekter som dominerer i alle de tre gruppene, men andelen av subjekter er en del større i innlærergruppene enn i den norske gruppen (pol. 83 %; eng. 82 %; nor. 66 %).

Figur 10-20 og tabell 10-18 (nederst forrige side) viser andelen av ubestemte og bestemte former i hver av gruppene. Alle tre gruppene har størst andel av bestemte former, men den polske gruppen har en noe lavere andel (pol. 75 %; eng. 83 %; nor. 96 %). Den norske gruppen har kun *fire* substantivfraser med ubestemt form, noe som er naturlig siden det er snakk om definitiv referanse, men begge innlærergruppene har en god del slike substantivfraser (pol. 29; eng. 16).

10.6.2 Sammenfalls- og markørkontekster for definitiv referanse

Sammenfalls- og markørkontekster for definitiv referanse ble beskrevet i 10.2.2. Ved definitiv referanse er det typiske formvalget i engelsk og norsk bestemte former, og siden den polske gruppen er predikert å ha bestemte former ved objekter, er dette sammenfallskontekst for definitiv referanse. Ved subjekter er derimot den polske gruppen predikert å ha ubestemt form, så dette er markørkontekst. Tabell 10-19 (under) viser de forventede formvalgene i de tre gruppene, med markørkonteksten (subjekt) uthevet.

Tabell 10-19 Sammenfalls- og markørkontekster for definitiv referanse

	POLSK	ENGELSK	NORSK
Definitiv referanse			
Subjekter	Ubestemt form	Bestemt form	Bestemt form
Objekter	Bestemt form	Bestemt form	Bestemt form

Med utgangspunkt i disse forskjellene og likhetene er forventningene at intergruppeheterogenitet mellom den polske og den engelske gruppen kommer til syne ved subjekter, og det samme gjelder forskjellene mellom den polske og den norske gruppen.

Som vi så i de deskriptive tallene over i 10.6.1, har alle gruppene, spesielt den polske og den engelske, en høy andel subjekter (pol. 83 %; eng. 81 %; nor. 66 %). Siden

det er subjekter som er markørkontekster ved definitt referanse, er dette positivt for undersøkelsen, og den tilsvarende lave andelen av objekter utgjør ikke et problem.

10.6.3 Samlet oversikt for definitt referanse

Tabellene på neste side viser fordelingen av former ved subjekter og objekter i de tre gruppene, og søylediagrammene på påfølgende side visualiserer de samme forholdene både innen kontekst og innen form for hver av de to delprediksjonene.

Prediksjonene for hypotese 2 er at subjekter vil ha ubestemt form, mens objekter vil ha bestemt form. Eksempelene under viser substantivfraser som er subjekter eller objekter.

De to første eksemplene viser et *subjekt* og et *objekt* med formvalg som er i tråd med prediksjonene. Subjektet «lege» i 10.6-d er definitt gjennom assosiasjon til sykehus, og har ubestemt form, noe som er i tråd med prediksjonen for subjekter. Objektet «vinduene» i 10.6-e ble også omtalt innledningsvis i 10.6, og er også definitt gjennom assosiasjon. Siden dette objektet har bestemt form, er formvalget også her i tråd med prediksjonen.

10.6-d Fire andre studenter er på sykehus men **lege** sier at de blir friske. (s0793)

10.6-e Det er også slitsomt med å vaske **vinduene** ofte fordi de er så mekete på grunn av trafikken. (s0921)

De to neste eksemplene viser et subjekt og et objekt med form som ikke stemmer med prediksjonene. Subjektet «legene» i 10.6-f er definitt ved assosiasjon til sykehus og har ubestemt form, noe som ikke er forventet for subjekter om hypotesen stemmer. Objektet «tradisjoner» i 10.6-g er definitt fordi det vises til andre menneskers spesifikke tradisjoner, som også har vært omtalt tidligere i teksten. Frasen har ubestemt form, noe som ikke stemmer med prediksjonen om at objekter skal ha bestemt form.

10.6-f Jeg har lest at kongen Harald 5 ble operert i går. Han ble operert fem og halv timmer i Rikshospital sykehuset. **Legene** sa at operasjon gikk bra. (s0745)

10.6-g Det er viktig føler jeg at man må forstå andre og hva de tenker og sier og respektere **tradisjoner** så man leve et komfortabel og stabil liv. (s0299)

Tabell 10-20 Polsk gruppe: Fordeling av former etter konteksttype (definit)

S1= POLSK (definit referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	23	6	29
Andel innen form	79 %	21 %	100 %
Andel innen kontekst	24 %	32 %	25 %
BESTEMT ENT/FLT	72	13	85
Andel innen form	85 %	15 %	100 %
Andel innen kontekst	76 %	68 %	75 %
Totalt	95 (83 %)	19 (17 %)	114

Tabell 10-21 Engelsk gruppe: Fordeling av former etter konteksttype (definit)

S1= ENGELSK (definit referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	14	2	16
Andel innen form	88 %	13 %	100 %
Andel innen kontekst	18 %	11 %	17 %
BESTEMT ENT/FLT	62	16	78
Andel innen form	80 %	21 %	100 %
Andel innen kontekst	82 %	89 %	83 %
Totalt	76 (81%)	18 (19 %)	94

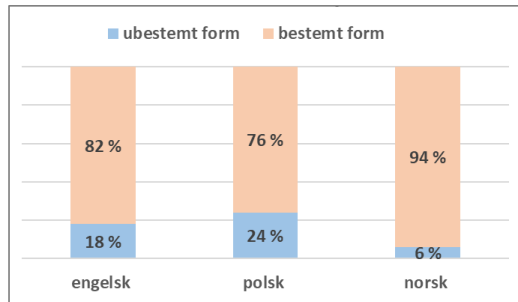
Tabell 10-22 Norsk gruppe: Fordeling av former etter konteksttype (definit)

S1= NORSK (definit referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	4	0	4
Andel innen form	100 %	0 %	100 %
Andel innen kontekst	6 %	0 %	4 %
BESTEMT ENT/FLT	65	36	101
Andel innen form	64 %	36 %	100 %
Andel innen kontekst	94 %	100 %	96 %
Totalt	69 (66 %)	36 (34 %)	105

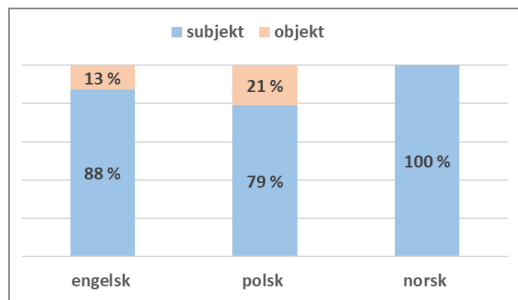
Delprediksjon 1:

subjekt og ubestemt form

Figur 10-21 Andel av form innen subjekter i gruppene



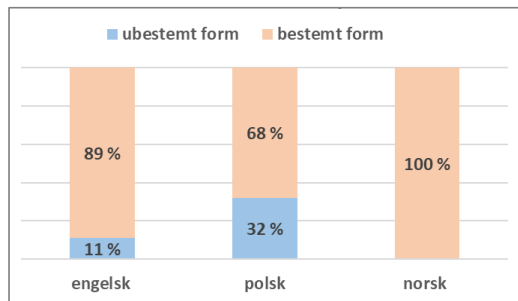
Figur 10-22 Andel av kontekst innen ubestemt form i gruppene



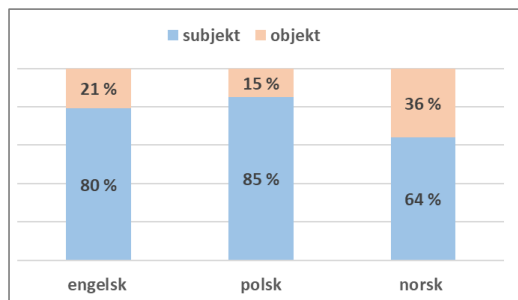
Delprediksjon 2:

objekt og bestemt form

Figur 10-23 Andel av form innen objekter i gruppene



Figur 10-24 Andel av kontekst innen bestemt form i gruppene



10.6.4 Form/kontekst-assosiasjoner på gruppenivå

De to øverste figurene på forrige side illustrerer assosiasjonen mellom *subjekt* og *ubestemt form*. Figur 10-21 viser hvilke former som brukes ved subjekter, og det er kun 24 % av subjektene som har ubestemt form, altså et relativt svakt morsmålssamsvar. Figur 10-22 viser hvilke kontekster ubestemt form brukes i, og her ser vi derimot et sterkere morsmålssamsvar, siden 79 % av frasene med ubestemt form er subjekter. For den første delprediksjonen er det dermed et ujevnt morsmålssamsvar.

De to nederste figurene på forrige side illustrerer assosiasjonen mellom *objekt* og *bestemt form*. Figur 10-23 viser at hele 68 % av alle objektene har bestemt form, mens figur 10-24 viser at av alle fraser med bestemt form er bare 15 % objekter. Vi ser altså her en lignende tendens som ved indefinitt referanse, med ujevnt morsmålssamsvar for begge delprediksjoner, men mønstret ser ut til å være motsatt i det definte materialet: *Subjekt/ubestemt form* har relativt høyt morsmålssamsvar innen form, men lavt innen kontekst, mens *objekt/bestemt form* har høyt samsvar innen kontekst, men lavt innen form.

Funnene for definitt referanse ser først og fremst ut til å være et utslag av utstrakt bruk av bestemt form både ved subjekter og objekter (subj. 76 %; obj. 68 %), heller enn valg av form i tråd med prediksjonene. Det er også veldig lik fordeling av former ved subjekter og objekter: 24 % av subjektene har ubestemt form, 32 % av objektene, men resten har bestemt form. Tilsvarende jevnhet kommer fram i form av at 21 % av frasene med ubestemt form og 15 % av frasene med bestemt form er brukt ved objekter og resten ved subjekter av den polske gruppen.

Når den polske gruppen sammenlignes med den norske, er det stor likhet mellom gruppene for begge de to predikerte form/kontekst-assosiasjonene. Av figur 10-21 og 10-22 (øverst forrige side) ser vi at i det norske materialet har bare 6 % av alle subjekter ubestemt form, mens samtlige fraser med ubestemt form er subjekter. Tilsvarende ser vi av figur 10-23 og 10-24 (nederst forrige side) at samtlige objekter har bestemt form, mens bare 36 % av frasene med bestemt form er objekter. Dette er i tråd med at bestemt form er det typiske valget i norsk ved definitt referanse. At gruppene er like både i sammenfallskonteksten (objekt) og markørkonteksten (subjekt), viser at den polske

gruppen i større grad har formvalg som er i tråd med målspråket enn med prediksjonene for morsmålssamsvar.

En tilsvarende likhet foreligger mellom den polske og den engelske gruppen. Den engelske gruppen har tilsvarende ujevne fordeling som den polske gruppen både for *subjekt/ubestemt* (12 % innen kontekst; 88 % innen form), og *objekt/bestemt* (89 % innen kontekst; 21 % innen form).

Det noe ujevne morsmålssamsvaret i den polske gruppen er altså utfylt av store likheter med de to andre gruppene, noe som viser at formvalgene i den polske gruppen ikke stemmer godt med det predikerte morsmålssamsvaret, og i den grad den gjør det, er det ikke spesifikt for den polske gruppen verken i forhold til den engelske eller den norske gruppen. At gruppene er like for objekter, er som forventet siden dette er en sammenfallskontekst, men om hypotese 2 stemmer, skal det foreligge forskjeller ved subjekter (markørkontekst), noe som ikke er tilfelle. At den polske gruppen ligner de andre to, svekker dermed argumentet for transfer innen definit referanse på gruppenivå. Det som derimot er felles for alle tre gruppene, er at de har stor andel av bestemt form ved definit referanse, uavhengig av om det er subjekt eller objekt, noe som er naturlig i norsk, og som også kan forventes hos engelske innlærere om de er påvirket av morsmålet.

10.6.5 Oppsummering av funn på gruppenivå

På gruppenivå ser det altså ut til at ubestemt og bestemt form ikke brukes til å skille mellom subjekter og objekter, slik hypotesen tilsier, og den polske gruppen skiller seg lite fra de andre to gruppene. Tendensen i det definte materialet er heller lik den som kom fram i gjennomgangen av det indefinte delmaterialet: Den polske gruppen bruker i stor grad den typiske formen for målspråket både ved objekter og subjekter, i dette tilfellet bestemt form. Den polske gruppen kan altså se ut til å være mer i tråd med målspråket enn med prediksjonene om morsmålssamsvar også innen definit referanse. Jeg vil nå se på de individuelle profilene.

10.6.6 Individuelle profiler

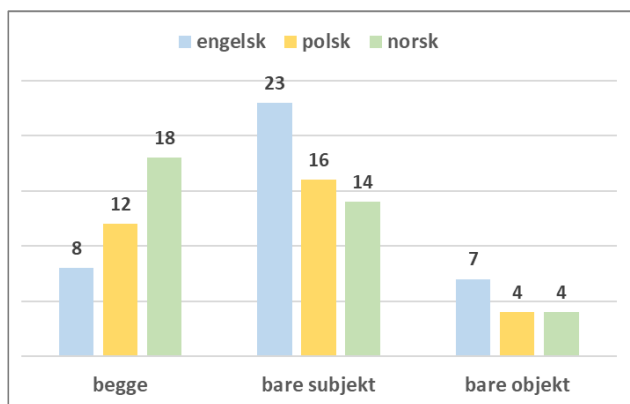
Det er til sammen 32 polske innlærertekster som har definte enkle substantivfraser, mens det er 38 i den engelske gruppen og 36 i den norske. Figur 10-25 (neste side) viser

hvor mange tekster som har både subjekt og objekt, og hvor mange som har enten bare subjekt eller objekt.

Den polske gruppen har 12 tekster med både subjekt og objekt, mens den engelske og den norske gruppen har henholdsvis 8 og 18, altså er det en viss forskjell mellom gruppene på dette området. Når det gjelder tekster som enten bare har subjekt eller bare har objekt, er det mest utbredt å bare ha subjekt i alle tre gruppene (pol. 16; eng. 23 nor. 14), mens bare noen få tekster kun har objekt (pol. 4; eng. 7; nor. 4).

Her ser vi altså en lignende tendens som for indefinitt referanse, bare motsatt. Mens det var stor andel profiler med bare objekter ved indefinitt referanse, er det profiler med bare *subjekter* som dominerer ved definitt referanse. I likhet med for indefinitt referanse er altså markørkonteksten, i dette tilfellet subjekt, godt representert i materialet. Dette er en fordel for den videre undersøkelsen ettersom det først og fremst er ved subjekter at eventuelle forskjeller mellom den polske gruppen og de andre to gruppene er ventet å komme til syne om hypotesen stemmer.

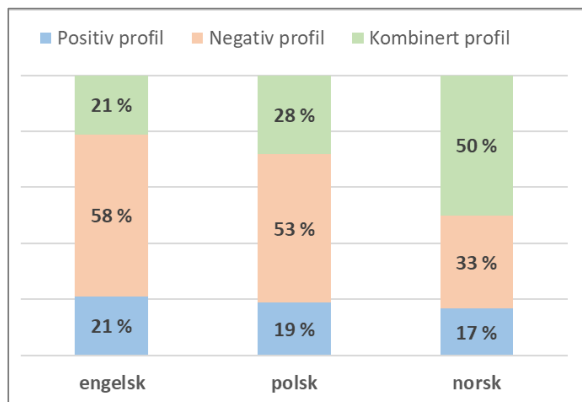
Figur 10-25 Antall tekster som har subjekt og objekt (definitt referanse)



10.6.6.1 Resultat - individuelle profiler

Tabell 10-23 og figur 10-26 (neste side) viser antall og andel av de ulike profiltypene i hver av gruppene. Det definte materialet har flere positive individuelle profiler i den polske gruppen enn det indefinte, men det er likevel bare 19 % av innlærerne (6 stk.) som velger form i tråd med prediksjonene, mens 53 % har en negativ profil og 28 % en kombinert.

Figur 10-26 Profiltyper i de tre gruppene (definitt referanse)



Tabell 10-23 Profiltyper i de tre gruppene (definitt referanse)

Profiltipe	Engelsk	Polsk	Norsk
Positiv	21 % (8)	19 % (6)	17 % (6)
Negativ	58 % (22)	53 % (17)	33 % (12)
Kombinert	21 % (8)	28 % (9)	50 % (18)
Total	38	32	36

De positive profilene står sterkere om de har markørkontekster (subjekter), og av de seks positive profilene har faktisk *fem* markørkontekster. Dette styrker de positive profilene ettersom de har det predikerte formvalget i kontekster som ikke sammenfaller med målspråkets formvalg, noe som er et positivt funn med tanke på morsmålssamsvar.

Av de ni kombinerte profilene i det polske materialet er åtte positive for assosiasjonen *objekt/bestemt form*, mens bare én er positiv for assosiasjonen *subjekt/ubestemt form*. De kombinerte profilene er dermed stort sett positive i sammenfallskontekster, og bare én er det i en markørkontekst.

Andelen av positive profiler i den polske gruppen ligger ganske jevnt med den engelske og den norske gruppen (pol. 19 %; eng. 17 %; nor. 21 %), altså er det ingen tydelige forskjeller mellom gruppene her.⁵⁴ Det viser seg likevel å være forskjeller når en undersøker hvilke form- og konteksttyper som er representert i disse profilene: De aller fleste av de engelske og norske positive profilene har *kun* bestemt form og objekter

⁵⁴ Her er gruppene så synlig like at jeg ikke har utført noen statistisk test.

(eng. 5; nor. 4), mens færre av dem har kun subjekt og ubestemt form (eng. 3; nor. 2), noe som vil si at de fleste positive profilene i disse to gruppene avspeiler utstrakt bruk av bestemt form generelt. Når det bare er brukt bestemt form, vil nemlig profiler som kun har objekter, nødvendigvis være positive. De fem polske profilene har derimot, som nevnt, også positive markørkontekster.

De kombinerte profilene i det polske materialet er i stor grad profiler som kun har bestemt form, en tendens en også finner i de to andre gruppene. Når kun bestemt form brukes, vil subjektene nødvendigvis være negative kontekster og objekter positive, noe som samlet gir en kombinert profil. Som vi så i de deskriptive tallene har den norske gruppen flere individuelle profiler med både subjekter og objekter enn de andre gruppene, og kombinert med kun bestemt form vil dette gi et høyt antall kombinerte profiler. Dette kan forklare hvorfor den norske gruppen har flere kombinerte profiler enn de andre to gruppene (pol. 28 %; eng. 21 %; nor. 59 %).

Konklusjonen basert på de individuelle profilene er dermed at det er få positive profiler i den polske gruppen, og at den polske gruppen skiller seg lite fra de andre to rent kvantitativt, i og med at det ikke er signifikante forskjeller i antall positive profiler i de tre gruppene. Det ble likevel funnet en mer kvalitativ forskjell mellom gruppene: Mens de engelske og norske positive profilene så ut til være en nødvendig konsekvens av en kombinasjon av høy andel bestemte former og mangel på markørkontekster, så hadde fem av de seks polske positive profilene faktisk markørkontekster. Dette settet med fem positive profiler i det polske materialet stemmer altså alle overens med prediksjonene for hypotese 2, i tillegg til å skille seg fra de positive profilene i de andre gruppene.

10.6.7 Oppsummering av funn for hypotese 2 ved definitte referanse

Tabell 10-24 (neste side) gir en oversikt over hovedfunnene på gruppenivå og individnivå i det definte materialet. På gruppenivå er det kun prosentandeler for de predikerte form/kontekst-assosiasjonene som er oppgitt, og på individnivå vises andel av positive versus negative/kombinerte profiler.

Tabell 10-24 Oversikt over hovedfunn for hypotese 2 ved definitivt referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
Subjekt/ubestemt form			
- Innen kontekst	24 %	18 %	6 %
- Innen form	79 %	88 %	100 %
Objekt/bestemt form			
- Innen kontekst	68 %	89 %	100 %
- Innen form	15 %	21 %	36 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	19 %	21 %	17 %
- Negative/kombinerte	81 %	79 %	83 %

Morsmålssamsvaret på gruppenivå var ujevnt for begge delprediksjonene, med høyt morsmålssamsvar innen *enten* kontekst eller form, men ikke innen begge deler. Det viste seg imidlertid at den polske gruppen var relativt lik både den engelske og den norske på dette området, både innen sammenfallskontekster, hvor likhetene var forventet, og innen markørkontekster, hvor det var forventet forskjeller.

Også på individnivå var det små forskjeller mellom gruppene, ca. 20 % av profilene var positive i alle de tre gruppene, mens de resterende var enten negative eller kombinerte. Undersøkelsen av markørkontekster avslørte imidlertid en kvalitativ forskjell mellom fem av de seks polske positive profilene og de positive profilene i den engelske og den norske gruppen: Mens de positive engelske og norske profilene i stor grad kun hadde bestemt form og objekter, noe som nødvendigvis gir en positiv profil, hadde de fem polske profilene alle markørkontekster. Disse fem profilene hadde altså morsmålssamsvar og de skilte seg fra profilene i den engelske og norske gruppen. Dette vil si at det for disse profilene er mulig å påvise tre av transfereffektene, både morsmålssamsvar, intergruppe-heterogenitet og intralingval kontrast, men det lave antallet av slike profiler tilsier ikke at dette er representativt for den polske gruppen, og dermed er det ikke godt grunnlag for den siste transfereffekten (intragruppe-homogenitet).

I det definitive materialet kan det se ut til at det er bestemt form som dominerer uavhengig av syntaktisk funksjon i alle tre gruppene, også den polske. Dette er en tendens som kom til syne allerede i den deskriptive redegjørelsen innledningsvis

(10.6.1.), der vi så at den polske gruppen bruker bestemt form relativt ofte både ved subjekter og objekter (subj. 76 %; obj. 68 %), noe som er likt med både den engelske gruppen (subj. 82 %; obj. 89 %), og den norske gruppen (subj. 94 %; obj. 100 %). Det er likevel en lavere andel av bestemt form i den polske gruppen, og noe lavere ved objekter enn ved subjekter.

Det kan se ut til at funnene ved definit referanse utfyller funnene fra indefinit referanse med tanke på å skape det mønsteret som ble observert i det samlede materialet: I det samlede materialet ble det observert tendenser til assosiasjon mellom *objekt/ubestemt form*, og også en viss assosiasjon mellom *subjekt/bestemt form*, mens det i det indefinite materialet først og fremst var bruk av *ubestemt form* uavhengig av syntaktisk funksjon. I det definite materialet kom det derimot fram utstrakt bruk av *bestemt form* uavhengig av syntaktisk funksjon. Samlet kan altså funnene ved indefinit og definit referanse bidra til å forklare det mønsteret som ble funnet i det samlede materiale, noe som illustrerer nytten av å dele opp i referensielle typer.

Funnene på gruppe- og individnivå ved definit referanse gir altså ingen sterke argumenter for transfer hos den polske gruppen. Morsmålssamsvaret er ujevnt og har en tendens til å sammenfalle med formvalg i begge de to andre gruppene, som vi så at det er lite intergruppe-heterogenitet og forskjell fra målspråket. Det er i tillegg lav intragruppe-homogenitet, men likevel fem polske innlærere som ser ut til å ha form/kontekst-assosiasjoner av den typen som hypotese 2 predikerer.

10.7 Generisk referanse

Den siste av de referensielle typene som skal undersøkes, er generisk referanse. Generisk referanse står i kontrast til indefinit referanse og definit referanse ved at det ikke refereres til én eller flere individuelle referenter, men heller til en hel klasse av referenter. (Se 8.4.4). Eksemplene under viser subjekter og objekter med generisk referanse.

Det første eksemplet (10.7-a) viser subjektet «læreren», som er generisk fordi det refereres generelt yrkesgruppen lærere, eller til det generelle begrepet «lærer», ikke til individuelle lærere. I 3.2.1 var jeg inne på at generisk referanse noen ganger kan innebære referanse til *alle* medlemmer i en klasse, mens det andre ganger er referanse til klassen i seg selv, men ikke nødvendigvis til alle medlemmer av klassen.

Substantivfrasen «læreren» er av den siste typen, siden det ikke nødvendigvis gjelder for alle individuelle polske lærere at de har stor autoritet.

10.7-a I Polen har **læreren** stor autoritet. (s0652)

Det neste eksemplet er et objekt, og «mennesker» er her generisk fordi det vises generelt til alle mennesker, ikke til individuelle mennesker.

10.7-b Han har hatet verden, han har hatet **mennesker**, han ville ikke snakke med noen. (s0794)

Jeg vil lenger ned komme med flere eksempler på generiske substantivfraser, da med tanke på om formvalget er i tråd med prediksjonene for hypotese 2 eller ikke.

Gjennomgangen av det generiske materialet vil skje på samme måte som for de to andre referensielle typene, bortsett fra at det ikke er framsatt noen klare sammenfalls- og markørkontekster på forhånd. I likhet med for hypotese 1 vil jeg basere dette på de faktiske forekomstene i den norske og engelske gruppen, og dette vil gjøres rede for i 10.7.2, etter en deskriptiv presentasjon av det generiske materialet.

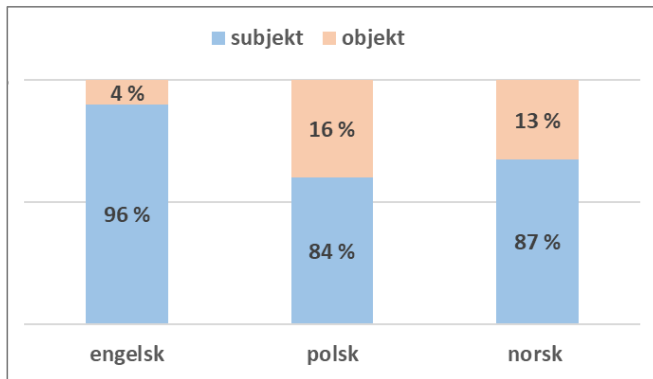
10.7.1 Deskriptivt for alle gruppene

Søylediagrammene og tabellene på neste side viser andelen av konteksttyper (subjekt og objekt) og formtyper (bestemt og ubestemt) i de tre gruppene.

Det er en del forskjell mellom gruppene når det gjelder antall enkle substantivfraser med generisk referanse, spesielt har den engelske gruppen en del færre slike fraser enn de andre to gruppene, mens den polske gruppen har flest (pol. 82; eng. 55; nor. 75). Det lave antallet i den engelske gruppen avspeiler nok at det bare er 22 tekster som har enkle substantivfraser med generisk referanse, mens det gjelder henholdsvis 28 og 31 tekster i den polske og den norske gruppen.

Figur 10-27 og tabell 10-25 (øverst neste side) viser andelen av subjekter og objekter, og det viser seg at gruppene er relativt like i den forstand at de alle har mest subjekter. Spesielt like er den polske og den norske gruppen, mens den engelske har en noe høyere andel av subjekter enn disse (pol. 84 %; eng. 96 %; nor. 87 %). Andelen av objekter er tilsvarende lavt i de tre gruppene, i antall er det snakk om 13 objekter for den polske gruppen, kun to for den engelske og 10 for den norske (pol. 16 %; eng. 13 %, nor. 4 %).

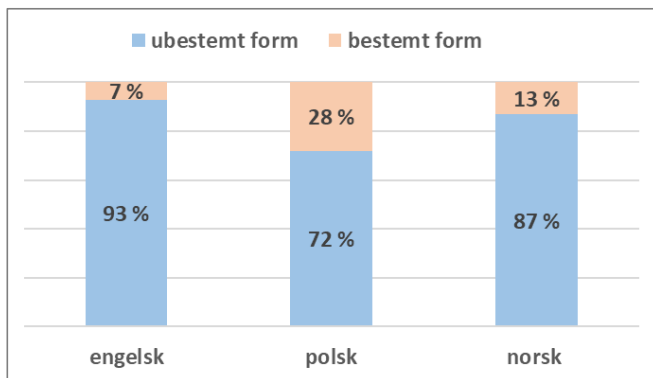
Figur 10-27 Andel av subjekt og objekt i de tre gruppene (generisk referanse)



Tabell 10-25 Andel av subjekt og objekt i de tre gruppene (generisk referanse)

Morsmål	Subjekt	Objekt	Totalt
Polsk	84 % (69)	16 % (13)	82
Norsk	87 % (65)	13 % (10)	75
Engelsk	96 % (53)	4 % (2)	55

Figur 10-28 Andel av ubestemt og bestemt form i de tre gruppene (generisk referanse)



Tabell 10-26 Andel av ubestemt og bestemt form i de tre gruppene (generisk referanse)

Morsmål	Ubestemt	Bestemt	Totalt
Polsk	72 % (59)	28 % (23)	82
Norsk	87 % (65)	13 % (10)	75
Engelsk	93 % (51)	7 % (4)	55

Figur 10-28 og tabell 10-26 (nederst forrige side) viser fordelingen av ubestemt og bestemt form i de tre gruppene. Alle gruppene har en overvekt av ubestemt form, men det er tydelig mindre i den polske gruppen (pol. 72 %; eng. 87 %; nor. 93 %). Videre er det den polske gruppen som har flest fraser med bestemt form, med 23 stykker, mens den engelske gruppen har bare fire og den norske 10.

Det generiske materialet består altså i stor grad av subjekter og substantiv med ubestemt form, og det er spesielt lite objekter og bestemte substantiv i det engelske materialet. Dette kan gjøre sammenligningen av formvalg ved subjekter og objekter mer problematisk enn for de andre referensielle typene, noe jeg vil komme tilbake til under.

10.7.2 Sammenfalls- og markørkontekster for generisk referanse

Som jeg var inne på i undersøkelsen av hypotese 1, er det ved generisk referanse vanskelig å presisere nøyaktig hvor prediksjonene om polsk morsmålsamsvar vil sammenfalle og ikke sammenfalle med formvalg i de andre to gruppene. Dette er fordi valg av form i engelsk og norsk er mindre fast enn ved definitiv og indefinitiv referanse, men som vi så ved hypotese 1, var den mest valgte formen både i den norske og den engelske gruppen ubestemt entall/flertall, selv om den norske gruppen hadde noe mer bruk av bestemt form enn den engelske (se 9.7.1). Siden materialet som de to hypotesene undersøkes i, bare delvis overlapper, er ikke tallene for hypotese 1 helt overførbare til denne delen av undersøkelsen, og jeg vil derfor basere vurderingen av sammenfalls- og markørkontekster på de deskriptive tallene som ble presentert over (10.7.1.). Der så vi at også i materialet for hypotese 2 har både den engelske og den norske gruppen først og fremst brukt ubestemt form, men norsk har en noe større andel av bestemt form enn den engelske (eng. 7 %; nor. 13 %) (tabell 10-26). Forskjellen i andel bestemt form er imidlertid noe mindre enn den var i materialet for hypotese 1 (eng. 10 %; nor. 23 %).

Siden andelen av bestemt form er såpass lav i både det engelske og det norske materialet, vil jeg ha som utgangspunkt for vurderingen av sammenfalls- og markørkontekster, at begge disse gruppene først og fremst bruker ubestemt form både ved subjekter og objekter. Siden den polske gruppen er ventet å velge ubestemt form ved subjekter, er dette sammenfallskontekst, mens objekter er markørkontekst ettersom den polske gruppen der er ventet å velge bestemt form. Tabell 10-27 på neste side viser de forventede formvalgene i de tre gruppene, og markørkonteksten er uthevet.

Tabell 10-27 Sammenfalls- og markørkontekster for generisk referanse

	POLSK	ENGELSK	NORSK
Generisk referanse			
Subjekter	Ubestemt form	Ubestemt form	Ubestemt form
Objekter	Bestemt form	Ubestemt form	Ubestemt form

Som vi så lenger opp, er de aller fleste generiske substantivfrasene *subjekter* for alle tre gruppene, spesielt den engelske gruppen har veldig få objekter (pol. 13; eng. 2; nor. 10) (tabell 10-25, s. 385). Dette betyr at materialet først og fremst inneholder sammenfallskontekster, altså subjekter, noe som begrenser muligheten til å bruke markørkontekster til å vise forskjeller mellom gruppene. Det er spesielt intergruppeheterogenitet som er vanskelig å undersøke, siden den engelske gruppen har så få objekter. Det er likevel så mye som 13 objekter i det polske materialet, noe som gir et visst utgangspunkt for undersøkelse av markørkontekster.

10.7.3 Samlet oversikt for generisk referanse

Tabellene på neste side viser fordelingen av former ved subjekter og objekter i de tre gruppene, og dette blir visualisert ved hjelp av søylediagrammer i gjennomgangen.

Prediksjonene for hypotese 2 er også ved generisk referanse at subjekter vil ha ubestemt form, mens objekter vil ha bestemt form. De følgende eksemplene viser subjekter og objekter med form som enten er i tråd med prediksjonene, eller bryter med dem.

De to første eksemplene viser et subjekt og et objekt med form som stemmer med prediksjonene. Subjektet «mobiltelefoner» i 10.7-c er generisk fordi det viser generelt til alle mobiltelefoner, ikke til individuelle mobiltelefoner, og substantivet har ubestemt form, noe som er forventet ved subjekter. Objektet «barna» i 10.7-d er generisk fordi det her vises generelt til å det å oppdra barn, ikke til spesifikke, individuelle barn, og siden det her er brukt bestemt form, stemmer formvalget med prediksjonen for objekter.

10.7-c Forskerne tror at **mobiltelefoner** kan bli årsaken til den alvorlige sykdommen som kreft. (s0770)

10.7-d Det er ikke lett å oppdra **barna**, det er våre vanskeligste oppgave i livet vårt. (s0078)

Tabell 10-28 Polsk gruppe: Fordeling av former etter konteksttype (generisk)

S1= POLSK (generisk referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	51	8	59
Andel innen form	86 %	14 %	100 %
Andel innen kontekst	74 %	62 %	72 %
BESTEMT ENT/FLT	18	5	23
Andel innen form	78 %	22 %	100 %
Andel innen kontekst	26 %	39 %	28 %
Totalt	69 (84 %)	13 (16 %)	82

Tabell 10-29 Engelsk gruppe: Fordeling av former etter konteksttype (generisk)

S1= ENGELSK (generisk referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	49	2	51
Andel innen form	96 %	4 %	100 %
Andel innen kontekst	93 %	100 %	93 %
BESTEMT ENT/FLT	4	0	4
Andel innen form	100 %		100 %
Andel innen kontekst	8 %		7 %
Totalt	53 (96 %)	2 (4 %)	55

Tabell 10-30 Norsk gruppe: Fordeling av former etter konteksttype (generisk)

S1= NORSK (generisk referanse)			
Kontekst:	Subjekter	Objekter	Totalt
Form:			
UBESTEMT ENT/FLT	58	7	65
Andel innen form	89 %	11 %	100 %
Andel innen kontekst	89 %	70 %	87 %
BESTEMT ENT/FLT	7	3	10
Andel innen form	70 %	30 %	100 %
Andel innen kontekst	11 %	30 %	13 %
Totalt	65 (87 %)	10 (13 %)	75

De neste to eksemplene viser formvalg som bryter med prediksjonene. Begge disse eksemplene ble omtalt og vurdert som generiske innledningsvis (10.7). Subjektet «læreren» i 10.7-e har bestemt form, mens objektet «mennesker» i 10.7-f har ubestemt form, noe som ikke er i tråd med prediksjonene.

10.7-e I Polen har læreren stor autoritet. (s0652)

10.7-f Han har hatet verden, han har hatet mennesker, han ville ikke snakke med noen. (s0794)

10.7.4 Form/kontekst-assosiasjoner på gruppenivå

De to øverste figurene på neste side illustrerer assosiasjonen mellom *subjekt* og *ubestemt form*. Figur 10-29 viser at 74 % av subjektene har ubestemt form i de polske tekstene, mens figur 10-30 viser at 86 % av alle fraser med ubestemt form er subjekter. Dette vil si at morsmålssamsvaret er ganske høyt for assosiasjonen *subjekt/ubestemt form*.

De to nederste figurene på neste side illustrerer assosiasjonen mellom *objekt* og *bestemt form*, og for denne er morsmålssamsvaret lavt for den polske gruppen: Av alle objekter har 39 % bestemt form (figur 10-31), mens kun 22 % av frasene med bestemt form er objekter (figur 10-32). Selv om det er litt mer bruk av bestemt form ved objekter, så er andelen bestemt form lav både ved subjekter og objekter i det polske materialet (subj. 26 %; obj. 39 %). I likhet med i det indefinitte materialet, er det altså også ved generisk referanse ubestemt form som dominerer hos de polske innlærerne.

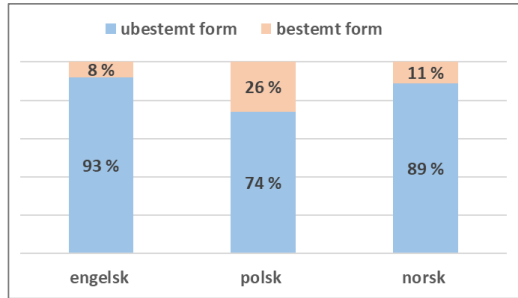
Om vi sammenligner den polske og den norske gruppen i de samme søylediagrammene, ser vi at den norske gruppen har tilsvarende tendens som den polske for begge assosiasjonene: Sterk assosiasjon for *subjekt/ubestemt* (89 % innen kontekst/89 % innen form), men svak for assosiasjonen *objekt/bestemt* (30 % innen kontekst; 30 % innen form). I den grad det er forskjell mellom gruppene, er det at den polske gruppen har en *lavere* andel av ubestemt ved subjekt enn den norske gruppen. Likheten med den norske gruppen, i tillegg til at det er høyest morsmålssamsvar i sammenfallskonteksten (subjekt) og lavest i markørkonteksten (objekt), gjør at den polske gruppen framstår som mer i tråd med målspråket enn med prediksjonene om morsmålssamsvar.

Når den engelske og den polske gruppen sammenlignes, er forskjellene derimot større. Den engelske gruppen har sterkere assosiasjon mellom subjekt og ubestemt form

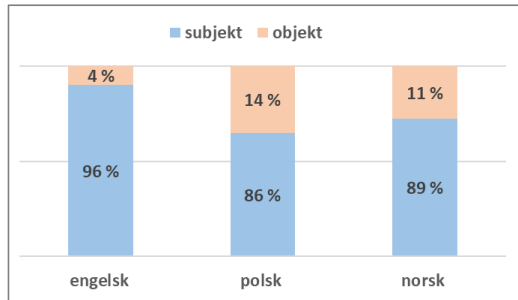
Delprediksjon 1:

subjekt og ubestemte former

Figur 10-29 Andel av form innen subjekter i gruppene



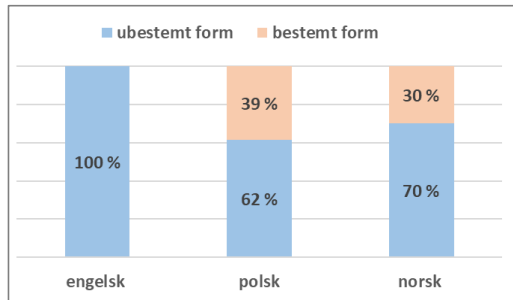
Figur 10-30 Andel av kontekst innen ubestemt form i gruppene



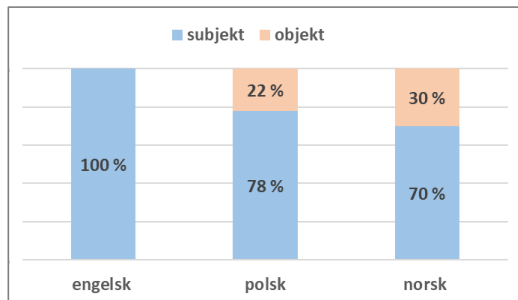
Delprediksjon 2:

objekt og bestemte former

Figur 10-31 Andel av form innen objekter i gruppene



Figur 10-32 Andel av kontekst innen bestemt form i gruppene



enn den polske har (93 % innen kontekst; 96 % innen form), og *ingen* assosiasjon mellom objekt og bestemt form (0 % innen kontekst; 0 % innen form). Forskjellen mellom de to gruppene er størst for assosiasjonen *objekt/bestemt*, men den engelske gruppen har som nevnt tidligere, bare to objekter.

Den engelske gruppen skiller seg dermed også fra den norske gruppen, og forskjellen handler om lavere andel av bestemt form, og høyere andel av ubestemt form hos de engelske innlærerne. At den engelske gruppen her skiller seg fra både den polske og den norske, kan tyde på at det er denne gruppen som har gruppespesifikke formvalg, ikke den polske. Dette var noe jeg også observerte i undersøkelsen av hypotese 1, hvor jeg knyttet det til forskjeller i formvalgene ved generisk referanse i engelsk og norsk, og transfer hos den engelske gruppen. (se 9.9.3).

10.7.6 Oppsummering av funn på gruppenivå

På gruppenivå er det altså høyt morsmålsansvar for assosiasjonen *subjekt/ubestemt* hos de polske innlærerne, men ikke for *objekt/bestemt*. Det samme mønsteret finner en imidlertid også i den norske gruppen, og den polske gruppen er dermed lik den norske både i sammenfalls- og i markørkontekster, noe som gjør at det som ser ut som morsmålsansvar like godt kan forklares som målspråkstilegnelse. Vi så også antydning til intergruppe-heterogenitet mellom den polske og den engelske gruppen, men det kunne virke som det i større grad var den engelske gruppen som hadde gruppespesifikke formvalg, og dermed kanskje transfer. Jeg vil nå se om en undersøkelse av individuelle profiler kan belyse funnene fra gruppenivå.

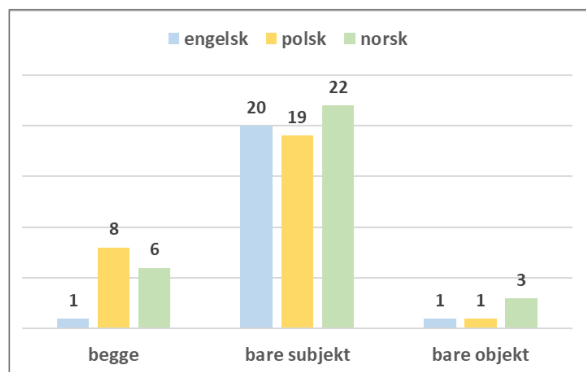
10.7.7 Individuelle profiler

Det er 28 polske innlærertekster som har enkle substantivfraser med generisk referanse, mens det er snakk om 31 tekster i den norske gruppen, og kun 22 i den engelske. Figur 10-33 (neste side) viser antall individuelle tekster som har både subjekt og objekt, eller som har enten bare subjekt eller bare objekt.

Det er ganske få tekster som har både subjekter og objekter, men det er flest i den polske gruppen (pol. 8; eng. 1; nor. 6). De aller fleste tekstene har kun subjekter (pol. 19; eng. 22; nor. 29), mens et fåtall bare har objekter (pol. 1; eng. 1; nor. 3). Dette er ikke uventet med tanke på at det generiske materialet stort sett består av subjekter. Som jeg

var inne på tidligere, gjør overvekt av subjekter og lavt antall markørkontekster det vanskelig å påvise eventuelle forskjeller mellom polsk og de andre to gruppene. Det er likevel positivt at det er en viss andel av polske innlærere som har begge deler, og at det er mulig å undersøke markørkontekster hos nesten en tredjedel av gruppen (9 stk.).

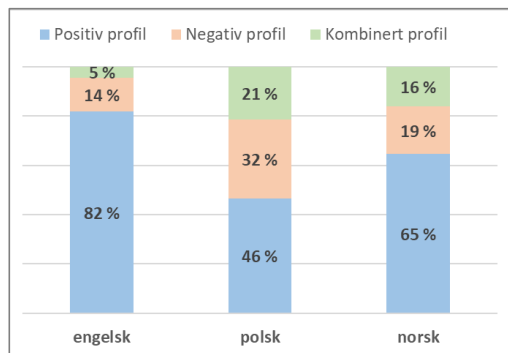
Figur 10-33 Antall tekster som har subjekt og objekt (generisk referanse)



10.7.7.1 Individuelle profiler sammenlignet med norsk og engelske gruppe

Figur 10-34 og tabell 10-31 (under) viser andelen av profiltyper ved generisk referanse.

Figur 10-34 Profiltyper i de tre gruppene (generisk referanse)



Tabell 10-31 Profiltyper i de tre gruppene (generisk referanse)

Profiltipe	Engelsk	Polsk	Norsk
Positiv	82 % (18)	46 % (13)	65 % (20)
Negativ	14 % (3)	32 % (9)	19 % (6)
Kombinert	5 % (1)	21 % (6)	16 % (5)
Total	22	28	31

I motsetning til ved indefinitt og definit referanse er det ved generisk referanse et relativt høyt antall positive profiler i den polske gruppen. Hele 46 % av de polske profilene (13) er positive, og har dermed form/kontekst-assosiasjoner som stemmer med prediksjonene. Sammenlignet med de to andre gruppene er dette imidlertid en *lavere* andel, ettersom den engelske gruppen har 68 % positive profiler, og den norske hele 82 %. Den polske gruppen har dessuten flest negative profiler (pol. 32 %; eng. 19 %; nor. 14 %). Forskjellen i antall positive profiler versus negative og kombinerte samlet, er signifikant mellom engelsk og polsk ($\chi^2= 6.5492$; $p=0.0104$), men ikke mellom polsk og norsk ($\chi^2= 1.9527$; $p=0.162$), og heller ikke mellom norsk og engelsk ($\chi^2= 1.8984$; $p=0.168$). Dette stemmer godt med funnene på gruppenivå, hvor den polske gruppen først og fremst skilte seg fra den engelske. Det at det er signifikant flere positive profiler i den engelske gruppen, stemmer også med at assosiasjonen *subjekt/ubestemt* var langt sterkere i denne gruppen på gruppenivå.

En nærmere undersøkelse av de 13 positive profilene i det polske materialet viser at 12 av disse *kun* har subjekt og ubestemt form, altså har bare én av de positive profilene markørkontekst. Disse 12 har dermed *ikke* formvalg som skiller seg fra målspåket. Tilsvarende tall for den engelske gruppen er 17 av 18, og for den norske gruppen 19 av 20. De positive profilene i den engelske og norske gruppen ser først og fremst ut til å være en avspeiling av utstrakt bruk av ubestemt form og mangel på objekter, og den polske gruppen skiller seg ikke nevneverdig fra dette, siden bare en av profilene har markørkontekst.

De individuelle profilene er altså i tråd med funnene på gruppenivå: Til tross for tendenser til morsmålsvarsvar, fins det stor likhet mellom alle tre gruppene, og den engelske er mer i tråd med prediksjonene for polsk morsmålsvarsvar enn den polske gruppen selv er. Positive profiler i alle tre gruppene er imidlertid i stor grad resultat av mangel på markørkontekster.

10.7.8 Oppsummering av funn for hypotese 2 ved generisk referanse

Tabell 10-32 (neste side) gir en oversikt over hovedfunnene ved generisk referanse både på gruppe- og individnivå. Tallene på gruppenivå viser bare tallene fra de predikerte

form/kontekst-assosiasjonene, og på individnivå er kombinerte og negative profiler slått sammen.

Tabell 10-32 Oversikt over hovedfunn for hypotese 2 ved generisk referanse

GRUPPENIVÅ	Polsk	Engelsk	Norsk
Subjekt/ubestemt form			
- Innen kontekst	74 %	93 %	89 %
- Innen form	86 %	96 %	89 %
Objekt/bestemt form			
- Innen kontekst	39 %	0 %	30 %
- Innen form	22 %	0 %	30 %
INDIVIDNIVÅ	Polsk	Engelsk	Norsk
Individuelle profiler			
- Positive	46 %	82 %	65 %
- Negative/kombinerte	54 %	18 %	35 %

Når det gjelder morsmålssamsvar for den polske gruppen var det tendenser på gruppenivå til en assosiasjon mellom subjekt og ubestemt form, men ikke en tilsvarende assosiasjon mellom objekt og bestemt form. På individnivå var det en stor andel positive profiler, da nesten halvparten av de polske innlærerne hadde en profil som stemte med prediksjonene, men en undersøkelse av sammenfalls- og markørkontekster viste at bare én av disse profilene hadde markørkontekst, og at andelen av profiler var høyere i de andre to gruppene enn i den polske. Den polske gruppen hadde dessuten større andel negative profiler enn de to andre gruppene.

Både på gruppe- og individnivå har det også kommet fram at den polske gruppen ikke skiller seg nevneverdig fra den norske, men at den engelske gruppen skiller seg fra begge de to andre gruppene, både i form av å ha en sterkere assosiasjon mellom subjekt/ubestemt, ingen assosiasjon for objekt/bestemt, og i å ha signifikant flere positive profiler enn den polske gruppen. Det er altså flere tegn på intergruppeheterogenitet mellom den polske og den engelske gruppen, men ettersom likheten mellom den norske og den polske gruppen er såpass stor, kan det virke som det først og fremst er den engelske gruppen som har gruppespesifikke formvalg, ikke den polske. Dette kan videre knyttes til at ubestemt form er den mest brukte formen ved generisk referanse i engelsk, og samlet sett kan dette være indikasjoner på transfer hos de

engelske innlærerne, noe jeg også argumenterte for i undersøkelsen av hypotese 1 (9.9.3).

Undersøkelsen av de fire transfereffektene har altså vist at det er sterkere tendens til morsmålssamsvar hos polske innlærere ved generisk referanse enn ved indefinitt og definitt, og de predikerte form/kontekst-assosiasjonene har vist seg som representative for cirka halve den polske gruppen (delvis intragruppe-homogenitet). Undersøkelser av intralingval kontrast i forhold til den norske gruppen gjennom sammenfalls- og markørkontekster viste likevel at de positive tendensene i det polske materialet stort sett også representerer målspåklarende formvalg, og derfor alternativt kan forklares som målspåklarende.

Gjennom undersøkelsen av hypotese 2 ved generisk referanse er det altså i likhet med for indefinitt og definitt referanse ikke funnet gode argumenter for å si at den polske gruppen har den type transfer som er predikert av hypotese 2.

Som jeg var inne på over, viser funnene for den polske gruppen at disse innlærerne i likhet med de andre to gruppene, ofte bruker ubestemt form ved generisk referanse uavhengig av om det er snakk om subjekter eller objekter. Likevel er det som vi så i de deskriptive tallene innledningsvis (10.7.1), brukt større andel bestemt form i den polske gruppen enn de andre to gruppene (pol. 28 %; eng. 7 %; nor. 13 %). Den polske gruppen bruker bestemt form oftere både ved subjekter og objekter, men mest ved objekter (subj. 26 %; obj. 39 %).

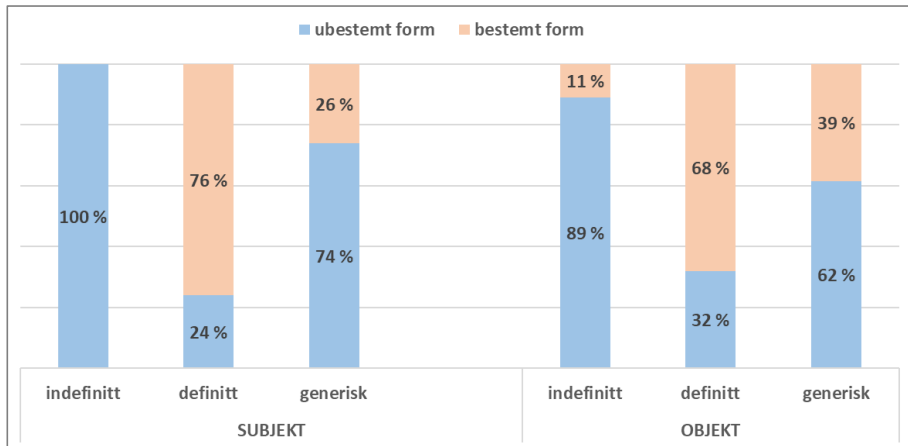
10.8 Polsk morsmålssamsvar på tvers av referensielle typer

Selv om det ikke ser ut til å være belegg for påstander om morsmålspåvirkning av den typen hypotese 2 handler om, vil jeg til sist se på formvalgene i det polske materialet på tvers av referensielle typer. Formvalgene for hver av typene har allerede blitt diskutert underveis i undersøkelsen, så dette er mest en oppsummering som presenterer form/kontekst-assosiasjonene innen de ulike referensielle typene i forhold til hverandre.

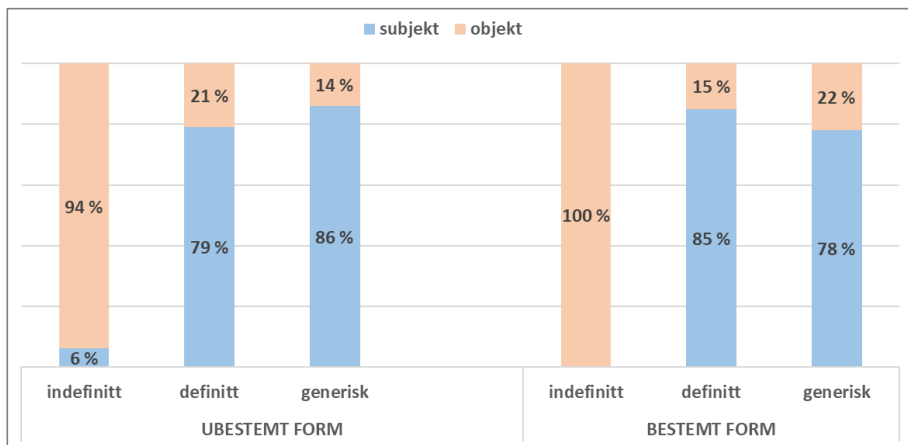
Hypotese 2 forutsetter at fordelingen av former er lik *på tvers av referensiell type*, altså at referanse ikke påvirker formvalget til de polske innlærerne, men at valg av form heller er styrt av om det er subjekt eller objekt. Det har kommet fram gjennom undersøkelsen så langt at dette ikke stemmer, men søylediagrammene på neste side visualiserer dette mer tydelig når de ulike referansetyperne står ved siden av hverandre.

Som ellers i undersøkelsen viser søylediagrammene på under til andel av form innen kontekst (figur 10-35), og andel av kontekst innen form (figur 10-36).

Figur 10-35 Andel av form innen kontekst etter referensiell type i den polske gruppen



Figur 10-36 Andel av kontekst innen form etter referensiell type i den polske gruppen



Det første som kommer tydelig fram av begge søylediagrammene, er at det ikke er spesielt stor likhet mellom det indefinite, definitive og generiske materialet, verken innen kontekst eller innen form. For eksempel varierer andelen av ubestemt form etter referensiell type både ved subjekter (ind. 100 %; def. 24 %; gen. 74 %) og objekter (ind.

89 %; def. 32 %; gen. 62 %).⁵⁵ Hovedtendensen er størst andel ubestemt form ved indefinitt referanse, lavest ved definitt og nest mest ved generisk, og dermed motsatt for bestemt form.

Også hvilke kontekster bestemt og ubestemt form brukes oftest i, varierer etter referensiell type (figur 10-35): Ubestemt form er sjeldnest subjekt ved indefinitt referanse, men er subjekt langt oftere ved definitt og generisk (ind. 6 %; def. 79 %; gen. 86 %). Den samme tendensen finner vi for bestemt form (ind. 0 %; def. 85 %; gen. 78 %). Disse tallene er nok også en naturlig avspeiling av liten andel av subjekter ved indefinitt referanse, og en langt større andel ved definitt og generisk (ind. 5 %; def. 83 %; gen. 84 %).

Det andre som kommer fram av figurene er at fordelingen av former er ganske stabil innad i hver referansetype. Ved indefinitt referanse er det en relativt lik andel ubestemt form både ved subjekter og objekter (subj. 100 %; obj. 89 %). Det samme gjelder ved definitt referanse (subj. 24 %; obj. 32 %), og generisk (subj. 74 %; obj. 62 %).

Innen form ser vi det samme (figur 10-36): Innen indefinitt referanse er både ubestemte og bestemte substantiv først og fremst objekter (ub. 94 %; be. 100 %). Tilsvarende ser vi innen definitt referanse, der begge formene først og fremst er subjekter (ub. 79 %; be. 85 %), noe som også gjelder ved generisk referanse (ub. 86 %; be. 78 %).

Formvalget i den polske gruppen er predikert til å skille mellom subjekter og objekter, og på samme måte innen alle de referensielle typene. Denne gjennomgangen viser imidlertid at det er store forskjeller *mellom* de referensielle typene, og at det er *små* forskjeller mellom subjekter og objekter *innen* hver av de referensielle typene.

Tallene innen kontekst, altså hvilken form som typisk brukes ved subjekter og objekter, stemmer derimot godt med målspråkets valg av form: ubestemt form ved indefinitt, bestemt form ved definitt, og mest ubestemt form ved generisk referanse. Tallene innen form avspeiler også til en viss grad at de ulike referensielle typene gjerne knyttes til ulike syntaktiske funksjoner i materialet: Indefinitt referanse er typisk

⁵⁵ Jeg oppgir her bare prosentandelen for ubestemte former for enkelhets skyld. Siden det bare er to muligheter her (bestemte eller ubestemte), vil prosentandelen av bestemte former gå fram indirekte av andelen ubestemte. Det samme gjelder for subjekter og objekter lenger ned.

objekter, mens definitt og generisk referanse typisk er subjekter. Dette finner en både i den polske gruppen, og i de andre to gruppene.

10.9 Oppsummering hypotese 2

I dette kapitlet har jeg undersøkt hypotese 2 om morsmålpåvirkning fra polsk, som gikk ut på at polske innlærere ville ta i bruk norsk substantivmorfologi for å markere skillet mellom nominativ og ikke-nominativ i innlærerspråket, slik de gjør i morsmålet. Hypotesen ble undersøkt i et materiale av subjekter og objekter, og hadde to delprediksjoner om *morsmålssamsvar* i form av assosiasjoner mellom form og språklig kontekst:

1. Assosiasjon mellom ubestemt form og subjekt
2. Assosiasjon mellom bestemt form og objekt

Hypotesen har blitt undersøkt ved å se på både gruppetall og individuelle profiler, og ved å se etter alle de fire transfereffektene *morsmålssamsvar*, *intragruppe-homogenitet*, *intergruppe-heterogenitet* og *intringval kontrast*. Intragruppe-homogenitet ble kun undersøkt gjennom individuelle profiler, mens intergruppe-heterogenitet og intralingval kontrast først og fremst ble undersøkt innen hver av de referensielle typene.

Den overordnede konklusjonen fra undersøkelsen er at det *ikke* foreligger transfer av den typen hypotese 2 handler om. Jeg vil avslutningsvis oppsummere de viktigste observasjonene og grunnene til at konklusjonen om transfer er negativ når det gjelder de polske innlærerne.

10.9.1 Morsmålssamsvar og intergruppe-heterogenitet i samlet materialet

Det har vært en antakelse for denne hypotesen at det predikerte morsmålssamsvaret gjelder innen alle de tre referensielle typene, og at et eventuelt morsmålssamsvar derfor skulle kunne komme til syne allerede i første del av undersøkelsen, der jeg så på hele materialet samlet. Det samlede materialet ga imidlertid ikke særlig støtte til hypotesen.

Det var for det første ikke funn av morsmålssamsvar for noen av de to predikerte form/kontekst-assosiasjonene, verken for *subjekt/ubestemt* eller *objekt/bestemt*, ettersom prosentandelene for begge disse lå *under* 50 % både innen kontekst og innen form. Undersøkelsen av individuelle profiler viste også at det ikke var en eneste positiv profil i den polske gruppen. Både på gruppenivå og individnivå var dermed funnene

negative med tanke på morsmålssamsvar i det samlede materialet. Den polske gruppen viste også stor likhet med både den engelske og den norske gruppen, som heller ikke hadde de aktuelle form/kontekst-assosiasjonene.

Alle disse observasjonene i det samlede materialet gjorde at det ikke var grunnlag for å argumentere for transfer. Jeg fant verken morsmålssamsvar, intragruppe-homogenitet, intergruppe-heterogenitet eller forskjeller fra målpråket.

10.9.2 Morsmålssamsvar og intergruppeheterogenitet i oppdelt materiale

Jeg vil nå oppsummere hva undersøkelsen av sammenfalls- og markørkontekster sammen med morsmålssamsvar fikk fram når materialet ble delt opp etter de tre ulike referensielle typene. Jeg vil se på hver av de to predikerte form/kontekst-assosiasjonene etter om de utgjør sammenfalls- eller markørkontekster. Subjekter og objekter bytter på å være markørkontekster: Subjekter ved definitiv referanse og objekter ved indefinitiv og generisk.

Prediksjon 1: assosiasjon mellom subjekt og ubestemt form

For assosiasjonen *subjekt/ubestemt form* fant jeg høyest morsmålssamsvar ved generisk referanse, der prosentandelen var over 70 % både innen form og kontekst. Utover dette så vi at morsmålssamsvaret var høyt innen kontekst ved indefinitiv referanse, mens det derimot var lavt innen kontekst ved definitiv referanse, som vi så at forholdet mellom form og kontekst var lite gjensidig innen disse referansetyperne. Siden subjekt er markørkontekst ved definitiv referanse, var det altså et mønster av høyt morsmålssamsvar i sammenfallskontekstene og lavt i markørkonteksten. Dette var videre supplert med en generell likhet med de to andre gruppene, både i sammenfalls- og markørkontekster. Ved generisk referanse var det noen forskjeller mellom den polske og den engelske gruppen, men på grunn av likheten mellom den polske og den norske, ble dette tolket å være gruppespesifikt for den engelske gruppen, og ikke den polske.

Disse observasjonene viser at den polske gruppen i det samlede materialet har lavt morsmålssamsvar for delprediksjonen om assosiasjon mellom subjekt og ubestemt form, og at dette er utfyllt med lav intergruppe-heterogenitet mellom den polske og den engelske gruppen, i tillegg til stor grad av målpråkslikhet. Tabell 10-33 (neste side) gir en oversikt over vurderingen av transfereffekter for denne delprediksjonen.

Tabell 10-33 Transfereffekter for assosiasjonen mellom subjekt og ubestemt form

	Indefinit referanse	Definit referanse	Generisk referanse
Subjekt =	Sammenfallskontekst	Markørkontekst	Sammenfallskontekst
Morsmålssamsvar	Innen kontekst	Innen form	Ja
Ulik engelsk gruppe	Nei	Nei	Ja
Ulik norsk gruppe	Nei	Nei	Nei

Prediksjon 2: Assosiasjon mellom objekt og bestemt form

For assosiasjonen *objekt/bestemt form* fant jeg en lignende, men motsatt tendens: Morsmålssamsvaret var lavt innen generisk referanse, med under 50 % både innen kontekst og form. Videre var morsmålssamsvaret høyt innen kontekst ved definit referanse, og lavt innen kontekst ved indefinit referanse, og dermed var det ikke noe gjensidig forhold mellom form og kontekst ved disse to referansetyperne. Siden objekter er markørkontekst innen generisk og indefinit referanse, ser vi altså at det også for denne delprediksjonen er høyt morsmålssamsvar i sammenfallskonteksten, men lavt samsvar i markørkontekstene. Det er også generell likhet med de andre to gruppene for assosiasjonen *objekt/bestemt*, uavhengig av om det er sammenfalls- eller markørkontekst, men også her var det forskjeller mellom den polske og engelske gruppen ved generisk referanse, som ble tolket som gruppespesifikt for den engelske gruppen, og ikke for den polske.

Samlet sett er det dermed også for denne delprediksjonen lavt morsmålssamsvar, lav intergruppe-heterogenitet og høy grad av sammenfall med målspråket.

Tabell 10-34 gir en oversikt over vurderingen av transfereffekter for denne delprediksjonen.

Tabell 10-34 Transfereffekter for assosiasjonen mellom objekt og bestemt form

	Indefinit referanse	Definit referanse	Generisk referanse
Objekt =	Markørkontekst	Sammenfallskontekst	Markørkontekst
Morsmålssamsvar	Innen form	Innen kontekst	Nei
Ulik engelsk gruppe	Nei	Nei	Ja
Ulik norsk gruppe	Nei	Nei	Nei

Individuelle profiler

De individuelle profilene viste hvor mange polske tekster som hadde formvalg i tråd med prediksjonene sammenlignet med de andre gruppene. Hovedtendensen var *få*

positive polske profiler både i det samlede materialet og innen indefinitt og definitt referanse, og dessuten stor likhet med de andre to gruppene.

Innen generisk referanse var det derimot en noe høyere andel positive polske profiler (46 %), som vil si at det til dels var morsmålssamsvar i gruppen, og at dette var representativt for nesten halve gruppen. Det var imidlertid bare én av de polske tekstene som hadde markørkontekster, så den høye andelen av positive profiler skyldtes at det stort sett bare var sammenfallskontekster i disse profilene. Innen generisk referanse hadde dessuten de to andre gruppene *flere* positive profiler enn den polske, men det bare *signifikant* flere i den norske gruppen. Tabell 10-35 gir en oversikt over transfereffektene for individuelle profiler, og parentesene viser hvor den engelske eller norske gruppen har flere positive profiler enn den polske (nor.>pol. og (eng.> pol.).

Tabell 10-35 Transfereffekter for individuelle profiler

	Samlet materiale	Indefinitt referanse	Definitt referanse	Generisk referanse
Morsmålssamsvar	Nei	Nei	Nei	Nei
Intragruppe-homogenitet	Nei	Nei	Nei	Nei
Ulik engelsk gruppe	Nei	Nei	Nei	Ja (eng.> pol.)
Ulik norsk gruppe	Nei	Nei	Nei	Nei (nor.> pol.)

Funnene for de individuelle profilene støtter altså i stor grad opp under funnene på gruppenivå: lite morsmålssamsvar og lite forskjell fra de andre to gruppene. Det var bare de individuelle profilene som kunne si noe om intragruppe-homogenitet, og også denne effekten er gjennomgående svak.

Til tross for lite morsmålssamsvar og små kvantitative forskjeller fra de andre gruppene ved definitt referanse, var det et interessant funn av *fem* polske profiler som hadde formvalg i tråd med prediksjonene og dessuten hadde markørkontekster. Jeg har sett nærmere på om dette kan ha en sammenheng med engelsknivået til disse polske innlærerne, men fant at *to* av dem hadde oppgitt å kunne engelsk på høyere nivå, *to* på begynnernivå og *én* kunne ikke engelsk i det hele tatt. Det vil si at det er en relativt jevn spredning i engelskferdigheter, noe som avkrefter en eventuell mistanke om at disse fem kan ha transfer fordi de ikke kan engelsk.

Konklusjon om transfer

En vurdering av både det samlede materialet og det som kom fram i analysen av sammenfallskontekster og markørkontekster i det differensierte materialet, peker altså stort sett i retning av en negativ konklusjon om transfer hos de polske innlærerne for hypotese 2. For begge de to form/kontekst-assosiasjonene er det svakere samsvar med prediksjonene enn det er med målspråkets formvalg, og formvalgene til de polske innlærerne er i liten grad spesifikke for innlærere med polsk som morsmål.

10.9.3 Målspråkslikhet

I likhet med det som kom fram i undersøkelsen av hypotese 1, har det også i undersøkelsen av hypotese 2 vist seg at den polske gruppen har stor grad av sammenfall med målspråksgruppen, og likheten har kanskje vært enda større her.

I det samlede materialet var den polske og den norske gruppen like i å ha en overvekt av ubestemt form (pol. 66 %; nor. 63 %). Denne tendensen var også til stede innen objekter (pol. 82 %; nor. 72 %), mens det innen subjekter var rundt 50 % av hver form i begge gruppene (pol. 48/52 %; nor. 55/45 %). Dette vil si at den polske gruppen hadde en noe lavere andel bestemt form ved objekter, og en litt høyere andel ved subjekter, men likheten mellom gruppene er likevel stor. I det samlede materialet var det også likheter mellom gruppene i å ha en sterk assosiasjon mellom objekt og *ubestemt form*, og mellom subjekt og *bestemt form*, altså motsatt av prediksjonene om polsk morsmålssamsvar. Dette viser også at mønstret i det polske materialet er relativt målspråkslikt.

Når materialet ble oppdelt i referensielle typer, ble det tydelig at også de polske innlærerne varierer formvalget etter om det er indefinitt, definitt eller generisk referanse, på samme måte som man gjør i målspråket. Det var størst andel ubestemt form ved indefinitt referanse (ub. 90 %; be. 10 %), størst andel bestemt form ved definitt referanse (ub. 25 %; be. 75 %), og ved generisk referanse var det også mest ubestemt form (ub. 72 %; be. 28 %). Dette distribusjonsmønsteret stemmer godt med det som er de vanlige formvalgene i norsk målspråk, og det var generelt store likheter mellom den polske og den norske gruppen, selv om det også var noen *kvantitative* forskjeller mellom de to gruppene innad i dette generelle distribusjonsmønsteret: Den polske gruppen hadde en noe lavere andel ubestemt form enn den norske ved indefinitt referanse (pol. 90 %;

nor. 98 %), en noe lavere andel bestemt form ved definitiv referanse (pol. 75 %; nor. 83 %), og dessuten *høyere* andel bestemt form ved generisk referanse (pol. 28 %; nor. 7 %). Den polske gruppen er altså mest lik den norske innen indefinitiv referanse, mens det er større forskjeller ved definitiv referanse og generisk referanse. For alle de tre referensielle typene slår forskjellene mest ut ved objekter.

At det er *minst* likhet med målspråksgruppen ved definitiv referanse, i form av *mindre* bruk av bestemt form i den polske gruppen, kan henge sammen med det som kom fram i undersøkelse 1, nemlig at -GB-innlærere oftere utelater bestemt form i obligatoriske kontekster enn det +GB-innlærere gjør. Materialet i undersøkelse 1 besto imidlertid både av substantivfraser med definitiv og med generisk referanse, og det vi ser her er at bruken av bestemt form hos de polske innlærerne er *lavere* enn i den norske gruppen ved definitiv referanse, men at den faktisk er *høyere* ved generisk referanse (pol. 28 %; nor. 7 %). Det er vanskelig å si om det er en sammenheng mellom funnene i disse to delundersøkelsene, men siden valg av form til en viss grad er mer fleksibelt ved generisk referanse enn ved definitiv, kan en spekulere i om mangel på bestemt form i større grad har ført til feil ved definitiv referanse enn ved generisk referanse i undersøkelse 1, et poeng jeg også var inne på i 9.9.3 for transfer hos de engelske innlærerne. Om det er slik, kan en mistenke at det er en sammenheng mellom utslaget med mindre bruk av bestemt form ved nettopp definitiv referanse i materialet for hypotese 2, og lavere grad av markering i undersøkelse 1.

10.9.4 Transfer hos de engelske innlærerne

Det har i undersøkelsen av hypotese 2 ikke kommet fram like tydelige utslag av gruppespesifikke tendenser og transfer i den engelske gruppen som det gjorde ved hypotese 1, men det ble påpekt noen tendenser innen materialet med generisk referanse.

Ved generisk referanse var det forskjeller i andelen av ubestemt form i de tre gruppene. Totalt sett hadde den engelske gruppen 96 % ubestemt form, mens både den norske og den polske hadde en lavere andel (pol. 84 %; nor. 87 %). Forskjellen fra de andre to gruppene er spesielt tydelig ved objekter, hvor den engelske gruppen ikke bruker bestemt form i det hele tatt, mens andelen av bestemt form er relativt høy i de andre gruppene (pol. 39 %; eng. 0 %; nor. 30 %). I tillegg til dette har den engelske gruppen lavere andel bestemt form ved subjekter, spesielt i forhold til den polske

gruppen (pol. 26 %; eng. 8 %; nor. 11 %). Som jeg var inne på i kapittel 7, og i forbindelse med gjennomgangen av hypotese 1, er det også slik at bestemt form brukes sjeldnere i engelsk enn i norsk ved generisk referanse, og det at den engelske gruppen skiller seg fra den polske gruppen nettopp på dette området, gjør at man kan argumentere for at det foreligger både et morsmålssamsvar som skiller seg fra målspåket, og intergruppe-heterogenitet i forhold til en annen innlærergruppe, noe som samlet sett kan tyde på transfer hos de engelske innlærerne. Dette er imidlertid kun basert på en skjønsmessig vurdering av et gruppemateriale uten hensyn til individuell variasjon, og uten bruk av statistiske tester. I tillegg har jeg heller ikke vurdert intragruppe-homogeniteten til dette trekket hos de engelske innlærerne, altså hvor representativt det er for gruppen.

At transferlignende tendenser ikke kom til syne i det engelske materialet ved *indefinit* referanse, slik det gjorde i undersøkelsen av hypotese 1, skyldes nok at materialet for hypotese 1 ikke inneholder artikkelfraser, som var det den engelske gruppen så ut til å ha overrepresentasjon av i materialet for hypotese 1. Materialet for hypotese 2 inneholder dessuten ikke-tellelige substantiv, noe som potensielt hever andelen av ubestemt entall i den engelske gruppen. Utslaget av transfer ved *generisk* referanse var heller ikke fullt så tydelig i denne delen av undersøkelsen som det var ved hypotese 1. Det var mulig å finne forskjeller mellom den engelske og den polske gruppen, men det var mindre forskjell mot den norske gruppen, noe som kan skyldes at ikke-tellelige substantiv var inkludert. Effekten av forskjellen mellom språkene norsk og engelsk på dette området vil dermed ikke være så tydelig som når ikke-tellelige er utelatt, siden andelen av ubestemt entall da vil øke i det engelske materialet.

10.9.5 Hypotese 2: avslutning

Det er lite som tyder på at hypotese 2 stemmer. Tvert imot, når materialet deles opp etter syntaktisk funksjon og referensielle typer, og det skilles mellom bestemt og ubestemt form, så utviser alle de tre gruppene et veldig likt mønster, et mønster med en sterk assosiasjon mellom indefinit og ubestemt form, og mellom definit og bestemt form, i tillegg til at generisk referanse domineres av ubestemt form.

Jeg vil i neste kapittel oppsummere hele avhandlingsstudien, og der også komme tilbake til en videre diskusjon av de negative funnene og gjennomføringen av undersøkelsen både for hypotese 1 og 2.

11 OPPSUMMERING OG DISKUSJON AV STUDIEN

I dette siste kapittelet vil jeg oppsummere studien som ligger til grunn for denne avhandlingen, og diskutere noen sentrale momenter knyttet til funnene og gjennomføringen av begge undersøkelsene. I 11.1 oppsummerer jeg de overordnede trekkene ved studien, og i 11.2 og 11.3 ser jeg tilbake på henholdsvis undersøkelse 1 og undersøkelse 2. I 11.4 diskuterer jeg mulig transfer i det engelske materialet, mens 11.5 gjør noen betraktninger rundt nullkontrast som utgangspunkt for transfer. I 11.6 runder jeg av med å oppsummere noen viktige erfaringer jeg har gjort med transfereffekter og bruk av korpus.

11.1 Forskningsspørsmål og studiens to delundersøkelser

Jeg har i denne avhandlingen undersøkt noen spørsmål knyttet til hvorvidt morsmålet påvirker bruk av bestemte og ubestemte former i tekster skrevet av norskinnlærere som ikke kjenner grammatisk bestemthet fra morsmålet. Jeg har gjort dette gjennom to større undersøkelser: *undersøkelse 1*, som har sett på den kvantitative siden av spørsmålet, altså om morsmålet påvirker hvor ofte bestemt form brukes i obligatoriske kontekster, og *undersøkelse 2*, som har sett nærmere på om morsmålet påvirker hvilke språklige kontekster bestemte og ubestemte former brukes i.

For begge disse undersøkelsene er det språklige datamaterialet hentet fra Norsk andrespråkkorpus (ASK). En metodisk tråd gjennom hele avhandlingen har dessuten vært Jarvis' (2000; 2010) rammeverk for transferstudier, der det mest sentrale metodiske grepet er påvisning av *transfereffekter*, altså språklige mønstre i innlærertekstene som ser ut til å være *spesifikke* og *representative* for en gruppe innlærere, og som i tillegg kan knyttes til språklige forhold i morsmålet. I denne studien er det grammatisk bestemthet det er snakk om, og derfor har en essensiell del av metoden vært sammenligning av tekster skrevet av innlærere *med* og *uten* grammatisk bestemthet i morsmålet (-GB/+GB-innlærere). I undersøkelse 1 så jeg derfor på tekster skrevet av innlærere fra til sammen *seks* ulike morsmålsgrupper, tre *med* og tre *uten* grammatisk bestemthet, mens undersøkelse 2 bare tok for seg tekster skrevet av innlærere med polsk (-GB) og engelsk (+GB) som morsmål, i tillegg til tekster skrevet av morsmålsbrukere av norsk.

I kapittel 1 presenterte jeg to overordnede forskningsspørsmål:

Forskningsspørsmål 1:

Påvirker morsmålet hvor ofte bestemte former brukes i obligatoriske kontekster i innlærerspråket til innlærere uten grammatisk bestemthet i morsmålet?

Forskningsspørsmål 2:

Påvirker morsmålet distribusjonen av bestemte og ubestemte former i innlærerspråket til polske innlærere?

Det første forskningsspørsmålet ble forsøkt besvart gjennom undersøkelse 1, der jeg sammenlignet markeringsprosenter for bestemt form hos -GB-innlærere og +GB-innlærere. Det andre spørsmålet ble forsøkt besvart gjennom undersøkelse 2, der jeg undersøkte to ulike hypoteser om hvordan polsk som morsmål kunne tenkes å påvirke valg av form på substantivet i ulike språklige kontekster. Begge disse hypotesene hadde som utgangspunkt at de polske innlærerne på grunn av mangelen på grammatisk bestemthet i morsmålet opererer med en antakelse om at målspråket norsk fungerer på samme måte som morsmålet, og at visse språklige innholdselementer fra morsmålet derfor også blir uttrykt i innlærerspråket ved hjelp av de norske bestemthetsmarkørene: For hypotese 1 dreide det seg om overføring av pragmatisk ordstilling som signal om substantivreferenters identifiserbarhet og at markering av identifiserbarhet oppleves som redundant når det går fram av ordstillingen. For hypotese 2 handlet det om at de polske innlærerne ville ta i bruk norsk substantivmorfologi for å uttrykke kasus, mer spesifikt skillet mellom nominativ og ikke-nominativ.

Mens det var positive funn for transfer i undersøkelse 1, var funnene i undersøkelse 2 i all hovedsak *negative* for begge hypotesene. Det var altså mulig å finne kvantitative utslag av transfer hos -GB-innlærere, men ingen av de to hypotesene knyttet til kvalitative utslag av morsmålpåvirkning fra polsk, hadde støtte i materialet. Jeg vil nå oppsummere og diskutere hver av de to delundersøkelsene.

11.2 Undersøkelse 1: Kvantitativt utslag av transfer

Undersøkelse 1 hadde som mål å finne ut om det er en sammenheng mellom det å ikke kjenne grammatisk bestemthet fra morsmålet og sjeldnere bruk av bestemte former i innlærerspråket i kontekster der det er obligatorisk i målspråket, noe jeg har omtalt som

et *kvantitativt* utslag av transfer. Et slikt kvantitativt utslag ble antatt å vise seg på to områder i materialet av innlærertekster: For det første i form av *lavere markeringsprosent* i -GB-gruppen enn i +GB-gruppen, og for det andre i form av *færre innlærertekster med 100 % markering* i -GB-gruppen enn i +GB-gruppen.

I denne delen av studien utgjorde -GB-gruppen innlærere med russisk, polsk eller bosnisk/kroatisk/serbisk som morsmål, mens +GB-gruppen besto av innlærere med engelsk, tysk eller nederlandsk som morsmål. Hele materialet besto av 1200 tekster, hentet fra begge de to testtypene som er representert i ASK, og utvalget representerte en blanding av tekster fra ulike ferdighetsnivå på skalaen til Det felles europeiske rammeverket for språk.

Undersøkelse 1 viste i stor grad positive funn for transfer. Det forelå statistisk signifikante forskjeller mellom de to gruppene både for markeringsprosent og antall tekster med 100 % markering, og dette kom til syne både i hele materialet og innen delkorpuserne fra hver av de to språktestene (SP og HN). Det ble også vist til en tidligere studie basert på samme materiale (Johansen 2010), der det kom fram at det også var forskjeller i markeringsprosenten mellom de to GB-gruppene innen både B1- og B2-nivå. I sammenligningen av innlærere med polsk og engelsk som morsmål, fant jeg bare signifikante forskjeller mellom gruppene på ett av de to områdene som ble undersøkt: Innen delkorpuset fra SP-testen var det signifikante forskjeller for markeringsprosent, mens det innen delkorpuset fra HN-testen var signifikante forskjeller for antall tekster med 100 % markering. Forskjellen i markeringsprosent på SP-testen ble i tillegg vurdert som *nært* signifikant ettersom p-verdien lå like over 0,05. Det ble også foreslått med utgangspunkt i Larsen-Hall (2010: 102) at dette kunne ses på som signifikant med en mindre streng holdning til statistiske grenseverdier i andrespråksforskning (se 6.5.1).

Disse signifikante forskjellene mellom gruppene innebærer positive funn for de to transfereffektene *intergruppe-heterogenitet* og *intragruppe-homogenitet*, da de statistiske testene tok hensyn til både forskjeller mellom gruppene og variasjon innad i hver av gruppene (se 6.1.4). Det ble derfor på denne måten vist at tendensene i -GB-gruppen var både spesifikke og representative for denne gruppen. Den siste transfereffekten som ble sett på i undersøkelse 1, *morsmålssamsvar*, var derimot oppe

til diskusjon i slutten av kapittel 6. I undersøkelse 1 var det nemlig fravær av grammatisk bestemthet i morsmålet som ble regnet som bakgrunnen for sjeldnere markering hos -GB-innlærerne, og morsmålssamsvaret var dermed ikke antatt å være knyttet til en konkret språklig struktur som ble overført fra morsmålet, men heller til det generelle fraværet av bestemthetsmarkører i morsmålet.

Ut fra funnene i undersøkelse 1 er svaret på forskningsspørsmål 1: Ja, morsmålet påvirker hvor ofte bestemte former brukes i obligatoriske kontekster i innlærerspråket til -GB-innlærere av norsk. Noe funnene *også* kan vise, er at morsmålet påvirker hvor ofte bestemte former brukes hos +GB-innlærere ettersom disse viste seg å oftere ha markering enn -GB-innlærere, og dermed også har tendenser som er gruppespesifikke og grupprepresentative. Funnene fra undersøkelse 1 kan dermed tyde på at når det gjelder korrekt bruk av bestemt form, kan morsmålet være fasiliterende for +GB-innlærere, mens det i større grad fører til utfordringer for -GB-innlærere, noe som også er funnet i andre studier som sammenligner -GB-innlærere og +GB-innlærere (Axelsson 1994; Jarvis 2002; Trenkic & Pongpairroj 2013; Nordanger 2017).

Siden undersøkelsen viste gruppespesifikke og grupprepresentative tendenser for begge innlærergruppene, kan man stille seg spørsmålet om dette betyr at begge gruppene faktisk har transfer, eller om bare én av dem har det. Når disse funnene av intragruppe-homogenitet og intergruppe-heterogenitet i tillegg kan knyttes til fravær eller tilstedeværelse av grammatisk bestemthet i morsmålet, kan man altså også vise til transfereffekten morsmålssamsvar for begge gruppene, noe som gir et sterkere argument for transfer hos begge. Som vi så i 4.2.1 fins det imidlertid argument for at fravær av en kategori i morsmålet ikke nødvendigvis fører til transfer, og det kan tenkes at funnene av lavere markeringsprosent i -GB-gruppen kan tilskrives andre forhold ved tilegnelsen, for eksempel at bruken av former er basert på språklige universalier eller målspåksinnputt, som begge også kan føre til gruppespesifikke og grupprepresentative mønster hos en gruppe med samme morsmål (Jarvis 2000: 256-7). I så fall kan det være at det er den engelske gruppen som er påvirket av morsmålet, og ikke den polske. Dette viser et problem med å koble transfereffekten morsmålssamsvar kun til fravær av grammatisk bestemthet i morsmålet, ettersom fraværet av denne kategorien (nullkontrast) ikke nødvendigvis fører til transfer, og transfer heller ikke er eneste

mulige forklaring på gruppespesifikke og grupperepresentative språklige valg. Dette handler kanskje også om hva som kommer inn under begrepet transfer: Om en nullkontrast mellom morsmål og målspråk fører til at innlærere baserer seg på språklige universalier når de velger form av substantivet, slik noen studier mener å finne (Ionin & Wexler 2003; Ionin et al. 2004; Ionin et al. 2008), er dette da en form for morsmålpåvirkning? Om en følger Odlin's (1989: 27) definisjon av transfer som et resultat av forskjeller og likheter mellom morsmål og målspråk (se 2.1), vil dette potensielt kunne regnes som en form for morsmålpåvirkning siden det er forskjellene mellom språkene som er utløsende. Som jeg også var inne på i 2.1 er denne definisjonen sett på som problematisk, nettopp fordi den muliggjør slike resonnement gjennom å ikke skille mellom det som fører til transfer og det som *er* transfer. Om en derimot skal legge til grunn en definisjon som den en finner hos Jarvis & Pavlenko (2008: 1), hvor transfer innebærer overføring av språklig kunnskap fra morsmålet til innlærerspråket, vil det bli mer tydelig at formvalg basert på universalier ikke kan regnes som morsmålpåvirkning, og at dette heller illustrerer at transfer ikke er eneste mulige konsekvens av en nullkontrast. Dette er et poeng som er avgjørende for om undersøkelse 1 har kunnet påvise transfer: Det foreligger gruppespesifikke og grupperepresentative språklige mønster i både -GB-gruppen og +GB-gruppen, men mangel på en operasjonalisering av transfereffekten *morsmålsvarsvar* som sier mer konkret *hva* i morsmålet som overføres, gjør at undersøkelsen ikke tydelig nok kan vise at det handler om transfer fra morsmålet framfor andre mulige kilder.

En undersøkelse som går nærmere inn på hva som styrer valg av form i de ulike substantivfrasene i undersøkelse 1, hadde kunnet avdekke om markeringsprosenten er en *sekundæreffekt* av mer kvalitative effekter av morsmålet, eller for eksempel formvalg basert på språklige universalier. Undersøkelse 2 hadde som mål å undersøke slike kvalitative effekter av morsmålpåvirkning, men utformingene til undersøkelse 1 og 2 er ikke like nok til at de to delundersøkelsene er helt sammenlignbare. Jeg vil likevel komme kort innom spørsmålet om hvorvidt funnene i undersøkelse 2 kan belyse markeringsprosenten i undersøkelse 1 lenger ned (11.3.8).

11.2.1 utfordringer med gjennomføringen av undersøkelse 1

I 6.5.3 nevnte jeg forbehold ved undersøkelse 1, som jeg vil oppsummere kort her. For det første undersøkte jeg kun substantiv og substantivenes bøyingsformer, uten å ta hensyn til at bestemthet kunne være markert andre steder i frasen med determinativer. Undersøkelsen sa dermed noe om hvorvidt *bestemt form var brukt*, ikke noe om hvorvidt *trekket bestemt var markert*. Undersøkelsen rettet seg med andre ord mot en spesifikk bestemthetsmarkør (bøyingsendelse), noe som ikke er tilstrekkelig for å kunne si noe om innlærernes generelle mestring av kategorien bestemthet uavhengig av hvordan det uttrykkes. Dette er et spesielt relevant forbehold med tanke på at +GB-morsmålene som var representert i materialet, alle uttrykker bestemthet ved hjelp av foranstilte artikler, ikke bøyning. Om +GB-gruppen hadde hatt morsmålpåvirkede tendenser til å foretrekke determinativ heller enn bøyning i innlærerspråket, kan forskjellen i markeringsgrad potensielt være større mellom -GB-gruppen og +GB-gruppen enn det som ble observert i undersøkelse 1, ettersom +GB-gruppen så ut til å oftere ha substantivfraser *med* determinativ og *uten* bøyning.

Det kom også fram i undersøkelsen at selv om -GB-innlærere hadde mindre markering enn +GB-innlærere, så hadde begge grupper relativt høy markeringsprosent. Om det er slik at det å ha substantivmorfologi i morsmålet, slik -GB-innlærerne har, eller å ha foranstilte artikler i morsmålet, slik +GB-innlærerne har, påvirker det strukturelle uttrykket i innlærerspråket, så kan dette forklare at gruppene ligger såpass nære hverandre.⁵⁶ Både Nordanger (2017) og Axelsson (1994) viste til slike forhold som relevante for hvilken type kontekster innlærerne mestret: -GB-innlærere utviste i begge studiene størst utfordringer med bruk av ubestemt artikkel, mens +GB-innlærere i tidlige faser hadde utfordringer med bestemthetsbøyningen (Axelsson 1994: 101; Nordanger 2017: 259). Siden min studie ikke ser på obligatoriske kontekster for både bestemte og ubestemte former, belyser ikke mine data disse spesifikke forskjellene, men Nordangers og Axelssons funn viser altså at det kan være nødvendig og hensiktsmessig å skille mellom hvorvidt obligatoriske kontekster er markert med bøyning eller determinativ i innlærerspråket.

⁵⁶ +GB-gruppen besto av både engelske, tyske og nederlandske innlærere, og selv om alle har bestemte og ubestemte artikler, er det ulik grad av substantivmorfologi i disse språkene: Engelsk og nederlandsk bøyer substantivene kun i tall, mens tysk bøyer i både tall og kasus.

Det andre forbeholdet med undersøkelse 1 var at jeg verken kunne belyse om substantivene krevde bestemt form av syntaktiske eller referensielle årsaker, eller om det var påkrevd fordi det var snakk om definitt eller generisk referanse. Dette var både fordi jeg kun hadde sett på substantivets form, og fordi materialet ikke var delt inn i referensielle typer. Obligatoriske kontekster for bestemt form kan være både defintitte og generiske, og i noen tilfeller også ikke-referensielle, og ved definitt referanse er bestemt form oftere obligatorisk enn ved generisk, siden det er en viss fleksibilitet i valget av form ved generisk referanse i norsk (se 7.2.4.2). Dette er forhold som hadde vært nyttig å kunne skille mellom, spesielt siden ulike +GB-språk i ulik grad bruker bestemte former ved generisk referanse. Som det kom fram i den kontrastive beskrivelsen i kapittel 7, er det for eksempel mer vanlig å bruke bestemt form ved generisk referanse i norsk enn i engelsk (se 7.5).

11.2.2 Konklusjon undersøkelse 1

Selv om det er noen forbehold ved undersøkelse 1, har det vist seg å være relevante og statistisk signifikante forskjeller mellom innlærere med og uten grammatisk bestemthet i morsmålet når det kommer til hvor ofte bestemt form brukes i obligatoriske kontekster. Det er nærliggende å tolke dette som transfer, men som diskusjonene over viser, er det også problematiske sider med en slik konklusjon.

11.3 Undersøkelse 2: kvalitativt utslag av transfer

Undersøkelse 2 rettet seg mot de mer kvalitative sidene av morsmålspåvirkning, altså hvorvidt morsmålet påvirker polske innlæreres valg mellom bestemte og ubestemte former i ulike språklige kontekster. Heller enn å måle kvantitativt hvor ofte formene blir brukt i obligatoriske kontekster, siktet undersøkelse 2 seg inn mot å finne ut om morsmålet påvirker innlærerne til å danne seg assosiasjoner mellom spesifikke substantivformer og konteksttyper, altså det jeg har omtalt som *form/kontekst-assosiasjoner*. Undersøkelsen tok utgangspunkt i to ulike hypoteser om morsmålspåvirkning hos polske innlærere av norsk, som begge predikerte spesifikke form/kontekst-assosiasjoner. I hypotese 1 var valg av form antatt å være styrt av hvorvidt markering av tekstuell identifiserbarhet på substantivet blir opplevd som

redundant eller ikke, og i hypotese 2 av hvorvidt substantivfrasen var subjekt eller ikke-subjekt.

Begge hypotesene ble undersøkt i et delkorpus av tekster skrevet av polske innlærere, i tillegg til engelske innlærere og norske morsmålsbrukere, med 41 tekster fra hver av gruppene. Tekstene fra alle tre gruppene var langsvarsoppgaver fra Språkprøven i norsk for voksne innvandrere, og alle innlærertekstene var vurdert til å ligge på B1-nivå.

Metoden for å identifisere de fire transfereffektene var mer eller mindre lik for begge hypotesene (se 8.5), og innebar for det første vurdering av form/kontekst-assosiasjoner i gruppen som helhet uten å ta i betraktning individuell variasjon, for det andre en vurdering av form/kontekst-assosiasjoner i individuelle profiler, og for det tredje en oppdeling av materialet etter om det var snakk om indefinit, definit eller generisk referanse. Dette ble gjort for å kunne skille ut områder der prediksjonene om polsk morsmålsamsvar tilfeldigvis sammenfaller med de ventede formvalgene for de to andre gruppene, og dermed synliggjøre de kontekstene hvor den polske gruppen kunne ventes å skille seg ut om hypotesen stemte. Alle forekomstene ble dermed delt inn etter om de var *sammenfallskontekster* eller *markørkontekster*. Markørkontekstene tydeliggjorde områdene hvor man kunne forvente *intergruppe-heterogenitet* mellom de to innlærergruppene, og de tydeliggjorde hvor prediksjonene for den polske gruppen ikke sammenfalt med målspåkets typiske formvalg, noe som bidro til å utelukke at språklige valg som var i tråd med prediksjonene ikke alternativt kunne forklares som målspåkstilegnelse.

11.3.1 Funn for hypotesene i undersøkelse 2

Undersøkelse 2 som helhet frambrakte ingen argument for transfer hos de polske innlærerne, verken av den typen som hypotese 1 eller hypotese 2 handler om. Derimot hadde den polske gruppen i stor grad *målspåkslike* formvalg innen hver av de tre referensielle typene, som i norsk målspåk innebærer ulike formvalg: ubestemte former ved indefinit referanse, bestemte former ved definit referanse og begge deler ved generisk referanse. Selv om det var sterke tendenser til målspåkslike valg, var det fremdeles ikke helt målspåkslikt, men undersøkelsen kastet ikke lys over om de forskjellene som forelå, kunne ha sitt opphav i morsmålet.

De negative funnene for transfer kan ha flere årsaker, og jeg vil under oppsummere funnene og trekke fram ulike forhold som kan forklare dem. Jeg vil først ta for meg hypotese 1 og de forholdene som er spesifikke for denne hypotesen, deretter hypotese 2, og til slutt vil jeg ta for meg det som er felles for de to hypotesene siden de i stor grad er undersøkt med samme datamateriale og samme metode.

11.3.2 Hypotese 1 – oppsummering og diskusjon av resultater

Hypotese 1 om morsmålpåvirkning fra polsk var formulert som følger:

De polske innlærerne vil overføre morsmålets bruk av pragmatisk ordstilling som signal om tekstuell identifiserbarhet, og en oppfatning av at markering av identifiserbarhet er overflødig når pragmatisk ordstilling brukes. De norske bestemthetsmarkørene vil bli tatt i bruk når identifiserbarhet ikke går fram av ordstillingen.

Denne hypotesen predikerte form/kontekst-assosiasjoner med utgangspunkt i om markering av tekstuell identifiserbarhet kunne oppfattes som redundant eller ikke. Dette innebærer at bestemt form og ubestemt artikkel blir utelatt når identifiserbarheten går fram gjennom at ordstillingen følger tema/rema-prinsippet, men at disse formene blir brukt når identifiserbarheten *ikke* går fram av ordstillingen fordi den bryter med tema/rema-prinsippet. Oppfatning av redundans var dermed et produkt av substantivenes identifiserbarhet og plassering i setningen, der identifiserbarhet (tema/rema) ble operasjonalisert som aktivert/ikke-aktivert, og setningsposisjon som forfelt/slutfelt. Det var dermed ventet å finne tre ulike form/kontekst-assosiasjoner i det polske innlærermaterialet om hypotesen stemte:

1. Ubestemt entall/flertall og aktiverte forfelt/ikke-aktiverte slutfelt
2. Bestemte former og aktiverte slutfelt
3. Artikkelfraser og ikke-aktiverte forfelt

Disse tre prediksjonene om form/kontekst-assosiasjoner ble undersøkt én for én, både i det samlede materialet og innen hver av de referensielle typene. Hovedfunnene var for det første at det *ikke* var indikasjoner på transfer av den typen som hypotese 1 handler om, og for det andre at det var stor grad av målspårslighet i det polske materialet.

Hypotese 1 ble avvist på grunnlag av *en kombinasjon* av manglende morsmålsamsvar og manglende forskjeller mellom den polske gruppen og begge de to andre gruppene. Den typiske tendensen var at i tilfeller der det var en viss grad av

morsmålssamsvar i den polske gruppen, altså at det forelå form/kontekst-assosiasjoner som var i tråd med prediksjonene, så hadde de andre to gruppene de samme tendensene. De områdene der den polske gruppen faktisk skilte seg fra de andre to gruppene, var det derimot *lavt* morsmålssamsvar, altså språklige valg som ikke var i tråd med prediksjonene. Det var også generelt sett få positive individuelle profiler i den polske gruppen. Det var med andre ord ingen av de predikerte form/kontekst-assosiasjonene som hadde positive funn for samtlige fire transfereffekter. Som jeg var inne på i 2.4.4, mener Jarvis (2010: 181) at det ikke er nødvendig å finne alle fire effektene, men at det er tilstrekkelig å støtte seg på en kombinasjon av minst én gruppebasert og én kildebasert transfereffekt.⁵⁷ I mine data fant jeg ingen slike kombinasjoner, da gruppebaserte og kildebaserte effekter bare forekom hver for seg, og aldri sammen.

Vurderingen av markørkontekster og sammenfallskontekster viste også at det typisk var *høyt morsmålssamsvar* og lite forskjell mellom gruppene i sammenfallskontekster, men *lavt morsmålssamsvar* og lite forskjell mellom gruppene i markørkontekster. Siden det var i markørkontekstene at den polske gruppen var ventet å skille seg ut, var et argument for transfer spesielt avhengig av at det var morsmålssamsvar og forskjeller fra de andre gruppene i nettopp disse kontekstene, noe det altså ikke var. Den samlede vurderingen av de fire transfereffektene ga dermed ikke argument for transfer hos de polske innlærerne, slik transfer ble definert gjennom hypotese 1.

11.3.2.1 *Diskusjon av hypotese 1*

Utgangspunktet for hypotese 1 var fenomenet *kommunikativ redundans*, som av flere har blitt trukket fram som en årsak til manglende bruk av artikler i engelsk innlærerspråk (Young 1996; Jarvis 2002; Trenkic 2009; Trenkic & Pongpaioj 2013). I min egen undersøkelse valgte jeg å ta utgangspunkt i Youngs (1996) studie fordi denne studien knyttet *både* tilstedeværelse og fravær av artikler i innlærerspråk til redundans: Artikler oppleves som redundant i noen språklige kontekster, men ikke i andre. Når jeg ikke har funnet noe i mitt materiale som tyder på at det er slike redundansforhold som styrer formvalgene hos de polske innlærerne, er det nærliggende å avvise transfer knyttet til

⁵⁷ De gruppebaserte transfereffektene er intragruppe-homogenitet og intergruppe-heterogenitet, mens de kildebaserte transfereffektene er morsmålssamsvar og intralingval kontrast (Jarvis 2010: 181).

pragmatisk ordstilling og oppfatning av kommunikativ redundans hos polske innlærere. Som jeg var inne på i 8.1.2, fant heller ikke Young (1996) støtte for alle delhypotesene sine.⁵⁸ Den eneste delhypotesen som stemte helt hos Young, var at artikler ikke ble brukt ved *finale rema*, altså den ene av de to kontekstene som skulle mangle artikkel. Utover dette fant Young at bestemte artikler ble brukt ved tematiske substantivfraser *uavhengig av plassering* i setningen, mens ubestemte artikler ikke kunne knyttes til forholdene som ble undersøkt i det hele tatt (Young 1996: 168).

Det er det samlede materialet mitt, altså når det ikke er delt opp i referensielle typer, som er mest sammenlignbart med Youngs materiale, siden Youngs studie ikke undersøker ulike referansetyper hver for seg (Young 1996: 150). Funnene mine fra det samlede materialet ser ut til å i stor grad være i tråd med funnene i Youngs studie: Jeg fant for det første at det tydeligste morsmålsvarslet var for assosiasjonen mellom UB-kontekst og ubestemt entall/flertall, men det viste seg å være ulikt for de to typene UB-kontekster: Det var først og fremst ikke-aktiverte slutfelt (*finale rema*) som manglet markering, mens aktiverte forfelt (*ikke-finale tema*) hadde bestemt form, noe som sammenfaller med Youngs funn. I likhet med Young fant også jeg at bestemte former var assosiert med aktiverte substantivfraser (*tema*) *uavhengig av plassering* i setningen. I tillegg var assosiasjonen mellom ikke-aktiverte forfelt (*ikke-finale rema*) og artikkelfraser den av prediksjonene som fant minst støtte i min undersøkelse, noe som også er i tråd med Young.

Metoden for å koble sammen form og kontekster i min og Youngs studie er ikke lik, da Young bruker det statistiske programmet VARBRUL, og det er også noen forskjeller i operasjonaliseringen av begrepene *tema/rema* og *final/initial* (se 8.1.2.1). Det er likevel grovt sett de samme kategoriene, og at funnene i stor grad peker i samme retning, er interessant, spesielt fordi jeg i motsetning til Young har konkludert med at funnene *ikke* viser transfer fra morsmålet. Young (1996: 171-172) antyder nemlig at hans funn viser transfer fra morsmålet, der bruk av engelske artikler ser ut til å fungere som signal om brudd på *tema/rema*-ordstilling, om enn et svakt signal (Young 1996: 172). Når funnene i min undersøkelse derimot ser ut til å tale imot både transfer og

⁵⁸ Som nevnt tidligere hadde Young (1996) *delhypoteser*, der jeg har hatt *delprediksjoner*, men Youngs hypoteser sammenfaller stort sett med mine prediksjoner.

varierende redundansoppfatning som styrende for formvalg, har dette ikke bare med funnene i det samlede materialet mitt å gjøre, som altså ligner på Youngs funn, men med at jeg gjorde ytterligere metodiske grep for å kunne identifisere transfer: Sammenligningen med kontrollgrupper både bestående av +GB-innlærere og norske morsmålsbrukere, i tillegg til en oppdeling i referensielle typer, bidro til å avdekke at de transferlignende tendensene hos de polske innlærerne ikke skilte seg særlig fra det en fant i målspråksgruppen, heller ikke i markørkontekster, hvor den polske gruppen var ventet å skille seg fra de andre to gruppene ved transfer. Disse metodiske grepene har kanskje kunnet sette funnene i et annet lys enn det Youngs metode gjorde, ettersom Young kun så på form/kontekst-assosiasjoner hos -GB-innlærere uten å sammenligne med andre grupper, og uten å dele inn i ulike referensielle typer.

De delene av undersøkelsen av hypotese 1 som så på referensielle typer hver for seg, avdekket videre at formvalgene som kom fram i det samlede materialet, kunne se ut til å bare være en avspeiling av at de polske innlærerne i stor grad valgte form i tråd med hvilken referanse substantivfrasen hadde, og dessuten i tråd med reglene som gjelder i norsk: ubestemte former ved indefinit referanse, bestemte former ved definit, og begge deler ved generisk. Sammenlignet med Youngs studie, har altså min undersøkelse klart å vise at tendensene som ble observert i tekstene til -GB-innlærere, heller kan være en avspeiling av målspråkslike formvalg enn av transfer.

Etter å ha oppsummert funn for hypotese 2, vil jeg komme nærmere tilbake til en diskusjon av den type transfer som hypotesen la til grunn, altså transfer som resultat av manglende tverrspråklig identifikasjon, eller *transfer to nowhere*.

11.3.3 Hypotese 2 – oppsummering og diskusjon av resultater

Hypotese 2 om morsmålspåvirkning fra polsk var formulert som følger:

De polske innlærerne vil overføre kasus fra morsmålet, og norsk substantivmorfologi vil bli tatt i bruk for å uttrykke kasusdistinksjoner.

Hypotese 2 innebar at norsk substantivmorfologi ville bli brukt for å uttrykke kasus i innlærerspråket, slik det gjør i morsmålet. I utformingen av prediksjoner baserte jeg meg på skillet mellom funksjonelt umarkerte og markerte kasustrekk, altså nominativ versus

ikke-nominativ. Siden jeg begrenset meg til å bare se på subjekter og objekter, var det to ulike delprediksjoner for det polske innlærermaterialet, nemlig assosiasjoner mellom:

1. Ubestemt entall/flertall og subjekter
2. Bestemt entall/flertall og objekter

Også denne hypotesen ble undersøkt i et samlet materiale og innen ulike referensielle typer, og i likhet med for hypotese 1 var den overordnede konklusjonen at det ikke foreligger transfer av den typen som hypotesen handler om. Det kom derimot til syne stor grad av målspråkslikhet i det polske materialet, slik det gjorde også ved hypotese 1.

Hypotese 2 ble avvist på grunnlag av både manglende morsmålssamsvar og manglende forskjeller mellom den polske gruppen og de to andre gruppene, der den typiske tendensen var at de form/kontekst-assosiasjonene som utviste tendenser til morsmålssamsvar, også forelå hos den engelske og den norske gruppen.

I det samlede materialet ble det observert liten grad av morsmålssamsvar og få positive profiler hos de polske innlærerne, og dessuten relativt stor likhet med de andre to gruppene. Når jeg vurderte morsmålssamsvaret i markør- og sammenfallskontekster innen hver av de referensielle typene, kom samme negative tendens fram, både for assosiasjonen *subjekt/ubestemt form* og *objekt/bestemt form*: Tendensene til morsmålssamsvar fantes kun i sammenfallskontekster, og ikke i markørkontekster. I tillegg var det lite forskjell fra den engelske gruppen og stort sammenfall med målspråksgruppen både i sammenfalls- og markørkontekster. Siden argumentet for transfer hviler spesielt på morsmålssamsvar og forskjeller fra de andre to gruppene i markørkontekster, er disse funnene negative for transfer.

Den samlede vurderingen av de fire transfereffektene ga dermed ikke argument for transfer hos de polske innlærerne, slik transfer ble uttrykt gjennom hypotese 2. Som nevnt over, kunne det vært tilstrekkelig å påvise en kombinasjon av minst én gruppebasert og én kildebasert transfereffekt, men selv om slike effekter ble påvist hver for seg, ble de ikke påvist i kombinasjon for noen av form/kontekst-assosiasjonene. De språklige mønstrene i den polske gruppen kunne derfor ikke tolkes som transfer.

11.3.3.1 *Diskusjon av hypotese 2*

Hypotese 2 tok utgangspunkt i grammatisk kasus som grammatisk kategori, og hvordan grammatiske kategorier legger føringer på hvilke innholdselementer man må ta stilling til når man formulerer seg på morsmålet. Ved transfer av et *kasus-perspektiv*, vil de aktuelle innholdselementene for kasus også måtte komme til uttrykk i innlærerspråket, og dette antas å skje ved hjelp av de morfologiske bestemthetsmarkørene i norsk.

Om hypotese 2 stemmer, var det ventet at de polske innlærerne ville ta i bruk ubestemte former i de språklige kontekstene som er forbundet med nominativ i polsk, og bestemt form i de øvrige kontekstene, i praksis kun objekt i min undersøkelse.

Når jeg ikke har funnet noe som tyder på at valg av form hos de polske innlærerne er styrt av om det er subjekter eller objekter, virker det naturlig å avvise denne hypotesen om morsmålpåvirkning. Det er imidlertid flere forbehold med måten undersøkelsen er lagt opp på, og jeg vil under spesielt trekke fram noen sider ved måten jeg har forholdt meg til den morfologiske og innholdsmessige kompleksiteten ved polske kasus på, og hvordan dette ble antatt å uttrykkes ved hjelp av norske bestemthetsmarkører.

Siden det polske kasussystemet ifølge hypotesen skulle komme til uttrykk gjennom de få formene som fins i norsk, måtte jeg basere prediksjonene mine på skillet mellom funksjonelt umarkerte og markerte kasus, altså mellom nominativ og ikke-nominativ. Dette ble i praksis undersøkt gjennom å sammenligne subjekter og objekter i materialet. Hvis det er slik at innlærere forsøker å uttrykke kasusinnhold i innlærerspråket gjennom tilgjengelige former i norsk, er det imidlertid flere forhold enn bare subjekt versus objekt som skal uttrykkes. For eksempel har man i polsk et skille mellom negerte og ikke-negerte objekter, som har henholdsvis akkusativ og genitiv (Swan 2002: 333). Dette har ikke blitt tatt hensyn til i min undersøkelse. Det er også forbehold knyttet til at jeg bare har sett på objekter der hypotesen viser mer generelt til *ikke-subjekter*. Siden det er flere andre kasus i polsk som er relevante, kunne kategorien ikke-subjekter med fordel ha vært mer omfattende. Dette kunne kanskje ha framhevet flere nyanser i formvalgene innen ikke-subjekter. Det kunne også ha styrket koblingen mellom *kasus* og subjekt/objekt-distinksjonen ettersom en inndeling i bare subjekt og objekt kan være for vid og generell til å bare skulle representere kasus. Subjekt og objekt sammenfaller nemlig gjerne med andre forhold også, for eksempel har subjekter tendens

til å være tematiske og setningsinitiale. Dette poenget berøres også av Jarvis (2002: 411), men da for å forklare det motsatte, nemlig at artiklene til finske innlærere markerer skillet mellom tema og rema eller initiale og finale substantivfraser, noe han mener alternativt kan forklares som markering av skillet mellom subjekt og ikke-subjekt.

Hypotesen var at norsk substantivmorfologi ville bli brukt til å uttrykke forskjellen på subjekt og objekt, og jeg konkluderte med at dette ikke virket å være tilfelle i mitt materiale. Dette er imidlertid en sannhet med modifikasjoner: Det er assosiasjonene *subjekt/ubestemt form* og *objekt/bestemt form* som ikke forekommer i det polske materialet, men det ble observert flere steder i undersøkelsen at det *motsatte* så ut til å gjelde, altså tendenser til en assosiasjon mellom subjekt og bestemt form, og mellom objekt og ubestemt form. Disse «motsatte» assosiasjonene kan egentlig også indikere at formene blir brukt til å uttrykke forskjellen mellom subjekt og objekt, men ikke med de prediksjonene om *formvalg* som ble gjort i min undersøkelse. En mer åpen prediksjon, der jeg ikke hadde predikert hvilken form som skulle assosieres med hvilken kontekst, men heller at én av formtypene ville vært assosiert med én av kontekstene, ville kunne satt dette i et annet lys. Slike retningsløse prediksjoner finner en i flere andre studier (for eksempel Young 1996: 155; Ionin & Wexler 2003: 249; Ionin et al. 2004: 18; Ionin et al. 2008: 559).⁵⁹

Noe som tilsier en mer åpen prediksjon, er at hypotese 2 egentlig bare sier at de polske innlærerne vil ta i bruk de tilgjengelige språklige midlene i norsk for å uttrykke kasus (nominativ – ikke-nominativ), men ikke hvilken form som vil bli brukt til hva. Når jeg koblet ubestemt form til nominativ og bestemt form til ikke-nominativ, var dette basert på en oppfatning av at det er henholdsvis ubestemt og bestemt som er umarkert og markert i norsk (Hagen 1998: 31). Dette har flere diskutabile sider, men viktigst her er at forholdene i norsk i prinsippet ikke skal spille noen rolle ettersom hypotesen ikke baserer seg på tverrspråklig identifikasjon.

Helheten i undersøkelsen min har likevel vist at de språklige valgene til den polske gruppen er mer i tråd med målspråkets typiske formvalg enn at formene brukes til å skille mellom subjekt og objekt. Den motsatte assosiasjonen mellom

⁵⁹ De retningsløse prediksjonene i disse studiene var imidlertid ikke knyttet redundans, som i min studie, men til universalierne ±HK og ±SR. Dette gjelder også Youngs studie ettersom slike prediksjoner kun ble brukt i den delen av studien som undersøkte universalier.

subjekt/bestemt og objekt/ubestemt sto dessuten sterkere i det samlede materialet enn når materialet ble delt inn etter referensiell type. I det differensierte materialet «forsvant» nemlig disse assosiasjonene, og det ble avdekket at tendensene i det samlede materialet så ut til å egentlig bare være en avspeiling av ulike formvalg innen ulike referensielle typer (se 10.9.3). Denne konklusjonen ville nok ha vært den samme selv med mer åpne prediksjoner som ikke spesifiserer hvilken form som blir brukt til hva. Under (11.5.2) vil jeg likevel komme tilbake til spørsmålet om hvorfor og på hvilken måte bestemthetsmarkørene i norsk kan få rollen med å uttrykke kasus, men da med utgangspunkt i tverrspråklig identifikasjon mellom polsk og norsk.

11.3.4 Konklusjon for forskningsspørsmål 2

Jeg har nå oppsummert og diskutert de to hypotesene, og når det gjelder forskningsspørsmål 2, må svaret være at morsmålet ikke påvirker distribusjonen av bestemte og ubestemte former i innlærerspråket til polske innlærere *på den måten som hypotese 1 og 2 antar at det gjør*. Undersøkelse 2 kan imidlertid kun si noe om akkurat den typen transfer som de to hypotesene tar for seg, og kan dermed ikke utelukke at morsmålet påvirker distribusjonen av former på andre måter.

11.3.5 Hypotesene og inaktiv transfer

Jeg vil nå vende tilbake til det teoretiske utgangspunktet for de to hypotesene i undersøkelse 2. Hypotesene dreier seg om to ulike former for transfer fra morsmålet, men har til felles at de begge bygger på de samme ideene om transfer ved nullkontrast mellom morsmål og målspråk: Nullkontrasten gjør at det ikke skjer en tverrspråklig identifikasjon mellom språklige elementer i polsk og norsk, og dette fører for det første til at innlærerne antar at norsk fungerer på samme måte som polsk, og for det andre til at de norske bestemthetsmarkørene blir tatt i bruk i innlærerspråket for å uttrykke språklige forhold som er sentrale i polsk.

Denne typen transfer har i løpet av avhandlingen blitt omtalt i forbindelse med flere ulike, men nært relaterte, fenomener fra faglitteraturen, nærmere bestemt *transfer som inaktivt utfall* (Jarvis 2000), *transfer to nowhere* (Kellerman 1995) og transfer som følge av *thinking for speaking* (Slobin 1996). Disse viser alle tre til at de perspektivene eller innholdselementene som ligger til grunn for visse strukturer i morsmålet også

ligger til grunn når man bruker et annet språk. Alle viser dessuten til at måten visse kategorier i målspråket fungerer på, er nærmest utilgjengelig eller forblir uanalysert av innlærerne, og at målspråkstrukturer får rollen med å uttrykke perspektivene som overføres fra morsmålet. Siden de tre fenomenene henger nært sammen, vil jeg videre vise til dem samlet under benevnelsen *inaktiv transfer*, som er ment å favne de to viktigste aspektene av disse tilnærmingene: For det første at morsmålsperspektivene ikke så mye overføres aktivt, men heller fortsetter å ligge til grunn selv om man bruker et annet språk, og for det andre at det ikke er noen aktiv prosess med tverrspråklig identifikasjon mellom morsmål og målspråk.⁶⁰

Når det i undersøkelsen av de to hypotesene ikke er funnet transfer, er dette kanskje en indikasjon på at det ikke foreligger noen slik inaktiv transfer, i hvert fall ikke av den typen som hypotesene predikerte. Likevel kan det selvsagt også være andre forhold på veien mellom hypotesedanning og tolkning av resultater som har bidratt til negative funn. Forhold som knytter seg til metode, vil jeg komme tilbake til senere (11.3.7); jeg vil først rette fokus mot selve hypotesene. Siden det ikke ble funnet transfer hos de polske innlærerne, er det vanskelig å gå inn i en drøfting av transfer relatert til selve resultatene, så jeg vil heller ta diskusjonen på et mer generelt grunnlag, ment for å belyse *hypotesene*, ikke funnene.

I de to hypotesene er det til sammen *tre* forhold fra polsk som antas å utgjøre morsmålspåvirkning i form av inaktiv transfer til innlærerspråket: pragmatisk ordstilling og oppfatninger om identifiserbarhet og redundans for hypotese 1, og innholdsdistinksjoner knyttet til grammatisk kasus for hypotese 2. Jeg vil først gjenta hvordan jeg i utformingen av hypotesene har sett for meg at disse er relatert til inaktiv transfer. Deretter vil jeg trekke fram noen spørsmål knyttet til dette, for eksempel om hypotesene egentlig handler om inaktiv transfer, og hvordan forholdet mellom inaktiv transfer og et siste relevant fenomen er, nemlig *konseptuell transfer*.

⁶⁰ Betegnelsen «inaktiv transfer» ligner på det som jeg tidligere har omtalt som «transfer som inaktivt utfall», som er en oversettelse av det Jarvis (2000: 250) kaller «transfer as inert outcome». Betegnelsen min er basert på samme tankegang, men er her ment som et felles navn for flere relaterte fenomener, deriblant transfer som inaktivt utfall.

Overføring av pragmatisk ordstilling og morsmålpåvirket identifiserbarhetsbegrep

Hypotese 1 er knyttet til inaktiv transfer på følgende måte: I polsk er bruk av pragmatisk ordstilling sentralt for å organisere setninger, og når ordstillingen er i tråd med tema/rema-prinsippet, vil det gå fram av plasseringen om en referent er *kjent/identifiserbar* eller *ny/ikke-identifiserbar* (se 3.3). Om en innlærer antar at norsk fungerer på samme måte som polsk, vil denne måten å organisere setninger på videreføres til og ligge til grunn for innlærerspråket også. Dette innebærer *to* antakelser om hvordan norsk er likt polsk: For det første at identifiserbarhet kan signaliseres gjennom ordstilling, og for det andre at referenter er identifiserbare kun når de er nevnt før. En overføring av pragmatisk ordstilling som signal om identifiserbarhet innebærer med andre ord også et morsmålpåvirket identifiserbarhetsbegrep.

Undersøkelsen av hypotese 1 var ikke innrettet mot pragmatisk ordstilling i seg selv, men heller mot to direkte konsekvenser av en slik overføring, nemlig at markering av identifiserbarhet oppleves som redundant når tema/rema-prinsippet følges, og for det andre at identifiserbarhetsbegrepet er definert ut fra om referenter er nevnt før eller ikke (se 8.1.2). Redundansoppfatningen vil dermed være styrt av det kriteriet for identifiserbarhet som ligger til grunn for pragmatisk ordstilling.

For hypotese 1 vil inaktiv transfer innebære at innlærerne tar i bruk norske bestemthetsmarkører for å opprettholde disse perspektivene som ligger til grunn for pragmatisk ordstilling og identifiserbarhet. Ifølge hypotesen vil dette komme til uttrykk gjennom en kombinasjon av pragmatisk ordstilling og bestemthetsmarkører: Når ordstillingen følger tema/rema-prinsippet, er dette et tilstrekkelig signal om identifiserbarhet og ytterligere markering oppleves som redundant, men når ordstillingen bryter med dette, tas bestemthetsmarkørene i bruk. Bruk av markører vil dermed styres av 1) om det er brudd på tema/rema og 2) om referenten er nevnt før eller ikke, og da er vi ved hypotesens prediksjoner.

Overføring av grammatisk kasus

Hypotese 2 var knyttet til inaktiv transfer på følgende måte: I polsk er grammatisk kasus en obligatorisk del av substantivfrasen, obligatorisk på den måten at språkbrukere må ta stilling til kasusinnhold og velge bøyningsform etter dette i enhver ytring. Slike obligatoriske grammatiske perspektiv er som jeg var inne på i kapittel 8.1.3, typiske

kandidater til inaktiv transfer i kraft av å være automatiserte og vanskelig å endre på eller å la være å uttrykke når en skal lære et nytt språk (Slobin 1996:89). Om en innlærer antar at norsk fungerer på samme måte som polsk, vil kasus kunne overføres og dermed ligge til grunn for formvalg i innlærerspråket også.

For hypotese 2 vil inaktiv transfer innebære at de polske innlærerne tar i bruk norske bestemthetsmarkører for å opprettholde kasusenes grammatiske perspektiv i innlærerspråket. De ulike formene av substantivet i norsk får dermed rollen med å uttrykke kasusdistinksjoner som er relevante for morsmålet. I operasjonaliseringen av hypotesen ble det framhevet at innhold som har å gjøre med kasus i polsk, har mange trekk og nyanser, og at det derfor virker usannsynlig å skulle kunne uttrykke alt slikt innhold ved hjelp av ubestemte og bestemte former. Jeg fokuserte derfor på skillet mellom funksjonelt umarkert kasus (nominativ) og funksjonelt markerte kasus (ikke-nominativ), som ble antatt å bli uttrykt ved hjelp av henholdsvis ubestemte og bestemte former i innlærerspråket.

Er hypotesene egentlig hypoteser om inaktiv transfer?

Over har jeg forsøkt å beskrive hvordan jeg har sett for meg at de to hypotesene er koblet til inaktiv transfer. Det er likevel grunn til å stille spørsmålsteget ved om denne vurderingen er riktig, altså om det faktisk er slik at disse hypotesene handler om inaktiv transfer, og om de er utformet på en måte som er egnet til å finne slik transfer.

Hypotesene ble i utgangspunktet formulert for å kunne besvare forskningsspørsmål 2, som handler om hvorvidt morsmålet påvirker distribusjonen av bestemte og ubestemte former. Inaktiv transfer har med andre ord blitt hentet fram for å forklare hvordan bestemthetsmarkører blir brukt i innlærerspråket. Dette er nærliggende ettersom inaktiv transfer blant annet er et resultat av at språklige kategorier i målspråket er utilgjengelige eller ikke oppfattes, og at én av konsekvensene av inaktiv transfer er at språklige strukturer i målspråket tas i bruk for å uttrykke inaktivt overførte perspektiv fra morsmålet. Bestemthetsmarkører kan ha både den første og den andre funksjonen. Det kan likevel være at et slikt utgangspunkt for undersøkelsen begrenser både selve fenomenet inaktiv transfer og mulighetene til å utforske det, sammenlignet med om jeg i stedet for å ta utgangspunkt i målspråkkategorien, hadde begynt med forhold i morsmålet. Dette kunne for eksempel ha vært å se etter kategorier eller andre trekk ved

morsmålet som kunne tenkes å bli overført som inaktiv transfer, og deretter undersøkt om og hvordan dette kommer til uttrykk i innlærerspråket. Min undersøkelse har rettet seg mot bare én spesifikk uttrykksmåte, mens inaktiv transfer like godt kan ventes å manifestere seg gjennom en kombinasjon av uttrykksmåter, og gjennom helt andre strukturer enn bestemthetsmarkører.

Bestemthet blir nevnt av både Kellerman (1995: 141) og Slobin (1996: 88-89) som en relevant kategori i forbindelse med inaktiv transfer fra -GB-morsmål. Det virker først og fremst å være kategoriens utilgjengelighet for -GB-innlærere som framheves og som anføres som utløsende for inaktiv transfer. Det er likevel forskjell på at grammatisk bestemthet kan være utløsende på denne måten, og det at bestemthetsmarkørene skal få funksjonen med å uttrykke det som blir overført fra morsmålet.

På den andre siden er det kanskje nettopp dette utilgjengelige ved bestemthetsmarkørene som kan gjøre dem til et mål for inaktiv transfer. Når man ikke vet hva markørene uttrykker, blir de stående uten funksjon og man trenger å fylle dem med innhold, for eksempel med sentrale morsmålskategorier eller -perspektiv som blir med over til innlærerspråket som inaktiv transfer. Om det er slik at disse to tingene, inaktiv transfer og en utilgjengelig målspråkskategori, opererer sammen, slik det ser ut til å være ifølge faglitteraturen, kan det være uklart hva som skjer først: Er det morsmålskategorien som overføres inaktivt (*transfer to nowhere*), eller er det strukturer i målspråket som trekker til seg et innhold? Kanskje er det her snakk om en type *transfer from elsewhere*, hvor det er målspråket som tar, heller enn morsmålet som gir.

At grammatisk bestemthet blir trukket fram av Kellerman (1995) og Slobin (1996), er kanskje et resultat av en engelskspråklig orientering innen andrespråksfeltet, eventuelt en orientering mot +GB-målspråk i studier som tar for seg bestemthet. Grammatisk bestemthet er en sentral kategori i engelsk og utgjør en velkjent utfordring for -GB-innlærere, som delvis står uforklart. Det samme gjelder denne kategorien i norsk. På grunn av dette blir nok bestemthet i denne sammenhengen først og fremst relevant som et nedslagsfeltet for inaktiv transfer fra et -GB-språk, men det motsatte er jo også aktuelt, nemlig at bestemthetskategorien selv overføres inaktivt når +GB-innlærere lærer *et -GB-språk*.

Det har ikke lyktes foreliggende studie å finne ut at det faktisk er inaktiv transfer som bidrar til utfordringer med grammatisk bestemthet eller til distribusjonen av former. At jeg har målspråkets kategori som startpunkt for hypotesene i studien er kanskje et argument for at hypotesene ikke egentlig retter seg mot inaktiv transfer. I alle fall kan man si at et slikt målspråkbasert startpunkt har begrenset treffsikkerheten til hypotesene hva angår både å *finne* inaktiv transfer og å kartlegge utslaget av det.

Litt lenger ned (11.5.2) vil jeg komme tilbake til spørsmålet om hvor rimelig det er å anta at det er nettopp bestemthetsmarkørene som får rollen med å uttrykke innhold som er overført inaktivt.

Handler hypotesene også om konseptuell transfer?

Inaktiv transfer berører gjennom sine ulike varianter det som i kapittel 2.3.6 ble omtalt som *konseptuell transfer*, en type transfer som står i motsetning til lingvistisk transfer, og som skjer som resultat av tverrspråklige *begrepsmessige* forskjeller og likheter heller enn tverrspråklige *språklige* likheter og forskjeller (Jarvis & Pavlenko 2008: 61). Selv om hypotesene mine berører forhold som kan komme inn under konseptuell transfer, har det ikke vært noe eksplisitt mål med undersøkelsen å utforske denne typen transfer, og hypoteseutformingen er heller ikke preget av dette. Likevel er det høyst relevant å spørre seg om ikke den typen transfer som hypotesene framstiller, kan kalles konseptuell.

Nullkontrasten knyttet til grammatisk bestemthet kan kanskje delvis innebære begrepsmessige forskjeller mellom norsk og polsk. Polsk har ikke grammatisk bestemthet, men signaliserer likevel identifiserbarhet gjennom ordstilling. Organisering i setninger etter tema/rema-prinsippet utgjør i seg selv ikke en begrepsmessig forskjell, både fordi de underliggende «begrepene» (tema/rema) fins i begge språk, men også fordi tema/rema-prinsippet bidrar til ordstillingen i norsk også, selv om den først og fremst er grammatisk styrt (Hagen 1998: 259-260). Forståelsen av *identifiserbarhet* kan derimot, som vi så i 3.3, være ulik i de to språkene da pragmatisk ordstilling innebærer et smalere identifiserbarhetsbegrep enn grammatisk bestemthet gjør. Dette *kan* utgjøre en forskjell på begrepsnivå hos norske og polske språkbrukere, noe for eksempel Ekiert (2010) foreslår (se 4.2.2.3). Samtidig representerer dette skillet også en *lingvistisk* forskjell mellom språkene siden det ene uttrykkes syntaktisk og det andre morfologisk.

Nullkontrasten knyttet til grammatisk *kasus* utgjør heller ikke nødvendigvis noen egentlig begrepsmessig forskjell mellom norsk og polsk, da innhold som uttrykkes av *kasus* også kan komme til uttrykk i norsk, bare ikke gjennom obligatorisk bøyning. Et eksempel på dette er at norsk gjennom å ha grammatisk ordstilling blant annet bruker ordstillingen til å eksponere syntaktisk funksjon (Hagen 1998: 259). Igjen representerer forskjellen mellom polsk og norsk vel så mye en *lingvistisk* forskjell. Denne lingvistiske forskjellen mellom polsk og norsk kan imidlertid knyttes til en forskjell som er relevant for konseptuell transfer. Det at *kasus* er en grammatikalisert kategori, gjør at den kanskje kan knyttes til det som Jarvis (2007) kaller '*conceptualization transfer*', som er en type konseptuell transfer som ikke handler om hvordan begreper er organisert i langtidsmminnet, men heller om hvordan de organiseres og hentes fram i arbeidsminnet for å formidle en spesifikk situasjon:

The primary difference between concept transfer and conceptualization transfer, again, is that the former results from the makeup of the inventory of concepts in a person's long-term memory, whereas the latter involves the process of selecting specific concepts from long-term memory, calling them up into working memory, and combining them dynamically in various orders[...] (Jarvis 2007: 54)

Forskjellen knyttet til identifiserbarhet i -GB-språk og +GB-språk har også potensial til å handle om slike forhold. Trenkic (2002: 114) omtaler forskjellen mellom identifiserbarhet ved pragmatisk ordstilling og grammatisk bestemthet som en forskjell mellom en «bevissthetskategori» og en «kunnskapskategori» (se 3.3). Dette innebærer at de to måtene å forholde seg til identifiserbarhet på også kan handle om hvor man henter informasjonen om en referents identifiserbarhet fra: Er det fra arbeidsminnet, gjennom at referenten nylig har blitt nevnt (bevissthetskategori), eller fra langtidsmminnet, der man har lagret kunnskap om hvilke referenter som er unike eller forekommer i visse situasjoner (kunnskapskategori)?

Det kan altså se ut til at det er noen aspekter av hypotesene mine som kan knyttes til mulige begrepsmessige eller kognitive forskjeller mellom norsk og polsk, og som dermed har potensiale til å kunne kobles til konseptuell transfer. Når jeg likevel har latt være å ta mine hypoteser om inaktivt utfall så langt at de også handler om konseptuell transfer, og heller har fokusert på det jeg har sett på som mer nøytralt med hensyn til om

det er lingvistisk eller konseptuelt, skyldes det at konseptuell transfer er vanskelig å få et fast grep på. Jarvis (2011: 4) medgir at det nok er umulig å vise helt sikkert at noe er konseptuell transfer: «[...] it is, of course, impossible to demonstrate conclusively whether cross-linguistic effects have indeed arisen from the conceptual level.» Jarvis skriver videre at man bør være forsiktig i sin analyse og alltid velge den mest konservative konklusjonen om man er i tvil, altså at det er lingvistisk transfer heller enn konseptuell. Jeg har med andre ord valgt et slikt *konservativt* utgangspunkt allerede ved utformingen av hypotesene, og jeg vil ikke bevege meg inn i et mindre konservativt alternativ for konklusjonen, spesielt ikke når det tross alt var negative funn for begge hypoteser.

Slik jeg ser det, er det minst ett forhold som må avklares før man kan begynne å snakke om hypoteser eller funn av konseptuell transfer, og egentlig også for å kunne snakke om inaktiv transfer: Er transferen man finner eller framsetter slike hypoteser om, faktisk et resultat av manglede tverrspråklig identifikasjon? Mangel på tverrspråklig identifikasjon har nemlig blitt framstilt som helt sentralt for inaktiv transfer, men hvorvidt det faktisk er dette som ligger til grunn for tilfeller av transfer, må regnes som et empirisk spørsmål i seg selv, og kanskje et som ikke er så lett å finne ut av.

Spørsmålet om hvorvidt det er manglende tverrspråklig identifikasjon som ligger til grunn for transfer eller ikke, vil jeg komme tilbake til lenger ned i sammenheng med en vurdering av om hypotesene i denne studien like godt kunne hatt sitt grunnlag i antakelser om tverrspråklig identifikasjon mellom strukturer i norsk og polsk. Om dette skulle være tilfelle, vil det være mindre sannsynlig at hypotesene handler om inaktiv transfer, og spørsmålet om konseptuell transfer blir da mindre relevant.

Før jeg kommer tilbake til dette, vil jeg se nærmere på funnene av målspråkslike formvalg hos de polske innlærerne, og deretter se på om noe ved metoden for undersøkelsen kan belyse de negative funnene.

11.3.6 Målspråksnære formvalg

De relativt målspråkslike tendensene som har kommet fram i det polske materialet ved begge hypoteseundersøkelsene, utelukker i prinsippet ikke transfer, da det kan være snakk om *positiv transfer*. I flere av kontekstene sammenfaller prediksjonene om polsk morsmålssamsvar med det typiske formvalget i målspråket, så når formvalget stemmer

med både prediksjonene og er målspråkslike, kan det bety at morsmålet og målspråket drar formvalget i samme retning, og at det kan være snakk om positiv transfer. Dette gjelder imidlertid bare i sammenfallskontekster. Det ble derimot observert stor grad av sammenfall med målspråksgruppen også i *markørkontekster*, hvor gruppene var forventet å ha ulike formvalg om hypotesen stemte. I markørkontekstene var formvalgene derfor først og fremst målspråkslike, og *ikke* i tråd med prediksjonene. I lys av dette ble det derfor naturlig å tolke morsmålssamsvaret i sammenfallskontekster som en generell målspråkslikhet som kan skyldes *målspråkspåvirkning*, og ikke som positiv transfer.

Sammenfallet med målspråket innebar også at det i noen deler av undersøkelsen har vært vanskelig å skille morsmålspåvirkning fra slik målspråkspåvirkning. Det har vært mulig å skille ut markørkontekster for alle de tre referensielle typene, men hvilke kontekster som har vært til stede i materialet, har i deler av undersøkelsene ikke lagt til rette for å undersøke slike markørkontekster. Dette gjaldt spesielt generisk referanse, hvor det var veldig få markørkontekster. Ved denne typen referanse var altså ikke materialet egnet til å isolere transfer som årsak til formvalgene. Sett i sammenheng med at funnene for markørkontekster innen de andre referansetyperne stort sett talte imot transfer, er det likevel grunn til å tro at det samme ville ha vært tilfelle om materialet for generisk referanse hadde inneholdt flere markørkontekster.

11.3.7 Metodiske forbehold for undersøkelse 2

Undersøkelse 2 har noen metodiske sider som bør problematiseres, både knyttet til operasjonalisering og vurdering av form/kontekst-assosiasjonene, og til utvalget av tekster og substantivfraser.

11.3.7.1 Operasjonalisering av konteksttyper

Begge hypotesene tok utgangspunkt i at visse former skulle bli brukt i visse språklige kontekster, og måten disse konteksttypene har blitt operasjonalisert eller delt inn på kan ha påvirket til negative funn i undersøkelsen. Mens hypotese 2 tok for seg konteksttypene *subjekt* og *objekt*, som er relativt standardiserte størrelser innen grammatikken, er utgangspunktet for konteksttypene i hypotese 1 mindre entydige.

Hypotese 1: kommunikativ redundans

Bestemthetsmarkørers kommunikative redundans var sentral for hypotese 1, og var et produkt av substantivfrasens plassering i setningen og referentens tekstuelle identifiserbarhet. *Setningsposisjon* (initial/final) ble operasjonalisert som henholdsvis *forfelt* og *slutfelt*, mens jeg hadde utelatt *midtfelt*, altså substantivfraser som står etter det finite verbet, men før eventuelle infinite verb. I motsetning til forfeltet, som alltid er initialt, og slutfeltet, som er finalt, er midtfelt mer varierende plassert i setningen avhengig av hvilke andre ledd som er til stede: Noen ganger plasserer midtfeltet seg før et uttrykt slutfelt og dermed *medialt* i setningen, men i setninger uten slutfelt og infinitt verb kan et midtfelt være *finalt*. Denne variasjonen for midtfelt var årsaken til at jeg valgte å utelate midtfelt fra materialet for hypotese 1, men midtfeltene kunne faktisk også belyst hypotesen nettopp på grunn av denne variasjonen, så å utelate midtfelt kan dermed ha potensielt fjernet interessante forekomster fra undersøkelsen. Det var imidlertid relativt få tilfeller av midtfelt, så det er tvilsomt at utslaget hadde vært spesielt stort.

Når det gjelder identifiserbarhet (tema/rema) ble dette definert med utgangspunkt i *tekstuell aktivering*, altså hvorvidt en substantivreferent er nevnt før eller ikke. Denne definisjonen var motivert av at jeg ville ha et *morsmålsbasert identifiserbarhetsbegrep* som dermed var basert på hvordan tekstuell identifiserbarhet styrer pragmatisk ordstilling i morsmålet. Dette mener jeg er viktig for hypotese 1 ettersom den handler om transfer, men måten tekstuell aktivering har vært operasjonalisert på, har derimot noen svake sider. Som jeg var inne på i 8.4.3, regnet jeg en frases referent som tekstuel aktivert om den hadde vært nevnt tidligere *i samme hovedsetning eller i hovedsetningen før*, en metode som først og fremst ble valgt for å effektivisere arbeidet uten å måtte nærlese tekstene. Jeg hadde heller ikke noe krav om at den aktuelle substantivfrasen og det tidligere nevnte uttrykket skulle være samreferensielle, det var tilstrekkelig om en frase med samme substantivleksem og referanse til *samme type av referent* var nevnt før. Tekstuell aktivering kan imidlertid ses på som et produkt av både *referensiell distanse*, *referensiell tvetydighet* og *referensiell framredenhet*, altså for det første hvor lenge det er siden referenten er nevnt, for det andre om det er andre intervenerende referenter, og for det tredje hvor framtrede referenten er i den videre teksten (Givón

1983: 13-15). Slike forhold kunne med fordel ha blitt bedre tatt hensyn til i klassifiseringen min. Min klassifisering ser kun på referensiell distanse, altså avstanden det er til forrige gang referenten ble nevnt, men jeg opererer med en relativt *kort* referensiell distanse på bare *to* hovedsetninger bakover i teksten. Til sammenligning viser Givon (1983:13) til en øvre grense på 20 setninger ('clauses') før referenten ikke lenger regnes som tilgjengelig (aktivert).

Det er mulig at det at den norske gruppen hadde en viss overrepresentasjon av ikke-aktiverte kontekster sammenlignet med de andre gruppene, kan ha sitt opphav i en kombinasjon av definisjonen på tekstuell aktivering og *teksttema*. Tekstene i den norske gruppen har som vi så i 8.2.1, nesten alle handlet om størrelser som «sunnhet og helse» og «penger og andre verdier» og «frihet og ansvar», og det er mulig at dette påvirker typen referenter i disse tekstene. Slike abstrakte begreper er ofte ikke-tellelige, og siden ikke-tellelige substantiv er utelatt i undersøkelsen av hypotese 1, kan dette ha ført til at de begrepene som går igjen i de norske tekstene, og dermed har potensial til å regnes som «nevnt før», har blitt utelatt. Til sammenligning er tema som «mobiltelefoner», «bomiljø» og «barneoppdragelse» sterkt representert i innlærertekstene, og umiddelbart framstår dette som tema som kan fremme en større andel av tellelige referenter, uten at dette er undersøkt nærmere.

Når en hypotese om transfer er at markering utelates fordi det framgår av konteksten og derfor oppleves som redundant å markere, krever det både en morsmålsbasert og presis operasjonalisering av nøyaktig *når* det framgår av konteksten og *når* det *ikke* gjør det. Dette kan bidra til at utforskningen av fenomenet *kommunikativ redundans* blir mer nyansert og potensielt kan avdekke effekter som verken foreliggende studie eller Youngs (1996) studie har fanget opp. Av de nevnte forholdene virker det som om det er operasjonaliseringen av *tekstuell aktivering* som har størst forbedringspotensial, og som eventuelt i størst grad kan ha bidratt til tilsløre eventuell transfer i materialet mitt.

Hypotese 2: Funksjonell markerthet

Når det gjelder hypotese 2, er subjekter og objekter relativt veldefinerte størrelser innen grammatikken, men subjekter og objekter var ment å representerte henholdsvis nominativ og ikke-nominativ, og påstanden om at nominativ er det *umarkerte* kasuset

og at de øvrige er *markerte*, kan problematiseres. Jeg tok utgangspunkt i *funksjonell* markerthet, og dermed har jeg sett bort ifra hvordan de ulike kasusene faktisk realiseres i polsk. Det kunne ha vært hensiktsmessig å heller ta utgangspunkt i *formell* markerthet, altså hvorvidt de enkelte substantivene har endelse eller ikke. Dette er noe jeg vil komme tilbake til i 11.5, der jeg diskuterer en alternativ hypotese om transfer av kasus basert på tverrspråklig identifikasjon mellom substantivmorfologi i norsk og polsk.

11.3.7.2 *Operasjonalisering av formtyper*

Over har jeg problematisert noen sider knyttet til *konteksttypene* i de predikerte form/kontekst-assosiasjonene, men også inndelingen i *formtyper* har diskutabile sider. Når det gjelder substantivformer, har jeg i undersøkelse 2 operert med tre grupper: 1) ubestemt entall/flertall, 2) bestemt entall/flertall, og 3) artikkel frase, hvorav den tredje bare var inkludert for hypotese 1. Motivasjonen for å dele inn på denne måten var ulik for de to hypotesene, men i begge tilfeller utelukkende basert på prediksjonene om polsk morsmålssamsvar, og ikke utformet med tanke på sammenligningen med kontrollgruppene. Den viktigste forskjellen er kanskje at prediksjonene om polsk ikke tilsa forskjeller i formvalg mellom entall og flertall, og at disse derfor ble slått sammen i undersøkelsen. Det viste seg at det likevel hadde vært hensiktsmessig å skille også disse med tanke på sammenligningen med den engelske og den norske gruppen, spesielt for ubestemte former siden engelsk og norsk skiller seg fra hverandre i bruken av ubestemt entall ved tellelige substantiv.

Et siste problematisk punkt når det gjelder form, handler om hvordan den ubestemte flertallsendelsen i norsk (-er) er blitt behandlet. Jeg har i undersøkelse 2 først og fremst sett på denne formen som en naken eller umarkert flertallsform, et substantiv uten (u)bestemthetsmarkør, og derfor gruppert ubestemt flertall sammen med ubestemt entall, som om det kun er kategorien tall som skiller dem. Det kan likevel være slik at -er best kan ses på som en endelse som uttrykker både trekket flertall og trekket ubestemt, og derfor ikke skulle vært gruppert sammen med ubestemt entall på denne måten.

11.3.7.3 *Inndeling i referensielle typer*

En siste side av måten materialet er klassifisert og operasjonalisert på, handler om oppdelingen i referensielle typer. Prediksjonene for polsk morsmålssamsvar var ikke

antatt å skulle være påvirket av referensiell type, så denne inndelingen var motivert av behovet for å framheve språklige kontekster hvor den polske gruppen ville skille seg fra de andre gruppene.

Inndelingen etter referensiell type var basert på hvordan disse fordeler seg i norsk og engelsk, men det kom fram i undersøkelsen at det kunne ha vært fordelaktig med en ytterligere inndeling for generisk referanse, ettersom språkene norsk og engelsk skiller seg fra hverandre her. Det ville imidlertid blitt mer omfattende å klassifisere disse undertypene, og som vi så i kapittel 7, er det ikke noe entydig en-til-en-forhold mellom form og generisk undertype verken i norsk eller engelsk, så det er usikkert om dette hadde gjort noen stor forskjell.

11.3.7.4 *Oppsummering av metodiske forbehold*

Det har altså vist seg å være to områder hvor oppdelingen av materialet i undersøkelse 2 kunne vært mindre grovinndelt, nemlig formtyper og referensielle typer. Hovedsakelig handler dette om at jeg har delt inn materialet etter prediksjonene for polsk morsmålssamsvar, og ikke i tilstrekkelig tatt hensyn til alle forhold som kan være relevante for sammenligningen med den norske og engelske gruppen. Mindre grove inndelinger kunne i større grad belyst spørsmålet om hvorvidt den polske gruppen er forskjellig fra de andre gruppene, fordi det enda tydeligere ville framhevet markørkontekster. Det er usikkert hvor stor forskjell det ville gjort i min undersøkelse, men det er uansett en nyttig erfaring å ta med videre at slike inndelinger bør ta hensyn til språklige forhold som er relevante både for den gruppen man studerer transfer i, og for kontrollgruppene som denne sammenlignes med.

Det en kan kalle en «ideell» undersøkelse av transfer, sammenligner gjerne *språkbruksdata* både fra innlærerspråk og morsmål, og helst hos de samme individene (Jarvis & Pavlenko 2008: 49). Denne undersøkelsen har ikke basert seg på et slikt design, men utfylt språkbruksdataene fra ASK med grammatiske beskrivelser av polsk engelsk og norsk. Det er mulig at noen av problemstillingene som er nevnt over, kunne ha vært unngått om jeg hadde brukt språkbruksdata til alle formål.

11.3.7.5 *Materiale og statistikk*

Jeg vil videre peke på noen utfordringer knyttet til utvalg av substantivfraser og tekster i materialet, i tillegg til måten jeg har undersøkt form/kontekst-relasjoner og transfereffekter på.

Utvalget av substantivfraser

Utvalget av substantivfraser var i undersøkelse 2 avgrenset til kun *enkle substantivfraser*, som for de fleste tekstene bare vil utgjøre en liten del av den totale mengden substantivfraser. Avgrensningen av utvalget hadde sine gode grunner, nemlig å isolere den type kontekstuelle kilder til formvalg som undersøkelsen har sett på. Tilstedeværelse av modifierende adledd av ulike slag kunne brakt inn andre alternative kilder til formvalg, som kunne ha påvirket resultatet. Likevel er utvalget som ble gjort, bare en liten del av de formvalgene som gjøres i hver av tekstene, og kanskje ville en mer helhetlig undersøkelse der også modifiserte substantivfraser er inkludert, ha gitt andre funn. Tilstedeværelse av definite determinativer er jo også et forhold som gjør bøyningssendelser redundante.

Statistikk og vurdering av styrken form/kontekst-assosiasjoner

En annen problematisk konsekvens av å bare inkludere enkle substantivfraser i materialet var at jeg endte opp med å ha *få substantivfraser per tekst*, noe som begrenset hvilke statistiske tester som kunne brukes. Hver av hypotesene forutsatte en vurdering av «styrken» til de predikerte form/kontekst-assosiasjonene, både innen den polske gruppen og sammenlignet med de andre to gruppene. Få substantivfraser per tekst gjorde det imidlertid vanskelig å bruke statistikk som tar hensyn til gruppe- og individnivå samtidig, slik jeg kunne i undersøkelse 1. Løsningen ble en kombinasjon av skjønnsvurdering for gruppen som helhet uten statistisk testing og bruk av statistisk test på individuelle profiler.

De statistiske testene som ble brukt på individuelle profiler (kjikvadrat), kunne på en grei måte vise om det var forskjeller mellom gruppene, men den skjønnsmessige vurderingen av «lavt» og «høyt» morsmålssamsvar på gruppenivå, var ikke noe mer enn nettopp en skjønnsvurdering. Skjønnsvurderingen ble ytterligere komplisert av ønsket om å dekke *gjensidigheten* mellom form og kontekst i de predikerte form/kontekst-

assosiasjonene, altså det som ble omtalt som «innen form» og «innen kontekst». Disse to sidene av assosiasjonen avspeilet i noen tilfeller hverandre på den måten at assosiasjonen framsto som enten sterk eller svak innen begge to, men i flere tilfeller var assosiasjonen svak innen kontekst og sterk innen form, eller motsatt. En statistisk løsning som hadde kunne bidratt til et forholdstall som indikerer styrken og gjensidigheten mellom form og kontekst, hadde absolutt vært hensiktsmessig for undersøkelsen.

Det er i seg selv også problematisk at det i prediksjonene for undersøkelse 2 har vært en forventning om at transfer skal komme til uttrykk gjennom nærmest *kategoriske* form/kontekst-assosiasjoner. Som vi har sett i løpet av undersøkelsen, er det sjelden at assosiasjonen mellom form og kontekst har nærmet seg 100 % verken innen kontekst eller innen form for den polske gruppen. Til tross for de tilsynelatende kategoriske prediksjonene, har jeg i praksis operert med en mer fleksibel og ikke-kategorisk vurdering av form/kontekst-assosiasjonene, både gjennom den skjønsmessige vurderingen på gruppenivå, og gjennom å åpne for variasjon i formvalget innen hver konteksttype i vurderingen av de individuelle profilene. I praksis har undersøkelsen dermed vært gjennomført med mindre kategoriske forventninger enn prediksjonene kan se ut til å tilsi. Forventningen om kategoriske funn er derfor ikke å anse som årsak til at jeg ikke har funnet støtte for hypotesene, og vi ser at den relativt liberale holdningen til variasjon, likevel *ikke* har ført til positive funn med tanke på transfer.

Utvalget av tekster

Det kom fram under undersøkelsen av begge hypotesene at den polske gruppen i stor grad hadde lignende formvalg som målspråksgruppen. Dette *kan* indikere at innlærerne som har skrevet tekstene i materialet, stort sett har tilegnet seg hvordan bestemte og ubestemte former brukes i norsk. Med andre ord ligger kanskje innlærertekstene i materialet på et for høyt nivå til å finne den type transfer jeg har vært ute etter. Alle tekstene i materialet for undersøkelse 2 er vurdert til å ligge på B1-nivå, og dette kan være med på å forklare de negative funnene. Young (1996: 172) påpeker for eksempel at det først og fremst var på et tidlig stadium i læringen at transfer virket, mens det heller var universalier som styrte formvalg på senere nivå. Young studie har imidlertid ikke presisert noen ferdighetsnivå som er sammenlignbare med de som er brukt min studie.

I beskrivelsen av grammatisk kompetanse for B1-nivå i Det felles europeiske rammeverket for språk kommer det fram at man på dette nivået har «[...] god kontroll, men uttrykksmåten er merkbart påvirket av morsmålet.» (Utdanningsdirektoratet 2011: 137). Dette skulle tyde på at transfer ikke skulle være utelukket på B1-nivå, og Nordanger (2012: 67) hevder på bakgrunn av en undersøkelse av markering av bestemthet hos engelske og russiske innlærere at morsmålpåvirket språkbruk kan utvides til å gjelde for B2-nivå også. Med tanke på dette virker det ikke naturlig å anta at det er ferdighetsnivået som er årsaken til at transfer ikke blir funnet i min undersøkelse.

I undersøkelse 1 viste jeg også at det var forskjeller mellom -GB-innlærere og +GB-innlærere både på B1- og B2-nivå, noe som ble tolket som at transfer foreligger på begge disse nivåene. Dette var imidlertid et *kvantitativ* utslag av transfer, og det er uklart om dette kommer inn under det som i sitatet over beskrives som «merkbart påvirket av morsmålet» (Utdanningsdirektoratet 2011: 137). Også Nordangers (2017) studie undersøker kun slike kvantitative utslag av morsmålet, og det kan derfor virke som denne problemstilling gjelder konklusjonene der også. Forholdet mellom kvantitative og kvalitative utslag av transfer som har vært nevnt før, er dermed relevant også her: Om de kvantitative utslagene som er funnet og tolket som transfer i Johansen (2010) og Nordanger (2012), er sekundæreffekter og avspeilinger av underliggende kvalitative effekter av morsmålet, så er de kanskje egnet til å si noe om hvorvidt transfer foreligger på de ulike nivåene. Om det derimot *ikke* er slik, og det kanskje er andre faktorer enn transfer som bidrar til markeringsprosenten, er det ikke sikkert slike funn kan si oss noe om hvorvidt innlærerspråk på B1- og B2-nivå er «merkbart påvirket av morsmålet». Dette er likevel en annen diskusjon enn den vi opprinnelig var inne i, altså om innlærerne i materialet mitt ligger på et for høyt nivå til at de har transfer. For å komme tilbake til dette, ser det altså ut til at innlærerne i materialet mitt er på et ferdighetsnivå hvor språkbruken fremdeles kan være påvirket av morsmålet (B1), om en skal basere seg på rammeverksbeskrivelsene av dette nivået. Hvorvidt og hvordan den type transfer jeg har sett etter, passer inn i disse beskrivelsene, er vanskelig å si så lenge funnene har vært negative for transfer.

Et interessant og relevant forhold som også bør nevnes, og som det bare så vidt har vært rettet fokus mot i undersøkelsen, er at en del av de polske innlærerne kan engelsk fra før av, om enn i ulik grad. Bare 7 av de 41 polske innlærerne har oppgitt å *ikke* kunne engelsk på et eller annet nivå. Det kunne selvsagt vært interessant å se på om disse spesifikke individuelle tekstene skiller seg fra de øvrige med tanke på prediksjonene i undersøkelse 2, men dette ble ikke prioritert i denne undersøkelsen. I forbindelse med oppsummeringen av individuelle profiler for hypotese 2 (10.9.2), sjekket jeg de selvrapporterte engelskkunnskapene til fem innlærere med positive profiler, men spredningen i ferdighetsnivå var for stor til at de positive profilene kunne se ut til å ha noe med engelsknivået å gjøre.

Et siste relevant punkt er at både tekst og skrivesituasjonen kan ha påvirket til å dekke over transfer: Tekstene er skrevet som del av en språktest, en situasjon der man ofte har spesielt fokus på korrekthet og gjerne korrigerer egne formvalg. En testsituasjon er dermed ikke en helt spontan situasjon, og mulighetene til å overvåke og korrigere språkb Bruken kan selvsagt fjerne spor etter transfer som ellers ville framkommet i mer spontan kommunikasjon. Dette er likevel neppe så utslagsgivende at det kan forklare de negative funnene i undersøkelse 2.

11.3.8 Oppsummering undersøkelse 2

Jeg har over oppsummert hypoteser og funn fra undersøkelse 2 og diskutert ulike sider av operasjonalisering og utvalg som utgjør utfordringer ved undersøkelsen, og som potensielt har kunnet bidratt til negative funn for hypotesene. Det er likevel usikkert i hvor stor grad de negative funnene har sammenheng med slike metodiske valg, eller om det er hypotesene i seg selv som ikke har vært egnet til å avdekke transfer der den gjør seg gjeldende. Svaret kan selvsagt også være at distribusjonen av substantivformer når alt kommer til alt *ikke* er påvirket av morsmålet. Målspråkslikheten som preger det polske materialet, kan for eksempel tyde på at det handler om målspråkstilegnelse. Målspråkslikheten er likevel ikke total, og et samspill mellom målspråkspåvirkning og morsmålspåvirkning kan være en naturlig kombinasjon for å forklare formvalgene til den polske gruppen. Et mulig scenario, er det som foreslås av Trenkic og Pongpairoj (2013: 161-162), nemlig at morsmålspåvirkningen står sterkere i starten av innlæringen i form av en preferanse for naken form, men at målspråksstrukturen med artikkel øker i

styrke over tid og gradvis fortrenger bruken av naken form i innlærerspråket. I mellomtiden virker morsmålet og målspråket *samtidig* på innlærerspråket, men drar formvalget i hver sin retning.

Uansett vet man at transfer er i samspill med andre forhold, og flere av undersøkelsene som er omtalt i kapittelet om andrespråksforskning, har sett på transfer sammen med ulike typer universalier. Universalier i seg selv, og ikke minst samspillet mellom transfer og universalier, har ikke vært en del av undersøkelsene i denne avhandlingen, så dette kan jeg ikke si noe om på grunnlag av mine funn. Jeg vil likevel bemerke at med tanke på den store overlappingen mellom det som behandles som språklige universalier og det som behandles som morsmålskategorier i forskningen på bestemthet, er det kanskje ekstra nødvendig å forsøke å avdekke forskjellene ved hjelp av en metode som både kan isolere morsmålet som kilde og samtidig trekke de nødvendige grenselinjer mellom hva som representerer universalier og hva som er morsmåls spesifikke kategorier.

11.3.8.1 *Sammenheng mellom undersøkelse 1 og undersøkelse 2?*

Et siste naturlig spørsmål å stille seg er om funnene fra undersøkelse 2 kan belyse noen av funnene fra undersøkelse 1. Funn fra undersøkelse 2 kan, som jeg har vært inne på flere ganger, potensielt bidra til å avgjøre om lavere markeringsprosent hos -GB-innlærere bare er et uttrykk for en generell utfordring knyttet til bruk av bestemt form, eller om det er en *sekundæreffekt* av en mer kvalitativ morsmålspåvirkning på distribusjonen av former. Ettersom hypotesene i undersøkelse 2 ikke kunne vise verken *om* eller *hvordan* morsmålet påvirker distribusjonen av former hos de polske innlærerne, kan ikke funnene fra denne delen av studien belyse markeringsprosenten direkte, men det kan være noen observasjoner som likevel kan si oss noe om sammenhengen mellom det jeg har omtalt som kvantitative og kvalitative utslag av morsmålet.

De to delundersøkelsene er ikke helt sammenlignbare siden de baserer seg på ulikt utvalg av morsmålsgrupper og substantivfraser, i tillegg til at bare undersøkelse 2 har delt opp i ulike referensielle typer. Materialet i undersøkelse 2 besto av tekster fra Språkprøven skrevet av engelske og polske innlærere, og dette var riktignok en delmengde av GB-korpuset i undersøkelse 1, men det var nettopp i delkorpuset av polske og engelske tekster fra Språkprøven det ble funnet svakest tegn til transfer i

undersøkelse 1. Siden materialet for undersøkelse 1 besto av obligatoriske kontekster for bestemt form, altså i hovedsak definitte og generisk referanse, er det disse delmaterialene fra undersøkelse 2, jeg vil fokusere på videre.

Undersøkelse 2 viste stor grad av målspråkslikhet i det polske materialet, men når det kom til bruk av bestemte former innen definitte og generisk referanse, var det noen konteksttyper hvor bruk av bestemt form i den polske gruppen skilte seg mer fra den norske gruppen: Innen definitte referanse brukte de polske innlærerne *mindre* bestemt form, og ved generisk referanse *mer* bestemt form enn den norske gruppen.

Disse tendensene var ikke like sterke i alle konteksttypene, og ved hypotese 1 var andelen av bestemte former ved definitte referanse aller høyest og mest målspråkslik i BE-kontekster (92 %), altså der hypotesen predikerte bestemt form, mens andelen var markant lavere i ART-kontekster (60 %). Dette viste seg å avspeile generelt mer bruk av bestemt form i *aktiverte* kontekster sammenlignet med *ikke-aktiverte*: Aktiverte substantivfraser hadde samlet sett bestemt form i over 85 % av tilfellene, mens ikke-aktiverte hadde bestemt form i under 60 % av tilfellene.

Den samme tendensen kom til syne i det generiske materialet: Selv om det var ubestemte former som dominerte i det generiske materialet, brukte den polske gruppen *mer* bestemt form enn de andre to gruppene, og det var aller mest i BE-kontekster og merkbart mindre i ART-kontekster. Dette vil si at også ved generisk referanse kunne mer bruk av bestemt form hos de polske innlærerne knyttes til aktiverte substantivfraser, og mens det var mindre ved ikke-aktiverte. Denne høyere andelen av bestemt form ved aktiverte referenter, kan minne om funnene til Ekiert (2010: 140), hvor bruk av bestemt form tidligere falt på plass ved substantiv som var tekstuelte aktiverte enn de som var situasjonelt aktiverte.

Det definitive og generiske materialet i undersøkelse 2 hadde i tillegg *få* substantivfraser med ikke-aktivert referent, da hoveddelen av de ikke-aktiverte referentene var indefinitte. Om disse tendensene som er funnet ved definitte og generisk referanse også gjelder for materialet i undersøkelse 1, altså *få* forekomster av og *mindre* bruk av bestemt form ved ikke-aktiverte substantivreferenter, kan det hende at markeringsprosenten til den polske gruppen er en avspeiling av dette. Dette kan imidlertid bare forklare hvorfor markeringsprosenten er såpass høy som den er, ettersom

materialet da vil bestå av mest aktiverte fraser, noe som fremmer bruk av bestemt form. Jeg har imidlertid ikke undersøkt nærmere hvilke egenskaper de umarkerte kontekstene i undersøkelse 1 har, så dette blir kun en spekulasjon.

11.4 Transfer i det engelske materialet

Selv om det er morsmålspåvirkning fra -GB-språk som har vært hovedtema for denne avhandlingen, har det i både undersøkelse 1 og undersøkelse 2 kommet fram tendenser hos +GB-innlærere som potensielt kan tilskrives transfer.

Undersøkelse 1 viste ikke bare at -GB-innlærerne *sjeldnere* markerte substantivene i obligatoriske kontekster, men også at +GB-innlærerne gjorde dette *oftere*, og dessuten hadde flere tekster med 100% markering i flere av delkorpusene. Jeg var tidligere inne på at dette vil si at *begge* gruppene har gruppespesifikke og grupprepresentative tendenser, og at det dessuten kan kobles til transfereffekten morsmålsamsvar i form av om morsmålet har eller ikke har grammatisk bestemthet.

Når det gjelder undersøkelse 2 var det også tendenser i det engelske materialet som så ut til å være spesifikke for den engelske gruppen, nemlig tendenser til mer frekvent bruk av artikkel frase ved indefinitt referanse, og mer frekvent bruk av ubestemte former ved generisk referanse.

I materialet for hypotese 1 var bruken av disse formene mer frekvent enn i både den polske og den norske gruppen, som var relativt like hverandre, noe som kunne vise at tendensene i den engelske gruppen var både gruppe-spesifikke og målspåks-ulike. Den mer utstrakte bruken av artikkel fraser kom kun til syne i materialet for hypotese 1, en konsekvens av at artikkel fraser ikke var inkludert når jeg så på hypotese 2. Den mer utstrakte bruken av ubestemt entall/flertall kom derimot til syne også i undersøkelsen av hypotese 2.

Disse tendensene hos de engelske innlærerne kunne knyttes direkte til konkrete forskjeller mellom språkene norsk og engelsk: I engelsk bruker man oftere indefinitte artikkel fraser enn i norsk, og man bruker sjeldnere bestemte former ved generisk referanse (se 7.5). Dermed er transfereffekten morsmålsamsvar også dekket for den engelske gruppen.

Så langt er det altså mulig å knytte språklige mønster i det engelske materialet til to av transfereffektene, nemlig *intergruppe-heterogenitet* og *morsmålsamsvar*, men det

er også mulig å si litt om transfereffekten *intralingval kontrast*: Språkene norsk og engelsk er mest like innen definitiv referanse, men har større forskjeller innen indefinitiv og generisk, og vi ser at det er på området der språkene er ulike, at forskjellene fra den norske gruppen er størst. Der språkene er like, altså definitiv referanse, er det ikke like stor forskjell mellom den engelske og den norske gruppen. Dette betyr ikke at det *ikke* er tegn til transfer ved definitiv referanse, men at den ikke kommer til syne som forskjeller mellom den norske og den engelske gruppen, nettopp på grunn av likheten mellom språkene. Det er dermed mulig at det er snakk om positiv transfer ved definitiv referanse, og at transfer innen indefinitiv og generisk referanse kan knyttes til det som ble sagt om *overantakelser* i kapittel 2.3.4: Om likhetene mellom morsmålet og målspråket er store, kan dette føre til overantakelser om at de er mer lik enn de faktisk er (Jarvis & Pavlenko 2008: 180; Ringbom & Jarvis 2009: 109).

Det som kommer fram over, er dermed at det i undersøkelse 1 er observert en fasiliterende effekt av morsmålet engelsk når det gjelder bruk av bestemt form i obligatoriske kontekster. I undersøkelse 2 er det også observert en fasiliterende effekt av likheter mellom engelsk og norsk innen definitiv referanse, men innen indefinitiv og generisk referanse ser det ut til å være snakk om *negativ transfer*, altså transfer der utfallet er avvik fra målspråksnormen (Odlin 1989: 36). I den grad observasjonene i undersøkelse 1 og 2 henger sammen, kan man anta at tendensen ved definitiv referanse i undersøkelse 2 bidrar til at markeringsprosenten i undersøkelse 1 blir høy, siden dette materialet i stor grad består av substantiv med definitiv referanse. Det kan da potensielt være den negative transferen ved generisk referanse som bidrar til at kontekster *ikke* er markert i undersøkelse 1. Dette er imidlertid heller ikke sjekket, siden materialet i undersøkelse 1 ikke er delt opp i referensielle typer, så dette er kun en spekulasjon.

Et siste moment i denne sammenhengen er også hvorvidt den typen negativ transfer vi har observert i det engelske materialet fører til *feil* eller bare til en noe annen formuleringsmåte enn det som er vanlig i norsk. Dette vil kunne ha konsekvenser for om frasene har blitt korrigert i ASK og dermed regnet som umarkerte fraser i undersøkelse 1. Verken artikkel frase ved indefinitiv referanse eller bruk av ubestemt flertall ved generisk referanse er nødvendigvis regnet som grammatisk *feil* i norsk. Som vi så i kapittel 7, vil valget mellom artikkel frase eller naken flertallsform handle om hva

en ønsker å framheve (type eller eksemplar), og dermed vil valget først og fremst føre til ulikheter i uttrykket for slike nyanser, og mer sjeldent til faktiske feil. Det samme gjelder bruk av ubestemt flertall ved generisk referanse, som også er en valgmulighet i norsk, og kun i visse undertyper vil dette føre til feil.

Jeg har tidligere vært inne på at jeg i oppsettet av sammenfalls- og markørkontekster for den polske og den engelske gruppen, hadde som utgangspunkt at også den engelske gruppen var påvirket av morsmålet, noe jeg ikke hadde empirisk belegg for å anta (se 8.5.3.3). Som et «heldig» biprodukt av undersøkelsen av transfer fra polsk, har jeg, som det går fram av drøftingen over, likevel kunnet vise at de engelske innlærerne ser ut til å være påvirket av morsmålet. Utgangspunktet for sammenfalls- og markørkontekstene virker derfor å ha vært legitimt.

11.5 Hypoteser om transfer ved nullkontrast

Hovedfokus i hele denne studien har vært transfer i innlærerspråk når det foreligger en nullkontrast mellom morsmålet til innlærerne og målspråket. Jeg har i den forbindelse operert med et skille mellom *nullkontrast* og *nullrelasjon*, der det første viser til den lingvistiske beskrivelsen av forholdet mellom språkene, mens det andre viser til en subjektiv antakelse som innlærere kan ha om forholdet mellom morsmål og målspråk (Ringbom og Jarvis 2009: 109-110). En hovedtanke i utformingen av undersøkelse 2, var at nullkontrasten mellom norsk og polsk ville føre til at innlærere oppfatter forholdet mellom norsk og polsk som en nullrelasjon, altså at språkene er for ulike til tverrspråklig identifikasjon mellom språklige elementer i de to språkene er mulig.

Jeg vil nå se nærmere på to sider av det å undersøke transfer ved nullkontraster, for det første at det kan være vanskeligere å påvise transfer når det foreligger en nullkontrast mellom morsmål og målspråk enn når språkene er mer like, og for det andre at nullkontrast ikke nødvendigvis utelukker tverrspråklig identifikasjon. I forbindelse med det sistnevnte, vil jeg foreslå noen alternative varianter av hypotese 1 og 2, varianter som ikke tar utgangspunkt i transfer som resultat av en nullrelasjon og *antatt* likhet mellom språkene, men derimot i likhetsrelasjon og *oppfattet* likhet.

11.5.1 Å finne transfer når det er nullkontrast

I undersøkelse 2 fant jeg ikke argument for at det var transfer hos de polske innlærerne, men det ble altså gjort funn som kan tolkes som transfer hos de engelske innlærerne, til tross for at dette ikke var målet for undersøkelsen. Dette betyr ikke nødvendigvis at de engelske innlærerne er påvirket av morsmålet, mens polske ikke er det, da funnene like gjerne kan være et resultat av at det er lettere å avdekke transfer når morsmål og målspråk er relativt like, enn når det foreligger en nullkontrast.

Som vi så over, kom transfer hos den engelske gruppen tydeligst til syne innen de referensielle typene der engelsk og norsk er mest ulike, altså indefinit og generisk referanse. Dette ble forklart med at den generelle likheten mellom språkene norsk og engelsk kan ha ført til positiv tverrspråklig identifikasjon og en overantakelse om at språkene er mer like enn de faktisk er. Oppdelingen av undersøkelsen i ulike referensielle typer var viktig for å avdekke forskjellene mellom den engelske og den norske gruppen ettersom språkene engelsk og norsk opererer innen de samme referensielle rammene selv om formvalgene kan være ulike enkelte steder.

Når det gjelder polsk og norsk, foreligger det derimot en grammatisk nullkontrast, og transfer hos polske innlærere kan i teorien skje på flere ulike måter, om i det hele tatt. Dette gjør det vanskeligere å sette fingeren på *hva* (om noe) som kan bli overført fra polsk, og de to hypotesene jeg formulerte, er kun to av flere mulige hypoteser. I denne studien har jeg bare tatt for meg forhold ved substantivfrasers plassering, identifiserbarhet og kasus, som det ikke var positive funn for, men det kan fremdeles finnes andre morsmålsbaserte forklaringer på forholdet mellom form og kontekst i det polske materialet.

Det er sannsynligvis en generell utfordring i møtet mellom nullkontraster og transfer at det er en rekke potensielle hypoteser å stille, og at en undersøkelse som denne bare kan ta for seg *noe* av dette hypoteserommet. Dette kan være en av årsakene til at det er vanskeligere å avdekke transfer ved nullkontraster enn når morsmål og målspråk er mer like. Dette handler kanskje også om at metoden jeg har brukt, baserer seg på å danne seg en idé om morsmålssamsvaret *på forhånd*, noe som er vanskeligere ved nullkontrast på grunn av mangelen på korrespondanser i morsmål og målspråk. En alternativ løsning kan være å gjøre dette i motsatt rekkefølge, dvs. først identifisere

mønster som er spesifikke og representative for en gruppe, og *deretter* forsøke å forklare hvordan disse mønstrene kan være motivert og påvirket av morsmålet. En slik tilnærming blir beskrevet av Granger (1996: 46-47) som del av det hun kaller *Den integrerte kontrastive modellen*, som består av både *predikerende* og *diagnostiserende* kontrastive analyser. Det siste innebærer at man først finner interessante feil eller mønster i innlærerspråket ved å sammenligne det med målspråksbruk og språkbruken til andre innlærergrupper, og deretter finner en forklaring gjennom kontrastive undersøkelser av morsmål og målspråk (Granger 1996: 46).

Det er likevel ikke nødvendigvis slik at det *ikke* eksisterer noen strukturelle korrespondanser mellom morsmål og målspråk selv om det foreligger en nullkontrast for en gitt grammatisk kategori, og under vil jeg se nærmere på noen potensielle strukturelle korrespondanser mellom norsk og polsk.

11.5.2 Nullkontrast og tverrspråklig identifikasjon – alternative hypoteser

Undersøkelse 2 har vært spesielt opptatt av nullkontrasten mellom polsk og norsk, som jeg i utgangspunktet antok at ville føre til en type transfer som *ikke* var basert på tverrspråklig identifikasjon. Begge hypotesene hadde som utgangspunkt at innlærerne antar at morsmål og målspråk fungerer på samme måte, uten at dette er motivert av målspråksinnputt, altså såkalt *antatte likheter* (se 2.3.2.1). Fravær av tverrspråklig identifikasjon er likevel ikke den eneste mulige konsekvensen av en slik nullkontrast: For det første behøver ikke nullkontrast for et bestemt språklig trekk lede til transfer i det hele tatt (Jarvis & Pavlenko 2008: 11), og for det andre kan det også finnes andre, oppfattbare likheter eller samvarierende trekk som kan være gjenstand for tverrspråklig identifikasjon. Det siste vil jeg se nærmere på under, og det er spesielt motivert av én bestemt utfordring som gjelder begge hypotesene i undersøkelse 2.

En sentral utfordring ved begge hypotesene slik de har blitt formulert, er hvordan de språklige forholdene som overføres fra polsk, kommer til uttrykk gjennom akkurat *bestemthetsmarkørene* i norsk. Dette var jeg også inne på i diskusjonen av inaktiv transfer tidligere i kapitlet. Begge hypotesene sier nemlig at noe fra polsk vil uttrykkes ved hjelp av de språklige midlene som fins i norsk, enten brudd på tema/rema-prinsippet eller kasus. Det hypotesene *ikke* klargjør, er hvordan disse språklige trekkene kan ende opp med å bli uttrykt nettopp gjennom norske bestemthetsmarkører om det ikke finner

sted en eller annen form for tverrspråklig identifikasjon mellom polsk og norsk som matcher form med innhold. Jeg vil videre gjøre rede for hvordan begge hypotesene om morsmålspåvirkning fra polsk alternativt kunne ha blitt formulert med utgangspunkt i en positiv tverrspråklig identifikasjon mellom strukturelle elementer i polsk og norsk.

Hypotese 1 – alternativ hypotese om tverrspråklig identifikasjon

For hypotese 1 ligger det etter mitt syn et forklaringsproblem i at innlærerne forventes å ta i bruk bestemthetsmarkørene i norsk for å uttrykke brudd på tema/rema-prinsippet, samtidig som den antatte transferen *ikke* skal være et resultat av tverrspråklig identifikasjon mellom strukturer i polsk og norsk. Fra et lingvistisk ståsted vet vi at det er en likhet mellom identifiserbarhet ved pragmatisk ordstilling og ved grammatisk bestemthet, men at innlærere skal oppfatte denne sammenhengen, har jeg allerede omtalt som mindre sannsynlig. Det er dermed kanskje like lite sannsynlig at de lar nettopp bestemthetsmarkørene i norsk ta rollen med å markere brudd på tema/rema-prinsippet. Young (1996) viser heller ikke til noen forklaring på denne overføringen fra syntaks til morfologi i sin studie, etter det jeg kan se.

En mulig forklaring kunne ha vært at innlærere blir mer oppmerksomme på bruk av bestemthetsmarkører akkurat i de språklige kontekstene som markerer brudd på tema/rema, og at disse språklige kontekstene fungerer som startsted for å forstå hva slike markører brukes til i målspråket. En annen forklaring, som jeg vil se nærmere på her, at en slik type transfer likevel *kan* være forankret i positiv tverrspråklig identifikasjon, noe som bedre kan forklare hvorfor det er nettopp bestemthetsmarkørene som ville fått en slik rolle i innlærerspråket.

Som jeg var inne på i 7.3.2, kan demonstrativer i polsk brukes til å signalisere brudd på tema/rema-prinsippet, og Szwedek (1976) hevder at dette ofte er *obligatorisk* om man vil ha rett fortolkning av frasen som definnitt eller indefinnitt (Szwedek 1976: 208). Dette betyr at det faktisk fins en strukturell og *oppfattbar* kobling mellom identifiserbarhet ved pragmatisk ordstilling og ved grammatisk bestemthet. Om innlærere gjør en tverrspråklig identifikasjon mellom disse demonstrativene i morsmålet og bestemthetsmarkører i målspråket, kan denne oppfattede likhetsrelasjonen føre til transfer i form av at bestemthetsmarkørene brukes i de samme språklige kontekstene i innlærerspråket som demonstrativene brukes i morsmålet, altså ved brudd på tema/rema-

prinsippet. Trenkic (2007: 312; 2008: 9) viser til en liknende tverrspråklig identifikasjon mellom demonstrativer i -GB-språk og engelske artikler, selv om det i Trenkics tilnærming har en annen konsekvens for innlærerspråket (se 4.2.2.4).

Basert på dette kan man se for seg en alternativ hypotese som ligner på hypotese 1, der antakelsen er at innlærerne observerer likheten mellom målspråkets bestemthetsmarkører og morsmålets demonstrativer, og på grunn av denne likheten overfører demonstrativenes funksjon i morsmålet til bestemthetsmarkørene i innlærerspråket. Siden demonstrativene i morsmålet brukes ved brudd på tema/rema-prinsippet, vil bestemthetsmarkørene kun bli brukt i slike kontekster, og utelates når ordstillingen følger tema/rema-prinsippet, altså tilsvarende prediksjoner som for hypotese 1. Fordelen med den alternative hypotesen er at den ikke forutsetter transfer på tvers av lingvistiske nivå, altså syntaks til morfologi, og i større grad kan forklare hvorfor det er akkurat bestemthetsmarkørene i norsk som får den rollen som de antas å få.

I prinsippet kan altså den alternative hypotesen ha de samme prediksjonene som hypotese 1, men det vil være snakk om en annen type transfer, nemlig en som har tverrspråklig identifikasjon som utgangspunkt. Denne endringen vil også kunne ha konsekvenser for om hypotesen er *målspråksavhengig* eller ikke. I 4.4.3 var jeg inne på at hypoteser som tok utgangspunkt i tverrspråklig identifikasjon kan være målspråksavhengige siden de baserer seg på oppfattede likheter mellom morsmål og målspråk. Siden oppfattede likheter helst knytter seg til *formelle* likheter mellom språk (Ringbom & Jarvis 2009: 108; Jarvis & Pavlenko 2008: 179), virker det naturlig å anta at demonstrativer, som er frie morfemer i morsmålet, helst blir jamført med bestemthetsmarkører som også er frie morfemer, altså artikler. I så fall vil denne alternative hypotesen være mer passende for engelsk som målspråk, siden norsk bare har ubestemte artikler, og ikke bestemte. Selv om norsk ikke har bestemt artikkel som sådan, har vi andre definte determinativer som kunne inngått i en tverrspråklig identifikasjonsprosess med polske demonstrativer, for eksempel demonstrativene *det*, *den* og *de*. Et materiale som inkluderer både enkle og modifiserte substantivfraser kunne ha belyst dette, men det er utenfor rekkevidde for mitt materiale. Hva angår ubestemte artikler i norsk, så vi i resultatene for hypotese 1 at det var spesielt for delprediksjonen

om bruk av artikkelfrase i ART-kontekster (ikke-aktiverte forfelt) at hypotesen hadde *aller minst* støtte. Dette vil si at den alternative hypotesen nok heller ikke hadde funnet støtte i datamaterialet som ble brukt i undersøkelse 2.

Et annet mulig område for tverrspråklig identifikasjon mellom polsk og norsk, er substantivmorfologi. I motsetning til engelsk, har norsk større likheter med polsk når det gjelder å ha substantivbøyning, så på samme måte som en tverrspråklig identifikasjon mellom polske determinativer og engelske artikler kan være mulig, kan en identifikasjon mellom bøyningssendelser i polsk og norsk være det. Dette bringer oss imidlertid over på en mulig strukturell korrespondanse mellom kasusbøyning og bestemthetsbøyning, som er mer relevant for en alternativ utforming av hypotese 2.

Hypotese 2: Alternativ hypotese om tverrspråklig identifikasjon

For hypotese 2 ligger ikke forklaringsproblemet i like stor grad på hvorfor det er nettopp bestemthetsmarkørene som tas i bruk for å uttrykke kasusinnhold i innlærerspråket, men kanskje heller det motsatte: Når det faktisk foreligger en formell likhet mellom polsk kasus og grammatisk bestemthet i norsk, altså at begge deler uttrykkes med substantivbøyning, er det ikke da mer sannsynlig at det skjer en tverrspråklig identifikasjon mellom disse strukturene, enn at det ikke gjør det?

Om bakgrunnen for transfer er tverrspråklig identifikasjon mellom strukturer i polsk og norsk, er det nødvendig å rette fokus mot hvilken *form* de enkelte substantiv har i ulike språklige kontekster i morsmålet. I utformingen av hypotese 2 så jeg bort i fra form, og forholdt meg kun til *funksjonell markerthet*, ikke til *formell* eller *morfologisk* markerthet. Om det skjer en tverrspråklig identifikasjon mellom morfologiske uttrykk i polsk og norsk, kan dette skje mellom morfologisk umarkerte og morfologisk markerte former i de to språkene, grovt sett mellom nullformer og bøyde former. Transfer kan komme til uttrykk gjennom at språklige kontekster som er assosiert med nullformer i polsk også har dette i innlærerspråket, mens de kontekstene som er assosiert med bøyde former i morsmålet også vil ha dette i innlærerspråket.

Denne tanken kan ha en parallell i ideen om morsmålpåvirkning i form av preferanse for naken form som ble omtalt kapittel 4. Trenkic (2009) og Trenkic og Pongpairoj (2013) hevder som jeg har vært inne på tidligere, at morsmålpåvirkning hos -GB-innlærere kan komme til uttrykk som en preferanse for naken form i

innlærerspråket, siden morsmålet ikke har artikler og derfor alltid har naken form. Studien til Trenkic og Pongpairoj (2013) ser imidlertid på innlærere som har thai som morsmål, og dette er et isolerende språk, noe som betyr at substantivet *aldri* vil bøyes. For polsk og andre -GB-språk med kasus, vil situasjonen være en annen ettersom man vil *veksle* mellom naken form (nullform) og bøyde former i ulike språklige kontekster. Om ideen om en preferanse for naken form skal kunne gjelde for polske innlærere, må den være mer nyansert og ta i betraktning den morfologiske variasjonen: Polske innlærere vil kunne ha preferanse for naken form kun i kontekster som er assosiert med nullform i morsmålet, men derimot en preferanse for *bøyde* former i kontekster som er assosiert med bøyning i morsmålet.

I enkelte språk med kasus vil det være kun nominativ som har nullform og dermed er morfologisk umarkert, mens øvrige kasus vil være morfologisk markerte (Wiese 2011: 102). Som jeg gjorde rede for i 7.3.1.2, har polsk ikke en slik entydig fordeling av nullformer og bøyde former: I polsk finner man i stedet et skille mellom gruppen av kasusene *instrumentalis*, *dativ*, *lokativ* og *vokativ*, som *alltid* har bøyningensendelse, og gruppen av kasusene *nominativ*, *akkusativ* og *genitiv*, som *veksler* mellom å ha nullform og bøyde form (7.3.1.2). En alternativ variant av hypotese 2 som heller tar utgangspunkt i tverrspråklig identifikasjon mellom substantivmorfologien i morsmålet og målspråket, må dermed ta i betraktning at bruksmønsteret for nullformer i polsk går på tvers av kasus og dermed også syntaktiske funksjoner. Dette krever en langt mer omfattende kartlegging av forholdet mellom form og kontekst i polsk og i innlærerspråket enn den oppdelingen mellom subjekt og objekt som jeg baserte meg på for hypotese 2. Funnene som foreligger fra denne studien, er heller ikke egnet til å si noe indirekte om en slik alternativ hypotese.

Oppsummering og diskusjon

Jeg har over sett på noen alternative versjoner av hypotesene i undersøkelse 2, som i stedet for å basere seg på at transfer er resultat av mangel på tverrspråklig identifikasjon og *antatte* likheter mellom morsmål og målspråk, heller baserer seg på en positiv tverrspråklig identifikasjon og *oppfattede* likheter mellom strukturer i polsk og norsk. Disse alternative hypotesene mener jeg har større kraft til å forklare hvorfor det er

nettopp bestemthetsmarkørene i norsk som kan få rollen med å uttrykke det språklige innholdet som antas å bli overført fra polsk.

Disse alternative variantene av hypotesene er ikke egnet til å forklare de negative funnene i undersøkelse 2 på noen som helst måte, men er heller ment å være forslag til hvordan lignende hypoteser kan formuleres i framtiden, det vil si hypoteser som ser på kommunikativ redundans eller overføring av kasus. Alternativet til hypotese 1 vil innebære å inkludere substantivfraser med foranstilte determinativer, men har ellers relativt like prediksjoner som hypotesen som ble undersøkt. Alternativet til hypotese 2 vil innebære prediksjoner som er ganske annerledes enn for hypotesen som ble undersøkt, men nettopp derfor er det kanskje større sjanse for at den alternative hypotesen kan finne eventuell transfer som ikke ble funnet i foreliggende undersøkelse.

De to alternative hypotesene innebærer en endring i hva som regnes som utløsende for transfer, en endring som kan illustreres gjennom forskjellen på *transfer to somewhere* (Andersen 1983) og *transfer to nowhere* (Kellerman 1995), som i korte trekk viser til transfer henholdsvis *med* og *uten* tverrspråklig identifikasjon (2.3.6). Hva som er utløsende for transfer, om det er innlærernes oppfattede likheter eller antatte likheter, er i prinsippet et empirisk spørsmål i seg selv, og det å ha antatte likheter og nullrelasjon som en *a priori* antakelse, slik jeg har gjort for mine hypoteser, kan i så måte være problematisk. Som vi så over, var prediksjonene for hypotese 1 og dens alternative variant nesten sammenfallende. Om en ønsker å finne ut om transfer er utløst av tverrspråklig identifikasjon eller ikke, ser det ut til å være nødvendig å kunne stadfeste i hvilke kontekster disse to hypotesevariantene vil ha ulike implikasjoner for innlærerspråket. Som jeg var inne på lenger opp, er dette også av relevans for å kunne si om eventuelle tilfeller av inaktiv transfer representerer lingvistisk eller konseptuell transfer.

Dette er også relevant for å avgjøre om funn og hypoteser i studier med ulike målspråk er sammenlignbare eller ikke, altså om morsmålsbaserte forklaringer er målspråkuavhengige eller ikke. I 4.4.3 hevdet jeg, i likhet med Nordanger (2017: 129), at hypoteser og funn fra ulike målspråk er mer sammenlignbare når de tar utgangspunkt i semantiske og pragmatiske forhold, ettersom transfer da ikke utløses av strukturelle korrespondanser mellom morsmål og målspråk. Om det er tverrspråklig identifikasjon

som er opphavet til transfer eller hypoteser om transfer, eller ikke, er dermed avgjørende for sammenlignbarheten mellom studier med for eksempel engelsk og norsk som målspåk. Det kan altså være vanskelig å vurdere den reelle sammenlignbarheten til funn og hypoteser for ulike målspåk uten å vite om tverrspråklig identifikasjon ligger bak innlærernes språklige valg eller ikke.

Et siste poeng knyttet til sammenlignbarhet, er at ved inaktiv transfer kan likevel strukturelle forskjeller mellom ulike målspåk ha betydning for sammenlignbarheten, til tross for at det ikke skjer noen tverrspråklig identifikasjon. Et sentralt element ved inaktiv transfer er jo nettopp at *målspåkets strukturer* tas i bruk for å uttrykke det som overføres inaktivt, og dermed vil slik transfer kunne manifestere seg på ulike måter i innlærerspråk med ulike målspåk, til tross for at det kan være det samme innholdet eller perspektivet som overføres inaktivt fra morsmålet. I så fall er det sannsynlig at det foregår et samspill mellom det som overføres inaktivt og de strukturene som fins i målspåket, i form av at målspåksstrukturene bidrar til å forme eller begrense det som overføres inaktivt. Hypotese 2 demonstrerte at om et stort antall innholdsnyanser knyttet til kasus skulle kunne uttrykkes gjennom et langt mindre antall bestemthetsmarkører i norsk, måtte prediksjonene begrenses. I min undersøkelse fungerte en reduksjon til nominativ og ikke-nominativ som et verktøy for å formulere prediksjoner (se 8.1.3.1), men dette kan også fungere som et eksempel på hvordan de tilgjengelige strukturene i målspåket kan tenkes å begrense inaktiv transfer om det forekommer. Inaktiv transfer *må* sannsynligvis begrenses på en eller annen måte for å kunne komme til uttrykk gjennom målspåkets markører, i så fall er det interessant om dette skjer på en tilfeldig eller en systematisk måte, og om det kanskje, når alt kommer til alt, kan være en eller annen form for tverrspråklig identifikasjon involvert likevel.

Dette markerer slutten på denne avhandlingens teoretiske betraktninger rundt fenomenet transfer, og jeg vil helt til slutt vende tilbake til de to mest sentrale metodiske sidene ved studien, transfereffekter og korpus.

11.6 Transfereffekter og korpus

I denne studien som helhet har både det å skulle påvise transfereffekter og bruk av korpus lagt avgjørende føringer på metoden. Avslutningsvis vil jeg oppsummere noen erfaringer og betraktninger jeg har gjort meg på disse to områdene.

11.6.1 Transfereffekter

Jeg har både i undersøkelse 1 og 2 hatt det som en sentral del av metoden å undersøke transfereffekter (Jarvis 2000, 2010). De fire transfereffektene *morsmålssamsvar*, *intralingval kontrast*, *intragruppe-homogenitet* og *intergruppe-heterogenitet* regnes av Jarvis som essensielle for å kunne argumentere for at noe faktisk er transfer ved å både vise at språklige mønster har sitt opphav i morsmålet, og at de er representative og spesifikke for en innlærergruppe.

Jeg vil på de neste sidene framheve noen erfaringer med og betraktninger rundt transfereffekter. En viktig metodisk grep som jeg vil framheve spesielt, er inndelingen i *markørkontekster* og *sammenfallskontekster*. Denne inndelingen viste seg nyttig for å belyse flere av transfereffektene, noe jeg vil gå nærmere inn på under.

11.6.1.1 Morsmålssamsvar og intralingval kontrast

Transfereffekten *morsmålssamsvar* har hatt som hensikt å vise at språklige mønster eller språklige valg i innlærerspråket har morsmålet som kilde eller på annen måte er motivert av morsmålet (Jarvis & Pavlenko 2008: 47). Denne transfereffekten har blitt definert ulikt i de to undersøkelsene: I undersøkelse 1 ble morsmålssamsvar definert med utgangspunkt i fravær av grammatisk bestemthet i morsmålet, mens morsmålssamsvar i undersøkelse 2 var definert ut fra mer spesifikke språklige forhold i morsmålet, som ga spesifikke prediksjoner om form/kontekst-assosiasjoner i innlærerspråket.

Jeg har tidligere i dette kapittelet vist til utfordringer ved begge disse måtene å definere morsmålssamsvar på, for det første spørsmålet om hvorvidt fravær av grammatisk bestemthet i morsmålet er en tilstrekkelig definisjon av morsmålssamsvar (11.2), og jeg var også inne på noen utfordringer med form/kontekst-assosiasjonen for de to hypotesene (11.3.7). Her vil jeg derimot fokusere på nytten av å vurdere morsmålssamsvar i sammenheng med transfereffekten *intralingval kontrast*, ettersom det satte tendenser som kunne se ut som morsmålssamsvar, i et noe annet lys.

Inndelingen i markør- og sammenfallskontekster gjorde det nemlig mulig å avdekke at språklige mønster som kunne se ut som morsmålssamsvar i den polske gruppen, i stor grad sammenfalt med typiske formvalg i målspåket. Dette gjorde det mer naturlig å peke på målspåket heller enn morsmålet som en kilde til formvalgene hos de polske innlærerne. Funn av transfereffekten morsmålssamsvar alene var dermed ikke tilstrekkelig for å isolere morsmålet som kilde til formvalget.

Som jeg var inne på over, *utelukker* ikke målspåkslikhet nødvendigvis morsmålet som kilde i disse sammenfallskontekstene, det kan også ses på som positiv transfer utløst av at morsmål og målspåk drar i samme retning, men sett i en større sammenheng, kom det fram at formvalgene var målspåkslike også i markørkontekster, altså der den polske gruppen var ventet å skille seg fra den norske om hypotesene stemte. Skillet mellom markør- og sammenfallskontekster viste seg på den måten å være veldig avgjørende for å konkludere med at det *ikke* forelå transfer av den typen hypotesene tok for seg.

11.6.1.2 *Intragruppe-homogenitet og intergruppe-heterogenitet*

Intergruppe-heterogenitet og *intragruppe-homogenitet* omtales av Jarvis som *gruppebaserte* transfereffekter. Transfereffekten *intergruppe-heterogenitet* handler om å vise at språklige mønster er spesifikke for en innlærergruppe, i dette tilfellet -GB-innlærere, mens *intragruppe-homogenitet* handler om å vise at disse språklige mønstrene er representative for denne innlærergruppen (Jarvis 2000: 253-254). Disse to transfereffektene ses gjerne i sammenheng med hverandre, der homogeniteten *innad* i en innlærergruppe skal være større enn homogeniteten *mellom* to ulike innlærergrupper.

Disse to transfereffektene ble sett i sammenheng med hverandre både i undersøkelse 1 og i undersøkelse 2, men på ulike måter: I undersøkelse 1 ble dette ivaretatt gjennom bruk av ikke-parametriske statistiske tester som tok i betraktning forholdet mellom gjennomsnittet og den individuelle variasjonen i markeringsprosent, mens det i undersøkelse 2 ble gjort gjennom statistisk testing av antall individuelle profiler som var i tråd med prediksjonene for hypotesene i de tre gruppene. Antallet av positive polske profiler kunne si noe om grad av morsmålssamsvar og hvor

representativt det var, mens sammenligning av antallet positive profiler i gruppene kunne si noe om intergruppe-heterogenitet og forskjeller fra målspråket.

Mens metoden i undersøkelse 1 er relativt etablert som en måte å vise disse transfereffektene og forholdet mellom dem på, innebærer metoden i undersøkelse 2 en viss «reduksjon» av dataene, der rent statistisk informasjon blir omgjort til profiler. Metoden har likevel stor nærhet til den individuelle teksten og et mer kvalitativt preg ettersom hver enkelt tekst faktisk er vurdert etter om den følger prediksjonene eller ikke.

I undersøkelse 2 var inndelingen i sammenfalls- og markørkontekster også sentral for å framheve intergruppe-heterogenitet mellom den polske og den engelske gruppen, noe som gjorde det mulig å vise at det i stor grad var likhet mellom disse gruppene. Et interessant funn var at i sammenligningen av alle de tre gruppene, kom det fram at det ofte var den engelske gruppen som skilte seg fra den norske og polske gruppen, som ofte var relativt like, noe som var avgjørende for å trekke en konklusjon om mulig transfer hos de engelske innlærerne.

11.6.1.3 *Individuelle profiler og sammenfalls- og markørkontekster*

Inndelingen i sammenfallskontekster og markørkontekster var altså nyttig både for sammenligningen med den norske og den engelske gruppen, men den viste seg også å være viktig for å utfylle den kvantitative vurderingen av individuelle profiler. I vurderingen av individuelle profiler ble intergruppe-heterogenitet og intragruppe-homogenitet målt kvantitativt ved å telle opp antall av *positive* og *ikke positive* profiler i de ulike gruppene. I noen tilfeller hadde den polske gruppen et relativt høyt antall av positive profiler, i andre tilfeller ikke, og i noen tilfeller var forskjellene mellom gruppene signifikante, mens de ikke var det i andre tilfeller. I tillegg var det en god del tilfeller der spesielt den norske gruppen hadde *flere* positive profiler enn den polske. Kartleggingen av om positive profiler hadde markørkontekster eller om de bare hadde sammenfallskontekster, avslørte at både høy andel av positive polske profiler og høy andel i de to andre gruppene, kunne tilskrives *mangel på markørkontekster* i disse profilene. Hvilke konteksttyper som var representert i materialet viste seg med andre

ord å slå ut på det kvantitative resultatet for individuelle profiler.⁶¹ Når markørkontekstene manglet, ble spørsmålet om morsmålssamsvar stående åpent siden det verken kunne utelukkes eller bekreftes uten markørkontekster.

Sammenfalls- og markørkontekster

Inndelingen i sammenfalls- og markørkontekster hadde som mål å forhindre at forventede likheter mellom gruppene skulle dekke over faktiske forskjeller som ville være avgjørende for å påvise transfer. Inndelingen bidro altså på flere områder: for det første med å belyse forholdet mellom morsmålssamsvar og målspråkslighet, for det andre for å utelukke forhold som er universelle for alle innlærere uavhengig av morsmål, og til å avdekke transfer hos de engelske innlærerne. I tillegg til dette ble sammenfalls- og markørkontekster viktige i vurderingen av *individuelle profiler*.

Selv om grepet med å skille ut markørkontekster fungerte godt og virker nødvendig for å si om det foreligger transfer eller ikke, har jeg også tidligere vært inne på noen utfordringer knyttet til å predikere hva som er markør- og sammenfallskontekster i sammenligningen mellom to innlærergrupper. Den første utfordringen lå i å komme ned på et detaljert nok nivå til å kunne si noe tydelig om hva som var forventet formvalg i kontrollgruppene, noe som spesielt slo ut ved generisk referanse. Den andre utfordringen er mer prinsipiell, nemlig at det for sammenligningen med den engelske gruppen lå en antakelse om at også disse innlærerne var påvirket av morsmålet i sine språklige valg, noe som ikke i seg selv var empirisk vist. I oppsummeringen av engelsk transfer (11.4), kom jeg fram til at antakelsen likevel kan ha vært legitim, da det i løpet av undersøkelse 2 ble funnet flere tegn på transfer hos de engelske innlærerne.

11.6.1.4 Intralingval kontrast eller intergruppe-heterogenitet med målspråket

Jeg var i 8.5.3.4 inne på at min måte å undersøke transfereffekten *intralingval kontrast* på ikke nødvendigvis er gjort slik det er ment ifølge Jarvis (2010), og at min metode heller kan være en type *intergruppe-heterogenitet* i forhold til *målspråksgruppen*. Når

⁶¹ Et lignende poeng gjaldt de «alternative form/kontekst-assosiasjonene som ble funnet i undersøkelsen av hypotese 1, der det i det samlede materialet så ut til å være tegn til assosiasjon mellom *ubestemte former* og *slutfeltet*, og mellom *bestemte former* og *aktiverte substantivfraser* (se 9.9.3). Dette viste seg imidlertid å være et utslag av utstrakt bruk av ubestemte former og høyt antall slutfelt ved indefinitt referanse, og ustrakt bruk av bestemte former og høyt antall forfelt ved definitt referanse. Dette var med andre ord ikke faktiske form/kontekst-assosiasjoner, men avspeiling av hvilke konteksttyper som var representert i de ulike delene av materialet.

jeg likevel valgte å kalle dette intralingval kontrast, var det fordi inndelingen baserte seg på et sentralt element i transfereffekten intralingval kontrast, nemlig en inndeling i kontekster hvor morsmål og målspråk er like eller ikke like. Det var også uklart hvordan intralingval kontrast ellers skulle vært undersøkt. Slik jeg tolket Jarvis (2010), så denne effekten ut til å rette seg mer mot mestring av ulike språklige trekk, og undersøkelse 2 så ikke på mestring, men på form/kontekst-assosiasjoner. Uavhengig av hva jeg har omtalt denne transfereffekten som, har den uansett vært avgjørende for å skille transfer fra andre kilder til formvalg.

11.6.2 Korpuserfaringer

Helt til sist vil jeg nå dra ut noen erfaringer og betraktninger rundt bruk av et elektronisk innlærerkorpus. Undersøkelsene i denne avhandlingen kan kalles *korpusbaserte*, men kombinerer også mer «ren» korpusbruk med manuell annotering og bearbeiding av dataene utenfor selve korpuset, spesielt i undersøkelse 2. Jeg vil under trekke fram noen måter det å bruke korpuset på har påvirket studien min utover å bidra med selve språkbruksdataene: Effektivisering av databehandlingen, tilrettelegging for transferstudie, store mengder data og mulighet til ekstra annotering tilpasset egne studiespesifikke behov.

Effektivisering av databehandling

Innlærerkorpuset ASK har i denne studien bidratt både med selve språkdataene, men også til å effektivisere undersøkelsene. Språkbruksdataene var allerede annotert for mange sentrale språklige trekk, for eksempel ordklasse, morfologiske kategorier og syntaktisk funksjon, og korpuset bidro dessuten med informasjon om tekstene og tekstskriverne, slik at jeg enkelt kunne lage mindre delkorpus hvor spesifikke variabler var kontrollert for. Selv om jeg i undersøkelse 2 måtte annotere visse kategorier manuelt, var den tilgjengelige infoen fra ASK med på å effektivisere denne annoteringen også, for eksempel kunne sortering av konkordansene etter syntaktisk funksjon effektivisere klassifiseringen av setningsposisjon siden det en sammenheng mellom subjekt/midtfelt og objekt/slutfelt. Et annet eksempel er at jeg kunne sortere ut alle substantiv med bestemt form som ikke hadde feilkoden F (gal form), noe som utelukket indefinit

referanse, dermed var det raskere å gjøre den videre klassifiseringen i definitiv og generisk referanse.

Tilrettelegging for transferstudie

I kapittel 5.3 framhevet jeg at ASK var godt egnet til en transferstudie, fordi det tilbyr språkbruksdata fra et relativt stort antall tekster skrevet av innlærere med ulike morsmål, og dessuten tekster fra morsmålsbrukere av norsk. ASK legger med dette til rette for å undersøke tre av transfereffektene (intragruppe-homogenitet, intergruppe-heterogenitet og intralingval kontrast), i tillegg til å tilby et utvalg relevante variabler som kan brukes for å sette sammen kontrollerte delkorpus eller på annen måte utforske forholdet mellom innlærernes bakgrunn og språkbruk. Den siste effekten, *morsmålssamsvar*, krever imidlertid at man finner andre korpus med morsmålsdata fra språkgruppene i ASK, eller slik jeg har gjort i denne undersøkelsen – bruke eksterne grammatiske beskrivelser av språkssystemene.

Manuell annotering av studie-spesifikke forhold

Det har underveis i avhandlingen blitt nevnt noen områder hvor korpuset ikke kunne bidra på forhånd med essensiell informasjon, men dette var først og fremst knyttet til egenskaper som var spesifikt relevant for den grammatiske kategorien bestemthet eller de spesifikke hypotesene i denne studien.

Som nevnt tidligere, ville det i undersøkelse 1 ha vært nyttig å kunne dele inn de ulike referensielle typene for bestemte former, altså definitiv og generisk (se 6.5.) Annotering av referensiell type kan anses som veldig relevant for framtidige ASK-studier siden man vanskelig kan komme unna dette i en studie av bestemthet i norsk innlærerspråk, eller substantivfraser generelt for den saks skyld. Annotering av referensiell type er imidlertid både tidkrevende og komplisert, så dette krever nok et større prosjekt og et fast rammeverk for annotering. Fordelen med dette hadde også vært at det ville vært et felles utgangspunkt for alle som senere bruker korpuset til å studere substantivfraser i innlærerspråk.

I undersøkelse 2 gjorde det begrensede utvalget av substantivfraser det mulig for meg å annotere referensiell type manuelt gjennom bruk av ASKs annoteringsfunksjon, som er et veldig nyttig verktøy fordi forskeren kan annotere direkte inn i korpuset slik

at annoteringene blir søkbare i ettertid. I senere tid har det også blitt lagt til rette for deling av slik annotering mellom ulike brukere.

Ulike studier vil ha ulike tilnærminger, og det er derfor uansett nyttig, og kanskje helt nødvendig, å ha annoteringsfunksjonen som bidrar til at man kan tilføre info som er spesifikk for det en undersøker. I en transferstudie der målet er å predikere morsmålssamsvar, som her, vil det gjerne kreve en studie-spesifikk annotering for å avdekke det man ser etter. Å implementere all slik ekstrainformasjon i ASK er nok ikke mulig, dermed er annoteringsfunksjonen et utrolig viktig verktøy. I kombinasjon med den informasjonen som er tilgjengelig i ASK, er det, som jeg var inne på over, også mulig å bruke ASKs grensesnitt strategisk for å effektivisere egen annotering.

Store mengder data - kvantitativ orientering

Bruken av korpus i denne studien har uten tvil gjort at studien har fått en relativt *kvantitativ orientering*, og i mindre grad dreid analysen i retning av nærlesning og analyser av faktiske eksempler fra tekstene. Mengden data og den kvantitative orienteringen, forutsetter også god bruk av statistikk, noe som til en viss grad ligger utenfor språkviterens område, og definitivt er et område med mange fallgruver og kompliserte vurderinger. Under vil jeg oppsummere noen problemstillinger som har dukket opp i møtet mellom korpus, utvalg og statistikk.

Undersøkelse 1 må sies å være den mest «rene» korpusbaserte delundersøkelsen i denne studien, ettersom dataene var hentet rett ut av ASK uten ekstra annotering eller bearbeiding annet enn i statistikkprogrammet. Her kom også tilgangen til et stort antall tekster til nytte slik at jeg kunne lage delkorpus som besto av relativt mange tekster fra flere morsmålsgrupper, og hvor mange substantiv per tekst gjorde det mulig å bruke statistiske tester for å finne ut om -GB-innlærere sjeldnere markerer obligatoriske kontekster enn +GB-innlærere.

Det å slå sammen flere morsmålsgrupper i undersøkelse 1 var nyttig for å gi mer data til undersøkelsen, men her er det alltid et problem at språkene som ble gruppert sammen, også har sine forskjeller. Et eksempel som har vært nevnt over, er en av forskjellene mellom morsmålene polsk og thai i Trenkic og Pongpairøj (2013): Begge språkene mangler grammatisk bestemthet, men polsk har rik substantivmorfologi der thai ikke har morfologi i det hele tatt. Vi så også i undersøkelse 1 at de tre +GB-språkene

var ulike i at tysk, i motsetning til engelsk og nederlandsk, har kasus i tillegg til grammatisk bestemthet, noe som gjør at de tyske innlærere også hadde noe tilfelles med *-GB-innlærerne* i undersøkelsen. Sammenslåing av grupper av innlærere med morsmål som deler språklige trekk, utgjør med andre ord en kraftig ressurs, men rekkevidden til funnene vil altså kunne være begrenset av den interne variasjonen innad i den sammenslåtte gruppen.

Kontroll på variabler og frasetype: mindre mengder data

I undersøkelse 2 var det ikke mulig å bruke slike statistiske tester som i undersøkelse 1, ettersom mengden tekster og mengden substantivfraser per tekst var lavere. Dette skyldtes både at jeg bare så på to innlærergrupper og hadde avgrenset variablene ferdighetsnivå og testtype, men først og fremst ble det nok et problem fordi jeg bare så på enkle substantivfraser, og dermed bare trakk ut bare et lite utvalg av de substantivfrasene som var i tekstene.

Et annet poeng jeg har vært inne på, er at tekstene i ASK er av ulik lengde og har ulikt antall substantiv, og dermed har disse tekstene også ulikt antall av de ulike konteksttypene som har vært undersøkt, enten det er obligatoriske kontekster i undersøkelse 1, eller de spesifikke konteksttypene som var relevante for undersøkelse 2. Det er positivt med autentiske innlærertekster, men varierende lengde kan være et problem med tanke på bruk av statistiske tester. I kapittel 6 var jeg inne på hvordan dette påvirket markeringsprosentene på en måte som gjorde dem mindre egnet til visse statistiske tester fordi de ikke var *intervalldata* (se 6.2.1). En løsning kan være at man bare ser på tekster fra ASK som har like mange kontekster, men det vil nok igjen begrense størrelsen på materialet. Denne ulempen må likevel ses i forhold til fordelene det er med å ha mer autentiske tekster, selv om gjenfortellingstekster av den typen Nordanger (2017) eller Jarvis (2002) bruker, er godt egnet i så måte, fordi de omtaler de samme forholdene og ikke minst de samme referentene i alle tekstene.

Bruk av korpus: oppsummering

Bruk av korpus i denne avhandlingsstudien har gått i flere retninger og hatt en varierende balansegang mellom store data og små data, og automatisk og manuelt. Det er ingen tvil om at ASK har skapt mulighet for en transferstudie som innebærer en kombinasjon av

effektiv behandling av store mengder data, som i undersøkelse 1, og et mer kontrollert delkorpus hvor jeg har kommet nærmere inn på dataene med egen annotasjon og større lingvistisk detaljnivå.

11.7 Utleddning

I denne avhandlingen har jeg belyst to ulike sider av morsmålets rolle for bruk av bestemte og ubestemte former av substantivet hos norskinnlærere som mangler grammatisk bestemthet i morsmålet: For det første et *kvantitativt* utslag av morsmålpåvirkning knyttet til bruk av bestemte former i obligatoriske kontekster, og for det andre et mer *kvalitativt* utslag av morsmålpåvirkning knyttet til distribusjonen av bestemte og ubestemte former.

Det *kvantitative* utslaget ser ut til å foreligge, selv om det er uavklart om mindre markering hos -GB-innlærere er uttrykk for en generell og *usystematisk* påvirkning fra morsmålet, eller om det eventuelt kan være en sekundæreffekt av en mer spesifikk og *systematisk* påvirkning på når bestemte former brukes. Måten de to undersøkelsene i studien var utformet på, ga ikke mulighet til at dette spørsmålet kunne utforskes inngående, men dette er et interessant spørsmål for fremtidig forskning.

Det *kvalitative* utslaget var det imidlertid ikke positive funn for, da ingen av de to hypotesene så ut til å kunne avdekke den typen transfer som de handlet om. Hypotesene var basert på en antakelse om det jeg mot slutten av avhandlingen har kalt *inaktiv transfer*, altså transfer som ikke har opphav i tverrspråklig identifikasjon mellom strukturer i morsmål og målspråk, og der strukturer i målspråket blir tatt i bruk for å uttrykke språklig innhold eller perspektiv som overføres fra morsmålet. Dette utgangspunktet for hypotesene utfordret jeg mot slutten av avhandlingen ved å vise at lignende hypoteser like gjerne kunne ha vært formulert med utgangspunkt i positiv tverrspråklig identifikasjon mellom språklige strukturer i norsk og polsk.

En undersøkelse av inaktiv transfer er interessant og relevant å få belyst empirisk, men slike undersøkelser byr på betydelige metodiske utfordringer, hvorav én er hvordan man skiller mellom transfer som er resultat av tverrspråklig identifikasjon og transfer som *ikke* er det. Mangel på direkte tilgang til innlærernes faktiske psykotypologi, altså subjektive antakelser om språklige forskjeller og likheter mellom morsmål og målspråk, gjør at det er vanskelig å finne ut hva som reelt sett ligger til grunn for ulike tilfeller av

transfer. Det vil si at vi ikke har noen direkte kilde til om det skjer tverrspråklig identifikasjon eller ikke, og derfor må vi basere oss på hva vi kan utlede fra studier som finner transfer. Disse studiene må i så fall ha presise nok prediksjoner til å kunne synliggjøre evidens som kan utelukke det ene eller det andre alternativet.

Et stadig tilbakevendende spørsmål i arbeidet med denne avhandlingen har vært nettopp *hvordan skiller man mellom*. For eksempel, hvordan skiller man mellom ...

- ... transfer og ikke-transfer
- ... en direkte kvantitativ effekt og en sekundæreffekt
- ... antatte likheter og oppfattede likheter
- ... morsmålsamsvar og målspråkslikhet
- ... lingvistisk og konseptuell transfer

Dette er kun et lite utvalg av de forholdene som må avklares før man kan nå den endelige konklusjonen om hvorvidt transfer foreligger, hva som utløser den og hva som kjennetegner et spesifikt tilfelle av transfer. Metodisk framgangsmåte og presisjonen til hypoteser og prediksjoner blir som nevnt over, viktig: Hva er de spesifikke funnene som vil skille morsmålspåvirkning fra målspråkspåvirkning eller universalier? Hvor vil transfer som resultat av tverrspråklig identifikasjon gi andre utslag i materialet enn transfer uten tverrspråklig identifikasjon? Og her passer det fint å avrunde med Karl Poppers ord: «Good tests kill flawed theories; we remain alive to guess again.»

LITTERATURKISTE

- Alonso, Maria. R. A. 2002. *The role of transfer in second language acquisition*. Vigo, Spania: Servicio de Publicacións da Universidade de Vigo.
- Andersen, Roger. 1983. Transfer to somewhere. *Language transfer in language learning*, 1:177-201.
- Ariel, Mira. 2008. *Pragmatics and Grammar*. Cambridge: Cambridge University Press.
- Axelsson, Monica. 1994. *Noun phrase development in Swedish as a second language: A study of adult learners acquiring definiteness and the semantics and morphology of adjectives*/ [Doktoravhandling, Stockholms Universitet].
- Berggreen, Harald og Kari Tenfjord. 1999. *Andrespråklæring (2.utg)*. Oslo: Gyldendal akademisk.
- Bickerton, Derek. 1975. *Dynamics of a Creole system*. London: Cambridge University Press.
- Bley-Vroman, Robert. 1983. The Comparative Fallacy in Interlanguage Studies: The Case of Systematicity. *Language learning*, 33(1):1-17.
- Borthen, Kaja. 2003. *Norwegian Bare Singulars*/[Doktoravhandling, Norges teknisk-naturvitenskapelige universitet].
- Carlsen, Cecilie. 2010. Å knytte ASK til rammeverket - hvorfor og hvordan. I *Systematisk, variert, men ikke tilfeldig. Antologi om norsk som andrespråk i anledning Kari Tenfjords 60-årsdag*, redigert av H. Johansen, A. Golden, J. E. Hagen og A. Helland. Oslo: Novus Forlag.
- Carlsen, Cecilie. 2012. Proficiency level—A fuzzy variable in computer learner corpora. *Applied Linguistics*, 33(2):161-183.
- Carlsen, Cecilie og Eli Moe. 2019. *Vurdering av språkferdigheter*. Bergen: Fagbokforlaget.
- Carlson, Gregory N. og Francis J. Pelletier. 1995. *The Generic Book*. Chicago: University of Chicago Press.
- Chaudron, Craig og Kate Parker. 1990. Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, 12(1):43-64.
- Chesterman, Andrew. 1991. *On definiteness: A Study with special reference to English and Finnish*. Cambridge: Cambridge University Press.
- Croft, William. 1990. *Typology and universals (2.utg)*. Cambridge University Press
- Declerck, Renaat. 1986. Two notes on the theory of definiteness. *Journal of Linguistics*, 22: 25-39.

- Diderichsen, Paul. 1946. *Elementær dansk grammatik*. København: Gyldendal.
- Dixon, Robert. M. W. 1979. Ergativity. *Language*, 55: 59-138.
- Ekiert, Monika. 2004. Acquisition of the English article system by speakers of Polish in ESL and EFL settings. *Working Papers in TESOL & Applied Linguistics*, 4(1):1-23.
- Ekiert, Monika. 2007. The acquisition of grammatical marking of indefiniteness with the indefinite article *a* in L2 English. *Working Papers in TESOL & Applied Linguistics*, 7(1):1-43.
- Ekiert, Monika. 2010. Linguistic Effects on Thinking for Writing: The Case of Articles in L2 English. I *Linguistic Relativity in SLA: Thinking for Speaking*, redigert av Z. Han og T. Cadierno. UK: Multilingual Matters.
- Ellis, Henry C. 1965. *The transfer of learning*. New York: Collier Macmillan Ltd.
- Europarådet. 2001. *The Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Faarlund, Jan Terje, Svein Lie og Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Falk, Yehuda N. 1999. Unmarked case. Ms. Hebrew University of Jerusalem.
- Feldstein, Ronald F. 2001. *A concise Polish grammar*. Slavic and East European Language Research Center (SEELRC). Duke University.
- Firbas, Jan. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.
- Givón, Talmy. 1983. Topic Continuity in Discourse: An Introduction. I *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, redigert av T. Givón. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Golden, Anne, Scott Jarvis og Kari Tenfjord. 2017. *Crosslinguistic Influence and Distinctive Patterns of Language Learning: Findings and Insights from a Learner Corpus*. Bristol, Blue Ridge Summit: Multilingual Matters.
- Granger, Sylviane. 2002. A bird's-eye view of learner corpus research. I *Computer Learner Corpora, Second Language Acquisition and Foreign language Teaching*, redigert av S. Granger, J. Hung og S. Petch-Tyson. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Granger, Sylviane. 2004. Computer learner corpus research: Current status and future prospects. *Applied Corpus Linguistics*, 52:123-145.
- Granger, Sylviane. 2008. Learner Corpora. I *Corpus Linguistics: An International Handbook vol. 1*, redigert av A. Taylor. Berlin: Mouton de Gruyter.
- Greenberg, Joseph H. 1966. *Language universals*. Haag: Mouton de Gruyter.

- Grice, H. Paul. 1975. Logic and Conversation. I *Philosophy of Language*, redigert av A. P. Martinich. New York: Oxford University Press.
- Hagen, Jon Erik. 1998. *Norsk grammatikk for andrespråklærere (2.utg)*. Oslo: Gyldendal akademisk.
- Halmøy, Madeleine. 2010. *The Norwegian Nominal System: A Neo-Saussurean Perspective* [Doktoravhandling, Universitetet i Tromsø].
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of linguistics*, 42(1):25-70.
- Hinton, Perry R. 2004. *Statistics Explained*. London: Routledge.
- Hinton, Perry R., Charlotte Brownlow, Isabella McMurray og Bob Cozens. 2004. *SPSS Explained*. London: Routledge.
- Huebner, Thom. 1983. *A longitudinal analysis of the acquisition of English*. Ann Arbor, Michigan: Karoma Publishers.
- Ionin, Tania og Kenneth Wexler. 2003. The certain uses of the in L2 English. I *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*, redigert av J. M. Liceras et al. Somerville, MA: Cascadilla Proceedings Project.
- Ionin, Tania, Heejeong Ko og Kenneth Wexler. 2004. Article semantics in L2 acquisition: The role of specificity. *Language acquisition*, 12(1):3-69.
- Ionin, Tania, Maria L. Zubizarreta og Salvador B. Maldonado. 2008. Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4):554-576.
- Jarvis, Scott. 2000. Methodological Rigor in the Study of Transfer: Identifying L1 Influence in the Interlanguage Lexicon. *Language Learning*, 50(2):245-309.
- Jarvis, Scott. 2002. Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24:387-418.
- Jarvis, Scott. 2007. Theoretical and Methodological Issues in the Investigation of Conceptual Transfer. *Vigo International Journal of Applied Linguistics*, (4):43-71.
- Jarvis, Scott. 2010. Comparison-based and detection-based approaches to transfer research. *Eurosla yearbook*, 10:169-192.
- Jarvis, Scott. 2011. Conceptual transfer: Crosslinguistic effects in categorization and construal. *Bilingualism: Language and Cognition*, 14:1-8
- Jarvis, Scott og Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. New York og London: Routledge.
- Jin, F, Tor A. Åfarli, og W. A. van Drommelen. 2009. Variability in the L2 acquisition of Norwegian DPs. I *Second Language Acquisition of Articles: Empirical*

- findings and theoretical implications*, redigert av M. d. P. García Mayo og R. Hawkins. Amsterdam/Philadelphia: John Benjamins Publishing.
- Johannessen, Janne B., Kristin Hagen, André Lynum og Anders Nøklestad. 2012. OBT+stat. A combined rule-based and statistical tagger. I *Exploring Newspaper Language. Corpus compilation and research based on the Norwegian Newspaper Corpus*, redigert av G. Andersen. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Johansen, Hilde. 2005. Å passe verden inn i en ytring. *Nordica Bergensia*, 33:97-111.
- Johansen, Hilde. 2007. *Kodebok for ASK*. Bergen: ASK-prosjektet, Universitetet i Bergen. <https://clarino.uib.no/korpuskel/page?page-id=ask-codebook>
- Johansen, Hilde. 2010. Kontroll, Analyse, Reliabilitet, Innlærerkorpus og Transfer – Slik Egner seg Korpus til Studier av Transfer hos Innlærere. I *Systematisk, variert, men ikke tilfeldig. Antologi om norsk som andrespråk i anledning Kari Tenfjords 60-årsdag*, redigert av H. Johansen, A. Golden, J. E. Hagen og A. Helland. Oslo: Novus Forlag.
- Johansen, Hilde. 2011. *En kort innføring i bruk av ASK*. Bergen: ASK-prosjektet, Universitetet i Bergen. <https://clarino.uib.no/korpuskel/files/ask-documentation/en-grunnleggende-innforing-i-bruk-av-ask.pdf%5d>
- Kellerman, Eric. 1978. Giving Learners a Break – Native language intuitions as a source of predictions about transferability. *Working papers on bilingualism* 15:59-92.
- Kellerman, Eric. 1983. Now you see it, now you don't. I *Language Transfer in Language Learning*, redigert av S. Gass og L. Selinker. Rowley, Massachusetts: Newbury House Publishers.
- Kellerman, Eric. 1995. Crosslinguistic influence: Transfer to nowhere? *Annual review of applied linguistics*, 15:125-150.
- Kellerman, Eric og Michael Sharwood Smith. 1986. *Crosslinguistic influence in second language acquisition*. Oxford: Pergamon.
- Król-Markefka, Agnieszka. 2008. Some theoretical considerations on the use of contrastive data in teaching English articles to Polish learners. *Studia Linguistica Universitatis Jagellonicae Cracoviensis*, 125:103-112.
- Larsen-Hall, Jenifer. 2010. *A Guide to doing Statistics in Second Language Research*. New York: Routledge.
- Lakoff, George. 1987. *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.
- Li, Charles N. og Sandra A. Thompson. 1976. Subject and Topic: A New Typology of Language. I *Subject and Topic*, redigert av C. N. Li. London/New York: Academic Press.

- Lindsay, Roger. 2003. Perception and Language. I *Handbook of pragmatics*, redigert av J. Verschueren, J. O. Östman, J. Blommaert og C. Bulcaen. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Lundeby, Einar. 1965. *Overbestemt substantiv i norsk og de andre nordiske språk*. Oslo: Universitetsforlaget.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Moraczewska, Agnieszka M. 2017. *Tilegnelse av bestemthet hos polske norskstudenter. En kasusstudie* [Masteroppgave, Høgskolen i innlandet].
- Nistov, Ingvild. 2001. *Referential choice in L2 narratives: A study of Turkish adolescent immigrants learning Norwegian* [Doktoravhandling, Universitetet i Oslo].
- Nordanger, Marte. 2009. *Keiserens nye klær? Lingvistisk og konseptuell transfer i markeringen av grammatikalisert definit referanse i russiskspråklige og engelskspråklige norske mellomspråk-en studie basert på ASK* [Masteroppgave, Universitetet i Bergen].
- Nordanger, Marte. 2012. Ferdighetsnivå og definit referanse. I *Norsk profil. Det felles europeiske rammeverket spesifisert for norsk. Et første steg*, redigert av C. Carlsen. Oslo: Novus Forlag.
- Nordanger, Marte. 2017. *The encoding of definiteness in L2 Norwegian: A study of L1 effects and universals in narratives written by L1 Russian and L1 English learners* [Doktoravhandling, Universitetet i Bergen].
- Nyqvist, Eeva-Liisa. 2013. Species och artikelbruk i finskspråkiga grundskoleelevers inläraarsvenska: En longitudinell undersökning i årskurserna 7-9/ [Doktoravhandling, Åbo Universitet]
- Odlin, Terence. 1989. *Language Transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- Parrish, Betsy. 1987. A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37(3): 361-383.
- Paulsen, Gro. 2009. *Bestemt om bestemthet? Om voksne innlæreres bruk av kategorien bestemthet i norsk* [Masteroppgave, Universitetet i Oslo].
- Premovic, Marija. 2018. *Bestemthet i norsk mellomspråk: Bruken av kategorien bestemthet i norsk mellomspråk blant voksne innlærere med serbisk som morsmål* [Masteroppgave, Universitetet i Oslo].
- Prince, Ellen. 1981. Toward a taxonomy of given-new information. I *Radical Pragmatics*, redigert av P. Cole. New York: Academic Press Inc.
- Ringbom, Håkan. 2007. *Cross-linguistic Similarity in Foreign Language Learning*. Cleveland: Multilingual Matters Ltd.

- Ringbom, Håkan og Scott Jarvis. 2009. The importance of cross-linguistic similarity in foreign language learning. I *The Handbook of Language Teaching*, redigert av M. H. Long og C.J Doughty. Chichester: Wiley-Blackwell.
- Ringbom, Håkan. 2011. Perceived redundancy or crosslinguistic influence? What L3 learners' material can tell us about the causes of errors. I *New trends in crosslinguistic influence and multilingualism research*, redigert av G. De Angelis and J-M. Dewaele. Bristol/Buffalo/Toronto: Multilingual Matters.
- Rowntree, Derek. 2000. *Statistics without Tears: An Introduction for Non-Mathematicians*. London: Penguin.
- Smólska, Janina og Jan Rusiecki. 1980. The generic noun phrase in English and Polish. *Papers and studies in contrastive linguistics*, 11:39-57.
- Schachter, Jacquelyn. 1974. An error in error analysis. *Language learning*, 24(2):205-214.
- Siewierska, Anna. 1993. Syntactic weight vs information structure and word order variation in Polish. *Journal of Linguistics*, 29(2):233-265.
- Slobin, Dan I. 1996. From "Thought and Language" to "Thinking for Speaking". I *Rethinking linguistic relativity*, redigert av J. J. Gumperz og S. C. Levinson. Cambridge: Cambridge University Press.
- Soboleva, Elena. 2016. Bestemthet i norsk: Utfordringer voksne innvandrere med russisk som førstespråk møter når de bruker bestemthet i norsk mellomspråk [Masteroppgave, Universitetet i Oslo].
- Stockwell, Robert. P., J. Donald Bowen og John W. Martin. 1965. *The grammatical structures of English and Spanish*. Chicago: University of Chicago Press.
- Swan, Oscar E. 2002. *A grammar of contemporary Polish*. Bloomington, Indiana: Slavica Publishers.
- de Swart, Henriëtte og Joost Zwarts. 2009. Less form—more meaning: Why bare singular nouns are special. *Lingua*, 119(2): 280-295.
- Szwedek, Aleksander. 1976. *Word order, sentence stress and reference in English and Polish*. Edmonton, Alberta: Linguistic Research.
- Tenfjord, Kari, Jon Erik Hagen og Hilde Johansen. 2006. The hows and whys of coding categories in a learner corpus (or "How and why an error-tagged learner corpus is not ipso facto one big comparative fallacy"). *Rivista di Psicolinguistica Applicata*, VI (3):93-108.
- Tenfjord, Kari, Jon Erik Hagen og Hilde Johansen. 2009. Norsk andrespråkskorpus (ASK) - design og metodiske forutsetninger. *NOA Norsk som andrespråk* 25(1):52-81.
- Tenfjord, Kari, Paul Meurer og Knut Hofland. 2006. The ASK Corpus - A Language Learner Corpus of Norwegian as a Second Language. I *Proceedings from 5th*

- Thomas, Margaret. 1989. The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10:335-355.
- Trenkic, Danijela. 2002. Establishing the definiteness status of referents in dialogue (in languages with and without articles). *Research Centre for English and Applied Linguistics Working Papers*, 7:142-162.
- Trenkic, Danijela. 2007. Variability in second language article production: Beyond the representational deficit vs. processing constraint debate. *Second Language Research*, 23(3):289-327.
- Trenkic, Danijela. 2008. The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, 11(1):1-18.
- Trenkic, Danijela. 2009. Accounting for patterns of article omissions and substitutions in second language production. *Second language acquisition of articles: Empirical findings and theoretical implications*, 49:115-143.
- Trenkic, Danijela og Nattama Pongpaibroj. 2013. Referent salience affects second language article use. *Bilingualism: Language and Cognition*, 16(1):152-166.
- Utdanningsdirektoratet. 2011. *Det felles europeiske rammeverket for språk: Læring, undervisning, vurdering*. Oslo: Utdanningsdirektoratet.
- Vannebo, Kjell Ivar. 2010. «Nød lærer naken kvinne å spinne» - litt om bruk av naken substantivfrase i norsk. I *Systematisk, variert, men ikke tilfeldig. Antologi om norsk som andrespråk i anledning Kari Tenfjords 60-årsdag*, redigert av H. Johansen, A. Golden, J. E. Hagen og A. Helland. Oslo: Novus Forlag.
- Venneri, Diana. 2011. Demonstrative Pronoun *ten* and Definiteness in Polish. I *Proceedings of the 2011 ESSLLI Student Session*. 23rd European Summer School in Logic, Language, and Information, Ljubljana.
- Weinreich, Uriel. 1963. *Languages in contact: Findings and Problems*. New York: Linguistic Circle of New York.
- Whorf, Benjamin Lee. 1956. Linguistics as an exact science. I *Language, thought, and reality: selected writings of Benjamin Lee Whorf*, redigert av J. B. Carroll. Cambridge: MIT Press.
- Wiese, Bernd. 2011. Optimal specifications: On case marking in Polish. I *Syntax and morphology multi-dimensional*, redigert av A. Nolda og O. Teuber. Berlin: Mouton de Gruyter.
- Young, Richard. 1996. Form-function relations in articles in English interlanguage. I *Second language acquisition and linguistic variation*, redigert av R. Bailey og D. R. Preston. Amsterdam/Philadelphia: John Benjamins Publishing Company.

VEDLEGG

Vedlegg 1: Personvariabler for delkorpuset i undersøkelse 2

PERSON-/BAKGRUNNSVARIABLER				
S1=	Engelsk	Polsk	Norsk	Totalt
Alder				
17-19 år	3		1	4
20-29 år	15	24	10	49
30-39 år	14	16	4	34
40-49 år	3	1	9	13
50-59 år	4		11	15
60-69 år	1		6	7
Kjønn				
Kvinne	27	36	25	88
Mann	14	4	16	34
Ikke oppgitt		1		1
Utdannelse				
Høgskole/universitet	27	31	29	87
Videregående skole	6	10	12	28
Grunnskole				0
Annet	7			7
Ikke oppgitt	1			1
Yrke				
Akademisk yrke	4	1	2	7
Helsearbeid	1	15	4	20
Hjemmeværende	2	2	1	5
Kontorarbeid	5	1	9	15
Kultur	1	1		2
Manuelt arbeid	3	3	1	7
Opplæring/underv	2	1	10	13
Servicenæring	3	5		8
Ikke oppgitt	20	13		33
Annet			6	6

VARIABLER KNYTTET TIL EKSPONERING FOR OG BRUK AV NORSK			
S1=	Engelsk	Polsk	Totalt
Oppholdstid (år)			
0-1	12	17	29
1-2	17	12	29
2- 3	5	6	11
3-4	2		2
4-5	2		2
Over 5	3	4	7
Ikke oppgitt		2	2
Sosial omgang			
Ja	41	24	65
Nei		6	6
Ikke oppgitt		11	11
Bruk av norsk på fritiden			
Ja	34	20	54
Nei			0
Ikke oppgitt	7	21	38
Bruk av norsk på jobb			
Ja	27	21	48
Nei			0
Ikke oppgitt	14	20	34
VARIABLER KNYTTET TIL FORMELL NORSKOPPLÆRING			
S1=	Engelsk	Polsk	Total
Kurslengde (timer)			
0-200	6	1	7
201-400	14	19	33
402- 5	2	8	10
502-850	13	6	19
851-1500	1	1	2
Ikke oppgitt	5	6	11
Type norskkurs			
AMO		2	2
Folkeuniversitet	9	2	11
Grunnskole voksne	1	1	2
Kommunale	27	30	57
Privat		2	2
Universitet		1	1
Videregående	1		1
Ikke oppgitt	3	3	6
Tid siden norskkurs (måneder)			
0-5	6	10	16
6-12	17	16	33
13-24	9	14	23
25-36	5		5
Over 37	3		3
Ikke oppgitt	1	1	2

Vedlegg 2: Tema som er valgt for tekstene til de tre gruppene

S1=	Engelsk	Polsk	Norsk	Totalt
Tema i polske og engelske tekster:				
Barneoppdragelse		6		6
Bomiljø	20			20
En sport du liker		1		1
Fjernsynsprogram for barn		1		1
Fremmedspråklæring		1		1
Frivillig hjelp i organisasjoner		2		2
Før og nå		3		3
Gjenbruk		1		1
Mobiltelefoner	9	10		19
Møte med norsk kultur	3			3
Nyheter	3	5		8
Trafikk og boligområder		3		3
Vaner og tradisjoner	4			4
Vennskap	1	2		3
Viktige verdier i livet	1	6		7
Tema i norsk kontrollkorpus:				
Sunnhet og helse			21	21
Penger og andre verdier			12	12
Frihet og ansvar			7	7
Spillemaskiner			1	1
Sum	41	41	41	123

Vedlegg 3: Søkeuttrykk for grunnsøk fra ASK (undersøkelse 2)

Grunnsøkene som er gjort for å hente ut enkle substantivfraser i korpuset tar utgangspunkt i den automatiske taggingen som er gjort på ordklasse- og morfologisk og syntaktisk nivå. Søkene går over flere korpusposisjoner (skilt fra hverandre med [hakeparentes]), og for hver posisjon har jeg spesifisert enten *tilstedeværelse* av ett trekk, eller *fravær* av et trekk (unntak).

Grunnsøk	Søkeuttrykk	
1 Substantiv ubestemt entall	[!<> & pos != "adj prep det" & syntax = (!"@<p-utfyll" !"@adj>" !"@adv>" !"@det>" !"@subst>")] {[!"@.*" & pos = "subst" & morph = ("ub" "ent") & path = (!"head")} [!<> & pos != "det" & syntax = (!"@<det")]	
2 Substantiv ubestemt flertall	[!<> & pos != "adj prep det" & syntax = (!"@<p-utfyll" !"@adj>" !"@adv>" !"@det>" !"@subst>")] {[!"@.*" & pos = "subst" & morph = ("ub" "fl") & path = (!"head")}] [!<> & pos != "det" & syntax = (!"@<det")]	
3 Substantiv bestemt entall	[!<> & pos != "adj prep det" & syntax = (!"@<p-utfyll" !"@adj>" !"@adv>" !"@det>" !"@subst>")] {[!"@.*" & pos = "subst" & morph = ("be" "ent") & path = (!"head")} [!<> & pos != "det" & syntax = (!"@<det")]	
4 Substantiv bestemt flertall	[!<> & pos != "adj prep det" & syntax = (!"@<p-utfyll" !"@adj>" !"@adv>" !"@det>" !"@subst>")] {[!"@.*" & pos = "subst" & morph = ("be" "fl") & path = (!"head")}] [!<> & pos != "det" & syntax = (!"@<det")]	
5 en/ei/et + substantiv	[!<> & pos != "prep"] [lemma = "en ei et"] {[!"@.*" & pos = "subst" & path = (!"head") "}]	



Grafisk design: Kommunikasjonseksperimenter, UIB / Trykk: Skjerve Kommunikasjon AS



uib.no

9788230856086 (print)

9788230869543 (PDF)