

Generisk anafor i tekst – en korpusundersøkelse av norsk¹

Anna Kurek-Przybilski

Adam Mickiewicz universitet, Poznań, Polen

Generic Anaphora in Text. A Corpus-Based Study of Norwegian

This paper examines generic anaphora in Norwegian. The study is corpus-based and consists of an analysis of 170 generic texts retrieved from an online encyclopaedia *Store norske leksikon*. The texts were divided into five thematic categories: people, animals, plants, tools and other. The analysis focuses both on generic nouns in each of the texts, as well as anaphoric references in them. Generic anaphora can be expressed with the use of pronouns as well as nouns that are synonymous to the main noun in each of the texts (e.g. *animal* as an anaphoric reference to *dog* etc.). The study shows that the forms of anaphoric expressions were often used throughout the text in accordance with noun forms. What is more, the type of nouns influenced the choice of a given noun form or a pronoun in anaphoric expressions. For instance, animals considered gregarious and plants that seem uncountable appeared in definite plural form (e.g. *the mosses* and *the plants* for the anaphoric reference).

Keywords: genericity, generics, anaphora, Norwegian, corpus linguistics

1. Generisk anafor og *donkey*-setninger

I faglitteraturen defineres anaforisk referanse som et uttrykk som peker på en antecedent nevnt tidligere i setningen (Lyons 1977: 659). Anaforiske setninger lar oss referere til en person eller et objekt ved bruk av pronomen eller andre substantiver (vanligvis synonyme til antecedenten). Et eksempel på anaforisk referanse er vist i (1):

(1) *Jon har en bil. Den er rød*

der pronomenet *den* peker tilbake til substantivet en bil. Dette kalles for direkte anafor, der antecedenten er nevnt tidligere i teksten eller i ytringen. Et mulig problem med tolkning av anaforisk referanse kan oppstå når et ubestemt substantiv ikke refererer til et enkelt objekt, men heller uttrykker en ikke-spesifikk referanse, slik som i eksempel (2).

¹ Undersøkelsen ble utført med støttet av National Science Center (NCN) i Polen. Prosjektnummer: 2017/25/N/HS2/00003.

- (2) *John wants to catch a fish and to eat it for supper.* (Lyons 1977: 675)
 'John vil fange en fisk og spise den til middag.'

Her kan man ikke tolke *en fisk* som en bestemt, spesifikk fisk som John skal fange, men heller som en hvilken som helst fisk han kanskje får på fiskekroken. Derfor uttrykker pronomenet *den* en referanse til noe ubestemt, noe som ikke er nevnt tidligere i setningen.

Anafor kan klassifiseres med hensyn til forskjellige aspekter, som for eksempel syntaktiske kategorier, kontekst, tolkning av anaforiske uttrykk eller sannhetsvilkår (Huang 2000: 2). I denne artikkelen analyseres anafor med hensyn til sannhetsvilkår og dens refererende funksjoner. Én av typene anafor er en referensiell anafor der man refererer til et objekt i den eksterne verden enten direkte eller indirekte. Antecedenten kan da være i samme setning eller diskurs (Huang 2000: 5). Andre anaforiske uttrykk refererer ikke til en bestemt enhet, men det er mulig å tolke dem slik siden de er avhengige av konteksten eller andre kvantifiserende uttrykk i samme setning/diskurs (Huang 2000: 6).

Generisk anafor fungerer på en lignende måte, den lar oss nemlig referere til noe abstrakt ved å bruke pronomen eller synonymer til antecedenten. I generiske setninger er antecedenten selv noe abstrakt og ubestemt, siden generiske substantiver og nominalfraser kan referere til hele arter eller kun til representanter av en gitt gruppe. Generisk referanse, og dermed også generisk anafor, uttrykker noe om *alle* eller *nesten alle* medlemmer av en gruppe (Lyons 1977: 454), og denne referansen kan gjengis ved bruk av forskjellige substantivformer. Man kan til og med referere til en art ved å bruke ubestemt entallsform, så vel som ubestemt flertallsform. Disse to tolkes litt annerledes,² men begge kan også forekomme i anaforiske setninger slik som i (3) og (4).

- (3) *Kelly is seeking a unicorn, and Millie is seeking it, too.* (Carlson 1977: 425)
 'Kelly leter etter en enhjørning og Millie leter etter **den** også.'³
- (4) *Queenie is seeking unicorns, and Phil is seeking them, too.* (Carlson 1977: 425)
 'Queenie leter etter enhjørninger og Millie leter etter dem også.'

Som vist i setningene (3–4) kan generisk anafor uttrykkes både ved bruk av første- og andrepersons pronomen. I begge eksemplene er det klart at antecedenten for *den* og *dem* er generiske substantiver, henholdsvis *en enhjørning* og *enhjørninger*. Ifølge noen av teoriene er det ikke mulig at et pronomen refererer til en hel art, slik som er tilfelle i eksemplene (se f.eks. Carlson 1979). Vi skal likevel betrakte slike pronomen i anaforiske setninger som generiske, uansett om de referer til hele arten eller til de fleste representantene av en gitt art. En mer detaljert diskusjon om generisitet finnes f.eks. hos Krifka *et al.* (1995), Chierchia (1998), Leslie (2007) m.fl.

² Tolkningen av enkelte substantivformer i generiske kontekster er ikke hovedtemaet i denne artikkelen. For flere opplysninger se f.eks. Borthen (2003) og Halmøy (2016).

³ Alle eksemplene fra utenlandsk faglitteratur er i min oversettelse og i original skrivemåte. Mulige feil er derfor kun mitt ansvar.

Nevner man anaforiske setninger, bør man også ta hensyn til såkalte *donkey*-setninger. Det er setninger som uttrykker anaforisk referanse uten at det finnes noen antecedent i ytringen. I den forstand ligner de veldig mye på generisk anafor. Antecedenten i *donkey*-setninger forstås ut av konteksten og dette er særlig relevant for generiske uttrykk. Navnet på den type setninger kommer fra Geachs (1980 [1962]: 143; 1972) som utviklet sin teori på Burleighs (1988 [originalt ca. 1328, nevnt Seuren 2010: 169]) eksempel fra det 13. århundre:

- (5) *Omnis homo habens asinum videt illum.*
'Hver mann som eier et esel, ser det.'

Det kan være naturlig å tenke at antecedenten for pronomenet *det* i setningen (5) er *et esel*, men det er ikke tilfelle. Substantivet *et esel* er en ikke-spesifikk referanse som også kan tolkes generisk. Derfor kunne man påstå at pronomenet ikke refererer til *et esel* direkte siden det ikke er mulig å referere til en abstrakt representasjon av en art. En av mulige tolkninger er at antecedenten i dette eksemplet er hele predikatet, altså *det å ha et esel* (Seuren 2010: 169), som Burleigh foreslo selv. Senere avviste forskere og logikere denne måten å tolke *donkey*-setninger på (se f.eks. Geach 1972, 1980). Uansett hva som karakteriserer forskjellige teorier og tilnærminger, forblir problemet med tolkning av *donkey*-konstruksjoner:

Donkey pronouns, therefore, behave like referring expressions even though they are not allowed to do so under the statuses of current logic. (Seuren 2010: 171)

Donkey-setninger er ikke særlig hyppige i daglig tale eller i skriftlig språk (f.eks. i korpuset), men uansett forårsaker de problemer for språkvitere og lingvister som ikke klarer å tolke dem på samme måte som direkte anafor. *Donkey*-setninger forstås her som en type generisk anafor.

1.1. Forskning på generisitet og generisk anafor

Generisitet som et språklig fenomen har vært til interesse for forskerne i årevis. Fenomenet inspirerer både språkforskere og filosofer (for filosofisk beskrivelse av generisitet se f.eks. Pelletier 2010b og Mari *et al.* 2013b m.fl.). Når det gjelder lingvistisk tilnærming til problemet finnes det også flere måter man kan undersøke dette fenomenet på. De første teoriene er basert på logikken og de viser til de abstrakte aspektene ved generisitet. Carlson (1977, 1982, 1983) var en av de første forskerne som analyserte generisitet i engelsk og laget en koherent teori. Den ble senere utviklet, kritisert, endret og forbedret mange ganger (også med hensyn til andre språk), til og med av Carlson selv, men også av andre lingvister (f.eks. Carlsson 2012 for svensk; Chierchia 1998 for engelsk, italiensk og kinesisk; Oosterhof 2008 for nederlandsk m.fl.).

En annen tilnærming til problemet baserer seg på kognitiv grammatikk og språkvitenskap. Også her er det engelsk som er beskrevet i de fleste publikasjonene, men man finner enkelte studier angående andre språk (se f.eks. Lakoff 1990; Raden 2009; Radden & Dirven 2007 m.fl. for engelsk og Karczewski 2016 for polsk).

Når det gjelder skandinaviske språk er saken litt mer komplisert. Det finnes enkelte opplysninger om generiske setninger og generisk referanse i grammatikkbøker, se f.eks. *Svenska Akademiens Grammatik* (Teleman *et al.* 1999) for svensk, *Grammatik over det Danske Sprog* (Hansen & Heltoft 2011) for dansk og *Norsk referansegrammatikk* (Faarlund *et al.* 1997) for norsk. Dessverre er beskrivelsene i grammatikkbøker ganske overfladiske og gjelder ikke en detaljert analyse av fenomenet. I faglitteraturen derimot, finner man flere detaljer, f.eks. hos Hansen (1994) for dansk, Pettersson (1976) og Carlsson (2012) for svensk. Generisitet i norsk ble beskrevet bl.a. av Borthen (2003), Halmøy (2016) og Rosén & Borthen (2017), men fenomenet var ikke hovedtema i disse studiene. En oversikt over norsk faglitteratur angående generisitet finnes også hos Kurek (2017).

Generisk anafor som et logisk og språkvitenskapelig problem utgjør en del av forskningen på generisitet i språket. Som det har blitt sagt før, har fenomenet skapt interesse hos forskerne i mange år. Hvordan burde man egentlig tolke generisk anafor? Hvilken funksjon har pronomenene og substantivene som er brukt anaforisk i slike setninger?

Generisk anafor kan uttrykkes på forskjellige måter, slik som generisitet selv. Bortsett fra sånne setninger som (2–4), finner man også følgende eksempler på generisk anafor:

- (6) *My mother hates raccoons because they stole her sweet corn last summer.* (Carlson 1977: 433)
'Mora mi hater vaskebjørner fordi de stjal maisen hennes sist sommer.'
- (7) *Dogs that have three legs are ill-formed. Such dogs (= dogs that have three legs) should be protected from cats.* (Carlson 1979: 71)
'Hunder som har tre bein er dårlig formet. Såne hunder (= hunder som har tre bein) skal forsvares fra katter.'
- (8) *The blueberry is so called because of its color.* (Ojeda 1991: 386)
'Blåbæret er kalt slik/dette (/sånn) på grunn av dets farge.'
- (9) *A carpenter needs his tools to build houses.* (Bennett-Kastor 1996: 299)
'En snekker trenger verktøyene sine for å bygge hus.'

Setningene (6–9) viser at generisk anafor kan ha forskjellige former og tolkninger. Carlson sitt eksempel med vaskebjørner skildrer et logisk problem med tolkning av en sånn setning. Vi vet at pronomenet *de* refererer til *vaskebjørner* nevnt tidligere i setningen, men det er også klart at uttrykket *å hate vaskebjørner* er generisk og det betyr ikke at hatet er rettet kun mot de vaskebjørnene som stjal maisen.

I eksempel (7) ser vi at generisk anafor også kan uttrykkes ved bruk av substantiver som refererer til hele arten. Arten er antecedenten i forrige setning. Her kan man også lure på om nominalfrasen *hunder med tre bein* er generisk eller ikke, men det er klart at den ikke refererer til en spesifikk gruppe sånne hunder, men uttrykker heller en generalisering om alle eller de fleste hundene med tre bein.

Ojedas og Bennett-Kastors eksempler viser derimot hvordan man kan uttrykke generisk anafor i setninger der generiske substantiver forekommer i entallsform, bestemt og ubestemt.

I faglitteraturen finner man flere eksempler på bruk av pronomen og substantiver som generisk anafor. Disse er i hovedsak engelske, og det mangler dermed en empirisk studie angående skandinaviske språk. I denne artikkelen skal det presenteres data vedrørende generisk anafor hentet fra et norsk korpus. Målet med studien er å skildre fenomenet i norsk ved bruk av en stor datasamling. De tidligere undersøkelsene innen feltet (f.eks. Borthen 2003; Halmøy 2016) er empiriske studier, men baserer seg kun på enkelte setninger. I tillegg er generisitet og generisk anafor ikke hovedtema i disse undersøkelsene.

2. Metode

Denne studien bygger på korpusmaterialet som ble samlet i perioden januar-juni 2018 som en del av et større prosjekt. Korpuset består av 170 generiske tekster (over 27 000 ord). Alle tekstene ble lastet ned fra *Store norske leksikon*. De ble så delt inn i fem kategorier: 1) mennesker, 2) dyr, 3) planter, 4) verktøy og 5) andre. Hver av tekstene hadde minst ett avsnitt. Målet var å samle tekster som omhandlet så ulike tematikk som mulig, og det var derfor krevende å samle inn tekster med lik lengde. En tekst om f.eks. en kjent dyreart kan være mye lengre enn en tekst om et hverdagslig objekt som saks.

Den første kategorien, altså ”mennesker”, inneholder 25 tekster som gjelder menneskegrupper, yrker og andre betegnelser som brukes når man refererer til mennesker. Den andre kategorien tekster, altså ”dyr”, er en samling av 25 beskrivelser av forskjellige dyrearter. Målet her var å samle inn tekster om både *kjente* arter som for eksempel hund, men også om noen eksotiske arter som ikke forekommer i Skandinavia. Den tredje kategorien, ”planter”, inneholder 25 tekster om plantearter. Kriteriene for valg av tekstene til denne kategorien var de samme som i tilfellet av dyr. De tre kategoriene tekster er også klassifisert som levende substantiver (eller *nesten-levende* i tilfellet planter).

Kategoriene 4 og 5 gjelder ikke-levende substantiver. Kategorien ”verktøy” inneholder 25 tekster om gjenstander som man kan bruke til å lage noe annet eller til å utføre forskjellige typer arbeid. Den siste kategorien, ”andre”, er litt spesiell av to grunner. For det første, inneholder kategorien 70 tekster, altså flere enn de andre kategoriene. For det andre, gjelder hver av tekstene substantiv som ikke passer til de andre gruppene. Det å lage en større kategori av diverse tekster har gjort det mulig å lage en detaljert datasamling.

2.1. Datautvalg og tagging av korpustekstene

Det å velge riktig materiale til analyse er et viktig steg, særlig når man forsker på et fenomen som ikke har blitt beskrevet på, eller angående norsk før. Generisitet

og generisk anafor er denne type fenomen og det mangler tilgjengelig materiale som kunne utnyttes til analysen. Korpus man finner på nettet (f.eks. aviskorpus) er sikkert en viktig kilde, men det er ikke mulig å laste ned et gitt antall tekster og tagge dem for generisitet etterpå. Derfor ble det bestemt at et spesialisert korpus av norske generiske tekster skulle bygges.

Hvilke sjangre burde man satse på? Det finnes undersøkelser der man utnytter skjønnlitterære tekster (f.eks. Behrens 2005), men de er veldig få og selve analysen er ganske begrenset når det gjelder kontekst. Flere forskere bruker generiske setninger laget for prosjektets skyld (se f.eks. Carlson 1977, 1982; Chierchia 1998; Krifka *et al.* 1995; Mari *et al.* 2013b m.fl.), men disse bygger hovedsakelig på forskerens intuisjon. En annen måte å drive med forskning av generisitet er å bruke to forskningsmetoder, slik som hos Oosterhof (2008) som utnytter både korpusdata og spørreundersøkelser for å analysere generisitet i nederlandsk og nederlandske dialekter. Man kan også bygge et korpus selv og velge bare den type tekster der man sikkert finner generiske uttrykk, slik som Carlsson (2012) har gjort i sin studie om generisitet i svensk.

Materialet til prosjektet er tofoldig: det består av encyklopediske tekster og to spørreundersøkelser (en kombinasjon av Oosterhoffs og Carlssons metodologi). I denne artikkelen fokuserer jeg kun på korpuset.

Valg av *Store norske leksikon* som en kilde for korpustekstene er både enkelt og komplisert. På den ene siden er det ganske enkelt å finne generiske tekster i et leksikon og man kan velge hvilke kategorier substantiver man vil ha. På den andre siden er et sånt materiale ganske begrenset fordi alle tekstene er faglige, altså at de tilhører samme sjanger. Tidligere undersøkelser har allerede visst at et sånt materiale kan gi tilstrekkelige resultater (Carlsson 2012), dermed ser jeg ikke det som en ulempe.

Alle tekstene i korpuset ble samlet inn i atskilte .xlsx-filer (170 filer totalt). Tekstene ble limet inn i én celle i hver av filene. Noe som hadde stor betydning for videre analyser utført i R-programmet. Filene ble så lastet opp i R-programmet som brukes til utføring av statistiske analyser, men også til håndtering av store datasamlinger. Ved bruk av UdPipe-pakken ble alle substantivene og pronomenene i korpuset filtrert. Alle substantivene og pronomenene ble dobbeltsjekket for mulige feil og korrigerert der det var nødvendig.⁴ På den måten var det mulig å gå gjennom materialet og tagge generiske substantiver og generisk anafor i tekstene. Tagging ble gjennomført direkte i .xlsx-filene som etterpå ble lastet inn i R-programmet på nytt.

⁴ UdPipe-pakken som ble brukt i denne undersøkelsen fungerte veldig godt og feilene i filtreringen av substantivene og pronomenene var veldig få. Bortsett fra det ble en tilfeldig prøve av filtrerte og taggede tekster dobbeltsjekket av en morsmålsbruker av norsk.

3. Resultater

I korpuset ble det funnet 228 anaforiske setninger. Det er verdt å nevne at både setningene med pronomen og synonymer til hovedsubstantivet (altså temaet) i hver av tekstene ble annotert som anaforisk referanse. Som sagt, forstås generisk anafor her både som *klassisk* anaforisk referanse med bruk av pronomen, men også som bruk av synonymer (f.eks. ordet *arten* i stedet for *hund* osv.). Tabell 1 viser korpus-elementer tagget som enten *generisitet* eller *anafor*.

Tabell 1. Generisitet og generisk anafor i korpuset

Korpus	Forekomst	Prosent
substantiver og pronomen	9319	100
generisitet	870	9.34
anafor	228	2.45

Den kvantitative analysen av innsamlet data viser at langt flere av alle substantivene og pronomenene i korpuset er generiske (9.34%), mens kun 2.45% av dem uttrykker anaforisk referanse. Generisk anafor virker ikke til å være den oftest brukte formen i generiske tekster.

Den deskriptive delen av analysen fokuserer på forholdene mellom substantivene i korpuset og generisk anafor. I noen av tekstene fantes det flere anaforiske referanser, mens i andre var det veldig få av dem eller ingen i det hele tatt. Det kan være flere grunner til det, både stilistiske (hver tekst er jo skrevet av en annen forfatter) og semantiske/pragmatiske.

3.1. Pronomen

Pronomen forekommer i 75 av 228 anaforiske setninger. Dette utgjør 32.89% av alle anaforiske referanser i korpuset. Bruk av pronomen virker til å være den mest naturlige måten å uttrykke anafor på, men i generiske tekster forekom anaforiske substantiver mye oftere (67.11% av tilfellene).

Bruk av pronomen i generisk anafor er ikke tilknyttet en bestemt form av substantivet. I eksempel (10) brukes *den* etter nakent substantiv, altså peker pronomenet på en generisk antecedent:

- (10) *Ukulele, firestrengt, lite instrument i gitar-familien. Den ble trolig utviklet på grunnlag av en portugisisk gitar-type. På 1800-tallet ble ukulelen et populært folkeinstrument på Hawaii, og ble i 1920-årene innført i amerikanske og europeiske danseorkestre.*⁵

⁵ Min framhevelse i alle eksemplene fra korpuset.

Etter det nakne substantivet og pronomenet i teksten forekommer *ukulele* i bestemt form. Noe som også kan være et resultat av den kontekstuelle sammenhengen. Tekstens tema blir introdusert ved bruk av naken form (det er altså et oppslagsord), så følger pronomenet og bestemt form. En annen mulig tolkning er at *ukulele*, når instrumentet ikke forekommer i teksten som oppslagsord, har en bestemt form fordi det er en underart av instrumentgruppen *gitar*. Dette blir eksplisitt sagt allerede i første setning i (10) der vi leser at *ukulele* er *et lite instrument i gitar-familien*.

Det at en slik bruk av substantivformene kan være stilistisk motivert ser vi i eksempel (11) der forfatteren ikke bruker bestemt form etter pronomenet:

- (11) **Krokodiller** er en orden av store krypdyr. **De** kjennes på den lange flate snuten, og halen og de korte beina. Det finnes 27 nålevende arter **krokodiller**.

Det er godt mulig å referere til en hel art ved å bruke ubestemt flertallsform, slik som man oftest gjør i engelsk (se f.eks. Carlson 1977; Chierchia 1998). Ubestemt flertall er derimot ikke den eneste formen som tolkes generisk i norsk (se delkapitlet 3.2 nedenfor).

Nakent substantiv blir brukt i korpuset ikke bare som oppslagsord, men også som subjekt, som vist i eksemplene (12–13).

- (12) **Solsikke** dyrkes i varmttempererte strøk på alle kontinenter som olje- og fôrplante, så langt nord som Sør-Sverige. De største produsentene er Russland, Ukraina og Argentina. I Norge dyrkes **den** som prydblant og finnes av og til forvillet, ofte spredt fra fuglebrett.
- (13) [Saksofon] har konisk resonator, klaffmekanisme og lignende munnstykke som klarinett. **Den** bygges i forskjellige størrelser fra sopranino til bass. Hovedtypene er sopran (i Bb), alt (i Ess), tenor (i Bb) og baryton (i Ess).

Pronomenet som følger etter det nakne substantivet i hver av setningene, altså *den*, peker igjen tilbake til artene, henholdsvis *solsikke* og *saksofon*.

I tekstene følger pronomen også substantiver i bestemt form flertall som i teksten i (14) om boremaskiner.

- (14) Har **boremaskinene** flere borspindler, kalles de flerspindel-boremaskiner. **Disse** benyttes mye ved seriefabrikasjon og kan ha 4–120 spindler. I større verksteder benyttes ofte radialboremaskiner. **De** egner seg bedre enn **søyleboremaskiner** for boring i tyngre gods og hvor det skal bores flere huller. **Disse** har en borspindel som er festet til en sleide, som kan beveges på en svingbar horisontal arm. Hvis spindelen dessuten kan stilles skrått, kalles **maskinen** universal radialboremaskin.

Påpekende pronomen *disse* peker tilbake til *søyleboremaskiner* som en underart av alle boremaskinene. Interessant nok virker norsk litt annerledes enn f.eks. engelsk der underarter oftest markeres med bestemt form (Mari *et al.* 2013b: 77) og hovedarter med ubestemt form av substantivet. Noen hevder likevel at formene kan brukes om hverandre uten at tolkningen blir annerledes i sånn kontekst (Lyons 1977: 196).

3.2. Substantiver

Generisk anafor i korpuset blir hovedsakelig uttrykt ved bruk av substantiver, oftest synonymer til hovedsubstantivet i hver av tekstene. Ordene som forekommer oftest i denne funksjonen er bl.a. *art*, *dyr*, *maskin*, *hunn/hann*, men også *plante*, *gruppe* og *slekt*.

Substantivene som blir brukt i anaforiske setninger forekommer oftest i ubestemt form flertall (60 forekomster), noe som tilsvarer teoriene om at nakne substantiver i flertall som regel refererer til arter i generiske setninger (se f.eks. Carlson 1979, 1982; Krifka *et al.* 1995; Mari *et al.* 2013b m.fl.).

Tabell 2. Substantivformer og pronomener i anaforiske setninger

Substantivform	Forekomst	Prosent
nakent substantiv	13	5.70
ubestemt entall	6	2.63
bestemt entall	36	15.79
ubestemt flertall	60	26.32
bestemt flertall	38	16.67
pronomen	75	32.89

Tabell 2 viser også at ubestemt entallsform, f.eks. *en art*, ikke var særlig hyppig i korpuset. Det samme gjelder for nakne substantiver brukt som anaforisk referanse, f.eks. *art*, *gruppe* osv. Tagging av noen av substantivene som nakne var utfordrende. Dette gjaldt ordene som har samme flertallsform som naken form, f.eks. *dyr*. Disse ble tagget med hensyn til kontekst og konsultert med to morsmålsbrukere av norsk. Likevel baserer taggingen seg på språklig intuisjon som varierer fra en person til en annen. Antallet sånne substantiver var relativt lavt sammenliknet med andre, så det er lite sannsynlig at en mulig feiltolkning skal påvirke hele undersøkelsen.

3.2.1. Nakent substantiv og entallsform

Nakent substantiv i korpuset var ikke den hyppigst brukte formen når det gjelder anaforisk referanse, men likevel kan man finne eksempler der naken form uttrykker generisk anafor.

- (15) **Sørgepil**, hybrid (kryssning) mellom *Salix alba* og *S. babylonica*. **Tre** i viferfamilien med lange, hengende grener.

I eksempel (15) er det både oppslagsordet *sørgepil* og anaforisk substantiv *tre* som er brukt i naken form. Det er lite sannsynlig at det er en kontekstuell sammenheng

her. Heller er det selve den encyklopediske skrivestilen som gjør at setningene ovenfor er skrevet på en sånn måte. Nakne substantiver som oppslagsord og ellipser er typiske for encyklopediske tekster. Ellipsen i (15) kan være en av årsakene til mangel på artikler og bestemthet, men den påvirker ikke selve tolkningen av eksemplet. Setningen *Tre i vierfamilien med lange, hengende grener* er generisk siden den inneholder et predikat som bare kan brukes når man refererer til hele arten – det er jo umulig å si at kun ett tre fra denne arten tilhører en viss familie. Nakent substantiv i (15) uttrykker derfor en generisk referanse som peker direkte tilbake til antecedenten *sørgepil*.

- (16) *Karakteristisk for alle maur er forbindelsen mellom for- og bakkropp. Ett eller to av bakkroppens forreste ledd er avsnørt og danner en stilk. Normalt har en maurart tre kaster; hanner, hunner og arbeidere, som er hunner med forkrøplede kjønnsorganer.*

I tekstene som (16) spiller ubestemt entallsform to roller: den uttrykker generisk anafor, men indikerer også at det kan finnes flere underarter av det omhandlede dyret, nemlig maur. Adverbet *normalt*, som er ett av de såkalte generiske adverbene (se f.eks. Lyons 1977; Krifka *et al.* 1995), antyder at det å ha tre kaster kan være sant for de fleste artene, mens ikke nødvendigvis for alle. En kan godt forestille seg en maurart som ikke danner kaster, men som likevel klassifiseres som maur.

- (17) *Oteren er utbredt over store deler av Europa, sør for tundragrensen, samt i Nord-Afrika. Den har imidlertid gått tilbake i antall over store deler av utbredelsesområdet. Arten forekommer i Norge over hele landet, men med hovedtyngden i de nordlige delene.*

Bestemt entallsform av substantivet forekommer betydelig oftere enn nakent substantiv og ubestemt entallsform analysert ovenfor. Den bestemte formen, som *oteren* for hovedsubstantivet og *arten* for generisk anafor, refererer til hele arten *oter*. Pronomenet *den* i (17), som også står i entall, kan være et resultat av kontekstuell og stilistisk tilpasning av substantivformene, men det kan også være slik at dyret *oter* er sett på som underart (sammenlign med eksempel (10) om ukulele). Her tar vi ikke hensyn til biologisk inndeling av artene, men heller til tolkning som kun baserer seg på språklig intuisjon. Det samme gjelder for når man ser på visse arter som selskapelige og dermed refererer til dem ved å bruke bestemt flertallsform (se seksjonen 3.2.2 nedenfor).

Eksempelene (10–17) viser at norsk ligner ganske mye på engelsk i visse tilfeller når det gjelder generisk referanse. Generisk anafor følger noen av mønstrene beskrevet i faglitteraturen, men mangfoldet av substantivformene i slike tekster indikerer at det muligens finnes flere retningslinjer i norsk når det gjelder generisitet og generisk anafor.

3.2.2. Flertallsform

En av forskjellene mellom generisk referanse i engelsk og norsk er bruk av flertallsform, særlig bestemt. Bestemt entallsform er ikke noe uvanlig i generiske kontekster i engelsk, heller ikke i en rekke andre språk som ble analysert med hensyn til dette (se f.eks. Kabatek & Wall 2013 for romanske språk).

- (18) *Skatene svømmer ved bølgende bevegelser av brystfinnenes frie kanter. Enkelte arter kan avgi elektriske støt. Bare et fåtall føder unger; de fleste legger egg.*

Ubestemt flertall *enkelte arter* i (18) påvirkes i en viss grad av kvantoren *enkelte*. Grammatisk sett er det mulig å si *de enkelte artene*, men dette ville innebære at det menes et bestemt antall arter som avgir elektriske støt. Det er klart her at ubestemt flertallsform betyr, slik som ubestemt entall i teksten i (16), at man kun refererer til noen av artene. Det er altså eksplisitt at noen av dem ikke har denne egenskapen.

Uttrykket *et fåtall* er også et interessant element i eksempel (18). Grammatisk sett står uttrykket i entall, men på det semantiske nivået uttrykker det en referanse i flertall. Også her kan vi forstå ut av konteksten at egenskapen til å føde unger eller legge egg kun gjelder noen av artene.

- (19) *Apebrødre, baobab, Adansonia digitata, tre i kattostfamilien. Afrikansk savannetre med mektig krone og kolossalt tykk stamme (inntil 9 meter i diameter). Fruktene, apebrød, som er langstilkete, agurklignende og syrlige, er et viktig næringsmiddel. De modnes når løvet er falt av treet. På eldre trær blir stammen hul, og den kan da brukes som bolig. Veden brukes til båter, barkfibrene til snøre og tauverk. I barken finnes et stoff som er virksomt mot feber. Treet spiller en sentral rolle i boken Den lille prinsen (Le petit prince) av Antoine de Saint-Exupéry.*
- (20) *Måkene er hvite, grå og svarte (svartbrune, rosenmåken svakt rosa) med lange vinger og til dels svevende flukt. Først ved 2–4 års-alder anlegges fullvoksen fjærdrakt, og fuglene blir forplantningsdyktige.*
- (21) *Edelgran, Abies, er en slekt i furufamilien med 40–46 arter i Asia, Europa, Nordafrika og Amerika. Ingen er viltvoksende i Norge, men mange arter blir plantet og finnes av og til forvillet. Trærne har flate nåler og opprette kongler som mister skjellene ved modningen.*

Som det ble nevnt tidligere, er norsk ganske spesielt i den forstand at det brukes bestemt flertallsform til å uttrykke generisk referanse og generisk anafor. Tekstene i (19–21) viser akkurat dette fenomenet.

I eksemplene (19) og (21) ovenfor står hovedsubstantivene i naken form, altså *apebrødre/baobab* og *edelgran*. Formen av setningene der de forekommer indikerer at de kun er oppslagsord, men også anaforisk referanse forekommer i naken form i (19) – *afrikansk savannetre*. Setningen har ingen verbfrase, men anaforen peker tydelig tilbake til *baobab*, altså hele arten.

De andre anaforiske referansene som *fruktene*, *veden* og *barken* i (19) viser fortsatt tilbake til treet baobab ved å referere til dets deler. Delene nevnes i bestemt form, noe som også er kjent fra spesifikk referanse i uttrykkene som *taket på huset* (se f.eks. Faarlund *et al.* 1997).

Bruk av anaforisk substantiv *trærne* i (21) viser tilbake til artene som nevnes i setningen før. Disse viser derimot tilbake til antecedenten i teksten, altså *edelgran*. Her forstås ikke *trærne* som et gitt antall trær, men heller som alle trærne av denne typen, altså referansen tolkes generisk slik som i tidligere eksempler.

Teksten om måkene i (20) er litt annerledes enn (19) og (21) fordi subjektet, som er et generisk substantiv, står i bestemt flertallsform. I engelsk brukes det kun til å referere til menneskegrupper (f.eks. *the Italians*, se f.eks. Radden 2009). I norsk er denne formen ofte brukt til å omtale selskapelige dyrearter eller arter som blir oppfattet som det. Den språklige intuisjonen behøver ikke å være i samsvar med biologisk kunnskap. Også anaforen i dette eksemplet er i bestemt flertallsform, altså *fuglene*. I korpuset ble det funnet flere eksempler der dyr ble beskrevet i bestemt flertallsform og disse hadde en tydelig generisk tolkning (f.eks. tekster om flaggermusene). Det samme gjelder for planter som kan betraktes som utelleglige (f.eks. mosene) og hverdagsobjekter som vanligvis forekommer i flere eksemplarer (f.eks. myntene).

4. Konklusjoner

Resultatene og eksemplene beskrevet ovenfor viser at generisk anafor og generisk referanse har mye til felles når det gjelder bruk av grammatiske former. Både anaforiske pronomen og substantiver blir brukt i entall og flertall, og i noen av tilfellene tilsvarer formene hovedsubstantivet i teksten (f.eks. ubestemt flertall i generisk referanse og ubestemt flertall i generisk anafor i samme tekst).

Kontekstuell sammenheng virker å spille en viktig rolle her. For å gjøre teksten koherent bruker forfattere ofte samme form av substantivet, eller begynner med naken eller ubestemt form for så å bytte til bestemt form, entall eller flertall. Dette tilsvarer også spesifikk referanse der man først nevner noe ukjent i ubestemt form og refererer til det senere ved å bruke bestemt form.

Denne empiriske studien viser at generisitet og generisk anafor i norsk ikke følger et bestemt mønster som kan observeres i andre skandinaviske språk (se f.eks. Hansen & Heltoft 2011 for dansk og Teleman *et al.* 1999 for svensk). Begge fenomenene kan uttrykkes på forskjellige måter og ved bruk av flere enn én grammatisk form. Selv om noen av formene kan brukes om hverandre (f.eks. noen av substantivformene i generisk anafor), kan man ikke si at de kan tolkes på samme måte. Kontekst, stilistikk og sjanger påvirker også måten generisitet og generisk anafor er uttrykt og tolket på.

Litteratur

- Behrens, L. 2005. "Genericity from a Cross-Linguistic Perspective". *Linguistics* 43(2): 275–344.
- Bennett-Kastor, T.L. 1996. "Anaphora, Non-Anaphora, and the Generic Use of Pronouns by Children". *American Speech* 71(3): 285–301.
- Borthen, K. 2003. *Norwegian Bare Singulars*. PhD-avhandling. Trondheim: Norwegian University of Science and Technology.
- Brown, K., Barber, A. & Stainton, R.J. (red.). 2010. *Concise Encyclopedia of Philosophy of Language and Linguistics*. Oxford: Elsevier.
- Burleighs, W. 1988. *Von der Reinheit der Kunst der Logik. Erster Traktat. Von den Eigenschaften der Termini. (De puritate artis logicae. De proprietatibus terminorum)*. Hamburg: Felix Meiner Verlag.
- Carlson, G.N. 1977. "A Unified Analysis of the English Bare Plural". *Linguistics and Philosophy* 1: 413–457.
- Carlson, G.N. 1979. "Generics and Atemporal *When*". *Linguistics and Philosophy* 3: 49–98.
- Carlson, G.N. 1982. "Generic Terms and Generic Sentences". *Journal of Philosophical Logic* 11: 145–181.
- Carlson, G.N. 1983. "Logical Form: Types of Evidence". *Linguistics and Philosophy* 6: 295–317.
- Carlson, G.N. & Pelletier, F.J. (red.). 1995. *The Generic Book*. Chicago–London: The University of Chicago Press.
- Carlsson, Y. 2012. *Genericitet i tekst*. PhD-avhandling. Stockholm: Stockholms universitet.
- Chierchia, G. 1998. "Reference to Kinds across Languages". *Natural Language Semantics* 6: 339–405.
- Faarlund, J.T., Lie, S. & Vannebo, K.I. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Geach, P.T. 1972. *Logic Matters*. Oxford: Basil Blackwell.
- Geach, P.T. 1980 [1962]. *Reference and Generality. An Examination of Some Medieval and Modern Theories*. Ithaca–London: Cornell University Press.
- Gårding, E. (red.). 1976. *Kontrastiv fonetik och syntax med svenska i centrum*. Lund: Liber Läromedel.
- Halmøy, M. 2016. *The Norwegian Nominal System – A Neo-Saussurean Perspective*. Berlin–Boston: De Gruyter Mouton.
- Hansen, E. 1994. "Generisk substantiv". Jørgensen, Platzack & Svensson 1994: 137–144.
- Hansen, E. & Heltoft, L. 2011. *Grammatik over det Danske Sprog*. Odense: Syddansk Universitetsforlag.
- Huang, Y. 2000. *Anaphora. A Cross-Linguistic Study*. Oxford: Oxford University Press.
- Jørgensen, N., Platzack, C. & Svensson, J. (red.). 1994. *Språkbruk, grammatik och språkförändring. En festskrift till Ulf Teleman 13.1.1994*. Lund: Institutionen för nordiska språk.
- Kabatek, J. & Wall, A. (red.). 2013. *New Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam–Philadelphia: John Benjamins.
- Karczewski, D. 2016. *Generyczność w języku i w myśleniu. Studium kognitywne*. Kraków: Universitas.
- Krifka, M. *et al.* 1995. "Genericity: An Introduction". Carlson & Pelletier 1995: 1–124.

- Kurek, A. 2017. "Hvordan skildres generisitet i norsk faglitteratur?". *Folia Scandinavica Posnaniensia* 22: 4–17.
- Lakoff, G. 1990. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago–London: The University of Chicago Press.
- Leslie, S.-J. 2007. "Generics and the Structure of the Mind". *Philosophical Perspectives, Philosophy of Mind* 21: 375–403.
- Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Mari, A., Beyssade, C. & del Prete, F. (red.). 2013a. *Genericity*. Oxford: Oxford University Press.
- Mari, A., Beyssade, C. & del Prete, F. 2013b. "Introduction". Mari, Beyssade & del Prete 2013a: 1–92.
- Ojeda, A.E. 1991. "Definite Descriptions and Definite Generics". *Linguistics and Philosophy* 14: 367–397.
- Oosterhof, A. 2008. *The Semantics of Generics in Dutch and Related Languages*. Amsterdam/Philadelphia: John Benjamins.
- Panther, K.-U., Thornburg, L.L. & Barcelona, A. (red.). 2009. *Metonymy and Metaphor in Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Pelletier, F.J. (red.). 2010a. *Kinds, Things, and Stuff. Mass Terms and Generics*. Oxford: Oxford University Press.
- Pelletier, F.J. 2010b. "Generics: A Philosophical Introduction". Pelletier 2010a: 3–15.
- Petterson, T. 1976. "Bestämda och obestämda former". *Gårding* 1976: 119–142.
- Radden, G. 2009. "Generic Reference in English. A Metonymic and Conceptual Blending Analysis". Panther, Thornburg & Barcelona 2009: 199–228.
- Radden, G. & Dirven, R. 2007. *Cognitive English Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Rosén, V. & De Smedt, K. (red.). 2017. *The Very Model of a Modern Linguist. In honor of Helge Dyvik*. Bergen: Universitetet i Bergen.
- Rosén, V. & Borthen, K. 2017. "Norwegian Bare Singulars Revisited". Rosén & De Smedt 2017: 220–240.
- Seuren, P. 2010. "Donkey Sentences". Brown, Barber & Stainton 2010: 169–171.
- Teleman, U., Hellberg, S. & Anderson, E. 1999. *Svenska Akademiens grammatik. Ord*. Stockholm: Norstedts.