

Ny forskning i grammatik

Marie Bojsen-Møller

Torben Juel Jensen

Sune Sønderberg Mortensen

Jan Heegård Petersen

Lennart Westergaard

Udgivet af:
Dansk Sprognævn.

Med økonomisk støtte fra:
Støtteforeningen for Grammatiknetværket.

Redaktion:
Marie Bojsen-Møller, Torben Juel Jensen, Sune Sønderberg Mortensen,
Jan Heegård Petersen, Lennart Westergaard

Serieredaktør: Eva Skafte Jensen, esj@dsn.dk.

© Forfatterne og Dansk Sprognævn, 2024

Sats: René Malchow
Omslag: Ulla Poulsen Precht
ISSN: 2446-1709 (onlineversion)

**Dansk
Sprognævn**
Adelgade 119b
DK-5400 Bogense

Indholdsfortegnelse

<i>Ken Ramshøj Christensen, Anne Mette Nyvad & Katrine Rosendal Ehlers: Vi er vild med Hestenettet! En korpusundersøgelse af alternativ kongruens</i>	5
<i>Klaus Geyer: Dansks ukendte søster – typologiske betragtninger om tysk og dansk</i>	20
<i>Anna Sofie Hartling: Sammenhængen mellem stærktbøjede verbers bøjningsgruppe og sandsynligheden for at opnå svag bøjning i dansk</i>	37
<i>Peter Juul Henriksen: Tekstordet som grammatisk domæne</i>	53
<i>Torben Juul Jensen & Tanya Karoli Christensen: Ledstilling i attributive ledsætninger</i>	71
<i>Oliver Rix Johannsen: Kasus i hypotagmer. En undersøgelse af flertalspronomenene i dansk</i>	88
<i>Michael Nguyen: Var det sin eller hans kone Jens kyskede? Refleksiv reference i sætningskløvning og pseudokløvning</i>	107
<i>Mads Nielsen, Naja Blytman Trondhjem & Kasper Boye: Leksikalske affikser</i>	125
<i>Peter Juul Nielsen: Det danske indirekte objekt i tre tekster fra 1500-tallet</i>	140
<i>Anne Mette Nyvad & Ken Ramshøj Christensen: Når en ø er fremmed land: Danskeres bedømmelse af ekstraktion fra relativsætninger i engelsk som andetsprog</i>	157

<i>Kathrine Thisted Petersen & Steffen Krogh: Ved Lejre græsse nu Faar paa Vold, hvor fordem Kæmperne drukke. Udviklingen af numeruskongruens mellem subjekt og finit verbum i dansk fra ca. 1500 til ca. 1900.....</i>	176
<i>Birutė Spraunienė & Emil Jonas Slott: Markering af bestemthed i litauiske studerendes indlærerdansk</i>	194
<i>Sten Vikner & Ken Ramshøj Christensen: Struktur og konstituens i forskellige grammatiske tilgange</i>	211
<i>Lennart Westergaard: Det middeldanske baggrundsfelt: Subjekts- og letledspladser</i>	234

Vi er vild med Hestenettet!

En korpusundersøgelse af alternativ kongruens

Ken Ramshøj Christensen, Anne Mette Nyvad
& Katrine Rosendal Ehlers

Abstract

Previous studies have shown that in colloquial Danish, a predicative adjective sometimes agrees with a fronted complement of a prepositional verb (P-Obj) instead of with the subject, which is normally the case. After reviewing the results of these previous studies, we present data from a study on written language using the *Danish Gigaword Corpus*, which consists of 228.9 million words. Our results show that even though there are thousands of examples with alternative agreement, in the vast majority of cases (99%), there is standard agreement with the subject. There is a significant interaction between fronting and grammatical number, whereas frequency of occurrence is not significant.

Nøgleord

fremflytning, numerus, kollokation, frekvens, performans

1. Indledning

Denne artikel handler om såkaldt **alternativ kongruens** i dansk, som findes i sætninger hvor der ikke er kongruens mellem subjektet og et adjektivisk subjektsprædikat, som det ellers normalt er tilfældet. I standardkongruens bøjes prædikative adjektiver i numerus (flertal *-e*) så de kongruerer med subjektet.¹ Det gælder også når det prædikative adjektiv har et **præpositionsobjekt (P-Obj)**:

1 Vi skærer et par hjørner her, for i ental bøjes de også for genus (intetkøn *-t*), *det er vildt*, og i nogle tilfælde spiller fonologi og semantik (animathed) også en vis rolle, se Christensen & Nyvad (2019: 103 fn. 16). Der er også visse undtagelser, fx *Den (kage) er ikke god for dig* vs. *Det er ikke god-t for dig*, men *Hun er vild med ideen* og ikke **Udvalget er vild-t med ideen*. Desuden bruges den ubøjede form af nogle adjektiver ofte i stedet for den bøjede flertalsform i talesprog og i uformelt skriftsprog, fx *Vi er ikke helt færdig/klar/parat* (Hansen 1967: 369). Det er specielt tilfældet for adjektiver dannet af præteritum participium, fx *interesseret*, *irriteret* eller *henrykt* (Sproget.dk, a).

- (1) a. Hun er ret **vild** med min mad.
 b. De er ret **vild-e** med min mad.

Men som Jensen (2004), Nyvad (2007) og Engberg-Pedersen & Poulsen (2010) bemærker, sker det at adjektivet tilsyneladende kongruerer i numerus med P-Obj i stedet for med subjektet i tilfælde hvor der er potentiel 'konkurrence' i numerus mellem subjektet og P-Obj. (De skal være forskellige for at man kan se om der er kongruens med det ene eller det andet.)

- (2) a. De der tærter er hun ret **glad-e** for.
 b. Den der tærte er vi ret **glad** for.

Jensen (2004: 166) beskriver fænomenet som "en ændring af rektionsforholdene således at prædikativet i stedet kongruerer med styrelsen i den tilknyttede præpositionsforbindelse", mens Engberg-Pedersen & Poulsen (2010) argumenterer for at der er tale om performansfejl.

I to elicitationsstudier har Christensen og Nyvad undersøgt sådanne tilfælde af **alternativ kongruens**, både når P-Obj er fremflyttet (2019), og når det står in situ efter adjektivet (2021). I forsøgene blev deltagerne bedt om skriftligt at fuldføre sætninger som i (3), hvor der blev vekslet mellem ental og flertal for hhv. subjekt og P-Obj, hvilket dermed eliciterede enten den bøjede eller den ubøjede form af det prædikative adjektiv. Alle sætningerne havde en kontekst der indeholdt standardkongruens, og der var således bias *imod* den alternative kongruens til fordel for standardkongruens.

- (3) a. Jeg er helt **vild** med kager. Og dem her, dem er jeg specielt ____ med.
 b. Vi er helt **vilde** med kager, især den her. Vi er specielt ____ med den her.

Resultaterne viste at der blev eliciteret væsentlig mere alternativ kongruens når der var fremflytning, og at det især var tilfældet når P-Obj var i ental og subjektet i flertal ($p < 0,001$). Når det fremflyttede P-Obj var i flertal, var der 17% alternativ kongruens (*Dem er jeg vilde med*), mens det var hele 48% for ental (*Den er vi vild med*). Med P-Obj in situ var det hhv. 1% (*Jeg er vilde med dem*) og 14% (*Vi er vild med den*). Den alternative kon-

gruens i resultaterne kan forklares som en performansfejl der skyldes to uafhængige faktorer: fejlanalyse og apokope. Eftersom dansk er et SVO-sprog, forventer vi at det første nominale led er subjekt. Vi analyserer derfor det fremflyttede P-Obj først som subjekt, og når vi kommer til subjektet, reanalyserer vi det som ikke-subjekt. Denne 'fejlanalyse' bliver ikke helt glemt (Christianson m.fl. 2001, Slattery m.fl. 2013) og påvirker efterfølgende valget af kongruens når vi kommer til adjektivet, hvilket gør at der er flere fejl ved fremflytning end ved P-Obj in situ. Apokope, eller bortfald af endelsen *-e*, der sandsynligvis er et udtryk for en generel tendens til reduktion (se diskussionen i Heegård & Thøgersen 2014), er med til at øge alternativ kongruens når det fremflyttede P-Obj er ental og subjektet flertal (*Den er vi vild med*). Omvendt trækker det i modsatte retning når P-Obj er flertal og subjektet ental, fordi fraværet af *-e* jo netop er standard her.

De to forsøg viste kort sagt at det er ret let at elicitere alternativ kongruens, selvom standardkongruens dog blev eliciteret i de fleste tilfælde. Overraskende nok var det lettere at elicitere alternativ kongruens når det fremflyttede P-Obj var ental. Baseret på egne observationer og eksempler i litteraturen, forventede vi det modsatte.

Nyvad & Christensen (2023) undersøgte hvordan position (fremflytning vs. in situ), numerus (P-Obj i flertal vs. ental), og kongruens (alternativ vs. standard) påvirker acceptabiliteten (på en skala fra 1 til 7) af sætninger hvor der er konkurrence om kongruensen, som i (2) og (3). De i alt 197 forsøgsdeltagere vurderede i gennemsnit standardkongruens signifikant højere end alternativ kongruens (1,9 point på 7-trinsskalaen, $p < 0,001$), og fremflytning gav en lille men signifikant forbedring af acceptabiliteten af alternativ kongruens (fra 3,49 in situ til 4,23 med fremflytning, $p < 0,001$), mens det ikke var tilfældet for standardkongruens. I denne undersøgelse gjorde det altså ingen forskel om det fremflyttede P-Obj var ental eller flertal.

Én ting er folks performans og vurderinger i forsøg, noget andet er forekomsten i naturligt sprogbrug. Engberg-Pedersen & Poulsen (2010) lavede en internetsøgning på *vild med*, *glad for* og *sur på* der viste 5,5% alternativ kongruens (mellem 1,6% og 11,1% afhængigt af adjektivet). Christensen & Nyvad (2021) gjorde det samme og fandt 9,6% alternativ kongruens. Sådanne ret uformelle undersøgelser skal tages med forbehold, for der er kun brugt tre adjektiver (eller rettere, PRÆP+ADJ-kollokationer (bigrammer), *vild med*, *glad for* og *sur på*), men de viser dog at det er ret let at finde

eksempler på internettet, hvor der til gengæld også er mange sider der er maskinoversat.

For at få et bedre indblik i hvor udbredt alternativ kongruens er, har vi lavet en korpusundersøgelse. De tidligere undersøgelser gav et mindre klart billede med tal fra ca. 5%–10% i de uformelle internetundersøgelser op til 48% i elicitationsstudierne. Vi ville også undersøge om resultaterne i elicitationsstudierne evt. kunne skyldes at det specifikke sæt af ADJ+PRÆP der indgår i studierne, ikke er repræsentativt for konstruktionen generelt. Dernæst ville vi undersøge om frekvensen af alternativ kongruens afhænger af om der er fremflytning af P-Obj eller ej, og om numerus spiller en rolle (P-Obj i ental og subjektet i flertal, eller omvendt). Desuden ville vi også teste om der skulle være en sammenhæng mellem kollokationsfrekvens (altså hvor hyppige kollokationer som *vilde med*, *glade for*, *sure på* etc. er) og alternativ kongruens.

Kort sagt, havde vi altså fire overordnede spørgsmål:

1. Hvor udbredt er fænomenet?
2. Er det sæt af ADJ+PRÆP (fx *vild med*) der indgår i de tidligere elicitationsstudier repræsentativt?
3. Afhænger frekvensen af den alternative kongruens af P-Objektets position og/eller numerus?
4. Er der en sammenhæng mellem kollokationsfrekvens og alternativ kongruens?

2. Metode

Data til artiklen er fra *The Danish Gigaword Corpus* (Strømberg-Derczynski m.fl. 2021). Vi har valgt at undersøge sproget i den del af korpusset som kommer fra Hestenettet, et dansksproget online-forum med uformelt, ”nutidigt” (eng. *contemporary*) skriftsprog og en bred vifte af emner skrevet mellem 1997 og 2020. Hestenet-korpusset indeholder 228,9 millioner ord og er dermed omtrent fire gange så stort som det 56 millioner ord store KorpusDK. Selvom størrelsen og typen af data helt klart er en fordel, er det til gengæld en ulempe at data ikke er tidskodet, så det er derfor ikke muligt at se hvad der fx er fra 1997, og hvad der er fra 2020.

Vi har med vilje søgt bredt efter kollokationer der består af et adjektiv efterfulgt af en præposition. Hestenettet indeholder 2.121.723 eksempler på ADJ+PRÆP-kollokationer fordelt på 66.071 forskellige kollokationer.

Vi har udvalgt de 1000 hyppigste kollokationer² og sorteret i alt 252 fra hvor adjektivet har samme form i ental og flertal (fx *bange for*). Denne mindre gruppe indeholder 630.544 eksempler. I denne gruppe har vi herefter søgt specifikt efter at finde eksempler på alternativ kongruens ved at specificere at adjektivet skal være bøjet i ental (henholdsvis flertal), og at der skal være et substantiv eller pronomen bøjet i flertal (henholdsvis ental) blandt de fem forudgående ord for at forøge chancen for at finde eksempler med alternativ kongruens (fx *men den skulle vi ikke være nervøs for* og *Dem er jeg ikke tilfredse med*). Denne søgning giver i alt 18.685 eksempler fordelt på 748 kollokationer som vi har gennemlæst manuelt og kodet for kongruens (standard vs. alternativ), position (fremflytning vs. in situ), og numerus (P-Obj i ental vs. flertal). Dette datasæt, som vi nedenfor kalder 'det fulde datasæt', indeholder i alt 3.096 reelle eksempler på alternativ kongruens fordelt på 63 forskellige kollokationer. (Der er ingen tilfælde af alternativ kongruens i de resterende 685 kollokationer.) Efterfølgende søgning i korpuset på disse 63 kollokationer giver 293.570 eksempler i alt, og her har vi så antaget at de resterende 290.654 (dvs. ud over de 3.096 eksempler med alternativ kongruens) har standardkongruens.³

Af de 3.096 eksempler på alternativ kongruens er 1.924 fordelt på 10 af de kollokationer der blev undersøgt i de tidligere artikler om alternativ kongruens (Christensen & Nyvad 2019, 2021). I disse studier var der godt nok 14 kollokationer, men *ked(e) af, flov(e) over* og *vred(e) på* optræder ikke med alternativ kongruens på Hestenettet, og de er derfor udeladt i vores analyse. Vi har desuden udeladt *sikker/sikre på* fordi alternativ kongruens evt. kunne skyldes 'r-problemer' (Sproget.dk, b)⁴. (Der er 21 eksempler med alternativ kongruens der svarer til *Dem er vi sikker på*, men nul der svarer til *Den er jeg sikre på*.) Der er således 10 kollokationer i det som vi kalder 'det reducerede datasæt' (reduceret i forhold til det fulde datasæt),

2 Langt størstedelen er meget sjældne. Ved meget lave frekvenser vil et enkelt, måske tilfældigt, tilfælde af alternativ kongruens potentielt skævvride statistikken, fordi det ser ud som om det er mere hyppigt end det i virkeligheden er.

3 Det kunne give en underestimering af den alternative kongruens fordi vi kun har kigget efter eksempler hvor afstanden mellem et muligt subjekt og adjektivet er mellem 0 og 5 ord. På den anden side fokuserer vores analyse udelukkende på den brøkdelen af data hvor der rent faktisk findes tilfælde med alternativ kongruens, hvilket nødvendigvis fører til en overestimering.

4 R-problemer viser sig også som usikker brug af nutids-r på verber, fx *at lære/beundre*, og flertal på substantiver der ender på r, fx *lærer*.

altså de oprindeligt 14 kollokationer minus *ked af, fløv over, vred på* og *sikker på* (se (5) nedenfor).

3. Resultater

3.1. *Det fulde datasæt (63 kollokationer)*

Tabel 1 indeholder frekvenserne for de i alt 63 ADJ+PRÆP-kollokationer i (4), både i entals- og flertalsform (fx *afhængig(e) af, allergisk(e) overfor*, etc.), der alle optræder i sætninger med alternativ kongruens.

- (4) afhængig af, allergisk overfor, ansvarlig for, behjælpelig med, betænkelig ved, bevidst om, dårlig til, dygtig til, enig i, enig med, enig om, færdig med, følsom for, forsigtig med, forskellig fra, fortrolig med, frisk på, fuld af, giftig for, glad for, glad ved, god mod, god til, god ved, hård ved, heldig med, interessant for, kendt for, klog på, ligeglad med, lun på, lykkelig for, misundelig over, misundelig på, modtagelig for, nervøs for, omhyggelig med, opmærksom på, opsat på, parat til, ræd for, rædselsslagen for, sikker på, sød ved, spændt på, stolt af, stolt over, superglad for, sur på, taknemmelig for, tilfreds med, træet af, tryk ved, uenig i, uenig med, uenig om, uheldig med, usikker på, utilfreds med, utryk ved, velkommen til, vild med, åben for

Kongruens	P-Obj fremflyttet		P-Obj in situ		Total	
Standard	36.701	98,2%	253.773	99,0%	290.474	98,9%
Alternativ	662	1,8%	2.434	1,0%	3.096	1,1%
Total	37.363	12,7%	256.207	87,3%	293.570	100,0%

Tabel 1: Frekvenser og procentvis fordeling i det fulde datasæt med 63 kollokationer.

Som det fremgår, er der i alt 293.570 sætninger, hvoraf der er standardkongruens i 98,9% af tilfældene. Selvom der således er hele 3.096 eksempler med alternativ kongruens, udgør de altså kun 1,1%. Der er fremflytning af P-Obj i 12,7% af alle sætningerne, hvilket er relativt mange i forhold til at Mikkelsen (2015) kun fandt 1% i hendes korpussøgning (en forskel der sandsynligvis skyldes genreforskelle, eftersom Hestenettet er meget uformelt, mens Mikkelsens data er sammensat af uddrag fra *Weekendavisen* maj 2006, romanen *Nordkraft* af Jakob Ejersbo 2002 og talesprogskorpuset *BySoc*, der er fra slutningen 1980'erne). Blandt disse sætninger med fremflyttet P-Obj er der kun 1,8% alternativ kongruens, hvilket er væsentlig færre end forventet i forhold til Christensen & Nyvad

(2019), hvor det blev eliciteret i hele 17% og 48% af gangene med P-Obj i hhv. flertal og ental. I sætninger med P-Obj in situ er der 1,0% alternativ kongruens, hvilket er tættere på (men stadig lavere end) tallene fra Christensen & Nyvad (2021), der eliciterede det i hhv. 1% (P-Obj i ental) og 14% (P-Obj i flertal) af tilfældene.

3.2. *Det reducerede datasæt (10 kollokationer)*

For at undersøge om de 10 (oprindeligt 14) kollokationer som blev brugt i Christensen & Nyvads (2019, 2021) forsøg, angivet i (5), afviger fra det overordnede mønster, har vi undersøgt dem i isolation. Resultatet er opsummeret i Tabel 2, og det er ret tydeligt at procentsatserne ikke er væsensforskellige fra dem i Tabel 1.

- (5) enig(e) med, glad(e) for, god(e) ved, ligeglad(e) med, stolt(e) af, sur(e) på, tilfreds(e) med, træt/trætte af, tryk/trykke ved, vild(e) med

Kongruens	P-Obj fremflyttet		P-Obj in situ		Total	
Standard	19.307	98,2%	136.849	98,9%	156.156	98,8%
Alternativ	362	1,8%	1.541	1,1%	1.903	1,2%
Total	19.669	12,4%	138.390	87,6%	158.059	100,0%

Tabel 2: Frekvenser og procentvis fordeling i det reducerede datasæt med 10 kollokationer.

I de i alt 158.059 sætninger er der altså kun 1,2% alternativ kongruens. De 1.924 eksempler er noget færre end forventet i forhold til elicitationsstudierne (Christensen & Nyvad 2019, 2021). I 12,4% af alle sætningerne (ca. 1/8) er der fremflytning, hvilket igen er relativt mange, og i disse sætninger er der kun 1,8% alternativ kongruens – igen væsentlig færre end forventet.

I Tabel 3 nedenfor er antal sætninger per kongruenstype fordelt efter numerus og position.

Numerus	Position	Standard	Alternativ	% Alternativ
P-Obj = flertal	fremflyttet	2.896	226	7,2%
P-Obj = flertal	in situ	18.618	680	3,5%
P-Obj = ental	fremflyttet	16.411	136	0,8%
P-Obj = ental	in situ	118.231	861	0,7%
Total		156.156	1.903	1,2%

Tabel 3: Antal sætninger med alternativ kongruens fordelt efter numerus og position.

De tidligere forsøg (Christensen & Nyvad 2019, 2021) viste at der ved fremflytning blev eliciteret 17% alternativ kongruens når P-Obj var flertal (*Dem er jeg vilde med*), mens det var hele 48% når P-Obj var ental (*Den er vi vild med*). Når P-Obj var in situ, var der meget færre tilfælde, men med den samme asymmetri: 1% alternativ kongruens med P-Obj i flertal (*Jeg er vilde med dem*) og 14% med P-Obj i ental (*Vi er vild med den*). Der var altså i det hele taget flere eksempler med P-Obj i ental. Resultatet i denne artikel viser den modsatte asymmetri. Her er det flertal der oftest udløser alternativ kongruens: 7,2% ved fremflytning (*Dem er jeg vilde med*) og 3,5% in situ (*Jeg er vilde med dem*) mod hhv. 0,8% og 0,7% for ental. Med andre ord: Der er flere tilfælde hvor endelsen *-e* tilføjes, end hvor den udelades, og hvor det fører til alternativ kongruens. Som det også fremgår af tallene i Tabel 3, er der flere tilfælde med alternativ kongruens med P-Obj in situ end ved fremflytning. Men fordi antallet af eksempler uden fremflytning overordnet er større, så er andelen af alternativ kongruens mindre ved P-Obj in situ (1,1%) end ved fremflytning (1,8%) (se Tabel 2).

Til statistisk analyse anvendte vi R (R Core Team 2023) i RStudio (Posit team 2023). Vi brugte en standard lineær model (basisfunktionen *lm*) med multipel regression som havde andelen af alternativ kongruens som afhængig variabel (*outcome*) og position (fremflyttet vs. in situ), numerus (flertal vs. ental), \log_{10} kollokationsfrekvens⁵, samt interaktionen mellem position og numerus som uafhængige variable (*predictors*).

Resultatet, se Tabel 4, viser at effekten af numerus er signifikant ($p = 0,001$), mens effekten af position i sig selv ikke er signifikant ($p = 0,073$). Effekten af kollokationsfrekvens (\log_{10}) er negativ, men heller ikke signifikant ($p = 0,178$). Til gengæld er der en signifikant, positiv interaktion mellem numerus og position ($p = 0,030$). (Selvom effekten af kollokationsfrekvens ikke er signifikant, forklarer den alligevel noget: Den beskriver en del af variansen. Hvis den udelades fra modellen, bliver effekten af position større og signifikant, $p < 0,05$.) Modellen tager flertal og fremflytning som udgangspunkt (*intercept*). Den samlede negative effekt på andelen af alternativ kongruens af et ental P-Obj in-situ er altså mindre end summen af numerus og position set hver for sig. Eller med andre ord: Der er flere tilfælde af alternativ kongruens ved fremflytning af et P-Obj i flertal end i ental, og der er flere tilfælde ved fremflytning af P-Obj sammenlignet med P-Obj

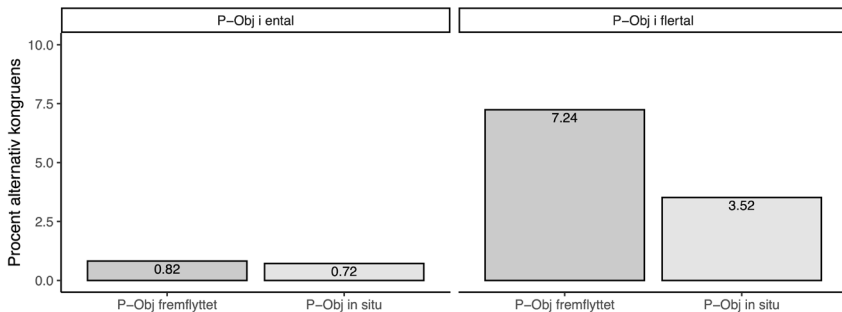
5 \log_{10} er en logaritmisk skala der har 10 som grundtal, hvilket betyder at når \log_{10} -frekvensen stiger med 1, stiger antallet med faktor 10. Fx hvis \log_{10} -frekvensen stiger fra 10 til 13, betyder det at antallet af eksempler stiger med faktor 1.000.

in situ, men den samlede effekt af fremflytning og flertal er altså mindre end summen af de to faktorer. Denne model forklarer 24% af variationen i kongruensen (hvilket er en ret stor effekt: justeret $R^2 = 0.24$, $p < 0,001$).

	Estimat	SE	t	p	
(Intercept)	0,15	0,04	4,28	0,000	***
Position (P-Obj in situ)	-0,05	0,03	-1,82	0,073	
Numerus (ental)	-0,09	0,02	-3,51	0,001	***
Log ₁₀ kollokationsfrekvens	-0,02	0,01	-1,36	0,178	
Position × Numerus	0,07	0,03	2,21	0,030	*

Table 4: Statistisk model for effekten på andelen af alternativ kongruens af position, numerus og log₁₀-transformeret frekvens.

Denne interaktion er også illustreret i Figur 1, hvor man tydeligt kan se forskellen mellem flertal og ental, mens forskellen mellem et fremflyttet og et in-situ P-Obj afhænger af numerus.



Figur 1: Procent alternativ kongruens som funktion af position og numerus.

4. Diskussion

Vi havde fire overordnede spørgsmål:

1. Hvor udbredt er fænomenet?
2. Er det sæt af ADJ+PRÆP der indgår i elicitationsstudierne, repræsentativt?
3. Afhænger frekvensen af alternativ kongruens af P-Obj's position og/eller numerus?
4. Er der en sammenhæng mellem kollokationsfrekvens og alternativ kongruens?

Hvor tidligere undersøgelser har fundet ca. 5%–10% alternativ kongruens i uformelle internetundersøgelser og op til 48% i elicitationsstudier, viser vores undersøgelse her at der er alternativ kongruens i 1,1% af de i alt 293.570 sætninger hvor der er potentiel konkurrence om kongruensen mellem subjekt og P-Obj. Hvis vi begrænser søgningen til de ADJ +PRÆP-kollokationer der blev brugt i elicitationsstudierne (Christensen & Nyvad 2019, 2021), er der alternativ kongruens i 1,2% af de i alt 158.059 sætninger. Der er ikke nogen nævneværdig forskel på procent-satserne i de to datasæt i Tabel 1 og Tabel 2, der udviser samme tendens. Det tyder på at de kollokationer der blev brugt i Christensen & Nyvads (2019, 2021) forsøg, udgør et repræsentativt sæt. Der er til gengæld relativt få tilfælde af alternativ kongruens i undersøgelsen, noget færre end forventet ud fra de tidligere forsøg. Formelle, kontrollerede forsøg er i sagens natur ikke 'naturlige', og det er velkendt at det påvirker forsøgsdeltagerenes adfærd. I elicitationsstudierne har selve forsøgsdesignet og det at deltage i et forsøg, sandsynligvis påvirket deltagerne, hvilket har ført til øget antal af kongruensfejl (den alternative kongruens).⁶ I det formelle forsøg stiger antallet af tilfældige fejl naturligt som følge af uafhængige faktorer.

Den statistiske analyse i Tabel 4 viser at der er en signifikant, positiv interaktion mellem numerus og position ($p = 0,030$). Kort sagt så er der mere alternativ kongruens når P-Obj er flertal, og når det er fremflyttet, og denne effekt er større end summen af de to faktorer (position og numerus). Der er til gengæld ikke nogen signifikant effekt af kollokations-frekvens.

Christensen & Nyvad (2019, 2021) og Nyvad & Christensen (2023) argumenterer for at alternativ kongruens i virkeligheden er performansfejl, snarere end en nyudvikling i grammatikken, fx topic-kongruens – altså hvor der er kongruens med det diskursfunktionelle topic i stedet for med det grammatiske subjekt. Topic har ikke nogen entydig definition. Ofte

6 Vi ændrer adfærd så snart vi ved at vi bliver observeret eller testet. Det er fx velkendt at vores puls stiger så snart der måles på den, at smerte intensiveres når vi mærker efter, at det er sværere at udføre en svær, fysisk opgave hvis nogen kigger, og at det kan være stressende at gå til eksamen. Vi husker også bedre at vaske hænder hvis vi tror at der er nogen der kigger. Bare det at vi observerer eller undersøger noget, ændrer på det vi observerer. Fænomenet kaldes *the observer effect* ('observatøreffekten') og er et grundlæggende princip i kvantefysik såvel som i psykologi. Måden vi undersøger nogen eller noget på, ændrer også på resultatet. Måden vi fokuserer på en sproglig opgave, kan således også have indflydelse på hvordan vi besvarer den. Se fx Baclawski (2018).

analyseres det element der er fremflyttet i en deklarativ hovedsætning, som topic eller fokus (Hoof 2006: 411, Lambrecht 1994: 117–205, Ross 1967: 209). I den traditionelle grammatik sidestilles topic ofte med subjektet, der ifølge Hansen & Heltoft (2011: 1729) er antifokus (men se Jensen 2011). Comrie (2003: 328–329) hævder at alternativ kongruens (som han kalder *trigger-happy agreement*) er afhængig af nominalfrasens *topicworthiness*, dvs. hvor sandsynligt det er at det pågældende element fungerer som topic, hvilket er forbundet med fx nominalfrasens bestemthed og *salience*, dvs. pragmatisk ‘prominens’. Hans definitionen på *topicworthiness* er dog temmelig vag: Det har at gøre med sandsynlighed for *topicality*, der bestemmes af *salience*, som også er meget vagt defineret (Comrie 2003 nævner begrebet uden at give nogen definition), og er desuden relateret til numerus, animathed og bestemthed. “Det er dog svært at se hvordan man skulle forklare den alternative kongruens med *topicworthiness* og topic-kongruens uden at det bliver temmelig ad hoc og cirkulært og dermed ikke falsificerbart. Hvis man hævder at der altid er topickongruens, så er der alternativ kongruens mellem fx P-OBJ og et subjektsprædikat når P-OBJ er topic, og det der viser at det er topic, er den alternative kongruens” (Christensen & Nyvad 2021: 130). De relativt få tilfælde af alternativ kongruens i undersøgelsen her er også svære at forklare med en ide om topic-kongruens. Det ville betyde at kun 1,8% af disse fremflyttede P-Objekter var topic. I de i alt 158.059 sætninger, hvor der er et P-Obj der ikke har samme numerus som subjektet, er det altså i 98,8% af tilfældene subjektet der er topic (og det er derfor der er standardkongruens), uafhængigt af om der er fremflytning, som det så ellers er tilfældet i 12,4% af sætningerne. Det virker i det hele taget ikke særligt sandsynligt. Det er selvfølgelig teoretisk muligt at der er variation, således at enkelte individer (1,2%), som har skrevet enkelte ting på Hestenettet, er ved at udvikle topic-kongruens, mens de 98,8% andre ikke er. (Alternativt kunne man antage at en større gruppe bruger alternativ kongruens i 1,2% af tilfældene, og altså dermed ikke laver kongruensfejl.) Det kunne signalere et tidligt stadie i en grammatikaliseringsproces, men det er der til gengæld ikke noget der tyder på, og det virker ikke rigtigt muligt at teste. Her er det videnskabelige princip kendt som *the Sagan Standard* relevant: ”Eksstraordinære påstande kræver ekstraordinære beviser”, og det må pålægge dem der mener at der findes topic-kongruens, at tilvejebringe evidens for det. Vi beskriver her et fænomen der optræder i ret få procent af tilfældene i korpusset, men som let kan eliciteres eksperimentelt. Denne adfærd

kan vi netop forklare med uafhængige faktorer (bortfald af endelsen *-e* og at man har tendens til tolke det første led som subjekt). Nul-hypotesen må være at der ikke er tale om forandring i systemet, og det er kompatibelt med alle vores resultater.

Bemærk også at der er flere tilfælde hvor endelsen *-e* **tilføjes** (alternativ kongruens med P-Obj i flertal), end hvor den udelades (alternativ kongruens med P-Obj i ental). Det går imod den gængse opfattelse af at danskerne reducerer (eller i folkemunde ”sluger”) alle endelserne, og det passer heller ikke så godt med en forestilling om at adjektivbøjningen i flertal generelt neutraliseres og er ved at blive reduceret til den ubøjede form (ligesom fx *parat* og *irriteret*, se fodnote 1).

Vi fandt ingen statistisk signifikant effekt af kollokationsfrekvensen på andelen af alternativ kongruens. Christensen & Nyvad (2019, 2021) rapporterede en statistisk signifikant, positiv korrelation mellem frekvens (dog frekvensen af kopula + prædikativt adjektiv, fx *er vild* eller *er vilde*) i KorpusDK og eliciteret standardkongruens – altså jo hyppigere, jo mindre alternativ kongruens. Men den nuværende undersøgelse viser at selv om der godt nok er en negativ korrelation mellem kollokationsfrekvens og alternativ kongruens i Hestenet-korpusset, så er den ikke statistisk signifikant.

5. Konklusion

I Nyvad & Christensens (2023) acceptabilitetsstudie blev standardkongruens vurderet signifikant højere end alternativ kongruens (1,9 point på en 7-trins Likert-skala, $p < 0,001$), og fremflytning gav en lille, men signifikant forbedring af acceptabiliteten af alternativ kongruens (0,74 point, $p < 0,001$), mens der ikke var nogen signifikant effekt af numerus. Det gjorde altså ingen forskel om det fremflyttede P-Obj var ental eller flertal. Sammenholdt med vores korpusdata, ser det kort sagt ud til at standardkongruens er langt den mest udbredte form (99%), og også den mest acceptable. Fremflytning øger både acceptabilitet og frekvens af alternativ kongruens. Dette forhold er også i tråd med Christensen & Nyvads (2023) undersøgelse af forholdet mellem acceptabilitet, frekvens og kompleksitet på tværs af konstruktioner. Numerus ser ud til at spille en signifikant rolle for alternativ kongruens i naturlig sprogbrug, men det ser dog ud til at afhænge af undersøgelsesmetoden, eftersom asymmetrien mellem ental og flertal er modsat i de kontrollerede forsøg. Helt overordnet er der stadig ikke belæg for at hævde at alternativ kongruens

er andet end en systematisk performansfejl⁷ der udløses af konkurrence om kongruensen, apokope og grammatisk fejlfortolkning af det første nominal. Der er i hvert fald ikke noget der tyder på at det skulle være en ny, udbredt form for topic-kongruens.

Om forfatterne

Ken Ramshøj Christensen, ph.d., lektor, Institut for Kommunikation og Kultur, Afdeling for Engelsk, Aarhus Universitet.

Anne Mette Nyvad, ph.d., lektor, Institut for Kommunikation og Kultur, Afdeling for Engelsk, Aarhus Universitet.

Katrine Rosendal Ehlers, ph.d., postdoc, Institut for Kommunikation og Kultur, Afdeling for Engelsk, Aarhus Universitet.

Litteratur

Baclawski, Kenneth (2018): The Observer Effect. *2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 83–89. <https://doi.org/10.1109/COGSIMA.2018.8423983>.

Christensen, Ken Ramshøj & Anne Mette Nyvad (2019): De her sætninger er jeg vilde med!: Sproglig variation i prædikative adjektivers kongruensmønster. I: *MUDS 17*. Yonatan Goldshtein, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). Aarhus: Aarhus Universitet, 95–113.

Christensen, Ken Ramshøj & Anne Mette Nyvad (2021): Vi er (stadig) vild med den her sætning! I: *MUDS 18*. Yonatan Goldshtein, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). Aarhus: Aarhus Universitet, 169–186.

Christensen, Ken Ramshøj & Anne Mette Nyvad (2023): Sætningskompleksitet, frekvens og acceptabilitet. *Ny forskning i grammatik* 30, 24–39. <https://doi.org/10.7146/nfg.vli30.137951>.

Christianson, Kiel m.fl. (2001): Thematic Roles Assigned along the Garden Path Linger. *Cognitive Psychology* 42, 368–407. <https://doi.org/10.1006/cogp.2001.0752>.

7 Det er helt almindeligt at folk af uafhængige årsager laver forskellige slags talefejl (se Harley 2014: kap. 13 for et overblik), og det samme gælder tilfældige skrivefejl. Tilsvarende laver folk også tilfældige fejl når de gennemfører et eksperiment. Tilfældige fejl er altså fejl der ikke systematisk kan udløses af bestemte faktorer i bestemte kontekster. Omvendt udløses systematiske fejl af bestemte faktorer i bestemte kontekster, som fx verbalkongruensfejl (Franck m.fl. 2002) og altså alternativ kongruens.

- Comrie, Bernard (2003): When agreement gets trigger-happy. *Transactions of the Philological Society* 101, 313–337. <https://doi.org/10.1111/1467-968X.00121>.
- Engberg-Pedersen, Elisabeth & Mads Poulsen (2010): Alternative agreement controllers in Danish: Usage or structure. I: *Language usage and language structure*. Kasper Boye & Elisabeth Engberg-Pedersen (red.). Berlin: Walter de Gruyter, 205–235.
- Franck, Julie, Gabriella Vigliocco & Janet Nicol (2002): Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes* 17, 371–404. <https://doi.org/10.1080/01690960143000254>.
- Hansen, Erik & Lars Heltoft (2011): *Grammatik over det Danske Sprog. bind I-III*. København: Det Danske Sprog- og Litteraturselskab.
- Hansen, Aage (1967): *Moderne Dansk. Bind II*. København: Grafisk Forlag.
- Harley, Trevor A. (2014): *The psychology of language: from data to theory*. 4. udg. Hove, East Sussex: Psychology Press.
- Heegård, Jan & Jacob Thøgersen (2014): Style-dependent rules for reduction: Mixed-model multiple linear regression analysis of reduction phenomena in Danish news readings. *Nordic Journal of Linguistics* 37, 325–365. <https://doi.org/10.1017/S0332586514000250>.
- Hoof, Hanneke van (2006): Split Topicalization. I: *The Blackwell Companion to Syntax*. Martin Everaert & Henk van Riemsdijk (red.). Malden, MA, USA: Blackwell Publishing, 410–465. <https://doi.org/10.1002/9780470996591.ch62>.
- Jensen, Eva Skaftø (2011): Er subjektet virkelig antifokus. I: *MUDS 13*. Inger Schoonderbeek Hansen & Peter Widell (red.). Aarhus: Aarhus Universitet, 195–208.
- Jensen, Jørgen Nørby (2004): Dem er jeg helt vilde med! Om fokuskongruens i moderne dansk. I: *På godt dansk. Festskrift til Henrik Galberg Jacobsen i anledning af hans 60 års fødselsdag den 4. februar 2004*. Henrik Jørgensen & Peter Stray Jørgensen (red.). Aarhus: Wessel og Huitfeldt, 163–171.
- Lambrecht, Knud (1994): *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620607>.
- Mikkelsen, Line (2015): VP anaphora and verb-second order in Danish. *Journal of Linguistics* 51, 595–643. <https://doi.org/10.1017/S002226715000055>.

- Nyvad, Anne Mette (2007): *Revisiting Agreement Paradigms: Predicative Adjectives in Danish*. Master-afhandling. London: University College London. https://www.researchgate.net/publication/238108068_Revisiting_Agreement_Paradigms_Predicative_Adjectives_in_Danish.
- Nyvad, Anne Mette & Ken Ramshøj Christensen (2023): ”Vi har fundet nogle sætninger på nettet, og dem er folk vild med!” Når performansen halter. I: *MUDS 19*. Kirstine Boas m.fl. (red.). Aarhus: Aarhus Universitet, 249–263.
- Posit team (2023): *RStudio: Integrated Development Environment for R*. Boston, MA: Posit Software, PBC. <http://www.posit.co/>.
- R Core Team (2023): *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ross, John Robert (1967): *Constraints on variables in syntax*. Ph.d.-afhandling. Cambridge, MA: Massachusetts Institute of Technology. <https://dspace.mit.edu/handle/1721.1/15166>. (16 april, 2018).
- Slattery, Timothy J. m.fl. (2013): Lingerings misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language* 69, 104–120. <https://doi.org/10.1016/j.jml.2013.04.001>.
- Sproget.dk (a). § 35. *Adjektiver dannet af præteritum participium*. <https://sproget.dk/raad-og-regler/Retskrivningsregler/retskrivningsregler/a7-20-39/a7-29-35-verbler/a7-31-35-preteritum-participium/a7-35-adjektiver-dannet-af-preteritum-participium>. (tilgået 2 februar, 2024).
- Sproget.dk (b). *R-problemer · uddybning*. <https://sproget.dk/raad-og-regler/typiske-problemer/r-problemer/r-problemer-uddybning.html/>. (tilgået 1 februar, 2024).
- Strømberg-Derczynski, Leon m.fl. (2021): The Danish Gigaword Corpus. I: *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*. Simon Dobnik & Lilja Øvrelid (red.). Linköping: Linköping University Electronic Press, 413–421. <https://aclanthology.org/2021.nodalida-main.46>.

Danske ukendte søster – typologiske betragtninger om tysk og dansk

Klaus Geyer

Abstract

Aus dänischer Perspektive gilt das Deutsche als eine schwierige Sprache, weil es so sehr verschieden sei vom Dänischen – vor allem im Vergleich mit dem Englischen. Diese Auffassung wird in diesem Artikel mit Hilfe zweier sprachtypologischer Datenbanken kritisch überprüft: (i) der *World Atlas of Language Structures* (WALS) mit seinem Fokus auf (morpho-)syntaktische Strukturen, und (ii) das *Automated Similarity Judgement Project* (ASJP), das Basiswortschätze vergleicht. Es zeigt sich, dass Dänisch und Deutsch weitaus größere strukturelle Ähnlichkeiten aufweisen als aus laienlinguistischer Sicht meist vermutet wird. Eine mögliche Schlussfolgerung ist, dass es nicht die Sprachstrukturen sind, die das Deutschlernen in Dänemark erschweren, sondern dass die Gründe anderswo gefunden werden müssen.

Nøgleord

tysk, dansk, sprogtypologi, WALS, ordforråd

1. Indledning

Fra et dansk folkelingvistisk perspektiv betragtes tysk som et vanskeligt sprog. Det kan overraske i betragtning af de mange genetiske og kontaktlingvistiske ligheder i disse to nært beslægtede sprogs leksikon og grammatiske struktur. Den danske tyskundervisnings fokus på ‘grammatik’ ser ud til at spille en afgørende rolle for denne opfattelse ved at understrege de eksisterende forskelle i morfologien (konjunktiv I er et godt eksempel her); dette sker ofte i sammenligning med det formentligt ukomplicerede engelsk, som er fremmedsproget *par excellence*. Figur 1 viser tre eksempler på udsagn af gymnasieelever fra en ny undersøgelse foretaget af Det Nationale Center for Fremmedsprog (NCFF 2022):



Figur 1: Eleudsagn om tysk i gymnasiet (NCFF 2022)

Lund et al. (2023) giver i deres store interviewundersøgelse (230 interviews med i alt ca. 1000 elever, om alle fremmedsprog i det danske skolesystem) et differentieret og detaljeret indblik i, hvor stærkt fokuseret på *grammatik* (figur 1) – og især på kontekstløse grammatikøvelser – danske elever oplever tyskundervisningen, og hvilke læringsproblemer det giver dem. Lund et al. dedikerer tilmed et særskilt afsnit (6.5) til emnet grammatik.

Denne artikel retter dog ikke sit fokus på grammatiske indlæringsproblemer og sprogdidaktikken i tysk; i stedet vil lighederne mellem de to sprog i leksikon og morfosyntaks skitseres ved hjælp af forskellige sprogtypologiske ressourcer. Hvor det er relevant, sammenlignes resultaterne med engelsk. Først vil jeg formulere nogle grundlæggende præmisser.

2. Præmisser

I dette afsnit skitseres forudsætningerne for sammenligningen af tysk og dansk. På den ene side skal det afklares, om den ovennævnte evidens for, at tysk ikke kan læres på grund af dets grammatik, kan underbygges med valide data, og på den anden side undersøges der, hvor ens eller forskellige tysk og dansk er med hensyn til deres morfosyntaks – ofte på en forkortende måde kaldt ‘grammatik’.¹ Jeg starter med grammatikken.

2.1. Dansk-tyske ligheder og forskelle i morfosyntaks

Med sprogtilegnelse og sprogundervisning i baghovedet er den følgende

1 En konstituerende del af ‘grammatikken’ udover morfosyntaksen er uden tvivl fonologien, jf. von der Gabelentz (1901).

fremstilling i høj grad orienteret mod skolegrammatiske begreber og er rent deskriptiv, dvs. uden videre sprogteoretiske ambitioner. Haspelmath (2010) påpeger dog med rette, at selv en intenderet rent deskriptiv præsentation ikke kan kaldes helt *framework*-fri. Til at starte med, så skelner Eisenberg (1989) mellem to niveauer af grammatisk / morfologisk beskrivelse: paradigmekategorier, som i tilfældet af dansk og tysk omhandler omfattende, men ikke særligt prominente, ligheder mellem de to sprog, og enhedskategorier (fleksionskategorier), som viser både ligheder og forskelle mellem de to sprog.² At orddannelse og syntaks tillige viser omfattende ligheder, adresserer jeg efter de morfologiske overvejelser.

En ret forenklet oversigt over morfologiske ligheder og forskelle følger her (se Geyer 2024 for en mere detaljeret redegørelse):

- Substantiver: Paradigmekategorier numerus og genus findes i tysk og dansk, tysk har dog 3 genera og dansk 2. Kun tysk har kasus ved substantiver, kun dansk har bestemthed.
- Verber: Tysk og dansk opviser de parallelle verbale kategorier (ordklasser) fuldverber, modalverber, hjælpeverber og kopulaverber; begge sprog har stærke (uregelmæssige) og svage (regelmæssige) verber (næsten samme inventar); paradigmekategorierne tempus, modus og diatese plus forskellen mellem finitte og infinitte verbalformer findes i begge sprog; kun tysk har dog numerus/person som kongruens-kategorier.
- Adjektiver: Bestemthed (stærk/ubestemt vs. svag/ubestemt) og gradbøjning findes i tysk og dansk; kongruens-kategorierne er numerus og genus for begge sprog; kasus som kongruens-kategori findes derimod kun i tysk.
- Orddannelse: I tysk og dansk forekommer der i stor omfang komposition³ af substantiver; derivation af verber sker hovedsageligt med hjælp af præfikser, derivation af substantiver hovedsageligt gennem suffikser (ofte samme affikser i begge sprog).

2 I senere udgaver af "Grundriß" bruger Eisenberg begreberne 'kategoriseringer' for paradigmekategorier og 'kategorier' for enhedskategorier. Min antagelse her er, at en overensstemmelse i en paradigmekategori, fx genus, skaber en mere grundlæggende lighed mellem sprog end overensstemmelsen mellem enhedskategorier (fx hvilke konkrete genera) som paradigmekategoriens værdier gør.

3 Substantivkomposition bruges også på dansk på en ekstensiv måde, jf. fx *speciallægepraksisplanlægningsstabiliseringsperiode* (www.danmarkdigitalt.dk/hvad-er-danmarks-laengste-ord).

- Syntaks: V-2, i transitive sætninger normalt subjekt før objekt (XV ^ S > O); verbets position i underordnet sætning som V-sidst dog kun i tysk.

Konklusionen på denne korte oversigt er, at tysk og dansk i vid udstrækning har parallelle grammatiske kategorier, især paradigmekategorier. Den vigtigste forskel mellem tysk og dansk ligger i den generelt større variation af fleksionsformer (enhedskategorier), jf.:

- kongruensbøjning i forhold til numerus og person på verbet (4 mulige fleksiver på tysk);
- kasusbøjning af substantivet (3-4 mulige fleksiver på tysk)
- kongruensbøjning ikke kun i forhold til genus og numerus, men også kasus på adjektivet (5 mulige fleksiver på tysk vs. 3 på dansk)

Mindre fremtrædende forskelle i tysk er derudover:

- den flekterende konjunktiv;
- det tredelte genussystem;
- manglen på bestemthedsfleksion af substantivet;
- verbets position i bisætningen.

Fraværet af kasus som paradigmekategori i de danske substantiver betragter jeg derimod ikke som en fundamental forskel mellem de to sprog. Den vigtigste funktion af kasus, dvs. distinktionen mellem subjekt og objekt (eller 'actor' og 'undergoer' i RRG-terminologi, jf. Van Valin & LaPolla 1997) i den transitive sætning er ganske vist til stede i dansk – dog ikke i substantivet, men i de personlige pronominer. Det kan ikke besvares her, hvorvidt de ovenfor skitserede forskelle trods de store systematiske ligheder kvalificerer den tyske grammatik som kognitivt særlig vanskelig og hele sproget måske endda som nærmest ulærbar set fra et dansk perspektiv.

Tværtimod deler de to sprog en række sprogtypologiske eksotismer, der ville lede til alvorlige læringsvanskeligheder for lørners med en anden sproglig baggrund. De vigtigste er nævnt i det følgende:

- V2-ledstilling (XV \wedge S før O);
- komplekse nominalkomposita;

- stærk vs. svag adjektivfleksion;
- afflyd-verber (stærke / uregelmæssige verber);
- tal 21-99: ener før tier (*fem-og-tyve*);
- omlyd (i pluralis og gradbøjning) og *Rückumlaut* (i nogle verber)⁴;
- runde ikke-åbne frontvokaler /y, ø/;
- komplekse (konsonantiske) stavelsesrande (hoved & koda; jf. *straks*, *fangst*);
- konsonantvokaliseringer af /ʁ/ (og endnu flere på dansk, fx /g/, jf. *steg*);
- stærk reduktion i ubetonede stavelser.

2.2. Tysk = grammatik!

I dette afsnit testes den tætte sammenhæng mellem ‘tysk’ og ‘grammatik’ i dansk sprogbrug ved hjælp af to korpusundersøgelser: For det første undersøges det danske korpus fra *Wortschatz Leipzig* (www.wortschatz.uni-leipzig.de), et webcrawl-korpus fra 2018 med næsten 80 millioner sætninger og godt 1,4 milliarder ord, og for det andet undersøges KorpusDKs Korpus 2000-korpus fra Det Danske Sprog- og Litteraturselskab (<https://ordnet.dk/korpusdk>). Korpus 2000 er et bevidst sammensat og balanceret korpus med tekster fra årene 1998 til 2002, der omfatter ca. 28 millioner ord og derfor er betydeligt mindre end det danske korpus fra *Wortschatz Leipzig*; til gengæld er det frit for ulemperne ved ”blind” webcrawling.

Jeg gennemførte en søgning efter signifikante højre kontekster eller naboer (dvs. de ord, der står umiddelbart til højre) til ordformen *tysk* såvel som til andre sprogbetegnende adjektiver (*engelsk*, *fransk*, *kinesisk*, *arabisk*, *spansk*, *italiensk*, *russisk*), dvs. der ledes efter de substantiver, der typisk modificeres af de sprogbetegnende adjektiver. Af særlig interesse her er, om adjektivet *tysk* hyppigere kombineres med substantivet *grammatik* end med det semantisk meget bredere substantiv *sprog*, dvs. om folk faktisk taler (skriver) mere om tysk grammatik end om tysk sprog – og hvordan dette ser ud sammenlignet med andre sprogbetegnelser. Resultatet fra *Wortschatz Leipzig* viser, at *tysk* faktisk er det eneste adjektiv, der signifikant hyppigere forekommer som modifikator til *grammatik* end til *sprog*, i modsætning til andre sprogbetegnende adjektiver.

Det langt mindre Korpus 2000 giver næsten ingen signifikante forskel-

4 *Rückumlaut* (terminus fra Jacob Grimm) betegner vokalvekslen (dvs. fravær af omlyd i præteritum-former) i nogle svage (regelmæssige) verber, fx da. *legen* – *lagde*, *tellen* – *talte*, *smøren* – *smurte* og ty. *rennen* – *rannte* ‘at løbe’, *senden* – *sandte* ‘at sende’.

le; *tysk* viser sig at være det eneste sprogbetegnende adjektiv, som er en betydningsfuld venstre nabo til *grammatik*.

Omvendt viser en supplerende søgning efter venstre naboer til *grammatik* i Korpus 2000, at *grammatik* alene er associeret signifikant med *tysk* og ikke med andre sprogadjektiver.

Kombinationen med *tysk* er altså den klart foretrukne for *grammatik* på dansk; tysk er så at sige reduceret til grammatik. Det følgende afsnit om ordforråd argumenterer for, at dette er en voldsom reduktion af den objektive lingvistiske situation.

4. Dansk og tysk – leksikon

Morfosyntaktiske – eller ”grammatiske” – regler og regelmæssigheder er helt bestemt nyttige og vigtige for sprogindlæringen, men ordforrådet er endnu mere nyttigt og vigtigt. Et simpelt tankeeksperiment illustrerer dette: Selv hvis man kendte alle de grammatiske regler i et sprog, men slet ikke havde noget ordforråd, så kunne man slet ikke opnå en vellykket (sproglig) kommunikation. Omvendt, hvis man havde et omfattende kendskab til ordforrådet, men ikke vidste noget om grammatik, ville en rudimentær, pidgin-lignende kommunikation i det mindste kunne lykkes. Forklaringen kan findes i en medfødt universel grammatik i stil med Bickertons bioprogram (1984), med visse typologiske karakteristika (som fx $S > O$), der overskrider individuelle sprog, eller i nogle andre, tredje forudsætninger for menneskeligt sprog.

Ingen tvivler på, at tysk og dansk har et stort fælles ordforråd, og at dette skyldes sprogligt slægtskab på den ene side og sproglig kontakt på den anden. I dette afsnit analyseres først de lingvistisk-genetiske årsager kvantitativt, og derefter diskuteres de kontaktlingvistiske årsager kvalitativt.

4.1. Sproglig affinitet og kerneordforråd

Et stort set låneresistent kerneordforråd er et godt grundlag for en kvantificerende ordforrådssammenligning, hvor svært dette kerneordforråd end måtte være at bestemme. I lingvistisk typologi er Morris Swadeshs 100-ordsliste (fx i Samarin 1967: 221-223) fra 1950’erne hyppigt blevet brugt. Swadeshs originale liste indeholder allerede tysk og engelsk, dog bruger denne artikel let moderniserede versioner, som findes i ASJP-databasen (*Automated Similarity Judgment Program*; jf. Wichmann et al. 2022). En dansk version af Swadesh-listen kan også findes i ASJP. Lighed eller ulighed bestemmes ved hjælp af Levenshtein-distancer (LD). Det betyder, at

for hvert ordpar bestemmes antallet af ændringer, der kræves grafemisk, dvs. på bogstavniveau, for at opnå den danske ordform, med afsæt i tysk eller engelsk. Indsættelser, sletninger og udskiftninger er de tre tilladte og mulige typer af ændringer. Resultaterne er behæftet med en vis usikkerhed, da beregningen af LD er bogstavbaseret, og de dybe grafematiske systemer (jf. Seifart 2006) i de tre sprog er kendt for at have deres idiosynkrasier. Resultaterne er vist i tabel 1:

tysk	dedk	dansk	engelsk	endk	tysk	dedk	dansk	engelsk	endk
<i>ich</i>	3	<i>jeg</i>	<i>I</i>	3	<i>Busen</i>	-	<i>barm</i>	<i>breasts</i>	-
<i>du</i>	0	<i>du</i>	<i>thou</i>	3	<i>Herz</i>	3	<i>hjerte</i>	<i>heart</i>	3
<i>wir</i>	2	<i>vi</i>	<i>we</i>	2	<i>Leber</i>	1	<i>lever</i>	<i>liver</i>	1
<i>dies</i>	3	<i>dette</i>	<i>this</i>	5	<i>trinken</i>	3	<i>drikke</i>	<i>drink</i>	2
<i>das</i>	2	<i>det</i>	<i>that</i>	3	<i>essen</i>	-	<i>spise</i>	<i>eat</i>	-
<i>wer?</i>	3	<i>hvem?</i>	<i>who?</i>	4	<i>beißen</i>	3	<i>bide</i>	<i>bite</i>	1
<i>was?</i>	3	<i>hvad?</i>	<i>what?</i>	2	<i>sehen</i>	3	<i>se</i>	<i>see</i>	1
<i>nicht</i>	-	<i>ikke</i>	<i>not</i>	-	<i>hören</i>	1	<i>høre</i>	<i>hear</i>	3
<i>alle</i>	0	<i>alle</i>	<i>all</i>	1	<i>wissen</i>	3	<i>vide</i>	<i>know</i>	-
<i>viele</i>	-	<i>mange</i>	<i>many</i>	2	<i>schlafen</i>	6	<i>sove</i>	<i>sleep</i>	4
<i>eins</i>	2	<i>en</i>	<i>one</i>	2	<i>sterben</i>	-	<i>dø</i>	<i>die</i>	2
<i>zwei</i>	4	<i>to</i>	<i>two</i>	1	<i>töten</i>	-	<i>dræbe</i>	<i>kill</i>	-
<i>groß</i>	-	<i>stor</i>	<i>big</i>	-	<i>schwimmen</i>	5	<i>svømme</i>	<i>swim</i>	4
<i>lang</i>	0	<i>lang</i>	<i>long</i>	1	<i>fliegen</i>	4	<i>flyve</i>	<i>fly</i>	2
<i>klein</i>	-	<i>lille</i>	<i>small</i>	-	<i>laufen</i>	4	<i>løbe</i>	<i>walk</i>	-
<i>Frau</i> ⁵	-	<i>kvinde</i>	<i>woman</i>	-	<i>kommen</i>	1	<i>komme</i>	<i>come</i>	2
<i>Mann</i>	1	<i>mand</i>	<i>man</i>	1	<i>liegen</i>	2	<i>ligge</i>	<i>lie</i>	2
<i>Mensch</i>	3	<i>menneske</i>	<i>person</i>	-	<i>sitzen</i>	3	<i>sidde</i>	<i>sit</i>	3
<i>Fisch</i>	2	<i>fisk</i>	<i>fish</i>	1	<i>stehen</i>	4	<i>stå</i>	<i>stand</i>	3
<i>Vogel</i>	3	<i>fugl</i>	<i>bird</i>	-	<i>geben</i>	3	<i>give</i>	<i>give</i>	0
<i>Hund</i>	0	<i>hund</i>	<i>dog</i>	-	<i>sagen</i>	2	<i>sige</i>	<i>say</i>	3
<i>Laus</i>	1	<i>lus</i>	<i>louse</i>	2	<i>Sonne</i>	3	<i>sol</i>	<i>sun</i>	2
<i>Baum</i>	-	<i>træ</i>	<i>tree</i>	2	<i>Mond</i>	2	<i>måne</i>	<i>moon</i>	2
<i>Saat</i>	3	<i>sæd</i>	<i>seed</i>	2	<i>Stern</i>	2	<i>stjerne</i>	<i>star</i>	3
<i>Blatt</i>	2	<i>blad</i>	<i>leaf</i>	-	<i>Wasser</i>	4	<i>vand</i>	<i>water</i>	3
<i>Wurzel</i>	-	<i>rod</i>	<i>root</i>	2	<i>Regen</i>	1	<i>regn</i>	<i>rain</i>	2
<i>Rinde</i>	-	<i>bark</i>	<i>bark</i>	0	<i>Stein</i>	1	<i>sten</i>	<i>stone</i>	2
<i>Haut</i>	2	<i>hud</i>	<i>skin</i>	-	<i>Sand</i>	0	<i>sand</i>	<i>sand</i>	0
<i>Fleisch</i>	-	<i>kød</i>	<i>flesh</i>	-	<i>Erde</i>	3	<i>jord</i>	<i>earth</i>	4
<i>Blut</i>	2	<i>blod</i>	<i>blood</i>	1	<i>Wölke</i>	-	<i>sky</i>	<i>cloud</i>	-

5 Swadesh: *Weib/Frau*

tysk	dedk	dansk	engelsk	endk	tysk	dedk	dansk	engelsk	endk
<i>Knochen</i>	-	<i>ben</i>	<i>bone</i>	2	<i>Rauch</i>	4	<i>røg</i>	<i>smoke</i>	-
<i>Fett</i>	1	<i>fedt</i>	<i>grease</i>	-	<i>Feuer</i>	-	<i>ild</i>	<i>fire</i>	-
<i>Ei</i>	2	<i>æg</i>	<i>egg</i>	2	<i>Asche</i>	2	<i>aske</i>	<i>ash</i>	2
<i>Horn</i>	0	<i>horn</i>	<i>horn</i>	0	<i>brennen</i>	3	<i>brænde</i>	<i>burn</i>	4
<i>Schwanz</i>	-	<i>hale</i>	<i>tail</i>	-	<i>Pfad</i>	-	<i>sti</i>	<i>path</i>	-
<i>Feder</i>	3	<i>fjer</i>	<i>feather</i>	5	<i>Berg</i>	1	<i>bjerg</i>	<i>mountain</i>	-
<i>Haare</i>	3	<i>hår</i>	<i>hair</i>	2	<i>rot</i>	2	<i>rod</i>	<i>red</i>	1
<i>Kopf</i>	-	<i>hoved</i>	<i>head</i>	3	<i>grün</i>	1	<i>grøn</i>	<i>green</i>	2
<i>Ohr</i>	3	<i>øre</i>	<i>ear</i>	3	<i>gelb</i>	2	<i>gul</i>	<i>yellow</i>	-
<i>Auge</i>	3	<i>øje</i>	<i>eye</i>	2	<i>weiß</i>	4	<i>hvid</i>	<i>white</i>	4
<i>Nase</i>	1	<i>næse</i>	<i>nose</i>	1	<i>schwarz</i>	6	<i>sort</i>	<i>black</i>	-
<i>Mund</i>	0	<i>mund</i>	<i>mouth</i>	3	<i>Nacht</i>	2	<i>nat</i>	<i>night</i>	3
<i>Zahn</i>	3	<i>tand</i>	<i>tooth</i>	4	<i>warm</i>	1	<i>varm</i>	<i>hot</i>	-
<i>Zunge</i>	1	<i>tunge</i>	<i>tongue</i>	2	<i>kalt</i>	2	<i>kold</i>	<i>cold</i>	1
<i>Fingernagel</i>	2	<i>fingernegl</i>	<i>fingernail</i> ⁶	2	<i>voll</i>	3	<i>fuld</i>	<i>full</i>	1
<i>Fuß</i>	2	<i>fod</i>	<i>foot</i>	2	<i>neu</i>	2	<i>ny</i>	<i>new</i>	2
<i>Knie</i>	2	<i>knæ</i>	<i>knee</i>	2	<i>gut</i>	2	<i>god</i>	<i>good</i>	1
<i>Hand</i>	1	<i>hånd</i>	<i>hand</i>	1	<i>rund</i>	0	<i>rund</i>	<i>round</i>	1
<i>Bauch</i>	-	<i>mave</i>	<i>belly</i>	-	<i>trocken</i>	6	<i>tør</i>	<i>dry</i>	3
<i>Hals</i>	0	<i>hals</i>	<i>neck</i>	-	<i>Name</i>	2	<i>navn</i>	<i>name</i>	2

Tab. 1: Tyske, danske og engelske ord fra Swadeshs 100-ord-liste, modificeret efter Wichmann et al. (2022); tallene indikerer grafemiske afstande beregnet med hjælp af Levenshtein-distancer. Største lighed tysk-dansk (dedk) hhv. engelsk-dansk (endk) markeret med gråt.

Analysen viser, at andelen af større overensstemmelse mellem dansk og tysk i kerneordforrådet ifølge Swadesh med 37 forekomster klart overgår den på engelsk med 29 forekomster. I 34 af tilfældene er afstanden dansk-tysk og dansk-engelsk lige.⁷

ASJP-projektet (<https://asjp.cldd.org>) stiller dog ikke kun lister med 100 hhv. 40 basale ord til rådighed (aktuelt > 10.000 for 5590 sprog), men tillader fonologiske sammenligninger mellem sprogene hhv. deres kerneordforråd – hvilket er dets egentlige mening som sprogtypologisk projekt. Der træffes nogle relativt barske fonologiske generaliseringer,

6 Wichmann et al. (2022) noterer *claw*, Swadesh derimod *fingernail* – hvilket synes mere passende her.

7 Det er værd at nævne – og kunne godt bruges i fremmedsprogsindlæringen – at man, når lekserne er forskellige mellem tysk og dansk, i mange tilfælde uden problemer kan finde formelt nærbeslægtede, genkendelige ord med lidt forskudt mening, jf. for dansk gammeldags *døde* i stedet for *døbe* (ty. *töten*), *klejn* i stedet for *lille* (ty. *klein*), *flæsk* i stedet for *kød* (ty. *Fleisch*) osv. Dette gælder i mindre grad for engelsk og dansk.

hvad der synes at være uundgåeligt i typologien. Fx noteres der /E/ for [a, æ, ε, œ, œ, e], /o/ for [ɤ, ʌ, ɑ, o, ɔ, ɒ] og /e/ for [e, ø], dvs. at blandt andet de forholdsvist eksotiske runde frontvokaler /ø/ og /y/ forsvinder. Derudover frasorteres ikke-kognate ord automatisk. Ved hjælp af normaliserede Levenshtein-distancer, der korrigerer for forskellige ordlængder, tydeliggør distanceberegningerne, at alle tre sprog (dansk, tysk og engelsk) viser store leksikalske ligheder, men at dansk og tysk med 0,6307 er lidt tættere på hinanden end dansk og engelsk med 0,6555. Tilmed er tysk og engelsk med 0,6465 tættere på hinanden end dansk og engelsk. At blokere frasorteringen af ikke-kognate ord burde føre til (endnu) tydeligere resultater.

4.2. Sproglige kontakter

Ud over sproglægtskabet, som genetisk giver de to sprog et omfattende fælles germansk arveordforråd, kan der identificeres (mindst) fire kontaktsproglige lag af leksikalske ligheder:

- På hansatiden kommer der mange lån fra middelnedertysk ind i dansk;
- Som et resultat af reformationen kom mange højtyske lån ind i dansk;
- Dansk og tysk konfronteres ligeledes med de samme låneordsgivende sprog i de respektive epoker, i det 19. århundrede primært med fransk og i dag med engelsk, men også med italiensk, arabisk og andre i visse områder;
- Begge sprog har i betydeligt omfang absorberet internationalismer (europæismer) fra latin eller oldgræsk, ofte formidlet via fransk og engelsk (som selv i høj grad er påvirket af fransk).

Der kan skelnes mellem to faser for de mange lån fra tysk: perioden under hansatiden (ca. 1200- til 1500-tallet), hvor middelnedertyske lån kom ubemærket ind i det daglige og mundtlige danske ordforråd, og perioden efter reformationen (ca. 1500- og 1600-tallet), hvor dansk overtog et stort antal højtyske ord, mestendels i skriftsproget og i mere specialiserede registre. Disse forbliver ret iøjnefaldende på grund af deres fonologiske struktur (jf. Winge 2005). Om kvantiteten kan der siges følgende:

Die mittelniederdeutschen Lehnwörter in den skandina-

vischen Sprachen waren Gegenstand zahlreicher Untersuchungen. Hinsichtlich des Prozentsatzes der Entlehnung konnte sich die Forschung bislang nicht einigen, die Zahlen bewegen sich zwischen 20 und 50 Prozent. Entlehnungen sind auf den verschiedensten Ebenen der Sprache zu beobachten: im Bereich des Wortschatzes, der Wortbildung und der Phraseologie (Winge 1992: 19)

[De middelnedertyske låneord i de skandinaviske sprog har været genstand for talrige undersøgelser. Forskerne har endnu ikke kunnet blive enige om, hvor stor en procentdel der er lånt, og tallene varierer mellem 20 og 50 procent. Lån kan observeres på forskellige niveauer i sproget: i ordforråd, orddannelse og fraseologi]

Uanset hvordan og hvad man tæller, ordformer eller leksemer, så kan der under alle omstændigheder konstateres en enorm andel af tysk leksikalsk indflydelse. Haugen (1984, 490) fastslår: ”In dieser Zeit [omkring 1600] war der größte Teil des mittelniederdeutschen Wortschatzes so vollkommen assimiliert, daß kaum jemand daran dachte, es könne sich dabei um Fremdwörter handeln. [I denne periode ... var det meste af det middelnedertyske ordforråd så fuldstændig assimileret, at næsten ingen troede, at det kunne være fremmedord.]” Nogle stemmer som svenskeren Samuel Columbus (omkring 1670) anser endda skandinavisk/svensk og tysk for at være søskende med et næsten fælles ordforråd, hvorfor man ikke skulle kunne tale om lån, jf. fig. 2:

At disputera huilket-dera är komit åf däd andre, Swenskan af Tyskan, eller Tyskan af Swänskan, är fåfängt, efter de äre ingendera tåf huar-ann. de äre både Systrar, huars Faar ok Moor äre längst döde. Nok af, at wij kunne frijtt ta den ene tåf den andre, huad som wij behage. Systrar ok Syskonbarn måtte låhna huar ann deras arf-gods.

Fig 2: Columbus (1963, 92)

Selv den fremvoksende nationalismes puristiske sproglige udrensning (fra midten af 1700-tallet og i 1800-tallet) var primært rettet mod franske, latinske og græske låneord, og man dannede aktivt nye ord efter (høj-) tyske forbilleder, der måtte erstatte fx franske lån: *admirere* – *be-undre* (ty. *be-wundern*), *passion* – *liden-skab* (ty. *Leiden-schaft*). De (høj)tyske lån blev sna-

rere set som næsten uundgåelige sproglige fejl, grundet det nære sproglige slægtskab; jf. Selmer (1861, 194f.) i fig. 3:

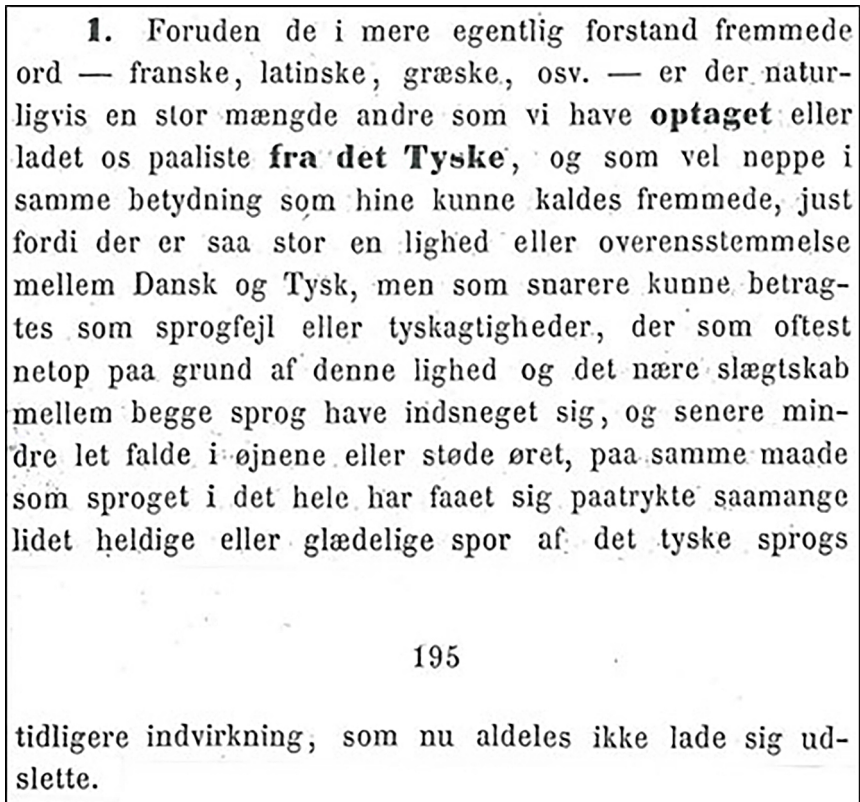


Fig. 3: Selmer (1861, 194f)

Indsigten at så mange tyske leksikalske lån er dybt forankret i dansk på det vigtigste område for fremmedsprogsindlæring, nemlig ordforråd, burde udnyttes meget mere i tyskundervisningen i Danmark, for eksempel gennem en tilpasset version af de ”Sieben Siebe [syv sier]” (Hufeisen / Marx 2014). Dette er en metode til at kunne dechifrere (store dele af) skriftlige tekster på et nyt men nærbeslægtet sprog gennem en følge af 7 interkomprehensionsprocedurer, startende med at identificere internationalismer, navn og tal (si 1), kognate ord (si 2), lydækvivalenser (si 3) osv. ved at udnytte de regelmæssige ligheder mellem sprogene.

5. Tysk og dansk i forhold til sprogtypologi

Den mest omfattende ressource der findes til typologisk sammenligning

af sprog er nok *World Atlas of Language Structures Online* (WALS; <https://wals.info>, Dryer & Haspelmath 2013). WALS indeholder oplysninger om ikke mindre end 2676 sprog på verdensplan, der maksimalt dækker 144 hovedkategorier (192 med underkategorier) inden for fonologi, morfologi, syntaks og leksikon. Fokus er klart på syntaktiske strukturer; ordforråd spiller en underordnet rolle.⁸

I denne artikel bruges WALS på to måder: For det første sammenlignes ligheder og forskelle mellem tysk-dansk og engelsk-dansk, hvor de er tilgængelige. For det andet sammenligner jeg det tyske sprogs *weirdness*, som er kommet frem i undersøgelsen af Schnoebelen (2013), med det danskes.

5.1. Dansk, tysk og engelsk i WALS

Kun 48 af de 144 hovedkategorier (58 af 192 med underkategorier) er i øjeblikket registreret for dansk, hvilket sandsynligvis skyldes de begrænsede tilgængelige kilder og WALS-teamets egne sprogfærdigheder. For (standard)tysk findes der oplysninger om 131 hovedkategorier (157 med underkategorier), for engelsk om 141 hhv. 159 kategorier. En sammenligning af dansk og tysk viser, at 35 af de 48 hovedkategorier, der overhovedet kan analyseres på basis af de beskedne danske data, har de samme værdier, svarende til 73 % (40 af 58 med underkategorier, 69 %). For engelsk viser 38 hovedkategorier, svarende til 79 %, samme værdierne (45 af 58 med underkategorier, 79 %). Når det gælder sprogtypologi med fokus på grammatik, er der i modsætning til hvad mange tror overordnet set således kun en lille forskel på ligheden mellem dansk-tysk og dansk-engelsk.⁹

5.2. Tysks og dansks weirdness

Undersøgelsen af Schnoebelen (2013) bruger WALS-data til at bestemme, hvilke sprog der er de mærkeligste og de mest normale i verden ud fra et typologisk perspektiv. Da tysk er et af de ti mærkeligste sprog i undersøgelsen, fandt den også sin vej ind i de danske medier (www.dr.dk; ikke længere tilgængelig) for at vise, at dette sprog slet ikke kan læres. For at kunne

8 Hovedkategorier fordeler sig som følgende: fonologi 19, morfologi 10, nominale kategorier 28, nominalsyntaks 7, verbale kategorier 16, ledstilling 19, enkel sætning 23, kompleks sætning 7, leksikon 10, tegnsprog 2, andet 2.

9 Der findes en kategori, hvor overensstemmelse ikke automatisk betyder større lighed og nemmere sprogindlæring. Denne er substantivgenus, hvor fravær af substantivgenus som i engelsk gør indlæringen nemmere ud fra et dansk perspektiv end at lære et andet genussystem som i tysk – selvom man har kendt denne paradigmekategori på forhånd.

analysere de komplekse og samtidig til tider ret rudimentære WALS-data på en meningsfuld måde, blev følgende skridt anvendt i tilpasset form i Schnoebelen (2013):

1. restriktion: udvælg kun egenskaber (dvs. kategorier; jf. fodnote 8), som der findes informationer om i mindst 100 sprog;
2. restriktion: udvælg kun sprog, hvor der findes informationer om mindst 10 egenskaber;
3. udvælgelse af de 21 egenskaber, der er mindst relaterede til hinanden;
4. restriktion: udvælg kun sprog, hvor der findes informationer om mindst 14 ud af de valgte 21 egenskaber, modsvarende 2/3;
5. resultat: reduktion til 239 sprog;
6. beregning af den relative hyppighed per egenskab;
7. beregning af normaliseret harmoniseret gennemsnit, subtraheret af 1, giver *Weirdness*-indexet.¹⁰

Jo højere den resulterende procentværdi er, desto mærkeligere er sproget således. Resultaterne for de ti mærkeligste sprog er vist i tabel 2 med tysk på 10. pladsen:

Rang	Sprog	Weirdness Index
1	Mixtekisk (Mexiko)	0.972
2	Nentsisk (Rusland)	0.935
3	Choctaw (USA)	0.924
4	Diegueo (USA)	0.920
5	Oromo (Etiopien)	0.919
6	Kutenaha (USA)	0.908
7	Iraqw (Tansania)	0.900
8	Kikongo (Kongo)	0.883
9	Armensk (Øst)	0.861
10	Tysk	0.858

Tabel 2: De ti mærkeligste sprog i verden ifølge Schnoebelen (2013)

10 At gennemsnittet trækkes fra 1 gør at højere tal giver et højere *Weirdness*-indeks, hvilket er kognitivt nemmere at processere.

En lille ændring af dataanalysen afslører imidlertid noget overraskende: Hvis der laves en minimal ændring i restriktionen i skridt 4 ovenfor – hvor man ud af de 21 valgte egenskaber kun medtager informationer om 13 vilkårlige egenskaber (kategorier) og ikke 14 – så er resultatet en stikprøve på 287 sprog; de resulterende 13 mærkeligste sprog er nu vist i tabel 3:

Rang	Sprog	Weirdness Index
1	Mixtekisk	0.972
2	Doyayo (Kamerun)	0.965
3	Nenzisk	0.935
4	Choctaw	0.924
5	Diegueo	0.920
6	Oromo	0.919
7	Kutenaha	0.908
8	Tashelhit (Marokko)	0.903
9	Iraqw	0.900
10	Kikongo	0.883
11	Armensk (Øst)	0.861
12	Tysk	0.858
13	Dansk	0.847

Tabel 3: De 13 weirdest sprog i verden, baseret på Schnoebelen (2013)

Det bemærkelsesværdige ved resultaterne i tabel 3 er ikke kun, at dansk tilsyneladende er næsten lige så mærkeligt som tysk ud fra et lingvistisk typologisk synspunkt, men frem for alt at tysk og dansk er de eneste nært beslægtede sprog på listen. Det kunne godt tolkes som en indikation af nogle af de sprogtypologiske eksotismer, som tysk og dansk deler.

6. Resumé og udsyn

På trods af de mange ligheder, især i ordforråd, men også i grammatikken, fokuserer tyskundervisningen i Danmark på de morfosyntaktiske og især morfologiske forskelle mellem de to sprog, hvor tysk har nogle flere fleksiver på substantiver, verber og adjektiver end dansk har. Dette bidrager væsentligt til, at tysk fra et dansk perspektiv synes at være meget vanskelig og endda næsten ulærbart – især når man sammenligner med det angiveligt så lette engelske sprog. Men som det fremgår af denne artikels analyser, viser denne vurdering sig at være problematisk i betragtning af

den lille sproglige afstand mellem tysk og dansk, både i grammatik og i leksikon. Men hvordan opstår et sådant indtryk?

Svaret på dette spørgsmål er vanskeligt og forbliver nødvendigvis spekulativt i mangel af empirisk evidens. De citerede gymnasieelevers udsagn tyder på, at tyskundervisningen i Danmark mildest talt ikke udnytter de muligheder, som så vidtrækkende ligheder giver for en kommunikativ og livlig undervisning. I stedet synes grammatisk korrekthed i alle detaljer, især de morfologiske, at være det altoverskyggende mål for tyskundervisningen.

En sprogdidaktisk tilgang, der fokuserer på eksisterende ligheder og forskelle og bygger videre på dem under hensyntagen til elevernes samlede sproglige repertoire og færdigheder, vil sandsynligvis være hensigtsmæssig. Tilgange som de førnævnte ”Sieben Siebe [syv sier]” af Hufeisen / Marx (2014) i en form, der er tilpasset dansk, synes at være yderst lovende med hensyn til undervisning for at gøre de tætte strukturelle relationer mellem de to sprog tysk og dansk håndterbare for eleverne og for at undgå unødvendige frustrationer som dem, der blev vist ovenfor.

Om forfatteren

Klaus Geyer, Dr. phil. (Kiels Universitet), Institut for Kultur- og Sprogvidenskaber på Syddansk Universitet i Odense.

Litteratur

- Bickerton, Derek (1984): The language bioprogram hypothesis. *The Behavioral Sciences* 7, 173-188.
- Columbus, Samuel (1963): *En Svensk Ordeskötsel. Utgiven med inledning och kommentar av Sylvia Boström*. Stockholm: Almqvist & Wiksell.
- Eisenberg, Peter (1989): *Grundriß der deutschen Grammatik*. Stuttgart: Metzler.
- Gabelentz, Georg von der (1901): *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. 2. udg. Leipzig: Tauchnitz.
- Geyer, Klaus (2024): Deutsch (lernen) aus dänischer Perspektive – sprachtypologische und kontaktlinguistische Überlegungen. I: *Language(s) – Sprache(n) – Sprog (sg./pl.): Ausgewählte Beiträge der 28. Linguistiktage der Gesellschaft für Sprache und Sprachen (GeSuS), Süddänische Universität in Odense, 16.-18. Mai 2022*. Klaus Geyer & Kerstin Pramstaller (red.). Hamburg: Kovač, 73-90.

- Haspelmath, Martin (2010): Framework-free grammatical theory. I: *The Oxford handbook of grammatical analysis*. Bernd Heine & Heiko Narrog (red.). Oxford: Oxford University Press, 341-365.
- Haugen, Einar (1984): *Die skandinavischen Sprachen. Eine Einführung in ihre Geschichte*. Hamburg: Buske.
- Hufeisen, Britta & Nicole Marx (2014): *EuroComGerm – Die sieben Siebe. Germanische Sprachen lesen lernen*. 2. udg. Aachen: Shaker.
- Lund, Sonja Barfod, Anne-Sofie Jakobsen, Mette Skovgaard Andersen & Hanne Wachter Kjærgaard (2023). *Elevperspektiver på fremmedsprog fra grundskole til gymnasium: En motivations- og barriereanalyse*. Online <https://viden.ncff.dk/ncff/elevperspektiver-paa-fremmedsprog-fra-grundskole-til-gymnasium>.
- Samarin, William J. (1967): *Field Linguistics: A Guide to Linguistic Field Work*. New York: Holt, Rinehart & Winston.
- Seifart, Frank (2006): Orthography development. I: *Essentials of Language Documentation*. Jost Gippert, Nikolaus P. Himmelmann, & Ulrike Mosel (red.). Berlin: Mouton de Gruyter, 275-300.
- Selmer, Hannibal (1861): *Om de i det danske sprog forekommende fremmede ord, samt tykskagtigheder, andre ufuldkomnetheder, sprog- og retskrivningsfejl. Et blik paa modersmaalets nuværende tilstand og umuligheden af dets fuldkomnere udvikling herefter*. København: Wroblewsky.
- Van Valin, Robert D. & Randy LaPolla (1997): *Syntax*. Cambridge: Cambridge University Press.
- Winge, Vibeke (1992): *Dänische Deutsche – deutsche Dänen. Geschichte der deutschen Sprache in Dänemark 1300-1800 mit einem Ausblick auf des 19. Jahrhundert*. Heidelberg: Winter.
- Winge, Vibeke (2000): *Pebersvend og poltergejst. Tysk indflydelse på dansk*. København: Gyldendal.
- Winge, Vibeke (2005): Language contact outside Scandinavia VI: with Germany. I: *The Nordic Languages*. Bind 2. Oscar Bandle, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, Ulf Telemann, Lennart Elmevik & Gun Widmark (red.). Berlin: de Gruyter, 2096-2104.

Internetkilder

- ASJP = Wichmann, Søren, Eric W. Holman & Cecil H. Brown (red.) (2022): *The ASJP Database* (version 20). Online <https://asjp.cld.org> (lokaliseret 18/4-2024)

NCFE 2022 = <https://ncff.dk/aktuelt/arrangementer/afholdte-konferencer/2022-ncffs-konference-om-barrierer-og-motivation-for-fremmedsprog> (lokaliseret 18/4-2024)

Schneebelen, Tyler (2013): *The weirdest languages*. Online <https://corplinguistics.wordpress.com/2013/06/21/the-weirdest-languages> (lokaliseret 18/4-2024)

WALS = Dryer, Matthew S. & Martin Haspelmath (red.) (2013): *The World Atlas of Language Structures Online*. Leipzig. Online <http://wals.info> (lokaliseret 18/4-2024)

Sammenhængen mellem stærktbøjede verbers bøjningsgruppe og sandsynligheden for at opnå svag bøjning i dansk

Anna Sofie Hartling

Abstract

This study delves into the shift from strong to weak inflection in Danish verbs, investigating the factors driving this evolution. The focus is on correlating inflectional groups with the likelihood of weak inflection acquisition. The strong verbs can be divided into five inflectional groups, and these are examined, with hypotheses on the influence of present and supine forms on preterite forms within the same verb. The material and method involves analyzing 24 transitioned verbs and 24 verbs with both strong and weak inflection, comparing those to 96 verbs with strong inflection alone. The statistical tests affirm a significant correlation, especially with the group of verbs that has a change of the root vowel in the preterite only.

Nøgleord

stærk bøjning, svag bøjning, verbalbøjning, sprogforandring

1. Indledning

De seneste århundreder er en række danske verber gået fra stærk til svag bøjning, og en række andre danske verber har i tillæg til den stærke bøjning fået svag bøjning. Men hvorfor tager den svage bøjning helt eller delvist over i nogle verber og ikke i andre? Tendensen til at verber med stærk bøjning i det hele taget går over til svag bøjning, har sin plausible forklaring i sprogbrugerens – til alle tider gældende – ubevidste tilbøjelighed til at få orden i systemet: For ikke at bruge unødigt kognitiv energi tilpasser vi formerne efter det system som gælder i størstedelen af tilfældene.

Analogidannelse er altså formentlig den bærende faktor i overgangen fra den mindre typiske præteritumsbøjning (stærk bøjning) til den mere typiske (svag bøjning). Analogi er kort sagt evnen til at få øje på strukturelle ligheder mellem eksisterende former og strukturer og lave forudsigelser på baggrund af disse ligheder (jf. fx Martinet 1966: 186): Det sprogtilegnende barn hører flest verber med *-ede*-præteritum og forudsiger at det næste verbum det skal producere i præteritum, også har den endelse.

Billedet er imidlertid langt mere komplekst end som så, for hvorfor er der så nogle verber der skifter bøjning, og andre der ikke gør? Hvorfor blev *gol* til *galede*, men *stjal* ikke til *stjælede*? Det peger på det faktum at der er flere medvirkende faktorer der dels samvirker og dels modvirker hinanden i sprogforandring. Ovennævnte analogifaktor og faktoren frekvens spiller givetvis begge en rolle (jf. Strik 2015 om udviklingen i svensk og frisisk verbalbøjning), men det er vanskeligt at kortlægge de måder det sker på, og samspillet mellem dem. Det er en almindelig antagelse at højfrekvente ord(former) er mere modstandsdygtige overfor at blive omstruktureret til at indgå i mere produktive bøjningsparadigmer (se også Hansen 2001: 23). Det kunne betyde at præteritumsformer af højfrekvente verber som *være* og *blive* er fredede. Det gælder imidlertid ikke at 'jo mere frekvent, jo mere beskyttet mod omdannelse', men derimod 'jo mere frekvent, jo mere beskyttet mod forandring ved analogidannelse'. Derfor ser man dels at sjældne ord får lov at "fossilere" i en bestemt form, dels at meget hyppige ord bliver "slidt ned" og fx får en kortere form, fx *hannem* > *ham*, *haver* > *har*.

I denne artikel undersøger jeg om der er en sammenhæng mellem den bøjningsgruppe de stærktbøjede verber tilhører (jf. afsnit 1.1 om bøjningsgrupperne), og verbernes sandsynlighed for (også) at opnå svag bøjning. Særligt interessant er det at se på om verberne, i en art intern analogidannelse, "smitter af" på sig selv: I de tilfælde hvor præsens og supinum ligner hinanden ved at have samme vokal, fx *drage*: *drager* (præsens) og *draget* (supinum), vil det så øge sandsynligheden for at præteritum får samme vokal som de andre, fx at *drog* bliver til *dragede*? Når jeg i det følgende skriver om at stærktbøjede verber opnår, får, udvikler, ... eller har sandsynlighed for at opnå, få, udvikle, ... svag bøjning, er det underforstået at sådanne verber oftest har begge bøjningsmuligheder i en lang periode indtil de eventuelt ender med kun at have svag bøjning (i overensstemmelse med principperne for sprogforandring i det hele taget, jf. fx Koch 2016).

Hvad angår afgrænsningen af svagt- og stærktbøjede verber, tager jeg

udgangspunkt i Retskrivningsordbogen (herefter også RO), vel vidende at der dermed bliver tale om en konservativ afgrænsning af stærk og svag bøjning i moderne dansk. De verber jeg regner for stærktbøjede, er altså dem RO regner for stærktbøjede, og de verber jeg har i gruppen af verber med både svag og stærk bøjning, er dem der i RO har begge bøjninger. Fordelen, som er rent metodisk, er at afgrænsningen af stærktbøjede verber giver sig selv (for en diskussion af ulemperne, se afsnit 5). En følge af mit afsæt i Retskrivningsordbogen er desuden at denne artikel beskriver det vi kan kalde fællesdansk, dvs. det dansk der ikke er bundet geografisk til noget bestemt sted i Danmark, vel vidende at der findes regionale forskelle – også indenfor verbernes præteritumsbøjning.

Beskrivelsen af svag og stærk bøjning tager her udgangspunkt i verbernes præteritumsform, sådan som der er tradition for (se fx Hansen & Heltoft 2019 [2011]: 651, Frederiksen 2019: 42), og inddelingen i svag og stærk er således baseret på om præteritumsformen har nulflexiv eller ej. Verber regnes for svagtbøjede når de har præteritumsendelserne *-(e)de* eller *-te* (med *-(e)de* som den langt mest produktive af de to¹), og for stærktbøjede når de har nulflexiv. Langt de fleste af de stærktbøjede verber har også vokalskifte i stammen.

1.1 De stærktbøjede verbers inddeling i bøjningsgrupper

Nedenstående inddeling af de stærktbøjede verber er baseret på Hansen & Heltofts gruppering i bøjningsmønstre, jf. Hansen & Heltoft (2019 [2011]: 654-659).

Jeg regner således med følgende mønstre:

- Mønster A: verber der kun har vokalskifte i præteritum, fx *bider-bedbidt*, *drager-drog-draget*.

1 Hos Hansen & Heltoft (2019 [2011]) hedder det at ”-te i moderne dansk ikke længere er produktiv” (s. 651). Formentlig tænker de her på nye verber, dvs. indlånte verber (fx *oute > outede*) og danske nydannelser (fx *ørle > ørlede*), for hvad angår eksisterende verbers omdannelse til *-te*-verber, forekommer det i hvert fald, fx *bedte* for *bad* og *landte* for *landede*; førstnævnte blandt både børn og unge (voksne); sidstnævnte fortrinsvis hos børn.

Omtrent 80 % af de svagtbøjede verber i Retskrivningsordbogen har i øvrigt *-(e)de*-endelsen (udregnet på basis af det samlede antal verber i RO (på <https://roplus.dk/#ordbog/>, januar 2024) fratrukket antallet af stærktbøjede verber i RO, jf. Hartling, under udarb.).

Til de svagtbøjede regnes også de verber som ender på *-te* eller *-de*, men som har vokalskifte i stammen i præteritum, fx *bringe > bragte*, *gøre > gjorde*.

- Mønster B: verber der har vokalskifte i både præteritum og supinum, med skifte til samme vokal, fx *flyver-fløj-fløjet*, *skriger-skreg-skreget*.
- Mønster C: verber der har forskellige vokaler i hhv. præsens, præteritum og supinum, fx *byder-bød-budt*, *skærer-skar-skåret*.
- Mønster D: verber uden vokalskifte, fx *græder-græd-grædt*, *falder-faldt-faldt*. (Hos Hansen & Heltoft har denne gruppe ikke noget navn).
- Udenfor kategori: To stærktbøjede verber passer ikke ind i nogen af de andre grupper, nemlig *vide* og *være*. Denne lille gruppe har jeg givet navnet ”gruppe O”.

Mønster A, C og D har undergrupper:

- Mønster A og C har undergrupper alt efter om supinum ender på *-t*: A.1. (fx *bidt*) og C.1. (fx *budt*) eller på *-et*: A.2. (fx *draget*) og C.2 (fx *skåret*).
- Mønster D har undergrupper alt efter om den vokalskifteløse præteritum har nullfleksiv: ”D.1.” (fx *græd*) eller ej (ender på *-t*): ”D.2.” (fx *faldt*).

Mønster B har kun supinum der ender på *-et* (fx *fløjet*, *skreget*).

2. Hypoteser

Den overordnede hypotese jeg undersøger i denne artikel, er:

- 1) Der er en sammenhæng mellem verbernes bøjningsmønster og deres sandsynlighed for at udvikle svag bøjning.

Jeg undersøger hypotesen både i forhold til mønstergruppe (A-O) og i forhold til vokalskiftemønster (præsens, præteritum og supinum).

Flere steder er tesen om at særligt mønster-A-verber med vokalvekslen *a-o-a* er sårbare for stærk-svag-bøjningsudvikling, blevet fremsat. Eksempelvis skriver Løkke (1855: 197) om denne ”5te Afdeling”:

Rodvokal i Præsens langt **a**; aflyder i Impf. til langt **o** (lukket); i Part. præt. vender den oprindelige Vokal (Lyden) tilbage, **a**. Denne Afdeling har især tabt mange af sine sterke Former. Selv Folkesproget har mistet flere, som Oldnorsk har (min fremhævelse).

Jensen (2002: 9-13) skriver:

Man kan ... finde støtte for en hypotese om at verber fra 6. aflydsklasse [svarende til Løkkes *5te Afdeling*, ASH] er i færd med at overgå fra det stærktbøjede til et svagtbøjet paradigme ved at se på disse verber hver for sig og tilsammen.

Verber med mønstret *a-o-a* hører under A-verberne, dvs. verber der kun har vokalskifte i præteritum, fx *bider-bed-bidt*, *drager-drog-draget*. Tillige motiveret af ovennævnte ide om en ”selvsmittende” analogi (de to ens vokaler smitter af på den afvigende vokal) undersøger jeg om sådanne A-verber i det hele taget har (haft) mere tendens til at opnå svag bøjning – udfoldet i hypotese 1.1):

- 1.1) Verberne i mønster A er overrepræsenterede blandt bøjningsskifterne.

Hypotesen om *a-o-a*-verbernes inklinations mod svag bøjning, som er fremsat ovenfor, undersøger jeg – specifikt formuleret som underhypotese 1.2):

- 1.2) Det er særligt verber efter mønstret *a-o-a* (fx *galer-gol-galet*) der har opnået svag bøjning.

3. Metode og datagrundlag

Når man tager udgangspunkt i Retskrivningsordbogen, finder man 120 simpleksverber (foruden 506 afledte og sammensatte verber) med stærk bøjning i moderne dansk (se den samlede liste i Hartling, under udarb.). For at finde frem til de stærktbøjede verber har jeg søgt på *præt. af* og 31-34 i RO i fuldtekstsøgningen på roplus.dk. 328 af de stærktbøjede verber i RO er opført med deres præteritumsform på opslagsplads, fx ”bandt *vb*, *præt. af* binde”. 300 af de stærktbøjede verber i RO (heraf en del overlappende med de ovennævnte 328) har en henvisning til § 31-34; det gælder de stærktbøjede verber som har mulighed for *n*-bøjning af participium, fx *binde* som altså kan bøjes med *-t* eller *-n* foran et substantiv i utrum: *en bundet eller en bunden hund*. Herudover har jeg lavet udførlige krydstjek på både præfikser, førsteled og rodverber i RO. Resultatet af disse søgninger har jeg i øvrigt sammenholdt med listerne over stærktbøjede verber i *Grammatik over det Danske Sprog* (med i alt 111 verber) (Hansen & Heltoft

2019 [2011]: 654-659): Udover visse uoverensstemmelser i de angivne bøjningsformer (fx supinum *sledet* i tillæg til *slidt* i Hansen & Heltoft 2019 [2011]) overfor kun *slidt* i RO) er der blot to verber som findes hos Hansen & Heltoft 2019[2011] og ikke i RO, nemlig *bides* og *rives*. De er ikke talt med her.

Af de 120 stærktbøjede simpleksverber har 24 tillige mulighed for svag bøjning ifølge Retskrivningsordbogen. Fra udgivelsen af den første retskrivningsnormerende tekst (Malling 1992 [1777]) til i dag, dvs. perioden 1777-2024, er – i hvert fald – 24 simpleksverber gået fra stærk til rent svag bøjning.² I artiklen fokuserer jeg på de verber der er gået fra stærk til rent svag bøjning, og de verber der i dag har mulighed for både svag og stærk bøjning, og ser på om der er særlige mønstre i deres bøjningsgruppetilhørsforhold sammenlignet med hele gruppen af verber der har rent stærk bøjning. Desuden fokuserer jeg på simpleksverberne og ikke de sammensatte og afledte verber da sammensætningerne og afledningerne vil mudre billedet, dels – og først og fremmest – fordi det er vidt forskelligt hvor mange sammensætninger og afledninger der findes for det enkelte simpleksverbum (fra 0 til 30), dels kan der være andre forhold der gør sig gældende hvad angår sammensatte/afledte verbers sandsynlighed for at skifte bøjningsmønster.³

	Simpleksverber	Sammensatte og afledte verber	I alt
Verber med stærk bøjning i moderne dansk, ifølge RO	120 (heraf 24 der tillige har svag bøjning)	506 (heraf 15 der tillige har svag bøjning)	626
Verber som tidligere havde stærk bøjning, men som nu har rent svag bøjning	24	4+ ⁴	28+

Tabel 1: Nøgletal for verber med stærk bøjning

2 De 24 verber der tidligere (også) har haft stærk bøjning og nu kun har svag bøjning, stammer fra lister i grammatikker fra 1700-1800-tallet, jf. tabel 2. For en grundig beskrivelse, se Hartling (under udgivelse).

3 Jf. fx den antydede hypotese i Hansen & Heltoft (2019 [2011]) om at stærktbøjede verber som er sammensætninger og afledninger, har større tendens til omdannelse til svag bøjning end stærktbøjede simpleksverber (s. 661). Denne tanke stammer sandsynligvis fra Hansen (1967, III: 17).

4 Plusset indikerer at der (naturligvis) er langt flere end fire, men på listerne i grammatikkerne nævnes som regel kun simpleksverber og herudover ofte disse fire: *befale*, *erfare*, *fornemme* og *overvære*.

I analyserne af sammenhængen mellem bøjningsgruppe og opnåelse af svag bøjning er de 24 verber der tidligere havde stærk bøjning, men som nu har rent svag bøjning, og de 24 verber der i moderne normeret dansk har mulighed for både stærk og svag bøjning, slået sammen. Det er de med den begrundelse at de unægtelig alle 48 enten har udvist eller udviser en vaklen i præteritumsbøjningen og en tilbøjelighed til også at tillade svag bøjning. Det er imidlertid ingenlunde sådan at der er vished for at de 24 verber med både svag og stærk bøjning i moderne dansk, er på vej mod kun at have svag bøjning; der er nok en vis sandsynlighed for at det vil gå sådan for en del af dem, men det er naturligvis også muligt at de vil bibeholde begge bøjninger, eller at udviklingen vil gå ”den anden vej” (igen) mod rent stærk bøjning.⁵

I det følgende refererer jeg til de ovennævnte 48 verber (simpleksverber der er gået fra stærk til svag bøjning, og simpleksverber med både stærk og svag bøjning) med fællesbetegnelsen ‘stærke > svage/svage + stærke verber’. Disse sammenholder jeg med gruppen af simpleksverber med rent stærk bøjning i RO, som jeg i det følgende benævner ‘stærke verber’.

4. Resultater

4.1. *Simpleksverber som tidligere (tillige) havde stærk bøjning, men som har rent svag bøjning i moderne dansk*

Bøjningsformerne i tabel 2 nedenfor bygger hovedsageligt på lister over verbalbøjningen hos Jessen (1868: 42-44) og hos Høysgaard (1747: 350-361). I eksempelvis *Ordbog over det Danske Sprog* (ODS) angives ofte flere forskellige mulige bøjningsformer for hvert verbum. Det ville imidlertid ikke være empirisk muligt at fastlægge hvilke af de angivne former der repræsenterer fællesdansk bedst muligt, og det ville være vanskeligt at godtgøre hvilke af formerne der er de ældste. Desuden ville det ikke være muligt at kategorisere verberne til dette formål hvis jeg medtog samtlige mulige former. Jeg har derfor valgt at basere det på ovennævnte grammatikker som har ganske fyldige lister over verbernes bøjning.

5 Udgangspunktet for de fleste af disse verber har formentlig været stærk bøjning, men nogle af dem kan have været født svage og siden haft en periode som stærke, evt. samtidig været svage.

Verbum	præsens-præteritum-supinum	Verbum	præsens-præteritum-supinum
age	ager-aag/og-aget	slynge	slynger-slang-slunget
brække	brækker-brak-brukken	smælde	smælder-smald-smældet
gale	galer-gol-galet	sprætte	sprætter-sprat-sprættet
gnave	gnaver-gnov-gnavet	svømme	svømmer-svam-sømmet
grave	graver-grov-gravet	syde	syder-sød-sydet/søden
grine	griner-gren-grinet	søbe	søber-søb-søbet/søb(b)en
hugge	hugger-hug-hugget	trives	trives-treves-treves
hvine	hviner-hven-hvinet	trænge	trænger-trang-trængt
knække	knækker-knak-knækket	tyde	tyder-tød-tydet/tydt
nemme/ næmme	næmmer-nam-næmmet/num- men	tærsk	tærsker-tarsk-torsken
skvatte ⁶	skvatter-skvat-skvattet	veje	vejer-vog-vejjet
skvætte ⁷	skvætter-skvat-skvættet	væve	væver-vov-vævet

Tablet 2: Verber som tidligere (tillige) havde stærk bøjning, men som nu har rent svag bøjning, jf. Høysgaard (1747: 350-361) og Jessen (1868: 41-44) (desuden Möbius 1871: 94-98; Jessen 1891; Matzen 1893: 290-297 – og Retskrivningsordbogen mht. konstateringen af den svage bøjning i dag).

4.2. Simpleksverber med både stærk og svag bøjning i moderne dansk

I tabel 3 ses de verber der if. RO har både stærk og svag bøjning.

Verbum	præsens-præteritum-supinum	Verbum	præsens-præteritum-supinum
briste	brister-bristede/brast-bristet	lade ⁸	lader-ladede/lod-ladet/ladt
fare	farer-farede/for-faret	nyse	nyser-nyste/nøs-nyst
fise	fiser-fisede/fiste/fes-fiset/fist/feset	skryde	skryder-skrydede/skrød-skrydet
fnise	fniser-fnisede/fniste/fnes-fniset/ fnist	skælve	skælver-skælvede/skjalv-skælvet
fnyse	fnyser-fnysede/fnyste/fnos-fny- set/fnyst	smyge	smyger-smygede/smog-smyget/ smøget

6 Som det fremgår af Hartling (under udgivelse), kunne *skvat* både være præteritum af *skvatte* og *skvætte*, og disse to verber havde flere overlappende betydninger i ældre dansk, jf. også ODS. Jeg har derfor inkluderet *skvatte* her uagtet at det ikke nævnes i nogle af de ældre grammatikker.

7 Om *skvætte* skriver Jessen (1891: 136): ”En usikkrere Existens fører enkelte Imperfecta som . . . skvat for skvættede.”

8 Der er her tale om 3. *lade* i Retskrivningsordbogen, dvs. *lade* i den betydning ordet har i udtrykket *lade vandet*.

fyge	fyger-fygede/føg-fyget/føget	smække	smækker-smækkede/smak-smæk- ket
gyde	gyder-gydede/gød-gydt	sprække	sprækker-sprækkede/sprak-spræk- ket/sprukket
gyse	gyser-gyste/gøs-gyst	stande	stander-standede/stod-standet
hænge ⁹	hænger-hængte/hang-hængt	svinge	svinger-svingede/svang-svinget/ svunget
jage	jager-jagede/jog-jaget	sværge	sværger-sværgede/svor-sværet/ svoret
klinge	klinger-klingede/klang-klinget	tie	tier-tiede/tav-tiet
kyse	kyser-kysede/kyste/køs-kysset/kyst	trine ¹⁰	triner-trinede/tren-trinet

Tabel 3: Verber med både stærk og svag bøjning i moderne dansk ifølge Retskrivningsordbogen.

4.3. Sammenhængen mellem verbernes bøjningsmønster og deres sandsynlighed for at udvikle svag bøjning

I tabel 4 nedenfor ses det hvordan fordelingen er på bøjningsmønstrene – for de stærke verber og de stærke > svage/stærke + svage verber. X, Y og Z symboliserer vokaler i præsens, præteritum og supinum.

Mønster	Stærke verber	Stærke > svage/stærke + svage verber
A (X-Y-X)	28 (29,2 %)	33 (68,8 %)
B (X-Y-Y)	33 (34,4 %)	1 (2,1 %)
C (X-Y-Z)	24 (25,0 %)	4 (8,3 %)
D (X-X-X)	7 (7,3 %)	3 (6,3 %)
O (vide, være)	2 (2,1 %)	0 (0 %)
Blandet	2 (2,1 %)	7 (14,6 %)
I alt	96 (100 %)	48 (100 %)

Tabel 4: Antal stærke verber og stærke > svage/stærke + svage verber, fordelt på mønstergrupper.

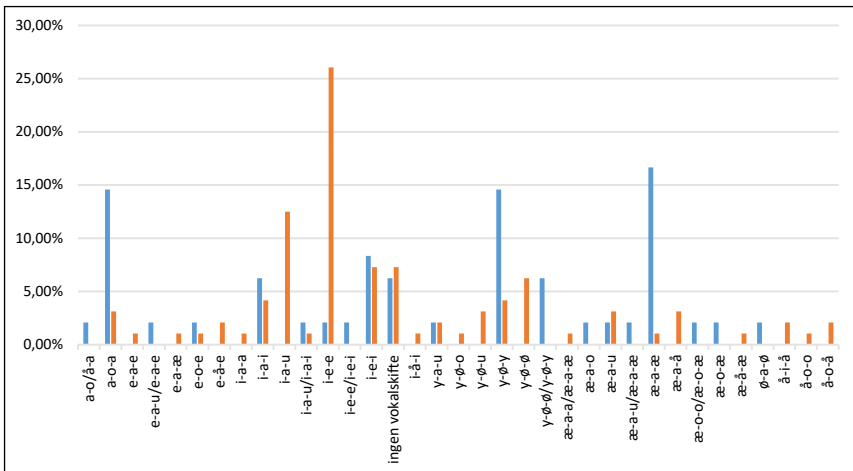
Den overordnede hypotese (1) jeg undersøger, er om der er en sammenhæng mellem verbernes bøjningsmønster og deres sandsynlighed for at udvikle svag bøjning, jf. afsnit 2. I tabel 4 kan vi se at der er nogle markante forskelle på hvordan de to sammenholdte grupper af verber fordeler sig på bøjningsmønstrene A-O. Især bemærker man at gruppen af stærke > svage/stærke + svage verber har en andel af A-verber som er over dobbelt så stor som de stærke verber. Omvendt er andelen af især B-verber,

9 Der er her tale om 2. hænge i Retskrivningsordbogen, dvs. det intransitive hænge, som modsat det transitive hænge (1. hænge) har stærk-svag-alternationen.

10 Dette verbum må nok siges at have en ekstremt marginal brug i dansk. Det har stilmarkøren ÆLDRE i Retskrivningsordbogen (foruden glosserne gå, træde).

men også C-verber langt større blandt de stærke verber end i den anden gruppe. Det vil altså sige at blandt de verber der i dag har både svag og stærk bøjning, og de verber der tidligere var stærktbøjede, men nu er svagtbøjede, er der relativt mange verber der kun har/havde vokalskifte i præteritum, fx *fare* og *grave*. Blandt de rent stærktbøjede verber er der langt flere verber der har én vokal i præsens og en anden i præteritum og supinum (mønster B), fx *rive*, og verber der har tre forskellige vokaler i præsens, præteritum og supinum (mønster C), fx *binde*. Med Fishers eksakte test (*Fisher's exact test*, FET), hvor signifikansniveauet regnes for $p < 0,05$ (gælder generelt for signifikansberegningerne i artiklen), kan det påvises at de umiddelbart markante forskelle i proportionerne i tabel 4 også er statistisk signifikante ($p < 0,001$). (Pga. det lave antal forekomster i kolonne 2 i tabel 4 kan en chi-i-anden-test ikke bruges i signifikanstestningen af disse proportioner). Testen peger altså på at der er statistisk signifikant forskel på fordelingen af stærke verber og stærke > svage/stærke + svage verber på bøjningsgrupperne og dermed en sammenhæng mellem verbernes fordeling på grupperne (A, B, ...) og bøjningen af verberne.

Også når man ser nærmere på de specifikke vokalskiftemønstre, kan man se store udsving, jf. figur 1. Det gælder både internt i grupperne 'stærke verber' og 'stærke > svage/stærke + svage verber' og imellem disse to grupper.



Figur 1: Fordelingen på vokalskiftemønstre. På y-aksen angives det hvor stor en procentmæssig andel hver verbaltype (fordelt på hvilke vokaler verbet har i præsens-præteritum-supinum) udgør indenfor sin egen gruppe, hlv. de 48 stærke > svage/stærke + svage verber (blå søjler) og de 96 stærke verber (orange søjler)

I figur 1 kan man se at de største forskelle findes i vokalskiftemønstrene *a-o-a*, *y-ø-y*, *æ-a-æ*, *i-e-e* og *i-a-u*, således at en markant højere andel af stærke > svage/stærke + svage verber har mønstrene *a-o-a*, *y-ø-y* og *æ-a-æ*, sammenlignet med de stærke verber, og at en markant højere andel af de stærke verber har mønstrene *i-e-e* og *i-a-u*, sammenlignet med de stærke > svage/stærke + svage verber. Det ser altså ud til at verber med vokalskiftemønstrene

- *æ-a-æ* (fx *kvæde: kvæder-kvad-kvædet*, *skælve: skælver-skælvede/skjalv-skælvet*, *knække: knækker-knak-knækket* (nu rent svagt))¹¹
- *a-o-a* (fx *drage: drager-drog-draget*, *fare: farer-farede/før-faret*, *gnave: gnaver-gnov-gnavet* (nu rent svagt))
- *y-ø-y* (fx *snyde: snyder-snød-snydt*, *skryde: skryder-skrydede/skrød-skrydet*, *tyde: tyder-tød-tydet/tydt* (nu rent svagt) (jf. Jessen 1868; Möbius 1871))

i særlig grad opnår svag bøjning. Man kan bemærke at alle tre vokalmønstre hører til gruppe A, hvilket stemmer fint overens med at andelen af verber fra netop gruppe A er stor blandt de verber der har fået svag bøjning, jf. tabel 4 og afsnit 4.4.

Samtidig ser det ud til at verber med vokalskiftemønstrene

- *i-e-e* (fx *gribe: griber-greb-grebet*, *skrive: skriver-skrev-skrevet*)
- *i-a-u* (fx *drikke: drikker-drak-drukket*, *slippe: slipper-slap-sluppet*, *vinde: vinder-vandt-vundet*)

i særlig grad er stabile overfor ændring i deres stærke bøjning.

4.4. A-verbernes sandsynlighed for at udvikle svag bøjning

Den første af de mere specifikke hypoteser (1.1) lyder ”Verberne i mønster A er overrepræsenterede blandt bøjningsskifterne”. Af tabel 4 fremgår det som nævnt at de stærke > svage/stærke + svage verbers andel af A-verber er over dobbelt så stor som de stærke verbers andel af A-verber. For at undersøge om denne forskel er statistisk signifikant og dermed understøtte hypotese 1.1, har jeg udført en z-test for proportioner af A

¹¹ Verber med mønstret *æ-a-æ* ser faktisk ud til i højere grad at tillade svag bøjning end verber med mønstret *a-o-a*.

(verber med vokalskifte kun i præteritum) i forhold til *I alt* (verber fra alle seks mønstergrupper), jf. tabel 4. Z-test for proportioner er en måde at analysere hvorvidt to uafhængige, ikkeoverlappende proportioner i virkeligheden er lig med hinanden. Testen viser at der er statistisk signifikant belæg for at proportionen af A-verber ift. det samlede antal verber er større i kolonnen for stærke > svage/stærke + svage verber end i kolonnen for stærke verber ($p < 0,001$). Man kan altså konkludere at verber fra bøjningsgruppe A, dvs. verber der kun har vokalskifte i præteritum og ikke i præsens og supinum, udgør en signifikant større procentdel af fordelingen for de stærke > svage/stærke + svage verber end for de stærke verber, og det peger altså på at mønster-A-verberne har større sandsynlighed for at få svag bøjning.

4.5. Fordelingen af verber mht. a-o-a-bøjning

Af tabel 5 nedenfor fremgår det hvor stor en andel af de stærke verber og af de stærke > svage/stærke + svage verber der har vokalmønstret *a-o-a* i præsens-præteritum-supinum.

	Stærke verber	Stærke > svage/stærke + svage verber
<i>a-o-a</i>	3 (3,1 %)	7 (14,6 %)
ikke- <i>a-o-a</i>	93 (96,9 %)	41 (85,4 %)
I alt	96 (100 %)	48 (100 %)

Tabel 5: Andel af de stærke verber og af de stærke > svage/stærke + svage verber med vokalmønstret *a-o-a* i præsens-præteritum-supinum

Anden specifikke hypotese (1.2) lyder ”Det er særligt verber efter mønstret *a-o-a* (fx *galer-gol-galet*) der har opnået svag bøjning”. Af tabel 5 fremgår det at 14,6 % af de stærke > svage/stærke + svage verber fulgte *a-o-a*-bøjningen overfor 3,1 % af de stærke verber (der stadig følger den), og det kunne således se ud som om der for verberne er sammenhæng mellem at have *a-o-a*-mønster og at have hhv. (tillige) svag bøjning og rent stærk bøjning. Også her kan FET vise at der er statistisk belæg for at sige at der er en sammenhæng mellem mønstret *a-o-a* og svag bøjning ($p < 0,05$). Der er imidlertid tale om en ensidet FET, så jeg har kun testet om der er sammenhæng i én retning (procenten stiger for *a-o-a* når man går fra stærke verber til stærke > svage/stærke + svage verber). (Pga. det lave antal verber med *a-o-a*-mønster kan en chi-i-anden-test ikke bruges for denne hypotese).

5. Diskussion

Foruden at have påvist en sammenhæng mellem verbernes bøjningsmønstre og sandsynligheden for at udvikle svag bøjning, evt. i tillæg til en eksisterende stærk bøjning (jf. afsnit 4.3.), har jeg også vist at der sandsynligvis er noget særligt på spil for mønster-A-verbene (jf. afsnit 4.4.-4.5.): Verberne i gruppe A, der kun har vokalskifte i præteritum, fx *bider-bed-bidt*, *drager-drog-draget*, har således især sandsynlighed for at opnå svag bøjning. De mulige forklaringer på hvorfor dette er tilfældet, er imidlertid langt fra udtømt. Man kan spekulere i om disse verbers præteritumsvokal har en øget sandsynlighed (sammenlignet med præteritumsvokalerne i verber fra de andre grupper) for at blive udskiftet, så den er den samme som de øvrige to hovedformers, dvs. fx fra *bed* til *bidede* og fra *drog* til *dragede*¹² – som en art intern afsmitning. En sådan vokalændring vil dog ikke nødvendigvis medføre at nulleksiven bliver til en svag endelse, men at præteritumsformer som fx *drog* ikke ser ud til at blive til den hypotetiske stærke form *drag*, kan måske skyldes at dette ikke er et særlig velrepræsenteret mønster i dansk. Stærktbøjede verber efter mønstret ”nulleksiv, men ingen vokalskifte i roden” (mønster D.1.) tilhører en meget lille gruppe i dansk, med i alt fem verber, og dermed er sandsynligheden for at andre verber i analogi med denne gruppe skulle danne præteritum ligesom dem, lille.

I denne artikel har jeg taget udgangspunkt i Retskrivningsordbogens angivelser af verbernes bøjning. Som nævnt betyder det at skellet mellem verber med svag og med stærk bøjning er sat relativt konservativt, forstået på den måde at ændringer i RO i såvel stave- som bøjningsmåder er længe undervejs og først føres ud i livet hvis det konstateres at de er udbredte blandt såkaldt gode og sikre sprogbrugere. Selvom den svage form af verbet *bede*, *bedte*, eksempelvis lader til at være temmelig udbredt i talesproget, vil der formentlig gå lang tid endnu før den bliver normeret alternativform til *bad*. Udgangspunktet i Retskrivningsordbogens skel betyder altså at jeg ser bort fra at flere af de verber der her regnes for rent stærktbøjede, også har en svag bøjning i sprogbrugen. Dermed får jeg altså ikke et helt retvisende billede af moderne dansk mht. verbernes bøjningsformer.

12 Hvad angår *dragede*, er denne omdannelse i fuld gang: Formen findes normeret i sammensætningerne/afledningerne *bedragede*, *bidragede* og *(gen)opdragede* (ved siden af *bedrog*, *bidrog* og *(gen)opdrog*). I sprogbrugen støder man også på *dragede*. Søger man i landsdækkende aviser, er forholdet imidlertid sådan at den normativt korrekte form klart er i overtal: *drog* = 7.748 forekomster (dog ikke korrigeret for *drog*, sb.) vs. *dragede* = 26 forekomster, jf. Sprognævnets korpus af landsdækkende aviser (Korp) på ca. 956,76 mio. løbende ord (januar 2024).

Hvis jeg skulle inddrage de verber der i sprogbrugen, men ikke i RO, har både svag og stærk bøjning, ville afgrænsningen imidlertid blive vanskelig at foretage, for hvor mange forekomster af en ikkenormeret bøjningsform skulle gælde som tilstrækkeligt til at den skulle tælle med? Og skulle det være forekomster i talesprog eller skriftsprog? Hvis man valgte skriftsprog, hvilken type skriftsprog skulle det da være? Og skulle man tælle eventuelle regionale forskelle med? I fremtidige undersøgelser ville det imidlertid være interessant at undersøge om denne artikels resultater også gælder for de verber der i disse år ser ud til at være ved at etablere en svag bøjningsform. Er det fx i særlig grad verber fra mønster A der i disse år mister stærk bøjning? I hvilken udstrækning vil analyser som den forhåndenværende (og lignende) kunne forudsige udviklingen for ændringer i verbalbøjningen? Med en udbygning af undersøgelsen af hvad det er for stærke verber der har fået svag bøjning i dansk, som dette er det første bidrag til, vil det blive muligt at pege på hvilke verber der har størst sandsynlighed for at miste den stærke bøjning. Det er muligt at sådanne analyser også kan være brugbare i Dansk Sprognævns normeringsarbejde, dvs. arbejdet med normeringen af verbernes bøjning.

6. Sammenfatning

En del verber er i dansk gået fra stærk til svag bøjning – eller i hvert fald til *også* at have svag bøjning. I denne artikel har jeg undersøgt et lille hjørne af spørgsmålet om hvorvidt nogle verber er mere ”eksponerede” for denne omdannelse end andre, og hvorvidt der kan kortlægges nogle mønstre i denne ”eksponeretthed”. Der er formentlig ikke tvivl om at sådanne omdannelser sker på baggrund af analogier mellem verberne, baseret på både stammer og bøjningsendelser. Det er også sandsynligt at verbernes frekvens spiller en rolle. I fremtidige undersøgelser vil det være interessant at granske disse faktorerers indvirkning. I denne artikel har jeg specifikt set på om der er en sammenhæng mellem verbernes bøjningsmønster og deres sandsynlighed for at udvikle svag bøjning, herunder også på om det særligt er mønster-A-verber der er skiftet, og om det blandt mønster-A-verberne særligt er dem der har vokalskiftemønstret *a-o-a*, der har fået svag bøjning.

Med støtte i statistiske tests (FET og z-test for proportioner) har jeg påvist at

- der er signifikant forskel på fordelingen af de stærke verber og de stærke > svage/stærke + svage verber på bøjningsgrupper;
- verber fra bøjningsgruppe A udgør en signifikant større procentdel af fordelingen for stærke > svage/stærke + svage verber end for stærke verber;
- de stærke > svage/stærke + svage verber havde/har en større procentdel verber med mønstret *a-o-a* end de stærke verber har.

Derudover antydes det af fordelingen på de specifikke vokalskiftemønstre at

- verber med mønstrene *y-ø-y* og *æ-a-æ* (foruden *a-o-a*) i høj grad (også) får svag bøjning, mens verber med mønstrene *i-e-e* og *i-a-u* i høj grad bevarer en rent stærk bøjning.

Bøjningsgruppertilhørsforhold blandt de stærktbøjede verber har altså, med stor sandsynlighed, betydning for om verberne overgår til (også) at have svag bøjning. Særligt verber fra mønster A og især dem med vokalskiftemønstrene *a-o-a*, *y-ø-y* og *æ-a-æ* har (også) fået svag bøjning. Omvendt ser det ud til at verber med mønstrene *i-e-e* og *i-a-u* i høj grad bevarer den (rent) stærke bøjning.

Om forfatteren

Anna Sofie Hartling, ph.d., forsker, Dansk Sprognævn.

Litteratur

- Frederiksen, Britta Olrik (2019): Bøjning. Gammeldansk. I: *Dansk Sproghistorie: Bøjning og bygning* (Bind 3). E. Hjorth, H. Galberg Jacobsen, B. Jørgensen, B. Jacobsen, M. Korvenius Jørgensen, & L. K. Fahl (red.). Aarhus: Aarhus Universitetsforlag, 41-72.
- Hansen, Erik og Lars Heltoft (2019 [2011]): *Grammatik over det danske sprog*. København: Det Danske Sprog- og Litteraturselskab og Syddansk Universitetsforlag.
- Hansen, Lise Horneman (2001): *Jysk -de-bøjning. En undersøgelse af svag præteritumbøjning*. Skrifter utgivna av Institutionen för Nordiska Språk vid Uppsala Universitet 54. Uppsala: Institutionen för Nordiska Språk vid Uppsala Universitet.

- Hansen, Aage (1967): *Moderne Dansk III*. København: Det danske Sprog- og Litteraturselskab.
- Hartling, Anna Sofie (Under udarb.): *Stærke verber i dansk. En samlet oversigt over stærktbøjede simpleksverber, sammensatte og afledte verber*.
- Hartling, Anna Sofie (Under udgivelse): De stærke som blev til de svage. Om stærkt- og svagtbøjede verber i 1800-tallet og om udviklingen fra stærk til svag bøjning. I: *Mange grublende Dage ere gaaede forud. Grammatik i 1800-tallet som nøgle til moderne normproblemer*. Eva Skafte Jensen & Jørgen Schack (red.). Bogense: Dansk Sprog-nævns skrifter 54.
- Høysgaard, Jens Pedersen (1747): *Accentuered og Raisonnered Grammatica*. København: Johann Christopher Groth.
- Jensen, Eva Skafte (2002): Historisk lingvistik. *NyS – Nydanske Sprogstudier* 31, 7-32.
- Jessen, Edwin (1868): *Dansk Sproglære*. København: Gyldendalske Boghandel. (F. Hegel & Søn).
- Jessen, Edwin (1891): *Dansk Grammatik*. København: Gyldendalske Boghandel. (F. Hegel & Søn).
- Løkke, Jakob (1855): *Modersmaalets Formlære i udførlig Fremstilling*. Kristiania: Johan Dahl.
- Malling, Ove (1992 [1777]): *Store & gode Handlinger af Danske, Norske og Holstenere samlede ved Ole Malling*. Udgivet af Erik Hansen. København: Det Danske Sprog- og Litteraturselskab og Gyldendal.
- Martinet, André (1966): *Elements of General Linguistics*. Chicago: The University of Chicago Press.
- Matzen, Matz (1893): *Modersmaalets Sproglære, udførligt og almenfatteligt fremstillet*. København: Gyldendal.
- Möbius, Theodor (1871): *Dänische Formenlehre*. Kiel: Verlag der Schwers'sche Buchhandlung.
- ODS = *Ordbog over det Danske Sprog* (<https://ordnet.dk/ods>).
- Retskrivningsordbogen* (2012): 4. udg.: A.Å. Jervelund, J. Schack, J.N. Jensen & M.H. Andersen (red.). København: Alinea.
- Strik, Oscar (2015): *Modelling Analogical Change. A history of Swedish and Frisian verb inflection*. University of Groningen.

Tekstordet som grammatisk domæne

Peter Juel Henriksen

Abstract

In this paper, we study the grammar of the individual tokens of ordinary Danish text. We use the formal framework CLINK, developed at the Danish Language Council (DSN) for use in language technology. Within the CLINK environment, all morphs (text segments with individual semantics) are lexicalized, including not only traditional word forms, but affixes, multi-word expressions, punctuation marks, and more. The grammar rules involved are purely abstract, viz. those of the Lambek calculus (categorial grammar). The topics discussed in this paper are spinoff of DSN's work with COR (det Centrale Ordregister).

Nøgleord

morfologi, kategorialgrammatik, tekstanalyse, Det Centrale Ordregister, CLINK

1. Indledning

Denne artikel beskriver DSN's nye værktøj til automatisk tekstannotation, applikationen CLINK. Skønt CLINK især er rettet mod sprogteknologiske anvendelser, har udviklingsprojektet også kastet ny viden af sig om dansk ortografi og morfologi. Denne artikel fokuserer på projektets lingvistiske sider, især inden for morfologisk analyse, og her spiller det Centrale Ordregister (COR) hovedrollen som leksikalsk basis. Artiklen begynder med en kort introduktion til COR og kategorialgrammatik (parserens formelle grundlag) fulgt af hovedafsnittene om CLINK, og til slut nogle bemærkninger om sprogteknologien i sparring med ortografien. Fodnoter kan springes over uden skade for sammenhængen.

I de sproglige eksempler man møder i lingvistiske artikler, fremstår ordene som regel i deres leksikalske nøgenhed, uden apostroffer, forkortelser, parenteser, citationstegn, streger og prikker, uden versal på første ord, og

uden punktum efter sidste. Hverdagens tekster, derimod, er et konglomerat af alle tastaturets tegn. For at kunne analysere tekst maskinelt må tekstordene først transformeres til leksemer, i en proces som sprogteknologien kalder ‘tokenizing’. Dette kan gøres mere eller mindre principfast, men mange tokenizere lider under en mangel på præcis orddefinition.

Vi har valgt en simplistisk definition af ‘tekstordet’ som en følge af teksttegn afgrænset af mellemrumstegn. For eksempel består teksten *ingen røg uden brand* af tekstordene [ingen], [røg], [uden] og [brand] mens *Ingen (af- og) pålæsning!* består af tekstordene [Ingen], [(af-], [og]] og [pålæsning!]. Denne definition er valgt især fordi den er enkel at implementere; i programmeringssproget Perl (populært blandt datalingvister) kan man for eksempel dele et tekstdokument op i tekstord ved hjælp af bare én kommando:

(1) `@tekstord = $tekstdokument =~ /(\S+)/g`

I øvrigt er definitionen ikke uden kognitiv relevans. Også vi læsere skanderer jo i første omgang en tekst i tegnbløkke adskilt af mellemrum.

2. Det Centrale Ordregister

Retskrivningsordbogen (RO) har en særlig betydning for dansk skriftsprog. Den definerer vores ortografi, ikke kun for de leksemer den dækker eksplicit, men også for tegnsætning og for regelmæssigt dannede komposita. Den er en nødvendig ressource for skolevæsen, offentlig administration, forfattere, forlag og, ikke mindst, sprogteknologi. Sidstnævnte er afhængig af RO som input til stave- og grammatikkontrol, automatisk oversættelse, men også – måske overraskende – til taleteknologi. Mere herom senere.

På opfordring fra de sprogteknologiske brancher (Kirchmeier et al. 2019) har Sprognævnet udarbejdet en maskinlæsbar version af RO kaldet Det Centrale Ordregister (COR). COR-registeret har en procedure til at forsyne ethvert dansk lemma med et unikt id der, ligesom CPR-nummeret, fungerer som en eksakt henvisning til en entitet med et kompleks af egenskaber, som ellers kan være vanskelige at referere til under ét. COR-registerets leksikalske grundressource COR₁ omfatter alle Retskrivningsordbogens lemmaer som fulde ordformer (godt 530.000 former). Hvert leksem har et COR-id i denne form:

(2) `COR. <lemmaindeks> . <bøjningsindeks> . <variantindeks>`

I denne artikel anvendes forkortelserne LIX, BIX og VIX for de tre indekser (bestående af hhv. 5, 3 og 2 cifre).¹ Databasen COR₁ og dens BIX-tabel (afbildning af BIX-værdier på ordklasse og bøjning) kan downloades fra ordregister.dk. De tre tekstciter (3) er fra korpus PAROLE (Keson 1999, Henrichsen 2023), her annoteret med leksikalske henvisninger ('COR-links').

(3a)	(3b)	(3c)
retten / COR. 43153.111	retten / COR. 43157.111	han / COR. 01880.991
frikassé / COR. 53868.110	til / COR. 00947.880	samlede / COR. 30151.206
af / COR. 00014.880	den / COR. 00267.910	et / COR. 00831.916
gedde / COR. 69689.110	samlede / COR. 30151.214	års / COR. 30151.206
og / COR. 00099.970	flåde / COR. 60753.110	junkmail / ?
ørred / COR. 64233.110		

Bemærk at de to forekomster af *retten* har forskellig LIX, men samme BIX (forskellige lemmaer, samme bøjning) mens det omvendte gælder *samlede*. Links til tekstord der ikke forekommer i COR₁, kræver særlig opmærksomhed (fx *junkmail*); mere om det herunder. Som det fremgår, er tekster annoteret med COR-links fri for homografi (modulo COR₁), hvad der i mange sammenhænge gør dem lettere at databehandle både for mennesker og maskiner.

COR₁ er, som allerede nævnt, Det Centrale Ordregister's leksikalske grundniveau; der findes også andre COR-ordbøger med fx ordsemantik (Nimb et al 2022), udtaleinformation (Henrichsen 2024), historisk ortografi (Widmann 2023) og – særlig relevant for denne artikel – specialordbøger med interpunktionstegn, affikser, fuger, numeraler m.v.

3. Kategorialgrammatik i morfologiens tjeneste

I dette afsnit gennemgås kategorialgrammatikkens grundtræk, med særligt fokus på den såkaldte 'sekventanalyse' som spiller en central rolle i CLINK. I (4a) og (4b) ses to komposita, *antiwoke* og *udrikkelighed*, analyseret som sekventer.

¹ VIX bruges bl.a. til ortografiske varianter såsom *ressource/ressurse* (hhv. COR. 95344.110.01, COR. 95344.110.02) og *af sted/afsted* (COR. 13533.900.01, COR. 13533.900.02). VIX udelades ofte i denne artikel.

<p>(4a)</p> <p>A/A A ==> A</p> <p>[anti] [woke] [antiwoke]</p>	<p>(4b)</p> <p>A/A V V\A A\N ==> N</p> <p>[u] [drikke] [lig] [hed] [udrikkelighed]</p>
---	--

En ‘sekvent’ består af to dele, ‘antecedenten’ til venstre for ‘sekventpilen’ (==>), ‘konsekventen’ til højre. Stammen [woke] har kategorien A (for adjektiv) mens affikset [anti] har en sammensat kategori (A/A). Skråstregen / bruges til funktioner der søger argument til højre, bagstregen \ når argumentet søges til venstre. Funktioner af type $x\backslash y$ og y/x siges at afbilde et x på et y (*mapping x to y*). Sekventanalyse består i at bevise at kategorien i konsekventen kan dannes af kategorierne i antecedenten gennem litter lovlig ‘applikationer’ (*applications*), altså skridt hvor en funktion får det argument den beder om. I (4b) må *drikkelig* komponeres før *u-* kan affigeres, og *-hed* til sidst.

Morfologiske flertydigheder får hver sit bevis. Det illustrerer vi her med en lille anekdote. For nylig fortalte forfatterens datter om sin nystiftede studenterkontorklub, og faderen troede det drejede sig om en kontorklub (kun) for studenter. Men nej, klubbens navn skyldes at den holder til i studenterkontoret, og den er således også åben for seniorer². Misforståelsen kan forklares med sekventeringen i (5) af navnet *studenterkontorklub*.

(5)

N	N\N/N	N	N\N/N	N	==>	N
[student]	[er]	[kontor]	∅	[klub]		[studenterkontorklub]

Hvis fugen [er] tages som udgangspunkt for delsekventen $N\ N\backslash N/N\ N\ N\ ==> N$, får man først komponeret *studenterkontor* og derefter tilføjet *klub*. Med fugen ∅ som basis komponeres *kontorklub* først.

Konstituenten med særlig grammatisk affinitet vil man ofte vælge at leksikalisere en bloc. For eksempel kan [student] og [er] komponeres vha. delsekventen $N\ N\backslash N/N\ N\ ==> N/N$ og leksikaliseres som en sammensætningsform (*studenter* i kategorien N/N). Denne teknik kaldes ‘partial deduction’. Retskrivningsordbogen har ca. 8000 sammensætningsformer af denne type. I tilfældet *studenter-* er det en fordel at den sjældne fuge [er] på denne måde bliver knyttet rent leksikalsk til de leksemer der skal bruge den, så den ikke er åben for orddannelse generelt. Bemærk

² Faderen bades holde sig væk alligevel.

i (6) at den strukturelle tvetydighed stadig er bevaret efter den partielle deduktion.

(6)

$N/N \quad N \quad N \setminus N / N \quad N \quad \implies N$
 [studenter] [kontor] [] [klub] [studenterkontorklub]

Hvert morfem i et sammensat ord bidrager til ordets sekventering med en eller flere kategorier. En stamme bidrager med én grundlæggende kategori mens alle andre morfemer er funktioner: præfikser, suffikser, infikser, fuger, osv. Man kan udtrykke det sådan at hvert morfems kategori deklarerer morfemets bidrag til hele ordets struktur. Selve bevisførelsen hviler på såkaldte ‘sekventregler’; det fører for vidt at forklare dem her, men der findes gode introduktioner til kategorialgrammatik (fx Wood 1993 for almen lingvister; Morill 1994 for formelle lingvister).³

I et kategorialgrammatisk perspektiv er de traditionelle termer præfiks og suffiks aliaser for henholdsvis y/x og $x \setminus y$ (hvor x og y er kategorier hver for sig) mens fugen, uanset sit materiale, er alias for $y \setminus x/x$. I dette afsnit bruger vi $\{V, N, A\}$ som grundlæggende kategorier og har altså ni forskellige fugekategorier til rådighed for orddannelse. Alle ni er rigt repræsenteret i moderne dansk (Henrichsen 2021), med $V \setminus V / V$ som den sjældneste (*løbetræne, tørretumble, sultestrejke*); se tabel 1.

Morfem	Kategorial funktion	Eksempler leksemer og sammensætninger	
stamme	grundlæggende kategori (svarende til en ordklasse)	[gul] : A [drikke] : V [ven] : N	-
suffiks	$x \setminus y$ (funktion fra kategori x til y)	[skab] : $N \setminus N$ [lig] : $V \setminus A$	[ven][skab] [drikke][lig]
præfiks	y/x (funktion fra x til y)	[over] : A/A [for] : V/V	[over][klog] [for][blænde]
fuge	$x \setminus y/y$ (funktion fra x og y til y)	[] : $A \setminus V / V$ [e] : $N \setminus N / N$ [s] : $N \setminus A / A$	[fin][pudse] [barn][e][vogn] [mand][s][høj]

Tabel 1. Udvalg af danske funktionsmorfemer (tabel tillempet fra Henrichsen 2021)

3 Lambek-kalkylen har én aksiomregel: $A \implies A$, samt fire bevisregler for operatorerne / og \ (kaldet ‘divisorer’), $L_0 A/B M_1 N_0 \implies C$ IF $M_1 \implies B$ AND $L_0 A N_0 \implies C$ OG $L_1 \implies A/B$ IF $L_1 B \implies A$ samt $L_0 M_1 B \setminus A N_0 \implies C$ IF $M_1 \implies B$ AND $L_0 A N_0 \implies C$ OG $L_1 \implies B \setminus A$ IF $B L_1 \implies A$ hvor A, B, C er vilkårlige kategorier, og L_n, M_n, N_n er lister af n -vilkårlige kategorier. Se detaljer i Wood (1993).

En fordel ved kategorialgrammatikkens koncise definitioner er at man i høj grad undgår dilemmaet mellem at udvikle en grammatisk nomenklatur afledt af sit eget sprog og at leve med one size fits all-termer fra et dominerende oversprog. I COR-sammenhæng er et sproguafhængigt regelsystem særlig vigtigt, alene fordi mange af COR₁'s kernebrugere (sprogteknologerne) hverken har lingvistisk baggrund eller dansk som førstesprog.

4. COR møder kategorialgrammatikken

Vi bruger termen 'morf' for et tekstelement som i en given analytisk sammenhæng er atomisk, altså ikke kan opdeles uden at miste sin leksikalske afbildning. Eksempler på morfer: [hest], [km/t.], [7], [!], [(.)], men ikke [he], [st], [km/t] og [(.). Herfra erstattes termerne substantiv, verbum og adjektiv (S, V og A) med BIX-indeks. Således er **123** kategori for sb.itk.pl.best (substantiv, intetkøn, pluralis, bestemt form). Også delvist instantierede kategorier forekommer, med punktum som vikar for et ciffer; fx står **1 . .** for en vilkårlig substantivform (sb), **12 .** for en neutrumsform (sb.itk) og **1 . 3** for en bestemt flertalsform (sb.pl.best). Den helt uspecificerede kategori noteres med et variabelsymbol (**X**, **Y**, **Z**, . . .), som illustreret i (7).

(7)

SEGMENTER: [studenter] [kontor] [] [klubben]
 SEKVENT: **Z/Z** **129** **Y\X/X** **111** ==> **111**

Valget af variabelsymboler er arbitrært. Fx kan **X** frit erstattes med **W**, forudsat at **W** ikke forekommer i forvejen (proceduren kaldes 'alfakonversion'). Sekventen i (7) finder to beviser, nemlig for hhv. **Z=129** (*studenterkontor | klubben*)⁴ og **Z=111** (*studenter | kontorklubben*). Bemærk at præfikskategorien er blevet generaliseret i forhold til kategorien N/N i afsnit 2. Dermed finder fx adjektivet *studenterløs* nu også en analyse (for **Z=300**, der er BIX for adjektiver som *løs*).

4.1. Segmentering af sammensatte tekstord

RO rummer en stor mængde hyppigt forekommende komposita, hvoraf langt hovedparten er regelmæssigt dannede (både morfologisk og se-

4 Kategorier med formen **1 . 9** svarer til RO's sammensætningsformer (sb.sms)

mantisk) og dermed i en vis forstand redundante. Men som bekendt har de mest benyttede komposita en tendens til at drive væk fra deres udgangspunkt mht. betydning og udtale (*hulvej, livmoder*), og det gør en vis overrepræsentation nødvendig. Som konsekvens står den praktisk virkende morfolog ofte over for et valg mellem flere segmenteringer. Tag for eksempel tekstordet *antifødselsdagssang-revolution*, der forekom i Politiken's stribe Ting Jeg Gjorde af Maren Uthaug (28/2 2024). I sin kontekst var ordets hensigt klar, nemlig at afsyngning af fødselsdagssange til voksne er en uskik og bør bekæmpes ved revolution. Segmentet [revolution] har altså skopus over resten af ordet, mens [anti] har skopus over [fødselsdagssang].

(8)

SEG: [anti] [fødselsdagssang] [-] [revolution]
 SEK: **z/z 119** **y\x/x 110** **==> 110**

Det viser sig imidlertid at COR₁ ikke har et opslag for *fødselsdagssang*. Hvis sekventanalysen skal dækkes ind af COR₁-ordbogen, må det lange token deles i mindst seks segmenter.

(9)

SEG: [anti] [fødselsdag] [s] [sang] [-] [revolution]
 SEK: **z/z 119** **v\w/w 119** **y\x/x 110** **==> 110**

Underdelingen kan naturligvis fortsættes.

(10a)

SEG: [fødsel] [s] [dag]
 SEK: **119** **y\x/x 119** **==> 119**

(10b)

SEG: [fød] [sel] [s] [dag]
 SEK: **200** **200\119** **y\x/x 119** **==> 119**

Somme tider kan man argumentere for én segmentering frem for en anden med henvisning til den almindelige sprogfornemmelse eller til de mest udbredte ordbøger; men oftest afhænger svaret mere af det sprog-

teknologiske formål⁵. I alle tilfælde vokser det morfologiske analyserum med antallet af segmenter (se tabel 2).

Segmenter	Analyser
[anti][fødselsdagssang][-][revolution]	2
[anti][fødselsdag][s][sang][-][revolution]	5
[anti][fødsel][s][dag][s][sang][-][revolution]	14

Tabel 2. Sammenhæng mellem segmentering og analyserum

Jo større dele af et kompositum der genkendes, des mindre flertydighed står tilbage. Et fyldigt kompositum som *medarbejdertilfredshedsundersøgelse* er nemt at tolke på grund af sine velkendte komponenter, modsat det ligedannede *udarbejderafpasningsoverfindelse*. Leksikalsk information spiller en afgørende rolle for disambiguering, både for mennesker og computere, men den overordnede diskurs spiller også en vigtig rolle. Vi vender tilbage til CLINK's parsestrategi og til Uthaug's stribe i afsnit 6.

5. Den leksikalske kategori som nøgle til semantisk komposition

Det er et karakteristisk træk ved kategorialgrammatikken at konstituentanalyse (morfologi, syntaks) og semantisk analyse går hånd i hånd. Denne isomorfi er også et hovedprincip i CLINK.

For leksemer med et COR₁-id bruges lemmaindekset som stedfortræder for lemmaets semantik. Adjektivet *grøn* har fx lemmaindekset **15974**, så i CLINK-sammenhæng er egenskaben 'grøn' repræsenteret med symbolet **15974**. Skulle nogen engang i fremtiden ønske at indføre en maskinlæsbar forklaring på farvekvaliteten *grøn* (hvordan den end kommer til at se ud), kan denne forklaring så træde i stedet for sin proxy. Indtil da nøjes vi med at ræsonnere med LIX-værdier som stedfortrædere for betydning. Morfer der ikke har en selvstændig betydning, men i stedet er funktioner af andre morfers betydning, kaldes somme tider 'synkategorematisk' i den semantiske litteratur. I CLINK kendes en funktionsmorf (fx affiks, fuge, bindestreg, interpunktionstegn) på at dens kategori har mindst én skråstreg. En sammensat kategori modsvares af en semantisk komposition, en såkaldt 'lambdafunktion'.

5 For eksempel behøver en maskinoversættelse en finere segmentering end en talesyntese. Ordet *antifødselsdagssangrevolution* kan læses op på samme måde uafhængigt af sin tolkning.

(11)

SEG:	[anti]	[woke]		
SEK:	X/X	300	==>	300
SEM:	la . a (02854)	27102		27102 (02854)

En lambdafunktion består af et hoved med et antal variabler (fx **λabc**) og en krop som de samme variabler indgår i. Funktionens argumenter indsættes i kroppen i den rækkefølge hovedet angiver. Når variablerne er brugt op, er formlen ikke længere en funktion, men et afsluttet semantisk udtryk der viser delenes relationer. Til *antiwoke* (funktionen **la . a (02854)**) for [anti], med argumentet **27102** for [woke]) svarer således udtrykket **27102 (02854)**, der kan parafraseres som: adjektivet *woke* modificeret af præfikset *anti*.

RO's præfikser (fx *anti*, *ærke*, *skide*) og sammensætningsformer (fx *rigs*, *lamme*, *lande*, *lands*, *land*) er i CLINK kategoriseret som **X/X**, med lambdaudtrykket **la . a (LIX)**, hvor **LIX** er lemmaindekset i COR_1 (fx **02854** for *anti*).

5.1. Det sammensatte farvenavn – præfikks versus fuge

De fleste sammensatte farvetermer følger et fast semantisk mønster, nemlig at det sidste led er hovedfarven og det første led en modifikation. Eksempler fra DDO ses i (12).

- (12a) lysegul "af en lys gul farve"
- (12b) karrygul "gul eller gulbrun som karry"
- (12c) rødgul "af en gul farve med en orange tone"
- (12d) orangerød "af en kraftig rød farve med en orange tone"
- (12e) bleg rød "af en svagt rød eller lyserød farve"
- (12f) blå rød "af rød farve med en blålig tone"

DDO forklarer desuden at formen *lys*+FARVE (*lysgul*, *lysrød*) også forekommer, dog "især fagligt". I RO finder man *lysegul*/-blå/-rød/-brun/-grøn og -grå, og dertil *lyslilla* og *lysviolet*; desuden sammensætningsformerne *lys*-, *lyse*- og *mørke*-, og dermed kan også *lysgult*, *mørkeviolet* osv. analyseres som præfikks+adjektiv.

(13)

SEG:	[lys]	[gult]		[lysgult]
SEK:	X/X	301	==>	301
SEM:	λa . a (40871)	15978		15978 (40871)

Sammensatte farvenavne hvis førsteled ikke er et præfiks, men fx et adjektiv eller et substantiv, samles i stedet med en fuge (oftest [], [e] eller [s]).⁶

(14)

SEG:	[blå]	[]	[rødt]		[blå+rødt]
SEK:	300	Y \ (X/X)	301	==>	301
SEM:	15302	λab . b (a)	15892		15892 (15302)

Som nævnt findes sammensætningsformerne *lys-*, *lyse-* og *mørke-* i COR₁ – men ikke *mørk-*. Med en nulfuge kan fx *mørkviolet* komponeres af [mørk]:COR. **15581** og [violet]:COR. **24305 . 301**. Bemærk at BIX er irrelevant i alle andre kompositumled end det sidste; den danske morfologi bøjer jo komposita efter kun det sidste led. Det er afspejlet i fugens kategori ved at venstreargumentet er uspecificeret, dvs. at de to argumenter ikke behøver at have samme kategori (både **X=Y** og **X≠Y** er tilladt). Nu kan vi forsyne de to sekventeringer af *studenterkontorklubben* fra afsn. 3 med semantisk annotation.

(15a)

SEG:	[student]	[er]	[kontor]	[]	[klubben]
SEK:	119	V \ (W/W)	129	Y \ (X/X)	111 ==> 111
SEM:	96483	λab . b (a)	48871	λcd . d (c)	47998

(15b)

SEG:	[studenter]	[kontor]	[]	[klubben]
SEK:	Z/Z	129	Y \ (X/X)	111 ==> 111
SEM:	λa . a (96483)	48871	λcd . d (c)	47998

5.2. Rødhvide farver, vi skal ud at slå

Visse sammensatte farvetermer er, modsat *karrygul* og *blårod*, ikke seman-

6 Bemærk at der er føjet parenteser til fugens kategori **Y \ (X/X)** for at vise at venstreleddet skal appliceres først da det leverer lambdafunktionens første argument (**15892** → **a**).

tisk velbeskrevne som et møde mellem en grundbetydning og en moderator.

- (16a) rødvid "ofte som symbol på den danske nation, fx i sportsbeklædning"
 (16b) blågul "blå og gul; fx om sportsbeklædning"
 (16c) rødgrøn "rød og grøn; som vedrører rød og grøn"

Ordforklaringerne i (16) er fra DDO. Der er lignende eksempler på komposita med ligestillede dele i mange andre ordklasser.

- (17a) sursød "Hjemmetlavet gammeldags sursød sovs"
 (17b) majjuni "I Danmark får den to kuld i majjuni og august"
 (17c) fredaglørdage "at score kassen på fire fredaglørdage med underholdning"
 (17d) totre "hvid skjorte, sorte bukser og totre dage gamle skægstubbe"
 (17e) f3 "hvid springer g1 til f3 kaldes Rétiåbningen"
 (17f) spaghettikødsovs "danskerne har fortjent deres spaghettikødsovs"

I den klassiske stilistik kaldes formerne i (17) for 'dvandvasammensætninger' (fx Albeck 1968: 79-80); de er karakteristiske ved at bestå af dele der er semantisk sideordnede, enten konjugerede (sursød = sur og sød), disjungerede (majjuni = maj eller juni) eller ubestemte mellem de to (fumlertumler = barn der fumler og/eller tumler).⁷ Dvandvakomposita har brug for en ny semantisk operator.

- (18)
 SEG: [sur] [] [sød]
 SEK: 300 $x \setminus (x/x)$ 300 ==> 300
 SEM: 27204 $\lambda_{ab} \cdot (a, b)$ 15282 (27204, 15282)

De to sideordnede led noteres som et par, **(a, b)**.

5.3. Fugen og de tre hængsler

Dansk morfologi har tre produktive fuger, [], [e], [s] (i ortografien suppleret af [-] og enkelte andre) mens de øvrige sjældent forekommer i fri form. I den klassiske morfologi har fugen [] som bekendt det mistænkelige navn

7 Der er en tendens til at dvandvaformer skrives med bindestreg; fx har RO og DDO bindestreger i (16)'s leksemer. Dette er dog ikke et gennemført ortografisk princip, og formerne i (16) og (17) er også hyppigt forekommende. Eksemplerne i 17 er googlet.

'nulfuge'. Det er en uheldig term. Nok er fugen usynlig i skriftbilledet, men ikke uhørlig i udtalen og altså ikke *nul* i lingvistisk forstand. I kompositummet *flagbajer* med segmenterne [flag] [] [bajer] fungerer det midterste segment som en fonetisk operator der dels selekterer for udtalen /flAw/ (frem for den leksikalske udtale /fla:?:/), dels styrer tryktabet i segmentet [bajer]. Regelmæssige udtaleændringer af denne type er vores ortografi dog for fattig til at repræsentere. Alle morfologiske fuger, inklusive [], kan derfor anses for materielle (i betydningen 'sansbare').

Vi bruger termen 'hængsel' for fugens semantik. Som omtalt i sidste afsnit kan to forskellige fuger godt have det samme hængsel (fx i *lysgul* og *mørkelilla*), og tilsvarende kan to forskellige hængsler være knyttet til den samme fuge (fx i *rødgul* og *rødavid*). Hængslet $\lambda_{ab}.b(a)$ kaldes 'prædikationshængslet'; det samler sine argumenter i en underordningsrelation. Det andet hængsel $\lambda_{ab}.(a,b)$ sideordner argumenterne og kaldes 'dvandvahængslet'. Der er også brug for et tredje hængsel, nemlig til sammensatte tekstord hvis segmenter slet ikke indgår en morfologisk relation, som i for eksempel teksten (*højt skum*) mellem segmenterne [] og [højt], samt [skum] og []. Da vi har defineret tekstordet som en gruppe af grafemer uden at skelne mellem bogstaver, cifre og andre tegn, er funktionstegnene nu kommet inden for den morfologiske analyses horisont – hvad der afslører en ny type af flertydighed i dansk ortografi.

(19a) Det var d. 19. feb. ph.d. T.H. Andersen betalte 1.200 kr. inkl. moms.

(19b) DSB overvejer Alstoms letbane-plattform Citadis til diesel el. el.

(19c) Hun er 5 år, han er 6. Hun bor på 7. sal, han bor på 8.

De fleste danske funktionstegn, især punktum, komma og bindestreg, har mere end én ortografisk funktion, normalt uden at det giver anledning til misforståelser (jf. 19a). I (19b) forekommer tekstordet *el.* i to strukturelt forskellige varianter, dels som forkortelse (et leksem) og dels som to samskrevne morfer (et leksem plus et slutpunktum med sætningsskopus). I sidstnævnte tilfælde er der ingen morfologisk forbindelse mellem segmenterne, og det kalder på et nyt hængsel som ikke relaterer sine argumenter til hinanden, men bare sætter dem på listeform [*x,y,z, ..*]. Dette tredje hængsel, 'liste-hængslet', har kategorien $Y \setminus ((Y * X) / X)$ med den semantiske form $\lambda_{ab}. [a, b]$. Bemærk at dets kategori har samme ydre struktur

som de andre hængsler, med et venstre- og et højreargument (hhv. \mathbf{Y} og \mathbf{X}), mens den indre funktor ($\mathbf{Y}*\mathbf{X}$) afspejler at de to kategorier ikke kombinerer.⁸

I (19c) forekommer tekstordene 6., 7. og 8.. Af konteksten fremgår det at de tre numeraler skal læses som hhv. seks, syvende og ottende.

(20a)

SEG:	[6]	[]	[.]		
SEK:	601	$\mathbf{Y} \setminus ((\mathbf{Y}*\mathbf{X})/\mathbf{X})$	x1	==>	601*x1
SEM:	NUM (6)	$\lambda a b . [a, b]$	f1		[NUM (6) , f1]

(20b)

SEG:	[7]	[.]		
SEK:	601	601 \ 602	==>	602
SEM:	NUM (7)	$\lambda a . \text{ORDINAL} (a)$		ORDINAL (NUM (7))

I (20b) tager punktumtegnet rollen som ordinalsuffixs, dvs. funktionen fra **601** (kardinaltal) til **602** (ordinaltal). I modsætning hertil er (20a) ikke en morfologisk komposition, men to uafhængige morfer og derfor forbundet med listehængslet. Man kan nu spørge om listehængslet kun er relevant for tekstord med funktionstegn, altså kun for rent ortografiske fænomener? Ikke nødvendigvis. Der findes som bekendt komposita med orddele der egentlig ikke indgår en morfologisk relation, men snarere minder om sætningskonstituentter, ord som *omsiggribende*, *italesætte*, *gudskelov*, *ovenikøbet* og *uladsiggørlig*. Hvis CLINK møder et sådant kompositum (og ikke har det leksikaliseret), kan en listehængselanalyse komme på tale. Dette emne, som også har bredere leksikologisk interesse, behandles dog bedst som del af en generel undersøgelse af grænsefladen mellem morfologi og sætnings syntaks. Herom i en senere publikation.

Med tekstordet 8. indtræffer en ortografisk krise. Ordet læses tydeligvis *ottende* i (19c), altså med punktummet som ordinalsuffixs; men ortografien forlanger et slutpunktum. Med andre ord, punktumtegnet er 'overlastet' med to funktioner hvoraf ingen kan undværes. For den almindelige læser er det måske en skønhedsplet, men for grammatikeren (og CLINK)

8 Operatoren * betegnes som Lambek-kalkylens 'produkt'. Der er to bevisregler for *:
 $L_0 \mathbf{A}*\mathbf{B}M_0 \implies C \text{ IF } L_0 \mathbf{A} \mathbf{B}M_0 \implies C$ og $L_1 \mathbf{M}_1 \implies \mathbf{A}*\mathbf{B} \text{ IF } L_1 \implies \mathbf{A} \text{ AND } \mathbf{M}_1 \implies \mathbf{B}$,
 hvor \mathbf{A} , \mathbf{B} , \mathbf{C} er vilkårlige kategorier, og L_n , M_n er lister af n + vilkårlige kategorier.

bryder overlastningen med det grundlæggende princip om én morf, ét semantisk bidrag. Den danske ortografi har mindst et andet eksempel på overlastning, nemlig i sammenhænge som *hunden der huler, katten og hønen*, hvor kommategnet nødvendigvis fungerer som både opremsnings- og slutkomma (Togeby 2000 har et lignende eksempel).

6. Automatisk tekstanalyse

Programmet CLINK ('COR-linker') kan beskrives som en input-output-automat der læser en tekst og annoterer hvert tekstord for kategori (BIX) og for semantik (λ -udtryk). Tabel 3 viser CLINK's output for teksten "Gud ske Lov for Sofahjørnets Fløj!" (fra Johs.V. Jensens digt "Ved Frokosten"). Bemærk at kategorier og semantiske udtryk udskrives med et løbenummer (0, 1, 2, ..) når de er genereret fra en flerordsforbindelse i COR₁ (i dette tilfælde *gud ske lov* med id **COR.14223.900.01**). Ud-råbstegnets kategori og semantik er let forenklet i tabel 3 (hhv. **x8** og **f8**).

Input (tekstord)	Segmenter	Kategori	Semantik
"Gud"	[gud]	900:0	14223:0
"ske"	[ske]	900:1	14223:1
"Lov"	[lov]	900:2	14223:2
"for"	[for]	880	00093
"Sofahjørnets"	[sofa] [] [hjørnets]	125	41222 (77215)
"Fløj!"	[fløj] [] [!]	110*x8	[48705, f8]

Tabel 3. CLINK's annotation af input "Gud ske Lov for Sofahjørnets Fløj!"

For tekstord der findes i COR₁, er annotationen direkte kopieret fra leksemets COR-id; i øvrige tilfælde er annotationen genereret med kategorial analyse (som omtalt i afsn. 4 og 5).

6.1. CLINK's algoritme – et overblik

CLINK's grundlæggende datastruktur er en 'template', en liste af leksikalske oplysninger af forskellig art der karakteriserer en morf.

- overflade (grafemisk form m.m.)
- kategori (ordklasse, bøjning, morfologisk funktion)
- betydning (proxy eller λ -formel)
- udtale (behandles i en senere artikel)

Templates for alle de morfer der er leksikalsk tilgængelige (i COR_1 og evt. andre COR -ordbøger), er samlet i CLINK's 'templatebase'. Basens tekniske specifikation er fremlagt andetsteds (Henrichsen 2024, ordregister.dk). Når CLINK-programmet kaldes med en inputtekst eller et tekstkorpus, hentes først alle relevante leksikalske oplysninger i templatebasen, og derefter begynder tekstordsanalysen. Algoritmen kan beskrives som en iteration i fem niveauer.

1. Teksten deles op i tekstord (tegnfølger adskilt af blanktegn)
2. Hvert tekstord deles op i segmenter (morfer med hjemmel i COR_1 m.fl.)
3. Hver segmentering omsættes til sekventer (med kategorier fra templatebasen)
4. Hver sekvent forsøges bevist (med Lambek-kalkylens bevisregler)
5. Hvert bevis lambda-reduceres (med λ -udtrykkene fra templatebasen)

For de enkleste tekstord (*hesten*, *er*, *gammel*) giver denne procedure kun én analyse. De fleste tekstord er dog flertydige, for det første hvis ordet kan segmenteres på flere måder (*5.*, *trækvinden*), for det andet hvis et segment er homograft i COR_1 (*sky*, *får*) og for det tredje hvis en sekvent har flere beviser (*studenterkontorklubben*). I tilfælde af flertydighed bruger CLINK en lang række forskellige heuristikker til at udpege en foretrukken analyse baseret på klassiske grammatikregler, kontekstanalyse og statistik (Henrichsen 2022). Output af analysen giver et godt udgangspunkt for fx stave- og grammatikkontrol, taleteknologi og maskinoversættelse – en computer finder BIX-værdier og lambda-formler lettere at læse end ord.

Og så er tiden kommet til at udsætte CLINK for Uthaug's *antifødselsdagssang-revolution*. CLINK finder fem forskellige betydninger (se tabel 4).

λ -udtryk	Parafrase
$\neg (\mathbf{SR} (\mathbf{F}))$	det modsatte af: sang-revolution ang. fødselsdag
$\neg (\mathbf{R} (\mathbf{FS}))$	det modsatte af: revolution ang. fødselsdagssang
$\mathbf{SR} (\neg \mathbf{F})$	sang-revolution ang. det modsatte af fødselsdag
$\mathbf{R} (\neg \mathbf{FS})$	revolution ang. det modsatte af fødselsdagssang
$\mathbf{R} (\mathbf{S} (\neg \mathbf{F}))$	revolution ang. sang om det modsatte af fødselsdag

Tabel 4. Semantiske analyser af *antifødselsdagssang-revolution*. λ -udtrykkene er let forenklet (\mathbf{F} , \mathbf{S} , \mathbf{R} og \neg erstatter LIX for hhv. fødselsdag, sang, revolution og anti)

Som før nævnt er den rette tolkning utvivlsomt **R (¬FS)**. Det er tankevækkende at verdens førende oversættelsesmaskine, Google Translate, selv med adgang til hele sribens tekst, konsekvent vælger en forkert morfologisk analyse blandt disse fem, uanset om målsproget er sat til engelsk, tysk, hollandsk, fransk, spansk, italiensk, svensk eller norsk.

6.2. CLINK i offentlighedens tjeneste

Programmet CLINK er her præsenteret i sin standardkonfiguration, men kan også indstilles til at aflevere andre former for output efter brugerens ønsker. Man kan fx vælge kun at få tekstord annoteret med en kategori (svarende til en almindelig PoS-tagger) eller at få samtlige analyser udskrevet (i CLINK's prioriterede rækkefølge). Endelig kan programmet kaldes med enhver COR-ordbog som er formelt velformet (dvs. holder COR's regler for indeksering m.m., se ordregister.dk). I begyndelsen af 2025 udgiver DSN ordbogen COR.TALE med udtaleinformation for alle morfer i COR₁, særligt rettet mod talesyntese. COR.TALE vil gøre det muligt at anvende CLINK til lydskrivning af tekst.

Denne korte artikel kan kun give et første indtryk af CLINK, og vi efterlader uden tvivl en mængde spørgsmål. Hvordan leksikaliseres RO's flerordsforbindelser, som jo går udover tekstordets domæne (*af sted, pro et contra*)? Hvordan kategoriseres et bindeord (*og, men, til*) for at afspejle kravet om kongruens mellem argumenterne? Hvordan leksikaliseres opremsningskommaet? Slutkommaet? Startkommaet??? Hvordan undgår CLINK at se nulfigurer overalt? Og sidst men ikke mindst: Hvilke selektionskriterier benytter CLINK-applikationen til flertydige tekstord (*5., el., sky, rødgul, antifødselsdagssang-revolution* osv.)? Disse og andre spørgsmål bliver diskuteret i DSN's løbende publikationer. I øvrigt kan man altid kontakte Sprognævnet på pjh@dsn.dk (om CLINK), tw@dsn.dk (om COR) og tha@dsn.dk (om rettigheder til Sprognævnets leksikalske ressourcer). Program CLINK er p.t. i version 0.9 og frigives til offentligheden som 'SaaS' (Software as a Service) fra og med version 1.0, efter planen i perioden 2024-25. Besøg COR-projektets site ordregister.dk, hvor der bringes nyheder om projektet, og hvor sprogressourcer omtalt i denne artikel kan tilgås.

7. Afsluttende bemærkning

Sprognævnets CLINK-arbejde har blotlagt flere træk ved den danske ortografi som hverken er beskrevet i Retskrivningsordbogen eller andre

steder. Den rige informationsstruktur som CLINK-templaten knytter til enhver morf (ikke kun ordbogens almindelige lemmaer, men også affikser, fuger, funktionstegn m.m.), gør det lettere at beskrive hvordan de forskellige sproglige lag, grammatik, betydning og udtale, interagerer med ortografien. For eksempel tror vi at RO's paragraffer – a fortiori kom-mareglerne – kan formuleres mere præcist i CLINK-termer end på eksempelbasis. Skulle det være rigtigt, går vi måske en fremtid i møde med pålidelig kommakorrektur i tekstbehandlingen. Den dag vil Sprognævnets spørgetelefon få frigjort en fjerdedel af sin kapacitet til mere fornuftige emner.

Om forfatteren

Peter Juel Henriksen, ph.d., seniorforsker, Dansk Sprognævn.

Litteratur

- Albeck, Ulla (1968): *Dansk Stilistik*. 6. udg. København: Gyldendal
- Henriksen, Peter Juel (2021): Glemte Ord. En undersøgelse af H.C. Ørstedes nyord og deres plads i nudansk. *NyS – Nydanske Sprogstudier* 60, 7-36.
- Henriksen, Peter Juel (2022): Det Centrale Ordregister. Et indeks for det danske ordforråd – en gave til dansk sprogteknologi. I: *Nordiske Studier i Leksikografi*. Louise Holmer et al. (red.). Lund: Lund Universitet, 113-126.
- Henriksen, Peter Juel (2023): Diktatoriske Befølelser: Om ord og uord i Det Centrale Ordregister. I: *19. Møde om Udforskningen af Dansk Sprog*. Kirstine Boas et al. (red.). Aarhus: Aarhus Universitet, 133-148.
- Henriksen, Peter Juel (2024): Make each morf count. A new approach to computational lexicography for text processing. I: Kristina Despot et al. (red.). *Proceedings of EURALEX-21*. Cavtat: Institute of Croatian Language and Linguistics (accepteret).
- Keson, Britt (1999): Vejledning til det Danske Morfosyntaktisk Taggede PAROLE-korpus. I: *Parole Dokumentation*. Jørg Asmussen (red.). København: Det Danske Sprog- og Litteraturselskab, 1-67.
- Kirchmeier, Sabine, Peter Juel Henriksen & Philip Diderichsen (2019): *Dansk Sprogteknologi i Verdensklasse. Rapport fra sprogteknologiudvalget under Dansk Sprognævn nedsat af Kulturministeriet*. København: Dansk Sprognævn
- Morill, Glyn Verden (1994): *Type logical grammar. Categorical logic of signs*. Dordrecht: Kluwer Academic Publishers.

Nimb, Sanni et al (2022): COR-S – den semantiske del af Det Centrale OrdRegister. *LexicoNordica* 29, 75-97.

Togeby, Ole (2000): 'Henret ikke benådet!'. Kronik i dagbladet Information, 28. februar 2000.

Widmann, Thomas (2023): Det Centrale Ordregister og dets leksikografiske anvendelser. I: *Nordiske Studier i Leksikografi*. Louise Holmer et al. (red.). Lund: Lund Universitet, 415-430.

Wood, Mary McGee (1993): *Categorial Grammars*. London: Routledge.

Internetlinks

<https://ordregister.dk>

Ledstilling i attributive ledsætninger

Torben Juel Jensen &
Tanya Karoli Christensen

Abstract

This article examines the use of declarative (XV(S)A) vs. neutral (SVA) word order in attributive clauses in spoken Danish. Overall, we find that they display a low proportion of declarative word order (11%) compared to nominal and adverbial subclauses, but this figure covers substantial variation. The word order alternation is thus significantly influenced by the factors syntactic function, conjunction, adverbial type, length, and factuality of the subclause. Further, qualitative analyses reveal that attributive clauses with declarative word order are very often governed by a cataphoric element in their matrix clause. We argue that the results support the hypothesis that declarative word order functions as a foregrounder, though assertivity (factuality) also seems to play a role.

Nøgleord

ledstilling, attributive ledsætninger, eksplikatkonstruktion, forgrund

1. Indledning

Dansk har to ledstillingsmønstre som på trods af de traditionelle betegnelser hel- og ledsætningsmønster (Diderichsen 1974) begge kan forekomme i ledsætninger. Af samme grund indføres i *Grammatik over det Danske Sprog* (GDS) betegnelserne deklarativ og neutral ledstilling for at undgå sammenblandingen med syntaktisk status (som hel- eller ledsætning) (Hansen & Heltoft 2011: 1677 ff.). Neutral ledstilling har altid subjektet placeret først, dernæst eventuelle adverbialer, og finittet er således på tredje plads (V3- eller AV-ledstilling; jf. eksempel 1), mens deklarativ ledstilling har et fundamentfelt hvori stort set alle led fra sætningen kan placeres, og det finitte verbal er altid placeret på pladsen umiddelbart derefter (V2- eller VA-ledstilling; jf. 2). I ledsætninger der starter med subjektet og ikke har

noget sætningsadverbial, kan ledstillingen ikke afgøres (det gælder hovedparten af ledsætningerne i talesproget; Jensen 2011: 128).

Det er også en af grundene til								
	K	S	SA			V _f	V _i	neutral (av)
(1)	<i>at</i>	<i>vi</i>	<i>næsten ikke</i>			<i>tør</i>	<i>flytte</i>	
	K	F	V _f	S	SA	V _i	deklarativ (va)	
(2)	<i>at</i>	<i>vi</i>	<i>tør</i>		<i>næsten ikke</i>	<i>flytte</i>		

Vi har i en serie af studier undersøgt brugen og betydningen af de to ledstillinger i ledsætninger (Jensen & Christensen 2013; Christensen, Jensen & Christensen 2015; Christensen, Christensen & Jensen 2016; Christensen & Jensen 2022). Undersøgelserne har haft som primært fokus at teste om ledstillingsmønstrenes distribution kan forklares ud fra funktionelle hypoteser: *Assertivitetshypotesen* hævder at ledstillingsforskellen har at gøre med illokutionær ramme, således at den deklarative ledstilling markerer at ledsætningen, trods sin syntaktisk underordnede status, er assertiv eller i hvert fald har ”konstativt potentiale” (Julien 2007; Hansen & Heltoft 2011: 1682 ff.; Julien 2015). *Forgrundshypotesen* hævder at deklarativ ledstilling er et signal til tilhøreren om at den vigtigste information i den komplekse sætning skal findes i ledsætningen, ikke i oversætningen (Wiklund m.fl. 2009; Jensen & Christensen 2013). Resultaterne understøtter hypotesen om at deklarativ ledstilling fungerer som et forgrundssignal, men også ledsætningens faktualitetsværdi spiller en rolle, hvilket tyder på at også assertivitet er knyttet til deklarativ ledstilling. Hvor vi hidtil har fokuseret på nominale og adverbielle ledsætninger, vil vi i denne artikel fremlægge resultaterne for attributive (også kaldet adjektiviske) ledsætninger. Disse har samlet set en væsentlig lavere andel af deklarativ ledstilling end nominale og adverbielle ledsætninger: Attributive ledsætninger har i LANCHART-korpusset deklarativ ledstilling i 11 % af tilfældene (de nominale: 64 % og de adverbielle: 58 %). Som vi vil vise i denne artikel, dækker dette tal dog over store forskelle, hvor visse typer af attributive ledsætninger har en væsentlig større forekomst af deklarativ ledstilling.

Attributive ledsætninger er kendetegnet ved ikke at udgøre et sætningsled i sig selv; de optræder derimod som efterstillet adled til et andet led (i visse tilfælde til hele sætninger). Vi definerer i vores undersøgelse attributive ledsætninger bredere end traditionen, som til kategorien

adjektiviske ledsætninger alene regner ledsætninger der modificerer et nominal (Hansen & Heltoft 2011: 1480). Vi følger derimod Sune Sønderberg Mortensen (Mortensen 2014) ved også at regne ledsætninger der modificerer adjektiver (jf. 3), adverbier (jf. 4) og prædikater/sætninger (jf. 5) som attributive, dvs. ledsætninger som ofte regnes som adverbielle.¹

- (3) jeg var bare heldig den ikke satte ind før
 (4) nu spiller jeg i Vinderup så meget jeg nu kan uden at overbelaste min ankel eller lyske
 (5) og der var han jo den type der hvis øh børnene de sang falsk ikke hvad de jo gør ikke så blev han meget ophidset

2. Data og metode

Undersøgelsen baserer sig på det samme datamateriale fra LAN-CHART-korpusset som vi tidligere har anvendt i vores undersøgelser af nominale og adverbielle ledsætninger. Det består af 132 samtaler med 92 informanter fra København og Vinderup optaget i perioderne 1987-88 og 2005-07.² Vi henviser til vores tidligere publikationer (fx Christensen & Jensen 2022) for en nærmere beskrivelse af design og informantdata, men vil ikke komme yderligere ind på de ikke-sproglige variable i denne artikel da vores analyser ikke tyder på at de har betydning for ledstillingsvariationen i attributive ledsætninger.

I datamaterialet forekommer 2022 attributive ledsætninger. Heri medregnes alene ledsætninger hvori ledstillingen kan afgøres som enten deklarativ eller neutral. Disse ledsætninger blev kodet for ledstilling samt en række syntaktiske og semantiske faktorer som fremgår af Figur 1. De første seks faktorer er medtaget på baggrund af vores tidligere undersøgelser og er relateret til henholdsvis forgrunds- og assertivitets-hypotesen.

Det drejer sig for det første om typen af sætningsadverbial, ledsætning-

1 I afgørelsen af om en given ledsætning er attributiv, blev bl.a. Mortensens topologiske kriterium anvendt: "To determine that a subclause is not a constituent in its own right, a topological guideline can be employed: Since a sentence constituent can normally only be fronted as a whole, attributive clauses, as opposed to nominal and adverbial clauses, cannot be fronted alone – they need their governing head" (Mortensen 2014: 411).

2 Tre af informanterne anvender dog ingen attributive ledsætninger, så de resultater der fremlægges i denne artikel, baserer sig reelt på 89 informanter og 130 samtaler.

gens placering i oversætningen samt ledsætningens længde. Disse er inddraget ud fra en antagelse af at forekomsten af dialogiske adverbialer i ledsætningen, final placering af ledsætningen samt ”indholdsmæssig tyngde” i ledsætningen operationaliseret som antal ord alt andet lige vil korrelere med at ledsætningen udgør forgrund information i den komplekse sætning den er del af, og dermed ifølge forgrundshypotesen er disponeret for deklarativ ledstilling.

Faktoren integration handler om hvor tæt ledsætningen er knyttet til oversætningen: Restriktive ledsætninger er kodet som integrerede mens parentetiske (ikke-restriktive) ledsætninger er kodet som ikke-integrerede. Pga. de ikke-integrerede ledsætningers selvstændige status må det antages at de ofte udgør forgrund information, men selvstændigheden kan også betyde at de udgør en selvstændig sproghandling (Hansen & Heltoft 2011: 1513 & 1520). Såfremt ikke-integration er korreleret med deklarativ ledstilling, kan det således siges at understøtte begge hypoteser.

Faktoren faktualitet, dvs. om taleren forpligter sig til om indholdet i ledsætningen er sandt eller ej, er derimod tæt knyttet til assertivitetshypotesen. Ud fra denne hypotese må man nemlig forvente at non-faktuelle ledsætninger forholdsvis sjældent, om nogensinde, forekommer med deklarativ ledstilling, eftersom sådanne ledsætninger ikke asserteres (hævdes). Også faktoren disfluens kan knyttes til assertivitetshypotesen på den måde at disfluens i starten af en ledsætning kan være udtryk for at taleren er ved at omprocessere sin ytring (en form for reparatur). Såfremt taleren omtolker sætningens status, er det muligt at den samtidig får en mere selvstændig illokutionær værdi og dermed gives deklarativ ledstilling. Vi henviser til vores tidligere publikationer for en nærmere beskrivelse af de hypotesebaserede faktorer (Christensen, Jensen & Christensen 2015; Christensen & Jensen 2022).

Endelig blev ledsætningerne kodet for konjunktion og syntaktisk funktion. I alt 13 forskellige konjunktioner indleder de attributive ledsætninger i datamaterialet; derudover forekommer en del af ledsætningerne enten uden konjunktion eller med et komplekst konjunktion som fx *end hvis*, eller *som at*. Af overskuelighedshensyn blev alle ledsætninger kategoriseret efter den første konjunktion (fx *som*), alle *hv*-indledte konjunktioner blev slået sammen til én kategori, og de sjældent forekommende konjunktioner *så*, *end*, *når*, *om* og *da* blev slået sammen til en restkategori (Other).

Hvad angår den syntaktiske funktion, drejer det sig primært om hvilken type af syntagma ledsætningerne fungerer som adled i. Adjektiv- og ad-

verbiums syntagmer blev slået sammen til én kategori, mens ledsætninger i nominalsyntagmer – som udgør 61 % af alle attributive ledsætninger i data – blev opdelt i ”relativsætninger” (NRel) og ”ikke-relativsætninger” (NOther) på baggrund af om de i sig selv er semantisk ukomplette hhv. komplette (dvs. enten kræver eller ikke kræver et korrelat i oversætningen). Dermed inkluderer kategorien relativsætninger i vores undersøgelse alene det der i GDS betegnes som nominalrelativer (Hansen & Heltoft 2011: 1516 ff.), mens NOther inkluderer såvel GDS’ appositioner som ekspliktkonstruktioner (Hansen & Heltoft 2011: 1506-1514) – mere om det senere. Kløvningsledsætninger udgør en kategori for sig selv, ikke mindst fordi de antages oftest at udgøre præsupponeret baggrundsinformation (Hansen & Heltoft 2011: 1793 ff.; Christensen 2019), og man derved ud fra begge semantiske hypoteser må forvente en lav forekomst af deklarativ ledstilling heri.

FORGRUNDSHYPOTEBSEN	Type af sætningsadverbial* i ledsætningen (Adv/Type)	Dialogisk (dialogiske partikler og bandeord), fx <i>jo, da, sgu, vel, fandeme, kraftedeme</i> Anden type, fx <i>ikke, lige, så, måske, ligesom, også, aldrig</i>
	Ledsætningens position i oversætningen	Initial , dvs. før finit verbal, fx <i>altså der <u>hvor det virkelig kneb</u> det var også det var matematikken ikke</i> Final , dvs. efter finit verbal, fx <i>men der er jo også de skilsmisser kan man sige <u>hvor moren prøver virkelig at P holde på sit og skubbe faren væk</u></i>
	Ledsætningens længde (Length)	Antal ord herunder ord i eventuelle indlejrede ledsætninger, fx <i>vi lavede sådan et håndboldhold P med lærerne <u>hvor vi spillede simpelthen i en i oh ganske almindelig turnering serie fire håndbold sammen med dem</u></i> (17 ord)
BEGGE HYPOTEBER	Integration	Integrerede (ikke-parentetiske) ledsætninger, fx <i>altså jeg har jo mødt mit en dreng <u>som aldrig har set en hvid mand før</u> ikke</i> Ikke-integrerede (parentetiske; NonInt) ledsætninger, fx <i>så det endte med at det blev Krebs’ Skole i oh oppe ved Stockholmsgade <u>som jo selvfølgelig var et lille stykke vej derfra</u> ikke</i>
ASSERTIVITETS- HYPOTEBEN	Faktualitet	Faktuelle ledsætninger (Factual), fx <i>altså jo jeg kan huske det på den måde <u>vi fik jo et nært forhold til vores mormor og morfar også</u></i> Non-faktuelle ledsætninger (NonFactual), fx <i>altså det er jo sådan noget <u>man lige kan gøre en time om aftenen</u></i>
	Disfluens på overgangen mellem ledsætning og matrixsætning	Disfluens (pauser, tøvelyde, ordgentagelser og selvrettelser, interjektionaler eller indskudt sætningsformet materiale; Disfl), fx <i>det sværeste er jo at skulle sige til folk de ikke skal være her mere specielt medarbejdere <u>som oh som egentlig gør det godt men som man ikke har brug for</u></i>

IKKE HYPOTEBASERET	Konjunktion (MainConj)	0, at, hv-ord (<i>hvor, hvad, hvilket, hvis, hvorfor, hvori</i>), som, Other
	Funktion	<p>NRel(ativsætning) <i>altså du får nogle salgstal som du simpelthen skal opfylde</i></p> <p>NOther <i>og så havde jeg fået den fidus jeg skulle bare sige jeg ville være oh militærnægter</i></p> <p>AdjAdv(-syntagmer) <i>og der er jo meget mere end man overhovedet aner</i></p> <p>Prædikater og sætninger (Pred) <i>jeg har altid siddet hjemme om aftenen ikke hvor jeg sådan set ikke har foretaget andet end at høre musik</i></p> <p>Kløvnings (Cleft) <i>der var en hel masse de ikke måtte</i></p>

Figur 1: Sproglige faktorer inkluderet i undersøgelsen. *I sætninger med mere end ét sætningsadverbial refererer typen til det første af dem.

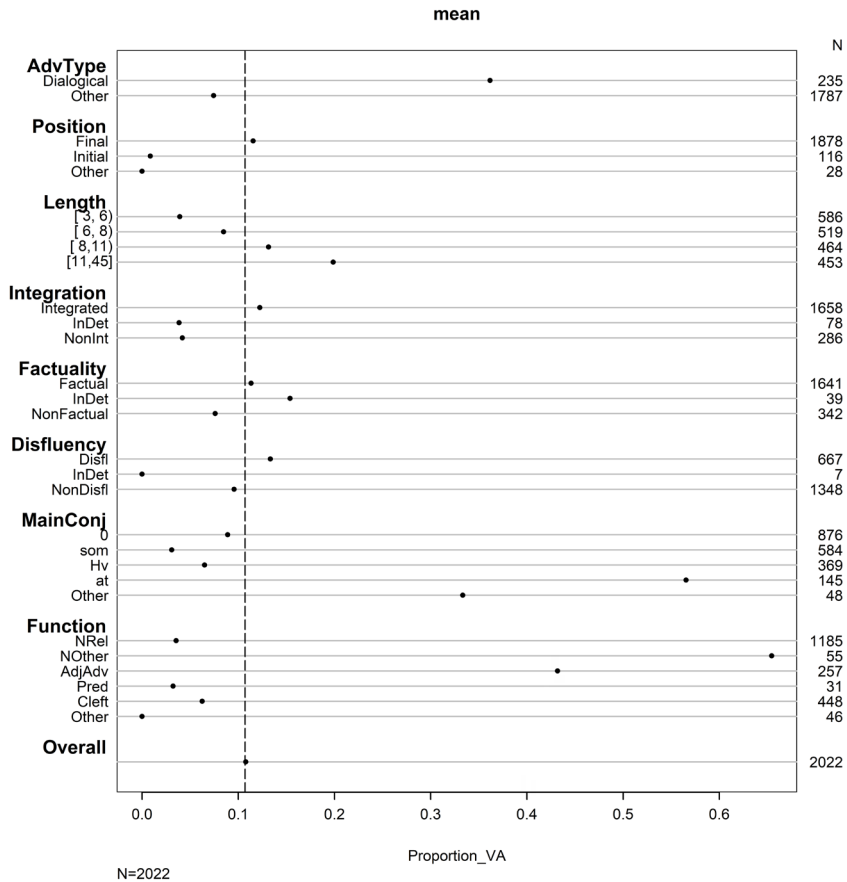
3. Resultater

Set under ét har de attributive ledsætninger i datamaterialet deklarativ ledstilling i 11 % af tilfældene, men som det fremgår af Figur 2, varierer andelen en del afhængigt af hvilken type ledsætning der er tale om.

Størst variation ses i forhold til ledsætningens syntaktiske funktion (Function): Ikke-relative ledsætninger i nominalsyntagmer (NOther) og ledsætninger i adjektiv- og adverbiumsnytagmer (AdjAdv) har relativt høje andele af deklarativ ledstilling, henholdsvis 65 % og 43 %, mens andelen af deklarativ ledstilling hos de øvrige typer ligger på mellem 3 og 6 % (i restgruppen, Other, er andelen 0). Som det fremgår af figuren, er relativsætninger (NRel) klart hyppigste type af attributiv ledsætning i datamaterialet (59 %), men de forekommer i under 4 % af tilfældene med deklarativ ledstilling.

Også ledsætningens konjunktion (MainConj) ser ud til at have stor betydning, og her skiller *at*-ledsætningerne sig ud med en andel af deklarativ ledstilling langt over gennemsnittet (57 %). Også Other-gruppen af konjunktioner har relativt høj andel (33 %), hvilket reelt alene skyldes konjunktionerne *end* og *især*, *så*, mens ledsætninger med *som* omvendt kun har deklarativ ledstilling i 3 % af tilfældene.

De øvrige variable er som nævnt medtaget i undersøgelsen på baggrund af hypoteser om funktionen af deklarativ ledstilling som markør for henholdsvis 'forgrund' og 'assertion'. Resultaterne understøtter i alle tilfælde på nær Integration begge hypoteser, på den måde at deklarativ ledstilling specielt forekommer i forbindelse med dialogiske adverbialer, final placering af ledsætningen, lange ledsætninger, faktisk værdi af ledsætningen og disfluens ved overgangen mellem matrixsætning og ledsætning.

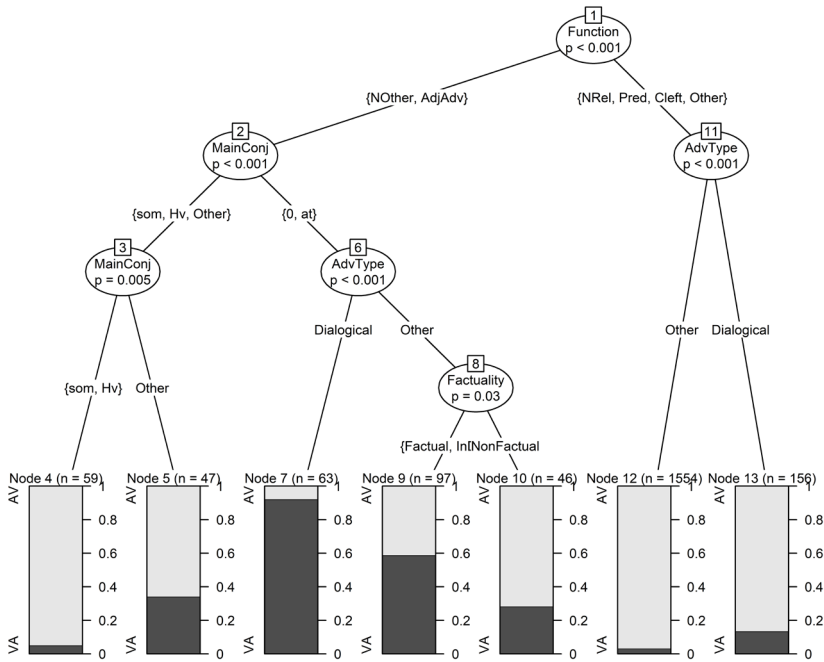


Figur 2: Distribution af deklarativ ledstilling (VA) i forhold til de undersøgte uafhængige variable. Other og InDet repræsenterer henholdsvis restkategorier og forekomster som er ubestemmelige i forhold til den givne variabel.

Det skal dog bemærkes at disse distributionelle resultater må tages med et gran salt idet variablene i flere tilfælde ikke er uafhængige af hinanden. Dette gælder ikke mindst for variablene funktion og konjunktion idet fx ledsætninger indledt med *som* altovervejende udgøres af relativsætninger og kløvningssætninger. Der er også afhængighed mellem funktion og flere af de hypotesebaserede variable idet ledsætninger i kløvninger, adjektiv- og adverbiumsntagmer samt ikke-relative adled i nominalsyntagmer stort set altid er integrerede, ligesom det kun er relativsætninger og ledsætninger i adjektiv- og adverbiumsntagmer der forekommer initialt i matrixsætningen.

Der er altså tale om kollinearitet mellem flere af variablene, hvilket gør at man ikke umiddelbart kan teste betydningen af de enkelte variable

i forhold til ledstillingsvariationen i en multivariat statistisk analyse, som fx logistisk regression. Vi har derfor i stedet anvendt metoden *conditional inference tree*, som på en anden måde kan give et indblik i hvilke variable der er vigtigst i forhold til ledstillingsvariation, og hvordan de interagerer med hinanden, uden i nær samme omfang at være følsom for kollinearitet (Hothorn, Hornik & Zeileis 2006; Tagliamonte & Baayen 2012). En conditional inference tree-analyse fungerer groft sagt således at datamaterialet opsplittes i to dele på basis af den af de inddragede uafhængige variable der har den største indflydelse på den afhængige variabel (som her er ledstillingsmønsteret deklarativ vs. neutral). Herefter opsplittes de to dele af datamaterialet hver især igen på basis af den variabel der har den største indflydelse, og sådan fortsætter opsplitningen rekursivt indtil ingen af de uafhængige variable har signifikant indflydelse på den afhængige variabel i den pågældende del af datamaterialet. Figur 3 viser resultatet af analysen, som blev gennemført ved hjælp af *ctree*-funktionen i pakken *party* i R (Hothorn m.fl. 2023; Team 2023).



Figur 3: Conditional inference tree-analyse. VA: deklarativ ledstilling; AV: neutral ledstilling

Som det fremgår, er den variabel der har størst indflydelse på ledstillingsvariationen, ledsætningens syntaktiske funktion, og analysen opsplitter datamaterialet i ledsætninger i adjektiv- og adverbiums syntagmer samt ikke-relative adled i nominalsyntagmer på den ene side (kendetegnet ved en samlet set høj andel af deklarativ ledstilling) og de øvrige typer på den anden side. Dette er ikke overraskende set i relation til distributionen i Figur 2; mere interessant er det at konjunktionen kun spiller en signifikant rolle for ledstillingen i den førstnævnte gruppe (NOther+AdjAdv). Her er konjunktionen afgørende for den efterfølgende opsplitning som først opdeler data i grupperne *som*+Hv+Other vs. 0+*at*, og herefter opdeler den førstnævnte gruppe i hhv. *som*- og Hv-sætninger (med en meget lav andel af deklarativ ledstilling; ca. 5 %) og Other-sætninger (med en høj andel; ca. 35 %).

Inden for de øvrige opsplitninger er det herefter adverbialtypen der er afgørende for ledstillingen: Ledsætninger med et dialogisk adverbial har en større tendens til deklarativ ledstilling end ledsætninger med andre adverbialer, svarende til resultaterne af vores tidligere undersøgelser. Den eneste øvrige variabel der herefter har signifikant indflydelse på ledstillingen, er faktualitet, men kun inden for sætninger med de syntaktiske funktioner NOther eller AdjAdv indledt med *at* eller uden konjunktion og med et ikke-dialogisk sætningsadverbial. Her ses det, i overensstemmelse med assertivitetshypotesen, at de non-faktuelle ledsætninger har en væsentlig lavere tilbøjelighed til deklarativ ledstilling end de faktuelle.

Inferenstræsanalysen viser at der ikke er flere faktorer der har signifikant indflydelse på ledstillingen; position, længde, integration og disfluens giver således ikke anledning til yderligere opsplitninger af datamaterialet.³

3.1. Kvalitative analyser

For at få en dybere forståelse for hvad der kan forklare ledstillingsalternationen, ser vi i dette afsnit nærmere på hvad der gemmer sig under nogle af de inddelinger som inferenstræet resulterer i. Det viser sig at særligt ét mønster går igen, nemlig det at attributive ledsætninger med deklarativ

3 Hvis kun de hypotesebaserede variable inkluderes i analysen (dvs. ikke funktion og konjunktion), bliver også længde og integration signifikante: Længere ledsætninger har større tendens til deklarativ ledstilling end kortere ledsætninger, og integrerede ledsætninger har større tendens til deklarativ ledstilling end ikke-integrerede (det sidste i modsætning til vores hypotese).

ledstilling som hovedregel er styret af et såkaldt eksplikat i oversætningen, dvs. et element der henviser kataforisk til ledsætningen (Hansen & Heltoft 2011: 1506). De følgende underafsnit fokuserer på inferenstræets forgreninger til venstre, dvs. vi kommer ikke ind på relativsætninger, kløvningsledsætninger mv., som er kendetegnet ved lav andel af deklarativ ledstilling. For nemheds skyld refererer vi til de relevante søjler (nederst i Figur 3) som boks 4-10 (eng. *node 4-10*).

3.1.1. Ledsætninger med gradsadverbiet *så* som eksplikat

Nederst til venstre i inferenstræet ses en forgrening der udskiller konjunktionerne *som/hv-* i boks 4 og *Other* i boks 5. Vi fokuserer her på boks 5, der har relativt høj tendens til deklarativ ledstilling (ca. 35 %). Under samlebetegnelsen *Other* gemmer sig konjunktionerne *da, end, når, om og så*, men det er kun *end* og *så* der forekommer med deklarativ ledstilling (henholdsvis 3 og 13 forekomster). Med andre ord bæres tendensen her af ledsætninger med konjunktionen *så*. Inspicerer man disse nærmere, viser det sig at samtlige har et eksplikat i form af gradsadverbiet *så*, der typisk modificerer et adjektiv, se (6-8).⁴

- (6) hendes område er **så** specielt så hun bliver aldrig arbejdsløs
- (7) der var **så** meget oppakning på så jeg sad faktisk *s-* spændt fast ikke
- (8) nu er de selvfølgelig **så** store så de kan godt være alene hjemme en aften ikke også

Kombinationen af eksplikatet *så* i oversætningen og deklarativ ledstilling i ledsætningen findes også uden for boks 5, nemlig i forekomster med konjunktionen *at* (og en enkelt med *som* (ikke vist her)). De fordeler sig på boks 4, 7, 9 og 10, se eksempler i (9-11).⁵

- (9) nu var man også **så** lille at man oplevede det jo på en helt anden måde ikke
- (10) det var forandret **så** meget at at han ville hellere tilbage til det # gode gamle Nyboder
- (11) hun er så nok **så** gammel at jeg ved ikke om hun kan kende mig

4 Blandt forekomsterne i boks 5 findes også ledsætninger med neutral ledstilling der er styret af gradsadverbiet *så*, fx *eller også har det været så normalt så jeg ikke har lagt mærke til det*.

5 Igen findes der også eksempler med neutral ledstilling, fx *der kan vi jo selvfølgelig godt blive så gamle og affældige at at # vi ikke har kræfter til det*.

GDS skriver at ledsætninger til eksplikateret *så* ”har værdi af følgesætning” (Hansen & Heltoft 2011: 1508), hvilket harmonerer fint med at der ikke umiddelbart er nogen betydningsforskel på om man anvender konjunktionen *så* eller *at* i denne kontekst. Der er dog forskel på tendensen til deklarativ ledstilling idet *så*-sætninger har deklarativ ledstilling i 68 % af tilfældene mod *at*-sætningernes 42 % (når man ser på alle 53 ledsætninger med eksplikateret *så*, har de deklarativ ledstilling i 45 % af tilfældene).

Boks 7, 9 og 10 indeholder også attributive ledsætninger der styres af andre eksplikater end gradsadverbiet *så*. Dem gennemgår vi herunder, organiseret efter hyppighed.

3.1.2. Ledsætninger med adverbiet *sådan* som eksplikater

Det mest frekvente eksplikater i disse ledsætninger er adverbiet *sådan*. Det forekommer 115 gange, hvoraf 80 af tilfældene er med deklarativ ledstilling, svarende til 69 %.⁶ Heraf kan en undergruppe på 25 forekomster diskuteres ift. deres status som ledsætninger idet *sådan* viser sig at fungere som citatmarkør (eng. ’quotative’; Rathje 2009), se (12), som både indeholder et interjektional og udviser deiksisskift ift. person og tempus.

(12) jeg var **sådan** jamen så kan I da ikke tage ham med

De resterende forekomster adskiller sig på forskellig vis fra citatfunktion, enten ved ikke at have interjektionaler eller deiksisskift, eller ved at taleren bruger underordningskonjunktion, se (13-14).

(13) vi har det **sådan** vi vil ikke have vores børn skal vokse op midt inde i København

(14) ja men det er jo **sådan** at en hest har jo en rytme ligesom en en en en anden sportsudøver ikke

Som vi ser det i (14), optræder eksplikateret *sådan* gerne som prædikativer efter et formelt subjekt, *det*.⁷ Det resulterer i nogle særdeles indholdslette

6 Der forekommer også i alt syv sætninger der ikke indledes med *at* eller *0* men med *hvad*, *hvor*, *som* eller *så* (dvs. i boks 4 og 5). Disse har med en enkelt undtagelse neutral ledstilling.

7 Vi tolker ikke disse *det*-er som foreløbige subjekter, for det ville kræve at (14) kunne omskrives således: *at en hest har en rytme ligesom en anden sportsudøver er jo sådan*. Dette svarer ikke indholdssyntaktisk til strukturen i (14), for der mangler et led der forklarer hvad der ligger i *sådan*.

oversættninger af strukturen *det VÆRE sådan* der har præsentativ funktion, dvs. angiver at den egentlige information kommer i ledsætningen. Oversættninger med denne struktur kombineres også hyppigt i materialet med ledsætninger med neutral ledstilling, se (15).

- (15) det er jo ikke **sådan** vi bare går og snakker heste heste heste heste
og sådan

At også neutral ledstilling optræder efter præsentative oversættninger, er helt som man ville vente det ud fra den grundantagelse at denne ledstilling er semantisk umarkeret og kan forekomme i alle kontekster (Christensen 2007: 111-112; Hansen & Heltoft 2011: 1679-1680).

3.1.3. Ledsætninger med *det/den* som *eksplikat*

Under kategorien NOther, dvs. attributive ledsætninger der er knyttet til et nominalt element, men ikke er relativitysætninger, finder vi to hovedtyper.⁸ I den ene består eksplikatet kun af pronomenet *det/den*, se (16), mens det i den anden er den bestemte artikel i et komplekst nominalsyntaxme (dvs. *det/den*), se (17).

- (16) så gjorde mine forældre også **det** at øh at hun skulle da ikke passe
mig

- (17) så havde jeg fået **den** fidus jeg skulle bare sige jeg ville være øh
militærnægter

Begge typer udviser en klar tendens til deklarativ ledstilling med hhv. 14 ud af 20 forekomster for pronomenet *det/den* og 16 ud af 23 for den bestemte artikel *den/det* (i begge tilfælde 70 %).⁹

Som vi så det for *sådan* ovenfor, er en del af oversættningerne med nominalt *det* som eksplikat meget indholdslette idet de udgøres af prædikativkonstruktioner med subjektet *det* eller *der*, dvs. med strukturen *det/der VÆRE det*, se (18-19).

- (18) det er så også **det** at på det tidspunkt sagde man faktisk Des endnu

8 Derudover er der en mere blandet restgruppe på 12 forekomster med *noget/nogle* eller et substantiv uden foranstillet bestemt artikel, hvoraf halvdelen har deklarativ ledstilling.

9 Disse konstruktioner findes også med konjunktionen *om* (dvs. i boks 5), i alt 3 forekomster, alle med neutral ledstilling.

- (19) der er jo **det** ved det at at det er jo en en en øh en bygningsmasse
der er i brug hele tiden ikke

Også i denne undergruppe ser vi altså en tendens til oversættelser med præsentativ funktion.

Den type hvor ekspliketet er bestemmer i et nominalsyntaxme, er typisk mere indholdstung, alene i kraft af det leksikalske indhold i syntagmets hovedord (jf. *den fidus* i (17)). Men også nominalsyntaxmerne kan være konstrueret så de fungerer kataforisk, som i udtrykket *på den måde*, hvoraf der i alt er seks, som alle har deklarativ ledstilling, se (20).

- (20) så blev han øh mekaniker ikke # hh øh på **den** måde at han skulle bare øh passe øh varmeapparater og kedler og sådan noget

Denne kvalitative gennemgang af mønstre og tendenser der går på tværs af nogle af inferenstræets forgreninger, inviterer til nogle mere overordnede overvejelser om vores undersøgelse og dens resultater. Dem tager vi hul på i diskussionen.

4. Diskussion og afrunding

Vi har ovenfor fremdraget en række eksempler på det GDS kalder eksplikatkonstruktion (Hansen & Heltoft 2011: 1508-1510). Om eksplikatkonstruktioner hvor ekspliketet kun består af pronomenet *det*, skriver GDS at de kun er mulige ”når den eksplikative ledsætning er præsupponeret og prædikatet altså er faktitiv” (Hansen & Heltoft 2011: 1509). Hansen og Heltoft giver bl.a. eksemplet i (21) som belæg.

- (21) **det** at du giver op, overrasker mig egentlig ikke

Selvom Hansen og Heltoft ikke omtaler topologiske forhold i denne sammenhæng, vil en karakteristik af ledsætningen som præsupponeret ifølge assertivitetshypotesen forudsige at den skal have neutral ledstilling. Det står i modstrid med vores resultater hvor altså 70 % af ledsætningerne til ekspliketet *det* har deklarativ ledstilling. Når man inspicerer prædikaterne i deres oversættelser, viser det sig da også at ingen af dem er faktive (de anvendte prædikater er *være, have, gøre, ske og prøve*), og ledsætningerne forekommer heller ikke at være præsupponerede. Tværtimod præsenterer de i konteksten ny og fremhævet information. Det harmonerer med den

nævnte tendens til at de undersøgte oversættelser er indholdslette og fungerer præsentativt.

Det leder os til det mulige kritikpunkt af vores undersøgelse at nogle af ledsætningerne måske ikke er egentlige ledsætninger, men enten helsætninger eller udtryk for det talesprogsfænomen som GDS kalder 'ytringsindbygning' (Hansen & Heltoft 2011: 1651-1657):

Det er et almindeligt træk ved talesprog at man indbygger en ny ytringshelhed ved at skifte til ny helsætningsstruktur inde i en allerede påbegyndt helsætning. ... Ofte finder denne indbygning sted som alternativ til en nominal ledsætning. (Hansen & Heltoft 2011: 1651)

Hansen og Heltoft understreger at der ikke kræves et anførende verbum for at der kan være tale om ytringsindbygning, men at samme funktion kan varetages af fx et kataforisk pronomen (Hansen & Heltoft 2011: 1655). Her begynder det at ligne eksplikatkonstruktion. Eftersom GDS opererer med "glidende overgang" (Hansen & Heltoft 2011: 1682) mellem ytringsindbygning og ledsætninger med deklarativ ledstilling, er det vanskeligt at afgøre hvor mange af vores ledsætninger der ifølge GDS burde kategoriseres som ytringsindbygning.

Grunden til at vi ikke har udgrænset ytringsindbygning fra ledsætninger i den kvantitative del af vores undersøgelse, er at vi ikke har fundet nogen generelle kriterier til udtømmende at afgrænse kategorien. I modsætning til i en beskrivelse som GDS' kan deklarativ ledstilling i en variationistisk undersøgelse som vores naturligvis ikke være et kriterium, da analysen så ville blive cirkulær.

Anderledes er det hvor der er selvstændige kriterier for at en sætning udgør ytringsindbygning. Som vi allerede har nævnt, er der en mindre gruppe af ledsætningerne i materialet som kan beskrives som direkte tale efter en anførende (citatermarkerende) brug af *sådan* fordi de indeholder interjektionaler eller deiksisskift. Det drejer sig om 25 forekomster. Det kan diskuteres om de burde være sorteret fra i undersøgelsen. Vi har fulgt det princip at vi ikke på forhånd frasorterer forekomster hvis de ellers lever op til kriterierne for hvad der tæller som en ledsætning. For attributive ledsætninger er et af kriterierne om der findes et element som de kan siges at lægge sig til. Eftersom *sådan* udgør en mulig 'vært' for et attribut, og eftersom anførende sætninger i øvrigt i traditionen er

blevet regnet for oversættninger (Diderichsen 1974: 203), er denne type sætninger medtaget i kodningen og i de kvantitative undersøgelser. Først en kvalitativ analyse kan vise hvilke forekomster der er direkte tale, og hvilke der ikke er.¹⁰

Vores mangeårige studie af ledstillingsalternationen i danske ledsætninger har bl.a. drejet sig om at undersøge hvorvidt funktionelle hypoteser kan forklare anvendelsen af deklarativ ledstilling i talesprogets ledsætninger. Det mener vi at dette delstudie af attributive ledsætninger giver solid støtte til. Som vist i de kvalitative analyser ser vi en udbredt tendens til præsentative oversættninger ved ledsætninger med deklarativ ledstilling, hvilket harmonerer med hypotesen om at deklarativ ledstilling signalerer forgrund. Samtidig ser vi de forventede effekter af dialogiske adverbialer og ledsætningens længde. Det er dog fortsat uafklaret præcis hvordan forgrund og assertivitet forholder sig til hinanden – også i dette studie finder vi at faktualitet er en statistisk signifikant variabel, som dog ser ud til at have relativ beskedne betydning. Den særlige rolle som underordningskonjunktionen *at* har, inviterer ligeledes til flere undersøgelser.

Om forfatterne

Torben Juel Jensen, ph.d., lektor i dansk sprog, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.

Tanya Karoli Christensen, ph.d., professor i dansk sprog, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.

Litteratur

Christensen, Marie Herget (2019): Sætningskløvnings i moderne talt dansk. *Danske talesprog* 19, 37-76.

Christensen, Marie Herget, Tanya Karoli Christensen & Torben Juel Jensen (2016): Forgrund af ledsætninger - et psykolingvistisk forsøg. *Ny Forskning i Grammatik* 23, 5-21.

¹⁰ Vores data er ikke opmærket sådan at man automatisk kan frasortere ledsætninger der udgør direkte tale eller tanke; det ville kræve en efterkategorisering. Sådan en foretog vi i forbindelse med undersøgelsen af nominale ledsætninger (Jensen & Christensen 2013: 45). Selvom vi ikke har gjort det her, vurderer vi på baggrund af de kvalitative analyser at kun en meget lille del af de attributive ledsætninger efter selvstændige, udtryksmæssigt baserede kriterier kan kategoriseres som ytringsindbygning.

- Christensen, Tanya Karoli (2007): *Hyperparadigmer – en undersøgelse af paradigmatiske samspil i danske modusystemer*. Ph.d.-afhandling, Roskilde Universitet.
- Christensen, Tanya Karoli & Torben Juel Jensen (2022): When Variants Lack Semantic Equivalence: Adverbial Subclause Word Order. I: *Explanations in Sociosyntactic Variation*. Tanya Karoli Christensen & Torben Juel Jensen (red.). Cambridge: Cambridge University Press, 171-206.
- Christensen, Tanya Karoli, Torben Juel Jensen & Marie Herget Christensen (2015): Adverbielle ledsætningers ledstilling i dansk talesprog. *MUDS* 15, 97-115.
- Diderichsen, Paul (1974): *Elementær dansk grammatik* (3. udg.). København: Gyldendal.
- Hansen, Erik & Lars Heltoft (2011): *Grammatik over det Danske Sprog*. København: Det Danske Sprog- og Litteraturselskab.
- Hothorn, Torsten, Kurt Hornik, Carolin Strobl & Achim Zeileis (2023). *party - A Laboratory for Recursive Partytioning version 1.3-14*. Tilgængelig på <http://party.R-forge.R-project.org>
- Hothorn, Torsten, Kurt Hornik & Achim Zeileis (2006): Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15(3), 651-674. <https://doi.org/10.1198/106186006x133933>
- Jensen, Torben Juel (2011): Ordstilling i ledsætninger i moderne dansk talesprog. *Ny Forskning i Grammatik* 18, 123-150.
- Jensen, Torben Juel & Tanya Karoli Christensen (2013): Promoting the demoted: The distribution and semantics of “main clause word order” in spoken Danish complement clauses. *Lingua* 137, 38-58. <https://doi.org/10.1016/j.lingua.2013.08.005>
- Julien, Marit (2007): Embedded V2 in Norwegian and Swedish. *Working papers in scandinavian syntax* 80, 103-161.
- Julien, Marit (2015): The force of V2 revisited. *The Journal of Comparative Germanic Linguistics* 18(2), 139-181. <https://doi.org/10.1007/s10828-015-9073-2>
- Mortensen, Sune Sønderberg (2014): Attributive clauses in Danish – and the relative irrelevance of relative clauses. *Nordic Journal of Linguistics* 37(03), 393-430. <https://doi.org/10.1017/s0332586514000274>
- Rathje, Marianne (2009): Definition af citater i talesprog. *NyS - Nydanske Sprogstudier* 37, 125-158.

- Tagliamonte, Sali A. & R. Harald Baayen (2012): Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(02), 135-178. <https://doi.org/10.1017/s0954394512000129>
- Team, R Core (2023). *R: A language and environment for statistical computing*. Tilgængelig på <https://www.R-project.org/>
- Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson & Thorbjörg Hróarsdóttir (2009): On the distribution and illocution of V2 in Scandinavian that-clauses. *Lingua* 119, 1914-1938. <https://doi.org/10.1016/j.lingua.2009.03.006>

Kasus i hypotagmer. En undersøgelse af flertalspronomenene i dansk

Oliver Rix Johannsen

Abstract

This paper investigates the distribution of case forms in two types of pronominal noun phrases: pronouns with a restrictive relative clause (e.g. *de/dem der har billet* ‘they/them who have a ticket’) and pronouns with an adverbial complement (e.g. *de/dem på tredje sal* ‘they/them on the third floor’) when functioning as subject, object, prepositional complement and predicative. The aim of the article is to 1) give a description of case distribution in more informal text than in previous studies, 2) identify factors influencing the distribution, and 3) discuss whether the proposed ‘common member factor’ can satisfactorily explain discrepancies from the norm.

Nøgleord

personlige pronomen, adverbial, relativsætning, register, skriftsprog, sociale medier

1. Indledning

Pronominalkasus er den sidste rest af kasusbøjning i dansk og angår kun otte pronomen: *jeg – mig, du – dig, han – ham, hun – hende, vi – os, I – jer, de – dem* og *man – en*. Alligevel volder det der er tilbage, en del sprogrigtighedsproblemer.

Ifølge det der er kendt som ‘hovedreglen’ (Diderichsen 1957: 108 f., Hansen 1972: 99, Hansen & Heltoft 2019: 438 m.fl.), anvendes nominativformen i subjekt, fx *jeg ser Albert*, og den oblique form i alle andre tilfælde, fx *Albert ser mig* (objekt); *Albert ser på mig* (styrelse); *Albert, det er mig* (prædikativ). Men da det er en hovedregel, findes der også tilfælde som hovedreglen ikke indfanger. Det er bl.a. i hypotagmer hvor pronomenet

ledsages af yderligere materiale, fx *Dem der skriver sådan her, påkalder sig sprogtørternes vrede*, *Tak til de der gør en indsats for vores sikkerhed* (restriktiv relativsætning) eller *Ham på fjerde sal ser godt ud* (adverbial). Der findes altså både nominativ og oblik hvor man if. hovedreglen ville vente det modsatte. Derfor er der både formuleret justeringer af hovedreglen og nye regler der søger at indfange denne brug.

Pronominalkasus er behandlet indgående i den grammatiske litteratur, ligesom der er fremlagt flere deskriptive regler der forsøger at redegøre for distributionen af nominativ og oblik form. Men "[f]orholdene ved 1p.pl. og 2p.pl. er generelt underbelyst. Der gøres observationer, men man mangler mere tilbundsgående undersøgelser, og man mangler dem for de enkelte konstruktionstypers vedkommende" (Jensen 2021: 64). Desuden er det hidtil undersøgte materiale begrænset til transskriptioner af sociolingvistiske interviews og af avistekster. Derfor vil jeg give en sådan tilbundsgående undersøgelse og det af andre typer tekster i de følgende konstruktioner:

- Pronomen med restriktiv relativsætning, fx
 - o *vi/os der har billet*
 - o *I/jer der har billet*
 - o *de/dem der har billet*
- Pronomen med specificerende adverbial, fx
 - o *vi/os på tredje sal*
 - o *I/jer på tredje sal*
 - o *de/dem på tredje sal*

Undersøgelsen dækker ikke kun første og anden person pluralis, men også tredje person, der opviser særlige metodiske problemer, og det i alle led-funktioner: subjekt, objekt, prædikat og styrelse for præposition. Desuden begrænser jeg undersøgelsen til at dække konstruktioner der består af én og kun én type bestemmende led. De spørgsmål jeg vil besvare med undersøgelsen, er som følger:

1. Hvordan er usus i andre typer data end de allerede undersøgte? Er der relevante forskelle?
2. Hvilke faktorer har indflydelse på distributionen af nominativ og oblik form?
3. Hvilke af de foreslåede regler kan forklare afvigelser fra hovedreglen?

I det følgende afsnit 2 vil jeg først beskrive undersøgelsens metode og den empiri der ligger til grund for undersøgelsen. I afsnit 3 præsenterer jeg nogle af de fremsatte forslag til forklaring af afvigelser fra hovedreglen, herunder 'fællesledsreglen' (Petersen, Schack & Jensen 2020: 109). Hovedvægten er lagt på analysen i afsnit 4. I afsnit 5 diskuterer jeg undersøgelsens resultater ift. litteraturen samt hvilke faktorer der har betydning for distributionen af kasus. I det afsluttende afsnit 6 samler jeg op og vurderer i hvor høj grad den såkaldte fællesledsregel kan forklare afvigelserne fra hovedreglen.

2. Empiri og metode

Korpusset jeg undersøger, består af en samling facebooktekster (opslag og kommentarer fra perioden 2015-2017) indsamlet af analysevirksomheden Analyse & Tal. Korpusset rummer ca. 12.000.000 *tokens* (dvs. ord, smileyer mv.). Korpusset dækker mere spontant skriftligt materiale end det hidtil undersøgte materiale, der fortrinsvis dækker avistekster og derfor kan tænkes at være underlagt (strengere) krav til korrektur (Petersen, Schack & Jensen 2020: 121). I eksempler fra korpusset gengivet i artiklen bibeholdes den oprindelige ordlyd og tegnsætning.

Undersøgelsens metode er todelt: Den består dels af en indledende automatisk databehandling, dels af manuel annotation og analyse. Til databehandlingen har jeg udviklet et pythonscript der anvender sprogteknologiske værktøjer såsom ordklassetagging og parsing snarere end simple søgestrengte, da sådanne giver for mange irrelevante resultater. Det er især konstruktionen {PRONOMEN + RELATIVSÆTNING} der volder problemer, dels pga. homografi mellem pronomenet *de* og artiklen *de*, dels pga. relativindlederne¹ *der* og \emptyset . Ved at anvende simple søgestrengte får man ved at kombinere et pronomen med relativindlederen *der* ikke kun relativsætninger, men også pronomen brugt demonstrativt, fx *Ræk mig de der æbler*. Endnu værre bliver det hvis man vil finde relativsætninger med \emptyset -indleder; her er man forvist til kun at søge efter pronomenet da det kan følges af alskens materiale. I tredje person pluralis nominativ (fx *De hun ikke kan lide, kommer om lidt*) falder dette problem endvidere sammen med homografiproblemet da man her kun kan søge efter *de*. Det har hidtil gjort det praktisk talt umuligt at undersøge konstruktionen {PRONOMEN +

1 "Relativindleder" er en praktisk terminologi jeg her indfører og anvender om det initiale element i en relativsætning, dvs. subjektvikaren *der*, subjunktionen *som*, \emptyset (når *som* er udeladt) og relativpronomenene.

RELATIVSÆTNING} med Ø-indleder systematisk (Jørgensen 2000: 188), om end der er taget tilløb hertil (Petersen, Schack og Jensen 2020: 112). Parsing kan imidlertid effektivisere sådanne søgninger: Med denne metode ekstraheres ikke blot tekststrengene (dvs. de 'nøgne' pronomener), men pronomener med bestemte syntaktiske egenskaber, in casu pronomener med restriktive relativsætninger.

Scriptet er udviklet med afsæt i pythonbiblioteket spaCy der rummer statistiske modeller for dansk (Explosion 2022), alle baseret på *The Danish UD treebank* (Johannsen, Alonso & Plank 2015). Jeg anvender den største model, hvor ordklassetaggeren har en oplyst nøjagtighed på 98 % og parseren en nøjagtighed på 84-86 % (Explosion 2022).

3. Forslag til forklaring

Det der er kendt som 'hovedreglen' for pronominalkasus i moderne dansk, er rent syntaktisk defineret: Nominativ anvendes i subjekt; i alle andre andre funktioner anvendes oblik kasus (se fx Diderichsen 1957: 108 f., Hansen 1972: 99, Hansen & Heltoft 2019: 438). Det er imidlertid velkendt at hovedreglen udfordres eller helt sættes ud af spil når pronomenet ledsages af yderligere materiale (se fx Hansen 1967: 242 ff., Jacobsen & Jørgensen 2013 123 ff., Becker-Christensen 2010: 207 f., Jensen 2019: 79, Petersen, Schack & Jensen 2020: 108). Derfor er der formuleret korrektioner til hovedreglen, fx for obligatorisk trykstærke (Hansen 1972) eller ikke-anaforiske pronomener (Jørgensen 2000: 312, Hansen & Heltoft 2019: 438 f.) eller pronomener med ikke-tydelig ledfunktion (Diderichsen 1957: 109, Schack 2013). Der er også fremsat forklaringer af særlige forhold: Navnlig nominativ i konstruktionen {PRONOMEN + RELATIVSÆTNING} uden for subjekt har påkaldt sig opmærksomhed (jf. afsnit 3.1 nedenfor). Det er også foreslået at oblik kasus er en slags defaultkasus der kan anvendes når hovedreglen udfordres af pronomener med yderligere materiale eller i særlige syntaktiske konstruktioner (Brink & Lund 1975: 663 f., Heltoft 1997: 233, Schack 2013: 67 f.).

3.1. Fællesledsreglen

Fællesledsreglen (Petersen, Schack & Jensen 2020: 109) er en særregel der specifikt angår {PRONOMEN + RELATIVSÆTNING}, fx *Tak til de der gør en indsats*. Som navnet på reglen antyder, er fællesleddet centralt; fællesleddet er den indholdsstørrelse som relativsætningen har til fælles med sin oversætning (Hansen & Heltoft 2019: 1481). I eksemplet ovenfor er

fællesleddet kun repræsenteret i oversætningen af *de*, mens det ikke er udtrykt i relativsætningen. Her er der tale om såkaldt **attraktion**². Med attraktion menes der at korrelatet (*de* i det anførte eksempel) tager den kasus som ventes af fællesleddets funktion i relativsætningen, frem for den kasus som ventes af dets funktion i oversætningen. If. hovedreglen ville man have ventet oblik *dem*, men fællesleddets funktion som subjekt i relativsætningen motiverer derimod nominativ *de*: ”The function within the relative clause, as it were, supersedes the primary rule” (Jensen, Schack & Petersen 2023: 194).

Tendensen er først beskrevet af Diderichsen som bisætningsattraktion (1941: 140 f.) og senere relativattraktion (1957: 278) og afprøvet for både gammeldansks (1941: 140 f.) og moderne dansks vedkommende (Jørgensen 2000: 188). I en række nyere artikler (Schack, Jensen & Petersen 2019, Petersen, Schack & Jensen 2020, Jørgensen 2021, Jensen, Schack & Petersen 2023) har fællesleddsreglen fået fornyet aktualitet, hvorfor jeg i det følgende efterprøver reglen.

4. Resultater

4.1. Pronomen med restriktiv relativsætning

I det følgende præsenterer jeg resultaterne for konstruktionen {PRONOMEN + RELATIVSÆTNING}, fx *Dem der gør sig umage, bliver belønnet derfor*. Jeg behandler hvert af pronomenerne efter tur, hhv. *vi – os* i afsnit 4.1.1, *I – jer* i afsnit 4.1.2 og *de – dem* i afsnit 4.1.3. De samlede resultater fremgår af tabel 1.

2 Kasusattraktion er et begreb der traditionelt anvendes ifm. de klassiske kasussprog latin og oldgræsk, hvor kasusreaktion er udbredt, i modsætning til moderne dansk.

	1. person		2. person		3. person	
	nominativ	oblik	nominativ	oblik	nominativ	oblik
	<i>vi</i>	<i>os</i>	<i>I</i>	<i>jer</i>	<i>de</i>	<i>dem</i>
Ledfunktion i oversætning						
subjekt	264 (43,3 %)	346 (56,7 %)	23 (15,3 %)	127 (84,7 %)	1580 (28,1 %)	4039 (71,9 %)
objekt	2 (2,5 %)	77 (97,5 %)		33 (100 %)	266 (9 %)	2675 (91 %)
styrelse	2 (22,2 %)	7 (77,8 %)			77 (43 %)	102 (57 %)
prædikativ				1 (100 %)	28 (47,5 %)	31 (52,5 %)
Ledfunktion i relativsætning						
subjekt	268 (38,5 %)	430 (61,5 %)	23 (14 %)	160 (86 %)	1919 (23 %)	6419 (77 %)
objekt				1 (100 %)	26 (7,5 %)	323 (92,5 %)
styrelse					6 (5,6 %)	101 (94,4 %)
prædikativ						
Kombination af subjektsforekomst og ikke-subjektsforekomst						
1*	264 (98,5 %)	346 (80,5 %)	23 (100 %)	126 (78,3 %)	1555 (79,7 %)	3806 (55,6 %)
2*				1 (0,6 %)	25 (1,3 %)	233 (3,4 %)
3*	4 (1,5 %)	84 (19,5 %)		34 (21,1 %)	368 (18,9 %)	2613 (38,2 %)
4*					3 (0,2 %)	195 (2,8 %)
Forekomster pr. kasusform	268	430	23	161	1951	6847
Forekomster pr. person	698		184		8798	
Forekomster i alt	9680					

Tabel 1: Resultater for pronomen med restriktiv relativsætning

Tabellen er inddelt i tre felter: *ledfunktion i oversætning*, *ledfunktion i relativsætning* og *kombination af subjektsforekomst og ikke-subjektsforekomst*. De fire mulige kombinationer af subjektsforekomst og ikke-subjektsforekomst i oversætning og relativsætning for fællesleddet (efter Jørgensen 2000: 184) er:

- 1* Subjekt i oversætning og subjekt i relativsætning
- 2* Subjekt i oversætning og ikke-subjekt i relativsætning
- 3* Ikke-subjekt i oversætning og subjekt i relativsætning
- 4* Ikke-subjekt i oversætning og ikke-subjekt i relativsætning

If. hovedreglen (jf. afsnit 3) ventes nominativ i 1* og 2* samt oblik i 3* og 4* mens fællesledsreglen (jf. afsnit 3.1) fordrer nominativ i 1* og 3* samt oblik i 2* og 4*.

4.1.1. *vi* – *os*

I første person pluralis forekommer nominativ (1) omtrent lige så ofte som oblik (2) i subjekt, hhv. 43,3 % og 56,7 % (jf. tabel 1). Forskellen er altså ikke statistisk signifikant³:

- (1) *Hvorfor skal vi der aldrig går i den Kongelige teater være med til at betale til det?*
- (2) *Os der arbejder i DK, betaler skat i DK*

Der er fire forekomster af nominativ uden for subjekt:

- (3) *Tænk på hvornår Danmark kommer på rette spor og begynder bl.a. at tilgodesee vi der har bidraget til velfærd*
- (4) *lad nu vi som bor her selv bestemme*
- (5) *de bliver jo ikke som, vi der er børn af 50-erne 60-erne, opdraget til at vise hensyn*
- (6) *„lær at blive mennesker „som vi der har (nogle) ansat jer,*

I (3) og (4) er {*vi* + RELATIVSÆTNING} direkte objekt, mens fællesledet *vi* i relativsætningen fungerer som subjekt. Her kan der være tale om attraktion.

I (5) og (6) udgør {*vi* + RELATIVSÆTNING} sammenligningsleddet i et sammenligningssyntagma⁴ indledt af *som*. Netop konstruktioner af typen {*som/ end* + PRONOMEN + RELATIVSÆTNING} er stadig genstand for

3 $p > 0,05$ ved simpel binominaltest (forventet sandsynlighed = 0,5) for nominativ og oblik.

4 Det er blevet foreslået at *som* her snarere fungerer som præposition (Mikkelsen 1975: 238, Diderichsen 1957: 109, Hansen 1967: 243, Hansen & Heltoft 2019: 442 f., Schack 2013 m.fl.). Derfor annoterer jeg sammenligningsled som styrelser; (5) og (6) opviser de eneste to eksempler på sammenligningssyntagmaer i første person. I anden person finder jeg ingen, og i tredje person finder jeg 10 eksempler, jf. afsnit 4.1.3.

diskussion: ”Her kræver en konservativ norm nominativ (...) Argumentet er at der foreligger en elliptisk sætning” (Hansen & Heltoft 2019: 442), fx *vi er som de [er]*. Derfor kan nominativen *vi* ikke uden videre tilskrives fællesledsreglen da der her muligvis er en ”konservativ norm” der påvirker kasusvalget. Af de 430 tilfælde af oblik kan de 84 (oblik i 3*) forklares med hovedreglen (og ikke fællesledsreglen), og de resterende 346 (oblik i 1*) hverken af hovedreglen eller fællesledsreglen (jf. tabel 1).

4.1.2. *I – jer*

I anden person finder jeg oftere oblik form i subjekt end i første person. Af alle forekomster i subjekt udgør *jer* de 127 forekomster (84,7 %, jf. tabel 1), mens nominativ kun udgør de 23 (15,3 %, jf. tabel 1). Langt de fleste konstruktioner er således af typen:

(7) *Jer der har stemt på dem har jo selv valgt DET*

Jeg finder ingen forekomster af *I* uden for subjekt, så de 23 forekomster er alle kasusvalg af typen nominativ i 1* i overensstemmelse med hovedreglen, fx

(8) *I der ved, hvad I stemmer behøver ikke (igen), at fortælle mig, at I stemmer ja*

Jeg finder én ytring (0,6 %) som kan forklares af fællesledsreglen:

(9) *Jer som jeg opfatter som seriøse politikere burde ikke ”ride med på den populistiske kynisme”.*

Her er fællesleddet *jer* subjekt for verbalet *burde* i oversættningen og direkte objekt for *opfatter* i relativsætningen.

34 af de oblikke forekomster kan kun forklares af hovedreglen (oblik i 3*, jf. tabel 1), mens 126 tilfælde hverken kan forklares af hovedreglen eller fællesledsreglen (oblik i 1*, jf. tabel 1).

4.1.3. *de – dem*

I tredje person pluralis er der i subjekt en klar præference for oblik form (4040 af 5620 tilfælde, 71,9 %, jf. tabel 1), fx

(10) *Dem der bakker tåbelighed op, gør sig til medskyldige.*

(11) *Dem vi kalder De Unge KAN, fordi de VIL <3*

I (10) fungerer fællesleddet *dem* som subjekt i både oversætningen og relativsætningen (oblik i 1*), mens fællesleddet i (11) er subjekt i oversætningen, men direkte objekt i relativsætningen (oblik 2*). Den type kasusvalg som (10) eksemplificerer, udgør 3806 af 6847 kasusvalgene (55,6 %) for *dem* mens den type (11) eksemplificerer, udgør 233 eller 3,4 % af kasusvalgene (jf. tabel 1). Den sidstnævnte type kan forklares af fællesledsreglen mens den første type ikke kan; i (10) kan der ikke være tale om attraktion.

Jeg finder dog også en del forekomster af nominativ i subjekt (1580 ud af 5620 tilfælde, 28,1 %), fx

(12) *De der skrev Molbo-historierne ville falde om af grin, hvis de vidste, hvordan en såkaldt borgerlig regering laver politik*

I (12) er fællesleddet *de* subjekt i både oversætningen og relativsætningen (nominativ i 1*). Dette gælder for de 1555 kasusvalg af de 1580 hvor fællesleddet er subjekt i oversætningen, hvad der svarer til 79,7 % af kasusvalgene for *de*. Denne type kan både hovedreglen og fællesledsreglen forklare. De resterende 25 kasusvalg i subjekt (1,3 %, jf. tabel 1) er således af typen nominativ i 2*, der kun kan forklares af hovedreglen, se fx (13).

(13) *Nu har jeg ikke læst alle kommentarer, men de jeg har læst, er totalt skræmmende menneskesyn.*

I funktion som objekt (direkte og indirekte) anvendes både *de* og *dem*. Dog er *dem* i overensstemmelse med hovedreglen langt det hyppigste med 2675 forekomster ud af 2941 (91 %, jf. tabel 1), fx

(14) *Tænk engang! Jeg skal både have nyt fjernsyn og kysse dem jeg holder af!*

Men nominativ anvendes også i objekter, og af de 266 forekomster har fællesleddet i de 263 tilfælde funktion som subjekt i relativsætningen og kan dermed forklares med fællesledsreglen. Det er altså kasusvalg af typen nominativ i 3* og udgør 98,9 % af alle kasusvalg for *de* i funktion som objekt, fx

(15) *De der er flygtninge skal vi selvfølgelig hjælpe (...)*

Jeg finder imidlertid tre eksempler på konstruktionen {*de* + RELATIVSÆTNING} i funktion som objekt der ikke kan forklares med attraktion, fx

(16) *(...) det gælder også de vi ikke kan lide.*

Det svarer til 3 kasusvalg ud af 266 for *de* i funktion som objekt (1,1 %) eller 3 ud af 1951 kasusvalg for *de* i alle funktioner (0,2 %). Det er således evidens for at konstruktioner af typen {*de* + RELATIVSÆTNING} hvor fællesleddet hverken fungerer som subjekt i oversætningen eller relativsætningen (nominativ i 4*), eksisterer, men ikke desto mindre er yderst sjældne.

Med funktion som styrelse for præposition er konstruktionen {PRONOMEN + RELATIVSÆTNING} mindre hyppig. I korpuset finder jeg 77 forekomster med nominativ (43 %) og 102 med oblik form (57 %, jf. tabel 1), så jeg finder i modsætning til konstruktioner med funktion som subjekt eller objekt her ingen signifikant⁵ præference for hverken nominativ eller oblik form. Oblik og nominativ forekommer altså omtrent lige hyppigt:

(17) *Dem som får støtte til kunst osv kunne man fjerne skatten fra*

(18) *LYT DOG TIL, DE DER VED NOGET OM TINGENE OG HAR BEVARET MEDMENNESKELIGHEDEN <3*

Eksempel (18) opviser hvad der er karakteristisk for samtlige 77 forekomster af *de* i styrelser: Fællesleddet *de* fungerer som subjekt i relativsætningen, altså nominativ i 3*, hvad fællesledsreglen kan forklare. Typen som (17) opviser, kan forklares med hovedreglen.

Endnu mindre hyppig er {PRONOMEN + RELATIVSÆTNING} med prædikativ funktion. Jeg finder 28 forekomster med nominativ (47,5 %) og 31 med oblik (52,5 %, jf. tabel 1). Her er der som i styrelser ingen signifikant⁶ præference for hverken nominativ eller oblik form:

5 $p > 0,05$ ved simpel binominaltest (forventet sandsynlighed = 0,5) for nominativ og oblik.

6 $p > 0,05$ ved simpel binominaltest (forventet sandsynlighed = 0,5) for nominativ og oblik.

- (19) *Nu er vi sygeplejersker de der observerer patienterne, ikke lægerne (...)*
 (20) *(...) de ældre er dem vi skylder tak for det liv vi har*

4.1.4. *Relativindlederens betydning for kasusvalget*

Vi kommer nu til relativindledernes distribution, og de fordeler sig således:

Relativindleder	Antal forekomster
<i>der</i>	7041 (80 %)
<i>som</i>	1307 (14,9 %)
Ø	439 (5 %)
<i>som der</i>	7 (0,1 %)
<i>hvis</i>	1 (< 0,1 %)
<i>hvor</i>	3 (< 0,1 %)

Tabel 2: Distribution af relativindledere

Relativindlederen *der* er altså langt den hyppigst forekommende, hvad der er forventeligt eftersom relativsætninger hvor fællesleddet er subjekt i relativsætningen også er langt det hyppigste (jf. tabel 1). Selvom *som* også kan anvendes her, foretrækkes der; *som* tegner sig for knap 15 % (1307 forekomster) af det totale antal forekomster, hvoraf 1285 er subjekt i relativsætningen (jf. tabel 3). Dernæst er der Ø-indleder i 5 % af forekomsterne, mens *som der* udgør blot 0,1 % af forekomsterne og *hvis* og *hvor* hver udgør mindre end 0,1 %. Jeg finder endvidere kun eksempler på konstruktionen {PRONOMEN + RELATIVSÆTNING} med Ø-indleder i tredje person pluralis. Da det i litteraturen (se fx Wivel 1901: 248; Mikkelsen 1975: 238; Jørgensen 2000: 188 og Jacobsen & Jørgensen 2013: 126) er blevet foreslået at (fra-)valget af relativindleder skulle have betydning for kasusvalget, fremlægger jeg derfor distributionen af kasusformerne i tredje person⁷ særskilt for relativindlederne *der*, *som* og Ø⁸. Resultaterne fremgår af tabel 3 nedenfor.

7 I første og anden person finder jeg kun forekomster med *der* og *som* som relativindleder, så første og anden person udgår.

8 Antallet af forekomster med *hvis*, *hvor* og *som der* som relativindleder er så lavt at de udgår.

	<i>der</i>		<i>som</i>		Ø	
	<i>de</i>	<i>dem</i>	<i>de</i>	<i>dem</i>	<i>de</i>	<i>dem</i>
Ledfunktion i oversætning						
subjekt	1290 (29 %)	3160 (71 %)	263 (28,7 %)	652 (71,3 %)	26 (10,4 %)	224 (89,6 %)
objekt	199 (8,2 %)	2234 (91,8 %)	64 (19,2 %)	269 (80,8 %)	3 (1,7 %)	169 (98,3 %)
styrelse	63 (52,5 %)	57 (47,5 %)	14 (34,1 %)	27 (65,9 %)		14 (100 %)
prædikativ	23 (60,5 %)	15 (39,5 %)	5 (29,4 %)	12 (70,6 %)		4 (100 %)
Ledfunktion i relativsætning						
subjekt	1575 (22,4 %)	5466 (77,6 %)	341 (26,5 %)	944 (73,5 %)	2 (100 %)	
objekt			4 (23,5 %)	13 (76,5 %)	22 (6,6 %)	311 (93,4 %)
styrelse			1 (25 %)	3 (75 %)	5 (4,8 %)	99 (95,2 %)
prædikativ						
Kombination af subjektforekomst og ikke-subjektforekomst						
1*	1290 (81,9 %)	3160 (57,8 %)	259 (74,9 %)	643 (67 %)		
2*			4 (1,2 %)	9 (1 %)	26 (89,7 %)	224 (54,5 %)
3*	285 (18,1 %)	2306 (43,2 %)	82 (23,7 %)	301 (31,3 %)		
4*			1 (0,3 %)	7 (0,7 %)	3 (10,3 %)	186 (45,3 %)
Forekomster pr. kasusform	1575	5466	346	960	29	410
Forekomster pr. person	7041		1306		439	
Forekomster i alt	8786					

Tabel 3: Kasusvalg for 3p.pl. med relativsætning efter relativindleder

Relativindlederen *der* anvendes kun hvor fællesleddet har funktion som subjekt i relivsætningen. Mellem *der* og *som* er der ingen synderlig forskel mht. præference for oblik form i subjekt, hhv. 71 % og 71,3 %. For Ø-indleder er der som forventet en stærkere præference for oblik form i subjekt: 89,6 % – men overraskende nok finder jeg ikke udelukkende oblik form

her (jf. tabel 3). Typen er allerede eksemplificeret i (13), men for overblikket skyld giver jeg yderligere et her:

(21) *De jeg har kendskab til kæmper en daglig kamp for at holde sig gående*

Dette er i stærk modsætning til Jacobsen og Jørgensen (2013: 126, min fremhævnings), der angiver at nominativ slet ikke er mulig:

Tendensen til at bruge akkusativ er særlig stærk i de tilfælde hvor det relative pronomen er andet end subjekt i relativsætningen. Den kan være så stærk at den næsten helt sætter hovedreglen ud af spil (...) **endnu værre bliver det hvis ordet som er udeladt. I den slags tilfælde er det kun akkusativ der kan bruges.**

En anden observation er at {*dem* + RELATIVSÆTNING} med Ø-indleder samlet set bidrager med flest tilfælde af oblik i 2* og 4*, nemlig 410; {*dem* + RELATIVSÆTNING} med *som*-indleder bidrager kun med 16 forekomster (jf. tabel 3). Fællesledsreglen fordrer netop oblik i 2* og 4*, så attraktion synes at virke stærkere (mht. hvor der ventes oblik form) hvor der anvendes Ø-indleder. Valget af relativindleder synes for så vidt at have betydning for kasusvalget idet Ø-indleder oftere kalder på oblik form end *der* og *som*.

4.2. *Pronomen med specificerende adverbial*

Resultaterne for konstruktionen {PRONOMEN + ADVERBIAL} fremgår af tabel 4 nedenfor:

	1. person		2. person		3. person	
	nominativ	oblik	nominativ	oblik	nominativ	oblik
	<i>vi</i>	<i>os</i>	<i>I</i>	<i>jer</i>	<i>de</i>	<i>dem</i>
subjekt	83 (94,3 %)	5 (5,7 %)	70 (92,1 %)	6 (7,9 %)	11 (17,2 %)	53 (82,8 %)
objekt				2 (100 %)		12 (100 %)
styrelse		10 (100 %)		6 (100 %)	1 (5 %)	19 (95 %)
prædikativ						8 (100 %)

Forekomster pr. kasusform	83	15	70	14	12	92
Forekomster pr. person	98		84		104	
Forekomster i alt			286			

Tabel 4: Resultater for pronomener med specificerende adverbial

Jeg finder 286 eksempler på konstruktionen {PRONOMEN + ADVERBIAL} i korpusset, så konstruktionen er ikke særligt hyppig korpussets størrelse taget i betragtning. De fleste forekomster (221 ud af 286 forekomster eller 77,3 %) er i overensstemmelse med hovedreglen, fx

- (22) *Nu er det vel sjældent I, på Borgen, går ud og spørger os vælgere til råds inden I vedtager love og regler!*
- (23) *Det sjovt hvergang sådanne noget sker, skriger befolkningen efter højere straffe. Men jeg ikke forstå hvorfor de inde på borgen ikke gør noget ved det*

Imod (23) kan man indvende at adverbialet kan opfattes som et frit adverbial der ikke knytter sig til pronomenerne, og at (23) derfor er tvetydig. Jeg har alligevel valgt at inkludere sådanne tvetydige tilfælde i resultaterne da de ikke i nævneværdig grad påvirker fordelingen af nominativ og oblik og altså *kan* opfattes som tilfælde af konstruktionen {PRONOMEN + ADVERBIAL}. De tvetydige resultater udgør omtrent en tredjedel af resultaterne.

Jeg finder også en del eksempler på oblik form i subjekt i alle personer:

- (24) *os her i huset stemmer ligesom vi plejer nemlig på DF*
- (25) *Jer inde på Christiansborg fatter jo ikk en skid alligevel*
- (26) *Hvornår bliver dem på tålt ophold smidt ud af landet?*

Det er interessant at der i tredje person er en klar præference for oblik form i subjekt (82,8 %, jf. tabel 4) hvor der i første og anden person er en klar præference for nominativ form (hhv. 94,3 % og 92,1 %, jf. tabel 4). Forskellen er statistisk signifikant⁹.

Omvendt finder jeg kun ét eksempel på nominativ uden for subjekt:

9 (df = 2, N = 228) = 132,16; $p < 0,0001$.

- (27) *Historier som hans giver energi til at kæmpe videre for en ordentlig alderdom - også til de, med de hårdeste job.*

5. Diskussion

Konstruktionen {PRONOMEN + RELATIVSÆTNING} har nydt størst opmærksomhed i litteraturen og er da også den grundigst undersøgte. Derfor sammenligner jeg i det følgende undersøgelsens resultater med tidligere undersøgelses resultater.

I subjekt finder jeg en stærk præference for oblik form i anden person (84,7 %) og tredje person pluralis (71,9 %) og omtrent lige så ofte nominativ som oblik form i første person, altså ingen klar præference (jf. tabel 1). Det er til forskel fra hvad Petersen, Schack og Jensen (2020: 113) finder, hvor der ”i skriftsproget [er] en stærk (og statistisk signifikant) tendens til nominativ i 1. og 3. person pluralis”. Anden person pluralis optræder så sjældent at de ikke konkluderer noget her (ibid.). Forskellene skyldes muligvis forskelle i register, for ”tale er jo ikke bare tale, og skrift ikke bare skrift” (Schack, Jensen & Petersen 2019: 453). De foreslår da også selv at undersøgelser af andre typer tale og skrift måske kan lede til andre resultater, og det må denne undersøgelse siges at have gjort.

I direkte og indirekte objekt finder jeg i skriftsprogskorpusset en meget stærk præference for oblik form i alle tre personer, hhv. 97,5 % i første person, 100 % i anden person og 91 % i tredje person (jf. tabel 1). Konstruktionen er ikke tidligere undersøgt i objektfunktion.

Som styrelse for præposition er der ingen klar præference for hverken nominativ eller oblik form i tredje person pluralis (hhv. 43 % og 57 %, jf. tabel 1). Her finder Schack, Jensen og Petersen (ibid.) derimod en klar præference for oblik form i avistekster. Første og anden person er for tyndt belagt til at jeg kan udtale mig håndfast om dem.

I funktion som prædikativ er forholdene omtrent som for styrelser for præpositioner. Jeg kan kun udtale mig om tredje person pluralis, og her er der heller ingen klar præference for oblik form (52,5 %). Konstruktionen er ikke tidligere undersøgt i funktion som prædikativ.

Konstruktionen {PRONOMEN + ADVERBIAL} er ikke behandlet indgående i litteraturen. Det skyldes formentlig at ”[k]onstruktionen regnes for stilistisk uheldig” (Jørgensen 2000: 302) og ”sjældent dukker op i skriftsprog” (Jensen 2021: 56). Konstruktionen er måske sjælden i nogle former for skriftsprog, men det er den ikke i facebooktekster; som citeret ovenfor er ”skrift ikke bare skrift” (Schack, Jensen & Petersen 2019:

453): Man må undersøge forskellige former for skriftsprog (og talesprog) hvis man vil opstille grammatiske regler, og denne undersøgelse giver et eksempel på den indsigt i grammatisk variation – og grammatiske regler – man kan opnå ved at inddrage andre genrer. Jeg finder at der i første og anden person er en stærk præference for nominativ i subjekt, hhv. 94,3 % & 92,1 %. I tredje person er der omvendt en stærk præference for oblik form i subjekt, 82,8%. I den henseende ligner forholdene i tredje person pluralis forholdene i tredje person singularis, hvor der dog ifølge Jensen (2021: 56 f.) formentlig slet ikke forekommer nominativ. Jeg finder kun ét eksempel på nominativ uden for subjekt i tredje person, nemlig *Historier som hans giver energi til at kæmpe videre for en ordentlig alderdom - også til de med de hårdeste job*. Beskrivelsen af tendensen til nominativ for oblik som ”meget udbredt” (Hansen & Heltoft 2019: 443) og ”særligt stærk” (Jacobsen & Jørgensen 2013: 126 f.) bekræftes altså ikke af de undersøgte data.

På baggrund af ovenstående er det muligt at pege på en række faktorer som er bestemmende for kasusvalget: register, syntaktisk funktion, syntagmestruktur, person, numerus og relativindleder. Mht. register er der ift. de tidligere undersøgelser generelt en stærkere præference for oblik form i subjekt; de tidligere undersøgelser af skrift baserer sig fortrinsvis på avistekster, mens denne undersøgelse baserer sig på opslag og kommentarer fra Facebook. Mht. syntaktisk funktion har jeg vist at der er klare forskelle i subjekt og objekt, mens der i prædikativ og styrelser for præpositioner ingen klare præferencer er for hverken nominativ eller oblik. Mht. syntagmestruktur har jeg vist at der er væsentlige forskelle i distributionen af kasus i konstruktionerne {PRONOMEN + RELATIVSÆTNING} og {PRONOMEN + ADVERBIAL}. {PRONOMEN + ADVERBIAL} har præference for oblik i subjekt – i tredje person. I første og anden person gælder hovedreglen næsten uden undtagelse. Men også anden person opfører sig på nogle punkter anderledes end første person: I konstruktionen {PRONOMEN + RELATIVSÆTNING} er præferencen for oblik i subjekt stærkere end i tredje person, mens der i første person ingen klar præference er for hverken nominativ eller oblik. Mht. numerus er der også forskelle: Jeg har heri kun undersøgt de plurale pronomener, men man skal blot kigge til konstruktionen {PRONOMEN + ADVERBIAL} for at finde forskelle; i tredje person singularis kan formentlig kun oblik anvendes (Jensen 2021: 58 f.), mens der i tredje person pluralis anvendes både nominativ og oblik, dog med en klar præference for oblik. Endelig har valget af relativindleder også betydning for kasusvalget i konstruktionen {PRONOMEN

+ RELATIVSÆTNING}: Når *som* er udeladt, anvendes oftere – men ikke altid – oblik form i subjekt end når relativsætningen indledes af *som* eller *der*.

6. Konklusion

I det ovenstående har jeg givet en korpuslingvistisk undersøgelse af kasusvalg i plurale personlige pronomener i moderne dansk i to konstruktioner: {PRONOMEN + RELATIVSÆTNING} og {PRONOMEN + ADVERBIAL}, og det i andre typer data end de allerede undersøgte. De tidligere undersøgte data omfatter navnlig sociolingvistiske interview og avistekster. Jeg har undersøgt opslag og kommentarer fra Facebook. Til formålet har jeg udviklet et pythonscript, der bl.a. løser et metodisk problem med at undersøge konstruktionen {PRONOMEN + RELATIVSÆTNING} med udeladt *som*.

I afsnit 3 har jeg givet en udførlig beskrivelse af usus i ovennævnte konstruktioner. På baggrund af beskrivelsen af usus har jeg vist at en række faktorer har indflydelse på distributionen af kasus: numerus, person, syntaktisk funktion, syntagmestruktur, register og valg af relativindleder.

Mht. reglerne kan hovedreglen ikke stå alene: Samlet set kan hovedreglen forklare knap halvdelen af tilfældene i konstruktionen {PRONOMEN + RELATIVSÆTNING} og 77,3 % i konstruktionen {PRONOMEN + ADVERBIAL}. I {PRONOMEN + ADVERBIAL} overholdes hovedreglen for langt det meste i første og anden person pluralis; i tredje person pluralis foretrækkes oblik form i subjekt. I konstruktionen {PRONOMEN + RELATIVSÆTNING} er oblik for nominativ meget udbredt og nominativ for oblik relativt udbredt (jf. tabel 1). En mindre del af tilfældene af oblik for nominativ kan forklares med fællesledsreglen, men i langt de fleste tilfælde af oblik for nominativ er fællesleddet hverken subjekt i oversætningen eller relativsætningen (oblik i 1*), hvor der ikke kan være tale om attraktion. Omvendt kan nærmest alle tilfælde af nominativ for oblik (nominativ i 3*, jf. tabel 1) forklares med attraktion; nominativ i 4*, hvor der ikke kan være tale om attraktion, er yderst sjælden. Det kan give god grund til at antage fællesledsreglen som en særregel mhp. at forklare nominativ for oblik i konstruktionen {PRONOMEN + RELATIVSÆTNING}, men der kræves i hvert fald yderligere én regel, fx en regel der tilskriver oblik kasus status som defaultkasus i komplekse pronominaler. En sådan regel ville kunne forklare den meget hyppige forekomst af oblik for nominativ i {PRONOMEN + RELATIVSÆTNING} (navnlig oblik i 1* hvor der ikke kan være tale om attraktion) samt i {PRONOMEN + ADVERBIAL}. Sup-

pleres hovedreglen af reglen om oblik som defaultkasus stiger forklaringskraften fra 49,5 % til 96,1 % i konstruktionen {PRONOMEN + RELATIVSÆTNING} og fra 77,3 % til 99,7 % i konstruktionen {PRONOMEN + ADVERBIAL}. Lægges fællesledsreglen også til, er forklaringskraften nærmest total: Kun de yderst sjældne tilfælde af nominativ i 4* og det ene tilfælde af nominativ for oblik i {PRONOMEN + ADVERBIAL} indfanges ikke at dette regelkompleks.

Om forfatteren

Oliver Rix Johannsen, cand.mag., videnskabelig assistent, Dansk Sprog-nævn.

Litteratur

- Becker-Christensen, Christian (2010): *Dansk syntaks. Indføring i dansk sætningsgrammatik og sætningsanalyse*. Frederiksberg: Samfundslitteratur.
- Brink, Lars & Jørn Lund (1975): *Dansk Rigsmål* (bind 2). København: Gyldendal.
- Diderichsen, Paul (1941): *Sætningsbygningen i Skaanske Lov*. København: Universitets-Jubilæets danske Samfund.
- Diderichsen, Paul (1957 [1946]): *Elementær Dansk Grammatik*. 2. udg. København: Gyldendal.
- Explosion (2022): Models for the spaCy Natural Language Processing (NLP) library. Github repository: <https://github.com/explosion/spa-cy-models> (lokaliseret 24/11-2023).
- Hansen, Aage (1967): *Moderne Dansk* (bind 2). København: Det Danske Sprog- og Litteraturselskab. Grafisk Forlag.
- Hansen, Erik (1972): Dr. Jekyll og Mr. Hyde i dansk grammatik. En filologisk skitse. I: *Glæden ved grammatik. Udvalgte artikler og afhandlinger*. Henrik Galberg Jacobsen & Henrik Jørgensen (red.). København: Hans Reitzels Forlag 2001, 98-110.
- Hansen, Erik & Lars Heltoft (2019 [2011]): *Grammatik over det Danske Sprog* 2. udg. København og Odense: Det Danske Sprog- og Litteraturselskab & Syddansk Universitetsforlag.
- Heltoft, Lars (1997): Hvem opslugte hvo. Et bidrag til beskrivelsen af det danske kasussystems udvikling. I: *Ord, Sprog oc artige Dict. Et overblik og 28 indblik 1500-1700. Festskrift til Poul Lindegård Hjorth*. Flemming Lundgreen-Nielsen, Marita Akhøj Nielsen & John Kousgård Sørensen (red.). København: Universitets-Jubilæets danske Samfund & C.A. Reitzel, 227-256.

- Jacobsen, Henrik Galberg & Peter Stray Jørgensen (2013 [1988]): *Håndbog i Nudansk*. 6. udg. København: Politikens Forlag.
- Jørgensen, Henrik (2000): *Studien zur Morphologie und Syntax der festlandskandinavischen Personalpronomina – mit besonderer Berücksichtigung des Dänischen*. Århus: Aarhus University Press.
- Jensen, Eva Skafté (2021): Et system i stadig bevægelse. Pronominalkasus i dansk. *Danske Studier*, 45-71.
- Jensen, Eva Skafté, Jørgen Schack & Jan Heegård Petersen (2023): The use of case forms in Modern Danish – an empirical study. *Acta Linguistica Hafniensia* 55(2), 182-215.
- Johannsen, Anders, Héctor Martínez Alonso & Barbara Plank (2015): Universal Dependencies for Danish. I: International Workshop on Treebanks and Linguistic Theories (TLT14), 157-167. https://bplank.github.io/papers/ud_danish_tlt15.pdf (lokaliseret 24/11-2023)
- Mikkelsen, Kristian (1975 [1911]): *Dansk Ordføjningslære*. København: Hans Reitzel.
- Petersen, Jan Heegård, Jørgen Schack & Eva Skafté Jensen (2020): Pronominalkasus i subjekt med bestemmende relativsætning. *Ny forskning i grammatik* 27. Bogense: Dansk Sprognævn, 108-124.
- Schack, Jørgen (2013): Dig og mig og vi to. Synspunkter på kasus i moderne dansk. *NyS – Nydanske Sprogstudier* 45, 48-72.
- Schack, Jørgen, Eva Skafté Jensen & Jan Heegård Petersen (2019): ”Tak til alle de der gør en indsats for vores sikkerhed”. Bagom dronningens kasusvalg. I: *Møde om Udforskningen af Dansk Sprog* 17. Yonathan Goldstein, Inger Schoonderbeek & Tina Thode Hougaard (red.). Århus: Aarhus Universitet, 443-456.
- Wivel, Hylling Georg (1901): *Synspunkter for dansk Sproglære*. København: Bogforlaget Ernst Bojesen.

Var det sin eller hans kone Jens kyssede? Refleksiv reference i sætningskløvning og pseudokløvning

Michael Nguyen

Abstract

A recurring issue in prescriptive Danish grammar is the reflexive use of *sin* and *hans*, e.g., *Han kyssede sin/hans* ('He kissed his wife'). As stated in many grammars and by the Danish Language Council (Dansk Sprognævn), the principal rule for standard Danish dictates that *sin* is used when referring to the subject of the clause ('He kissed his own wife'), while *hans* is used in all other cases ('He kissed another man's wife'). This distinction also applies to the use of *sig* and *ham* ('him(self)'). Notably, clefts and pseudoclefts are exceptions to the principal rule. Simplifying somewhat, grammars and the Danish Language Council recommend the use of *sin* and *sig* in such cleft constructions when these pronouns refer to the subject of the subordinate clause, e.g. *Det var sin kone han kyssede* ('It was his own wife he kissed'). Based on the data from two text corpora (one comprising novels from 1870-1899 and the other comprising modern national newspapers), I show that the recommendations of the grammars and of The Danish Language Council do not align entirely with actual written language usage. Consequently, I propose that it is reasonable to reconsider these recommendations.

Nøgleord

possessive og refleksive pronomener, sprognormer, præskriptivisme, korpuslingvistik

1. Indledning

Eksemplet i titlen peger på et klassisk sprogrigtighedsproblem: Skal man i standarddansk bruge *sin* eller *hans* når Jens kysser den kvinde han er gift med? Det er ikke kun lingvister der har været opmærksomme på dette problem (se fx referencerne i Diderichsen 1939), men også ikkelingvister (se fx de mange diskussioner og sider med sproglige råd på nettet). Dansk Sprognævn får desuden ofte spørgsmål om *sin* og *hans* fra selv erfarne kor-

rekturlæsere. Selvom det især er brugen af *sin* og *hans* der forbindes med det klassiske sprogrigtighedsproblem, hører brugen af *sig (selv)* overfor *ham (selv)* hører også til problemkomplekset.¹

Den klassiske prækriptive hovedregel for standarddansk er: Når pronomenet refererer til et singularissubjekt, bruges *sin* og *sig (selv)* (fx Dansk Sprognævn u.å.; Diderichsen 1962: 55-56); *hans* og *ham (selv)* bruges i andre tilfælde. Se følgende eksempler:

- (1) Jens tog **sin/hans** hat og gik. (Jens tog sin egen eller en andens hat).
- (2) Jens vaskede **sig selv/ham selv**. (Jens vaskede sig selv eller en anden).

Hovedreglen er ikke undtagelsesløs; i 'skjulte sætninger' anbefales det at man fraviger den (se afsnit 2). Det samme gælder ved 'sætningskløvning' (*Det var sin hat han tog*) og ved 'pseudokløvning' (*Det han tog, var sin hat*) (se afsnit 2 og 3). Ved reference til pluralissubjekt anbefales brugen af *deres* og *sig (selv)* i alle grammatiske kontekster.

I denne artikel undersøges det hvilke pronomener der bruges i sætningskløvning og pseudokløvning. Sprogbrugen undersøges i to skriftsprogskorpusser: (i) korpuset MeMo, som indeholder data fra romaner fra det moderne gennembrud (1870-1899), og (ii) et sprognævnskorpus som indeholder data fra landsdækkende aviser i perioden 1999-2023. Dette korpus kalder jeg i kortform LA (for landsdækkende aviser). Formålet med artiklen er dels at redegøre for sprogbrugen, dels at give et konstruktivt bidrag til hvilke anbefalinger Dansk Sprognævn bør udstikke for skriftsproget.

Sætningskløvning generelt er blevet indgående behandlet (fx Christensen 2019; E. Hansen 1995; Togeby 2009), og det samme gælder alternationen mellem *sin* og *hans* og alternationen mellem *sig (selv)* og *ham (selv)* udenfor begge typer kløvning (fx Jensen 2009; Ehlers 2024). Der er derimod blevet skrevet meget mindre om pseudokløvning (se dog Dyhr 1978). Artiklen bidrager således til at dække et hul i forskningslitteraturen.

Artiklen er struktureret således: Først gennemgås kort litteraturen om

¹ Når jeg i det følgende taler om *sin*, taler jeg om alle dets bøjningsformer (*sin, sit* og *sine*). Når jeg taler om *ham (selv)*, taler jeg også om *hende, den, det* og *én (selv)* – og tilsvarende om *hans*, som også dækker over *hendes, dens, dets* og *ens*. Det vil fremgå klart når jeg kun taler om ordformerne *ham (selv)*, *hans* osv. Parentesen rundt om *selv* bruges når noget gør sig gældende uanset om *selv* er til stede, modsat når *selv* står uden parentes.

sin og *sig (selv)* i skjulte sætninger og i sætningskløvning og pseudokløvning (afsnit 2), og dernæst præsenteres sætningskløvning og pseudokløvning (afsnit 3). Derefter gennemgås undersøgelsens data og metoder (afsnit 4) samt dens resultater (afsnit 5). Endelig diskuteres det hvilke ændringer i Sprognettets anbefalinger disse resultater kan give anledning til (afsnit 6).

2. Litteraturen om *sin* og *sig (selv)* i skjulte sætninger og i sætningskløvning og pseudokløvning

Som nævnt anbefales det at hovedreglen fraviges ved såkaldt skjulte sætninger. Betegnelsen 'skjulte sætninger' omfatter både en række ikkefinitte og ikkeverbale konstruktioner, og fælles for dem er at de uden større betydningsændring kan parafraseres til en enkel sætning. Fravigelsen fra hovedreglen består i at *sin* og *sig (selv)* bruges til at referere til subjektet i en skjult sætning (Jacobsen & Jørgensen 2013: 541-545; sproget.dk). Se følgende eksempler med *sin* (kursiveringen er beholdt fra originalen):

- (3) Hun bad *ham holde sin mund*. (Han holdt *sin* mund).
- (4) Vi så *præsten gå tur med sin hund*. (Præsten gik tur med *sin* hund).
- (5) De fandt *ham siddende i sin bil*. (Han sad i *sin* bil).
- (6) *Rektors modtagelse af sine gamle elever* var rørende. (Rektor modtog *sine* gamle elever).
- (7) Vi vil gøre *hende ansvarlig for sine udtalelser*. (Hun er ansvarlig for *sine* udtalelser).
- (8) Kammeraterne hjalp *ham op på sit værelse*. (Han blev hjulpet op på *sit* værelse).
- (9) Jeg traf *direktøren på sit kontor*. (Direktøren var på *sit* kontor).
(Jacobsen & Jørgensen 2013: 542)

At det er *sin* der anbefales, er ret klart – på trods af at det kan gøre sætningerne tvetydige (se Diderichsen 1962: 56-57 for en mere nuanceret diskussion):

”I de fleste af den slags sætninger [dvs. dem i (3)-(9)] er det naturligst at bruge *sin (sit, sine)* sådan som det er gjort her – også selvom sætningerne nogle gange er dobbelttydige, og selvom man kunne gøre dem entydige ved at bruge *hans* eller *hendes* i stedet.”

(Jacobsen & Jørgensen 2013: 542)

”Og uanset om meningen er den ene eller den anden [i tve-tydige sætninger som *Kaj bad Lis ringe til sin mor*], så er sætningerne med *sin* naturlige og korrekte. Man kunne selvfølgelig fjerne risikoen for misforståelser ved at bruge *hans* eller *hendes* i stedet for *sin* [fx i *Kaj bad Lis ringe til hans/hendes mor*] [...]. Humlen er blot at disse sætninger dårligt kan anses for naturlige eller ægte danske sætninger, men snarest må betragtes som lidt fortænkte eller kunstige nødformuleringer.”

(Jacobsen & Jørgensen 2013: 543)

Ved skjulte sætninger træder det jeg kalder ’parafraseprincippet’, i kraft:

(10) **Parafraseprincippet**

Hvis man ud fra hovedreglen skal bruge *sin* eller *sig* (*selv*) i den omformulerede, enkle sætning, skal man ligeledes bruge *sin* eller *sig* (*selv*) i den ikkeomformulerede sætning.

Mere konkret bruges *sin* i (3)-(9) fordi man ifølge hovedreglen skal bruge *sin* i parafraaserne. Parafraseprincippet går tilbage til i hvert fald Baden (1789: 189) og Dichmann (1800: 403 f.) (se Diderichsen 1939: 14 ff. for flere detaljer).

Ved sætningskløvning træder parafraseprincippet også i kraft, se fx By-skov (1943: 86), Diderichsen (1939: 51; 1962: 96), H. Hansen (1932: 59), Mikkelsen (1975: 258-260; 1894: 329-330) og Aa. Hansen (1965: 112). I grammatikkerne er der bred enighed om at *sin* og *sig selv* bruges i sætningskløvning – helt fra 1800-tallet til nyere tid. Uden forbehold skriver Lefolii (1871: 98) således:

”Det hedder: *Det er sig selv, han snarest skulde vredes paa. [...]*
– *Det var ikke til sit Fædrelands men til sin egen Fordel, at han førte den ene Krig efter den anden og sejrede i dem alle.*”

(Lefolii 1871: 98)

I nyere sprogbeskrivelser er synspunkterne mere moderate, men der er ingen tvivl om at det er *sin* og *sig selv* der er hovedformerne (E. Hansen & Heltoft 2019: 589-590, 595; se også E. Hansen 1995: 128, Nølle 1984: 77-78). Blandt nyere beskrivelser indtager Tøgeby (2009: 443-444) det

mest radikale synspunkt: Han betragter det som et **definerende** træk ved sætningskløvning at man bruger *sin* og *sig selv*.

Dansk Sprognævn har lagt sig i slipstrømmen af litteraturen og med afsæt i parafraseprincippet anbefalet *sin* (og vel dermed også *sig selv*) i sætningskløvning. Det fremgår af to skriftlige svar fra hhv. 2012 og 2018. I begge svar var Sprognævnet ikke helt afvisende overfor brugen af *hans*.

Så vidt pronomenerne i sætningskløvning. Så mange bemærkninger finder man ikke om pseudokløvning, formentlig fordi pseudokløvning er meget mindre udforsket. Dette har den konsekvens at den reflektive reference i pseudokløvning vist også er gået under radaren. Sprognævnet er dog blevet spurgt om følgende sætning:

(11) [...] så det eneste hun kunne bevæge, var **sine/hendes** øjne.

I Sprognævnets svar (det førnævnte fra 2018) er det *sin* der anbefales. Anbefalingen hviler på parafraseprincippet; man ville bruge *sin* i *Hun kunne bevæge sine øjne*.

Opsummerende kan det siges at *sin* og *sig (selv)* er de anbefalede former i skjulte sætninger, sætningskløvning og pseudokløvning når der er tale om reference til et singularissubjekt. Ved reference til et pluralissubjekt anbefales *deres* og *sig (selv)*, uafhængigt af konstruktion.

3. Om sætningskløvning og pseudokløvning

I dette afsnit vil jeg kort beskrive sætningskløvning og pseudokløvning. Fælles for de to konstruktioner er at pronomenet i det udkløvede led kan referere til subjektet på tværs af en sætningsgrænse (hvilket det normalt ikke kan, se fx Wiwel 1901: 250, 343-344), og at konstruktionerne typisk kan parafraseres til en enkel sætning.

Sætningskløvning består typisk af *det + være + det udkløvede led + en ledsætning*, fx *Det er sofaen der står derude* (se fx Christensen 2019: 21; E. Hansen 1995: 126). Det udkløvede led er typisk fokuseret, mens informationen i ledsætningen er præsupponeret.

Det ser ud til at der er en vis enighed om at pseudokløvning (typisk) består af (i) en ledsætning indledt af et *hv*-ord + *være + det udkløvede led*, eller (ii) *det/den + et eventuelt adjektiv + en ledsætning + være + det udkløvede led* (se bl.a. Christensen 2019: 13, 44; Patten 2013: 7, 62-64):

(12) Hvad der står derude, er sofaen.

- (13) Det der står derude, er sofaen.
 (14) Det eneste, man garanteret taber her, er **sine** penge.
 (15) Det næste, Frida husker, er **hendes** mosters ansigt [...].
 ((14) og (15) er fra LA).

Da pseudokløvning som sagt er langt mindre udforsket end sætningskløvning, har man i litteraturen vistnok ikke taget stilling til om eksempler som følgende involverer pseudokløvning:

- (16) Ikke engang [sic] det tøj, hun har på, er **hendes** eget. (Fra LA)

(16) adskiller sig fra (13)-(15) ved at ledsætningen indledes af *det* + et substantiv. Eksempler som (16) tælles her med som forekomster på pseudokløvning.

Man antager normalt at sætningskløvning og pseudokløvning kan føres tilbage til en ukløvet sætning (Hansen 1995: 127). (12) og (13) kan således føres tilbage til *Sofaen står derude*. Men dette ikke altid er muligt, især ved pseudokløvning (Patten 2013: 62). Sætningen i (14) kan således ikke uden videre tilbageføres til en ukløvet sætning; man er nødt til at indsætte et *kun*: *Man taber garanteret kun **sine** penge her*. I (15) skal man fx indsætte *som det næste*: *Frida husker som det næste **hendes** mosters ansigt*. I (16) skal der omstruktureres helt: *Hun har tøj på som ikke en gang er **hendes** eget*. Det afgørende er dog ikke hvad konstruktionerne i (12)-(16) kaldes, men blot at pronomenet her kan referere til subjektet på tværs af en sætningsgrænse, og at konstruktionen kan føres tilbage til en (evt. modifieret) enkel sætning.

4. Data og metoder

Som nævnt bruger jeg to korpusser i undersøgelsen: (i) MeMo, som indeholder ca. 64,6 mio. ord fra romaner fra det moderne gennembrud (1870-1899), og (ii) et korpus som indeholder ca. 894 mio. ord fra landsdækkende aviser (LA). De to genrer er ganske vist forskellige, men har det til fælles at de er undergået en form for redigeringsproces. Forfattere, journalister, redaktører og andre der har været inde over teksterne, har altså haft tid til at omformulere og fx vælge et andet pronomener end det der blev valgt i første omgang.

For at finde eksempler med sætningskløvning og pseudokløvning brugte jeg søgestrengene der består af følgende elementer:

(17) Sætningskløvning

- a. *Det/det* + en form af *være* + *sig/ham/hende/en/én/een/det/den/dem* + *selv* + komma
- b. en form af *være* + *det* + *sig/ham/hende/en/én/een/det/den/dem* + *selv* + komma
- c. *Det/det* + en form af *være* + *sin/sit/sine/hans/hendes/dens/dets* + 1-2 ord + komma
- d. en form af *være* + *det* + *sin/sit/sine/hans/hendes/dens/dets* + 1-2 ord + komma

(18) Pseudokløvning

- a. *Hvad/hvad/Det/det/Den/den* + 2-6 ord + komma + 0-1 ord + en form af *være* + *sig/ham/hende/en/én/een/selv/det/den/dem* + *selv* + komma
- b. *Hvad/hvad/Det/det/Den/den* + 2-6 ord + komma + 0-1 ord + en form af *være* + *sin/sit/sine/hans/hendes/ens/dens/dets*

Formerne *Det, det, Den, den, Hvad* og *hvad* skal indfange eksempler hvor disse ord har store og små begyndelsesbogstaver. Kommaet i søgestrengene i (17) skal indfange eksempler med kommaet før ledsætningen i sætningskløvning, mens kommaet i søgestrengene i (18) skal indfange eksempler med komma efter ledsætningen i pseudokløvning. I (17)a og (17)c søges der efter eksempler med ligefrem ordstilling, og i (17)b og (17)d søges der efter eksempler med omvendt ordstilling.

Da jeg er interesseret i alternationen mellem de anbefalede og ikkeanbefalede former, har jeg kun medtaget eksempler hvor en form er anbefalet, og en anden ikke er. Ved singularissubjekter der betegner kollektiver, kan man både bruge et pronomen i singularis og et pluralis, fx *Hvad Dansk Folkeparti derimod havde indflydelse på, var **dets** rolle som parlamentarisk grundlag, [...]* (fra LA). Sådanne eksempler har jeg ikke medtalt fordi der ikke er nogen anbefaling, hverken den ene eller den anden vej (det fremgår af et skriftligt svar fra 2006). Under alle omstændigheder var der ganske få eksempler af den slags.

Endelig udelades også tvivlsomme forekomster. Interessant nok var dette problem særligt relevant for søgeresultaterne med *hans* og *ham selv* ved sætningskløvning. Hvis der er eksempler der fejlagtigt er blevet sorteret fra, så er det især dem med *hans* og *ham selv*. Det gør det forholdsvis store antal medtalte forekomster af *hans* og *ham selv* endnu mere overbevisende.

5. Resultater

I dette afsnit præsenteres og diskuteres resultaterne for sætningskløvning og pseudokløvning. Der i alt kun 450 eksempler. Dette er et ret lille tal i betragtning af at det samlede korpusmateriale er ganske stort.

I det følgende vil jeg præsentere fordelingen af de forskellige pronomi- nale former. For at finde ud af om en given forskel er statistisk signifikant, vil jeg bruge binomialtests, chi i anden-tests og Fischer's exact tests. Jeg opererer med et signifikansniveau på 0,05; dvs. en p-værdi under 0,05 betyder at en forskel er statistisk signifikant. Da jeg dog i alt laver 23 stati- stiske tests, er der en forøget risiko for at nogle af testene er falsk-positive. Jeg har derfor anvendt Benjamini-Hochberg-metoden for at kontrollere for antallet af falsk-positive tests (False Discovery Rate, forkortet FDR). Jeg sætter FDR til 0,05, dvs. at 5 % af alle signifikante tests i gennemsnit alligevel ikke er signifikante.

I de følgende tabeller angives kun de modsætninger hvor der er en eller flere relevante forekomster, fx *sin* vs. *hans/hendes*. I mange af tabellerne angives fx ikke modsætningen *sin* vs. *ens*, ganske enkelt fordi der dér ikke er nogen forekomster der angår denne modsætning.

5.1. Sætningskløvning

5.1.1. Resultater for MeMo

Nedenfor ses resultaterne for sætningskløvning i MeMo (de anbefale- lede former markeres med **fed**, de ikkeanbefalede former er i ordi- nær):

	Subjektstype	Refleksive former	Nonrefleksive former	I alt
Possessive pronomener	Han/hun	Sin (26)	Hans/hendes (42)	68
	De	Sin (0)	Deres (2)	2
	I alt	26 (37 %)	44 (63 %)	70 (100 %)
Personlige pronomener	Han/hun	Sig selv (6)	Ham/hende selv (0)	6
	Den/det	Sig selv (1)	Den/det selv (0)	1
	I alt	7 (100 %)	0 (0 %)	7 (100 %)

Table 1. Fordeling af pronomener i sætningskløvning (MeMo).

Hvad de possessive pronomener angår, bruges *hans/hendes* mere end *sin* (62 % *hans/hendes* vs. 38 % *sin*). Forskellen ville have været statistisk signi- fikant hvis testen var blevet lavet i isolation ($p = 0,03406$), men brugen af Benjamini-Hochberg-metoden betyder at forskellen alligevel ikke er signi-

fikant. I øvrigt er der kun to eksempler med reference til et pluralissubjekt, og i begge eksempler bruges *deres*. Se eksempler nedenfor:

- (19) [...] som om det var **sine** egne Følelser, hun tolkede.
 (20) Men at det var **hans** Hustru, han skulde tale saaledes med, det vakte pludselig en mørk Tanke, [...].
 (21) [...] men saa tænkte han paa, at det var **deres** Fritid, de brugte, [...].

Hvad *sig selv* vs. *ham/hende selv* angår, forekommer *sig selv* signifikant mere ($p = 0,01563$). Og der var ét eksempel med *sig selv* vs. *den/det selv*. Eksempler:

- (22) [...], som om det var **sig selv**, hun anklagede, fortalte hun hele den Samtale, [...].
 (23) Det unge skyder ud af det gamle [...] og bærer friskt Liv tilbage til dets Kveld. Men saa kommer det rigtignok an paa, [...] om dette skal bli til Velsignelse, og det beror paa det gamle selv, om det var **sig selv**, det elsked i det ny, som skød ud fra det, [...].

5.1.2. Resultater for LA

Nedenfor ses resultaterne for sætningskløvning i LA:

	Subjektstype	Refleksive former	Nonrefleksive former	I alt
Possessive pronomener	Han/hun	Sin (39)	Hans/hendes (66)	105
	Man	Sin (2)	Ens (16)	18
	De	Sin (0)	Deres (42)	42
	I alt	41 (25 %)	124 (75 %)	165 (100 %)
Personlige pronomener	Han/hun	Sig selv (14)	Ham/hende selv (12)	26
	Den/det	Sig selv (1)	Den/det selv (0)	1
	Man	Sig selv (0)	En/én/een selv (1)	1
	De	Sig selv (3)	Dem selv (5)	8
	I alt	18 (50 %)	18 (50 %)	36 (100 %)

Tabel 2. Fordeling af pronomener i sætningskløvning (LA).

I andenøverste række ses det at der er stor variation mellem *hans/hendes* og *sin*; *hans/hendes* forekommer signifikant mere end *sin* ($p = 0,005409$).

Ens forekommer også signifikant mere end *sin* ($p = 0,0006561$). Der er yderligere 42 eksempler med reference til pluralissubjekt, alle med *deres*. Eksempler:

- (24) Det var **sin** egen ensomhed, han tænkte på.
- (25) Det var **hans** egen morfar, at den nu 58-årige musiker stjal penge fra.
- (26) Hvorfor er det så vigtigt, at det er **sine** gener, man bidrager med?
- (27) I solomoderskabet er det **ens** egen enlighed, man kroner: [...].
- (28) På videoen fortæller Henrik Vindfeldt, at det er **deres** egen afføring, de er smurt ind i.

Hvad de personlige pronomener angår, er andelen af *sig selv* og andelen *ham/hende selv* næsten lige store, og forskellen er ikke statistisk signifikant ($p = 0,845$). Det samme gælder *sig selv* overfor *dem selv* (3 forekomster overfor 5; $p = 0,7266$). De to forekomster ved subjektstyperne *den/det* og *man* kan man ikke sige noget statistisk meningsfuldt om, men de indgår i eksempel-materialet nedenfor:

- (29) [...] eller var det **sig selv**, hun truede?
- (30) Hvad enten gud eller djæveln fylder naturen, er det **sig selv**, mennesket ser.
- (31) Måske er det én **selv**, man er vred på, [...].
- (32) [...], og derfor er det **sig selv**, de kigger på, [...].
- (33) De opfatter simpelthen ikke, at det er **dem selv**, de kigger på.

5.2. Pseudoklønning

5.2.1. Resultater for MeMo

Nedenfor følger resultaterne for MeMo:

	Subjektstype	Refleksive former	Nonrefleksive former	I alt
Possessive pronomener	Han/hun	Sin (4)	Hans/hendes (20)	24
	De	Sin (0)	Deres (2)	2
	I alt	4 (15 %)	22 (85 %)	26 (100 %)
Personlige pronomener	Han/hun	Sig selv (1)	Ham/hende selv (0)	1
	I alt	1 (100 %)	0 (0 %)	1 (100 %)

Tabel 3. Fordeling af pronomener i pseudoklønning (MeMo).

De ikkeanbefalede former *hans/hendes* bruges signifikant mere end *sin* ($p = 0,0007719$). Ved modsætningen *sin* vs. *deres* er der 0 forekomster på *sin* og 2 på *deres*. Eksempler:

- (34) [...] det eneste, han muligvis kan glemme, er **sin** Koffert og **sine** Penge, [...].
 (35) Det eneste, han beklagede sig over, var **hans** Jomfruer, [...].
 (36) Den eneste Lyd, de hørte, var **deres** egne Fodtrin [...].

Der var kun ét eksempel med personlige pronomener i pseudoklønning:

- (37) Det første, hun blev sig bevidst, var **sig** selv staaende midt paa Landevejen [...].

5.1.2. Resultater for LA

Nedenfor følger resultaterne for LA:

	Subjektstype	Refleksive former	Nonrefleksive former	I alt
Possessive pronomener	Han/hun	Sin (42)	Hans/hendes (39)	81
	Den/det	Sin (0)	Dens/dets (4)	4
	Man	Sin (7)	Ens (2)	9
	De	Sin (0)	Deres (33)	33
	I alt	49 (39 %)	78 (61 %)	127 (100 %)
Personlige pronomener	Han/hun	Sig selv (6)	Ham/hende selv (0)	6
	Man	Sig selv (6)	En/én/een selv (0)	6
	De	Sig selv (1)	Dem selv (5)	6
	I alt	13 (72 %)	5 (28 %)	18 (100 %)

Tabel 4. Fordeling af pronomener i pseudoklønning (LA).

Der er stor variation mellem *sin* og *hans/hendes*, 42 forekomster med *sin*, 39 med *hans/hendes*. Forskellen er ikke statistisk signifikant ($p = 0,8243$). Forskellen ved modsætningen *dens/dets* vs. *sin* er heller ikke statistisk signifikant ($p = 0,0625$). Det samme gælder forskellen ved *sin* vs. *ens* ($p = 0,08984$). De ikkesignifikante forskelle skyldes det lave antal forekomster.

I pluralis er *deres* enerådende overfor *sin*, og forskellen er statistisk signifikant ($p = 2,328 \cdot 10^{-10}$). Se eksempler nedenfor:

- (38) [...] og det, hun møder, er **hendes** egne forestillinger [...].
 (39) Det dyrebareste, man har som borgmester, er **sin** integritet [...].
 (40) En ufatteligt narcissistisk fisk, der forelsker sig så meget i sit eget spejlbillede [...], at den til sidst dør af det. Måske lykkelig, fordi det sidste, den ser, er **dens** eneste ene.
 (41) I terapien kan det ofte vise sig, at det, man kæmper med, er **ens** egen fordømmelse [...].
 (42) [...] så det eneste, man selv skal medbringe, er **sin** eventyrlyst.
 (43) Hvad de færreste dog tager med sig, er **deres** affald [...].

Ved de personlige pronomener er der ganske få eksempler, og man skal derfor være varsom med at drage for stærke konklusioner. *Sig selv* bruges dog signifikant mere end *ham/hende selv* ($p = 0,01563$) og signifikant mere end *én selv* ($p = 0,01563$). Ved pluralissubjekt er det ikke statistisk signifikant at *dem selv* bruges mere end *sig selv* ($p = 0,1094$). Eksempler:

- (44) [...] fra begyndelsen fandt hun ud af, at det, hun ville undersøge i kunsten, var **sig** selv.
 (45) Det eneste, man kan være, er **sig** selv.
 (46) Det, de kommer med, er **sig** selv, [...].
 (47) Det eneste, politikerne tænker på, er **dem** selv.

5.3. Overordnede resultater

I det følgende gennemgås de overordnede resultater. Se resultaterne for sætningskløvning nedenfor:

Korpus	Numerus	De anbefalede former	De ikkeanbefalede former	I alt
MeMo	Sg.	33 (44 %)	42 (56 %)	75 (100 %)
	Pl.	2 (100 %)	0 (0 %)	2 (100 %)
LA	Sg.	56 (37 %)	95 (63 %)	151 (100 %)
	Pl.	45 (90 %)	5 (10 %)	50 (100 %)
MeMo + LA	Sg.	89 (39 %)	137 (61 %)	226 (100 %)
	Pl.	47 (90 %)	5 (10 %)	52 (100 %)

Tabel 5. Den samlede fordeling af anbefalede og ikkeanbefalede former i sætningskløvning

I singularis er der overordnet set stor variation mht. de anbefalede og de ikkeanbefalede former, 39 % overfor 61 % (se den andennederste ræk-

ke). De ikkeanbefalede former bruges signifikant mere end de anbefalede former ($p = 0,0008544$). I pluralis er der stort set ingen variation, 90 % overfor 10 % ($p = 6,424 \cdot 10^{-10}$).

Se resultaterne for pseudoklønning nedenfor:

Korpus	Numerus	De anbefalede former	De ikkeanbefalede former	I alt
MeMo	Sg.	5 (20 %)	20 (80 %)	25 (100 %)
	Pl.	2 (100 %)	0 (0 %)	2 (100 %)
LA	Sg.	61 (58 %)	45 (42 %)	106 (100 %)
	Pl.	34 (87 %)	5 (13 %)	39 (100 %)
MeMo + LA	Sg.	66 (50 %)	65 (50 %)	131 (100 %)
	Pl.	36 (88 %)	5 (12 %)	41 (100 %)

Tabel 6. Den samlede fordeling af anbefalede og ikkeanbefalede former i pseudoklønning

I singularis er der overordnet set stor variation mellem de anbefalede og de ikkeanbefalede former, 66 forekomster overfor 65 (forskellen er ikke signifikant: $p = 0,9017$). I pluralis er der derimod kun mindre variation, 88 % overfor 12 %, og forskellen er signifikant ($p = 3,921 \cdot 10^{-7}$).

Hvis vi lægger alle tallene sammen for begge typer klønning, får vi et samlet billede af sprogbrugen:

Korpus	Numerus	De anbefalede former	De ikkeanbefalede former	I alt
MeMo	Sg.	38 (38 %)	62 (62 %)	100 (100 %)
	Pl.	4 (100 %)	0 (0 %)	4 (100 %)
LA	Sg.	117 (46 %)	140 (54 %)	257 (100 %)
	Pl.	79 (89 %)	10 (11 %)	89 (100 %)
MeMo + LA	Sg.	155 (43 %)	202 (57 %)	357 (100 %)
	Pl.	83 (89 %)	10 (11 %)	93 (100 %)
	Sg. + pl.	238 (53 %)	212 (47 %)	450 (100 %)

Tabel 7. Den samlede fordeling af de anbefalede former og ikkeanbefalede former i sætningsklønning og pseudoklønning

Den nederste række viser at der er stor variation i sprogbrugen: I 53 % af tilfældene vælger sprogbrugerne en anbefalet form; i 47 % af tilfældene vælger de en ikkeanbefalet form. Forskellen er ikke signifikant ($p = 0,08304$). Variationen er især at finde i singularis (se den tredjenederste

række); de ikkeanbefalede former forekommer her signifikant mere end de anbefalede ($p = 0,01727$).

I pluralis er der stort set ingen variation: Her vælger sprogbrugerne altovervejende de anbefalede former (se den andennederste række). Det er de mange forekomster af *deres* der får balancen til at tippe (79 eksempler med *deres*, 0 med *sin*). Der er ikke mange forekomster der er relaterede til modsætningen *sig selv* vs. *dem selv* (4 eksempler med *sig selv*, 10 med *dem selv*), og forskellen er ikke statistisk signifikant ($p = 0,08978$). I disse data er der altså ikke noget der tyder på at den anbefalede form *sig selv* bruges mere end *dem selv*.

Resultaterne i pluralis er interessante at sammenholde med hvordan (i) *sig (selv)* og *dem (selv)* og (ii) *sin* og *deres* opfører sig **udenfor** begge slags kløvning. Hvad frekvensforholdet mellem *sig (selv)* og *dem (selv)* angår, findes der, så vidt jeg ved, endnu ikke nogen undersøgelse heraf, og en sådan undersøgelse er ønskværdig.

Hvad *sin* vs. *deres* angår, finder Jensen i sin undersøgelse af talesprogs-korpusset LANCHART at der er 7 % *sin* og 93 % *deres* ud af 644 eksempler (Jensen 2009: 137). Og Katrine Rosendal Ehlers informerer mig om at der i hendes bearbejdede (men upublicerede) data fra KorpusDK er en fordeling på ca. 2 % *sin* og 98 % *deres*. I mine data er der ingen variation (0 eksempler med *sin*, 79 med *deres*), men det kan skyldes at der er for lidt data. Ehlers fandt kun følgende eksempel på sætningskløvning med *sin* (præsenteret ved hendes ph.d.-forsvar i april 2024): *Det var fra sine østafrikanske kystforter, at tyskerne ekspanderede mod vest* (KorpusDK).

6. Diskussion

Som nævnt bygger undersøgelsen på blot 450 eksempler. Det er svært at vide hvorfor der er så få eksempler. Måske forsøger sprogbrugeren at undgå sætningskløvning og pseudokløvning i forbindelse med et pronomen der refererer til subjektet i ledsætningen – ganske enkelt fordi man er i tvivl om hvilken form man skal bruge. Men selv dér hvor der ikke er nogen variation, fx ved modsætningen *deres* vs. *sin*, er det ikke sådan at der er tusindvis af eksempler.

Det mest bemærkelsesværdige resultat er at de ikkeanbefalede former i singularis samlet set forekommer signifikant mere end de anbefalede former ($p = 0,000435$); se den tredjenederste række i tabel 7. I slutningen af 1800-tallet brugte man de ikkeanbefalede former i singularis i 62 % af

tilfældene, mens det tilsvarende tal i moderne dansk er 54 %. Men selvom forskellen ikke er statistisk signifikant ($p = 0,1977$), er der ikke noget der tyder på at man i mindre grad følger anbefalingerne i dag.

Resultaterne for singularis viser at både den århundredlange anbefaling om at bruge *sin* og *sig selv* i sætningskløvning og Sprognævnets anbefaling om at bruge *sin* og *sig selv* i pseudokløvning ikke er helt i overensstemmelse med sprogbrugen, hverken i slutningen af 1800-tallet eller i moderne dansk. Dette giver anledning til at genoverveje disse skriftsproglige anbefalinger. Sprognævnet tager nemlig i sin rådgivning ikke kun udgangspunkt i det der står i grammatikkerne, men også i sprogbrugen hos den gode og sikre sprogbruger (se også Diderichsen & Schack 2015).

Én mulighed er at ophøje **alle** de nu ikkeanbefalede former til også at være anbefalede former. Det har fordelen at reglerne bliver forenklet. En væsentlig ulempe vil være at former der sjældent forekommer eller slet ikke forekommer, vil være 'tilladte'. Man vil således kunne skrive *ens* ved *man* som subjekt, se (48), og *sin* og *dem selv* ved pluralissubjekt, se (49) og (50) (måske er det eksemplerne i (49) der er mindst gangbare):

- (48) a. Det er **ens** mor man skal lytte til.
 b. Den man skal lytte til, er **ens** mor.
- (49) a. Det var **sin** mor de besøgte.
 b. Den de besøgte, var **sin** mor.
- (50) a. Det var **dem selv** de bebrejdede.
 b. Dem de bebrejdede, var **dem selv**.

En anden løsning er at man **kun** i singularis ophøjer de ikkeanbefalede former til anbefalede former (fx *Det var hans egne bøger han bedst kunne lide*). Dette har fordelen at af de problematiske eksempeltyper i (48)-(50) er det kun brugen af *ens* i eksempler som (48) der bliver anbefalet. Til gengæld vil sprogbrugeren skulle holde styr på **to** regelsæt: Ved singularissubjekt vil der være valgfrihed, og ved pluralissubjekt skal man kun bruge *deres* og *sig selv*.

Interessant nok er den præskriptive regel om *sin* og *sig selv* i sætningskløvning ikke nævnt i nyere behandlinger af sprogrigtighed (Jacobsen & Jørgensen 2013; sproget.dk), i modsætning til nyere grammatiske beskrivelser (se afsnit 2). På Grammatiknetværksmødet 2023 spurgte

jeg tilhørerne om de havde hørt om reglen i deres skolegang, og det var ikke tilfældet for de fleste tilstedeværende. Hvis det samme gælder for lægmænd, falder dette fint i tråd med at der er stor variation i sprogbrugen.

Hvis reglen lever og altid har gjort det, ville vi ved singularissubjekt forvente at andelen af de anbefalede former ved sætningskløvning (39 %) ville være større end andelen af anbefalede former ved pseudokløvning (50 %) (se tabel 5 og 6), men det er ikke tilfældet. Tværtimod er andelen af de anbefalede former ved pseudokløvning signifikant større end ved sætningskløvning ($p = 0,02163$). (De tilsvarende andele ved pluralissubjekt er 90 % for sætningskløvning og 88 % for pseudokløvning, og forskellen er ikke statistisk signifikant ($p = 0,7448$)).

Noget kunne altså tyde på at der ikke er mange der kender (eller kendte) og bevidst anvender (eller anvendte) reglen. *Sin* og *sig selv* bruges ganske vist i sætningskløvning, men næppe fordi sprogbrugeren følger en tillært præskriptiv regel. Også i dét lys giver det god mening at indlægge mere frihed i anbefalingerne.

7. Sammenfatning

I denne artikel har jeg behandlet to tilfælde hvor et pronomen refererer til et subjekt på tværs af en sætningsgrænse: i sætningskløvning og pseudokløvning. Sprogbrugen blev undersøgt i to korpusser: (i) MeMo (romaner fra 1870-1899) og (ii) et korpus bestående af landsdækkende aviser (LA). Overordnet set er der stor variation i sprogbrugen ved singularissubjekt. De anbefalede former er ikke særlig dominerende, tværtimod: De ikkeanbefalede former forekommer signifikant mere end de anbefalede. Ved pluralissubjekt er den anbefalede form *deres* derimod enerådende overfor *sin*.

Dette giver anledning til at overveje om Dansk Sprognævn bør ændre sine anbefalinger. Her diskuterede jeg to muligheder: (i) **Alle** de ikkeanbefalede former ophøjes til anbefalede former, og (ii) **kun** de ikkeanbefalede former ved singularissubjekt ophøjes til anbefalede former. Endelig diskuterede jeg om sprogbrugerne kender til reglen om *sin* og *sig selv* ved sætningskløvning og bevidst bruger den.²

2 Tak til den anonyme fagfælle, Jan Heegård Petersen og Katrine Rosendal Ehlers for diskussion og brugbare kommentarer. Også tak til Christian Aaris Lærkedahl Ørnkov for hjælp til de statistiske beregninger.

Om forfatteren

Michael Nguyen, ph.d., forsker, Dansk Sprognævn.

Litteratur

- Baden, Jacob (1789): *Forelæsninger over det danske Sprog, eller resonneret dansk Grammatik*. København: Fr. Brummers Forlag.
- Byskov, Jens (1943): *Dansk Sproglære*. 4. udg. Gjedved – Horsens: Schönberg.
- Christensen, Marie Herget (2019): *Kløvninger - hvad er det det er? Danske sætningskløvningsers struktur, funktion og brug*. København: Københavns Universitet.
- Dansk Sprognævn (u.å.): Pronomener. <https://dsn.dk/wp-content/uploads/2021/01/pronomener-1.pdf>
- Dichmann, Carl Frederik (1800): *Forsøg til en dansk Sproglære*. København: Joh. Fred. Schultz.
- Diderichsen, Paul (1939): *Om pronominerne sig og sin*. København: Ejnar Munksgaard.
- Diderichsen, Paul (1962): *Elementer dansk grammatik*. 3. udg. København: Gyldendal.
- Diderichsen, Philip & Jørgen Schack (2015): Jagten på den gode og sikre sprogbruger. *Nyt fra Sprognævnet* 2015/3, 1-8.
- Dyhr, Mogens (1978): *Die Satzspaltung im Dänischen und Deutschen: eine kontrastive Analyse*. Tübingen: Narr.
- Ehlers, Katrine Rosendal (2024): *Number-sensitive reflexive pronouns in Danish. Optionality, microvariation, and cyclic change*. Aarhus: Aarhus Universitet.
- Hansen, Aage (1965): *Vort vanskelige sprog*. København: Grafisk Forlag.
- Hansen, Erik (1995): Sætningskløvning i moderne dansk. *Danske Studier* 90, 126-146.
- Hansen, Erik & Lars Heltoft (2019): *Grammatik over det danske sprog*. 2. udg. Århus: Syddansk Universitetsforlag.
- Hansen, Hans (1932): *Dansk Sproglære*. København: Gyldendalske Boghandel - Nordisk Forlag.
- Jacobsen, Henrik Galberg & Peter Stray Jørgensen (2013): *Håndbog i nutidens dansk*. 6. udg. Letland: Ordbogen A/S.
- Jensen, Torben Juel (2009): Refleksivt anvendte pronomener i moderne dansk. *Ny forskning i grammatik* 16, 131-151. <https://doi.org/10.7146/nfg.v17i16.23735>

- Lefolii, Hans Henrik (1871): *Sproglærens Grundbegreber, som de komme til Udtryk i Modersmaalets Sprogform: Bidrag til dansk Sproglære*. København: C. A. Reitzel.
- Mikkelsen, Kristian (1975): *Dansk ordføjningslære med sproghistoriske tillæg Håndbog for viderekomne og lærere*. 2. udg. København: Lehmann & Stages Forlag.
- Nølke, Henning (1984): Clefting in Danish? *NyS – Nydanske Studier & Almen kommunikationsteori* 14, 72-112. <https://doi.org/10.7146/nys.v14i14.13332>
- Patten, Amanda L. (2013): *The English it-cleft. A constructional account and a diachronic investigation*. Berlin/Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110279528>
- sproget.dk. Hans, hendes eller sin -uddybning. *sproget.dk*. <https://sproget.dk/raad-og-regler/typiske-problemer/hans-hendes-eller-sin/hans-hendes-sin-uddybning/>.
- Togebj, Ole (2009): Ordkløveri. I: *Dramatikken i grammatikken*. Rita Therkelsen & Eva Skafte Jensen (red.). Roskilde: Roskilde Universitet, 439-455.
- Wiwel, Hylling Georg (1901): *Synspunkter for Dansk Sproglære*. København: Det nordiske Forlag Ernst Bojesen.

Leksikalske affikser

Mads Nielsen, Naja Blytmann Trondhjem &
Kasper Boye

Abstract

Affixes are traditionally regarded as elements that are grammatical as opposed to lexical. Yet some have argued that certain affixes are lexical. We consider the existing conceptions of lexical affixes inadequate and provide a new cross-linguistic definition of lexical affixes based on a theory of lexical status as potentially discourse primary (Boye & Harder 2012) and on a cross-linguistic definition of affixes (Haspelmath 2023). We identify morphemes in West Greenlandic that qualify as lexical affixes according to the definition. The existence of lexical affixes questions the relevance of drawing a line between syntax and morphology. We discuss this and other theoretical as well as practical implications of the study.

Nøgleord

affiks, leksikalsk, grammatisk, polysyntese, vestgrønlandsk

1. Traditionel opfattelse af affikser

Affikser opfattes traditionelt som grammatiske elementer, mens rødder opfattes som leksikalske. Synspunktet er ikke associeret med en bestemt skoledannelse, men kan findes mange steder. I en introduktionsbog til lingvistik står der eksempelvis: "[u]nlige roots, affixes do not belong to a lexical category..." (O'Grady, Achibald & Katamba 2011: 119).

Sondringen mellem leksikon og grammatik forklares ofte med henvisning til åbne og lukkede klasser. Åbne klasser som substantiver, verber, adjektiver og adverbier regnes som leksikalske, mens lukkede klasser som tempusbøjning og kasusmarkering regnes som grammatiske (se Martinet 1960; Bybee, Perkins & Pagliuca 1994; Segalowitz & Lane 2000; Harley 2006).

I det følgende vil vi vise hvordan den traditionelle opfattelse af affikser som grammatiske er blevet udfordret af forskellige forfattere, indtil videre

på utilstrækkelig vis. Herefter vil vi give vores bud på hvordan leksikalske affikser kan identificeres, og illustrere dette med eksempler fra vestgrønlandsk. Til sidst vil vi se nærmere på nogle af de implikationer definitionen af leksikalske affikser har for sprogbeskrivelsen og for sprogbrugere i praksis.

2. Tidligere behandlinger af leksikalske affikser

Andre før os har argumenteret for at ikke alle affikser uden videre kan betragtes som grammatiske. Harley (2006) kalder fx engelske afledningsaffikser som *-fy* og *-ly* for 'leksikalske affikser' fordi de danner ord som hører til i åbne ordklasser, som hun betragter som leksikalske. Omvendt foreslår Hopper og Traugott (2003) at afledningsaffikser som *un-* og *-ling* er leksikalske fordi de ikke ændrer ordklasse, dvs. de hører til i de åbne, leksikalske paradigmer. De to bud på hvad der i engelsk gælder som leksikalske affikser, er uforenelige, men de er begge udtryk for en fornemmelse af at alle affikser ikke kan slås over en kam og regnes som grammatiske.

I beskrivelsen af en række wakashanske, salishanske og chimakuanske sprog, der tales i det nordvestlige Nordamerika på grænsen mellem USA og Canada, er der tradition for at kalde nogle affikser leksikalske. Disse affikser er af en helt anden type end de engelske idet de med Langackers (1987) terminologi har ting-agtige betydninger ligesom substantiver i sprog som dansk og engelsk. I (1) ses et eksempel på et sådant affiks i halkomelem, nemlig *-əyət*, der betyder 'baby'.

- (1) Halkomelem, salishansk (fra Gerdts 2003: 347)
- | | | | |
|-----------------------------|-------------------|-----------|-------------|
| <i>niʔ</i> | <i>ʃkʷ-əyət</i> | <i>tə</i> | <i>Mary</i> |
| AUX | bade- baby | DET | Marie |
| 'Marie badede baby ' | | | |

Det er værd at bemærke at (1) er en intransitiv sætning, og at *-əyət* altså ikke fungerer som objekt til verbet. Hvis begivenheden i (1) skal udtrykkes i en transitiv sætning, vil det kræve at verbet forsynes med et såkaldt 'transitivt' suffiks og en ergativmarkering, og i stedet for suffikset *-əyət* ville man bruge den frie rod *qeq*, der også betyder 'baby', til at betegne objektreferenten. Konstruktionen i (1) minder om inkorporering idet suffikset ikke har argumentstatus, hvilket i den danske oversættelse svarer til at læse *badede baby* som et komplekst prædikat med enhedstryk (jf. Nedergaard Thomsen 1996).

Halkomelem og de fleste andre salishanske sprog har over 100 suffikser som *-əyət*. For mange af disse suffiksers vedkommende eksisterer der rødder der er semantisk parallelle. Her er tre andre eksempler på suffiks-rodpar i halkomelem (Gerdts 2003: 346):

<i>-cəs</i>	<i>celəš</i>	'hånd'
<i>-šən</i>	<i>sšənə</i>	'fod'
<i>-wit</i>	<i>ləwəx</i>	'ribben'

Som det fremgår af eksemplerne, kan der være fonologiske ligheder mellem suffikserne og rødderne, og det har fået bl.a. Gerdts (2003) til at foreslå at suffikserne har udviklet sig fra frie rødder anvendt i inkorporeringskonstruktioner. Tolkningen underbygges af at suffikserne bruges med lidt bredere betydning end rødderne, hvilket er forventeligt hvis førstnævnte er udviklet fra sidstnævnte ved grammatikalisering. Således betyder *-əyət* også 'barn', *-cəs* 'finger', *-šən* 'ben', og *wit* 'beholder'. I de tilfælde hvor der ikke er fonologisk lighed mellem suffiks og rod, som fx *-əyət* og *qeq*, kan det være fordi nogle af de oprindelige rødder har mistet deres evne til at forekomme i fri form, og sprogbrugerne med tiden er begyndt at bruge nye rødder med beslægtede betydninger som frie former (Gerdts 2003: 355).

Denne forståelse af leksikalske affikser kan for så vidt være god som udgangspunkt for at identificere dem sprogspecifikt, men den kan ikke bruges til tværsproglig sammenligning fordi rod-affiks-korrespondenser typisk er sprogspecifikke. Det er dog ikke alle de affikser der er blevet kaldt leksikalske, der har semantisk parallelle rødder inden for sproget. I beskrivelsen af affikserne skeler mange til hvad der udtrykkes med rødder i andre sprog, typisk analytiske, nærmere bestemt dem der undertiden (fx Whorf 1960 [1941]; Haspelmath 2001) kaldes *Standard Average European* (SAE). Affikser med semantisk parallelle rødder i SAE siges at have rodliggende betydning (se fx Mithun 1997; Gerdts 2003; Mattissen 2017). Eksemplerne herunder viser to affikser med verbum-lignende betydninger:

- (2) *-sae* 'direktiv' i yaqui, uto-aztekisk (fra Escalante 1990: 89)
- | | | |
|------------------------------|-----------------|------------------------|
| <i>Peo</i> | <i>Maria-ta</i> | <i>bwiik-sae-n</i> |
| Peo | Maria-ACC | syng-DIREKTIV-PST.IMPF |
| 'Peo bad Maria synge' | | |

- (3) *-bal* 'ambulatív' i thulung, sino-tibetansk (fra Lahaussais 2020: 73)
- | | | | |
|-----------|------------|--------------------------------------|-------------|
| <i>me</i> | <i>pel</i> | <i>sam-bal-lu</i> | <i>ʔe</i> |
| DIST.DEM | skygge | fange- AMBULATIV -3SG>3SG.PST | RAPPORTERET |
- 'han **gik** vist og fangede den der skygge'

Mattissen (2017) kalder affikser med rodliggende betydning for *non-root bound morphemes*, og hun anser dem som karakteristiske for polysyntetiske sprog. Affikserne angiver typisk betydninger som bevægelse, retning, måde, grad og modalitet (p. 72). Mattissen fortæller imidlertid ikke hvorfor hun mener at rødder i mere analytiske sprog som SAE er den rette målestok for hvilke betydninger der skal regnes for leksikalske på tværs af sprog. Og der er i hvert fald to gode grunde til ikke at sætte lighedstegn mellem SAE-rødder og leksikalske tegn. For det første er det ikke alle rødder i SAE der anses for leksikalske. Eksempelvis omtales hjælpeverber og artikler ofte som grammatiske. For det andet, og dette er et mere vægtigt argument, er konceptuel betydning et dårligt kriterie for at skelne leksikalsk og grammatisk status idet den samme konceptuelle betydning undertiden kan være kodet grammatisk i et tegn og leksikalsk i et andet tegn. I dansk er 'fortid' fx kodet grammatisk i tempusbøjningen *-ede* og leksikalsk i adverbialiet *i går*, og vi kan tage epistemisk forbehold med både en leksikalsk matrixsætning som *jeg tror* og en grammatisk modalpartikel som *vel*.

Fordi litteraturen ikke tilbyder en definition af leksikalske affikser der kan bruges tværsprogligt uden skelen til SAE som standard, vil vi i det følgende foreslå en ny definition af leksikalske affikser, der forhåbentlig kan være et skridt på vejen mod en fælles forståelse af begrebet.

3. Leksikalske affikser kan modificeres

3.1. Definition

Vores definition af leksikalske affikser kombinerer Boye og Harders (2012) definition af leksikalsk status med Haspelmaths (2023) tværsproglige definition af affikser. Boye og Harders teori behandler leksikon og grammatik som et resultat af opmærksomhedsprioritering. Mennesker kan ikke være lige opmærksomme på alle betydningsdele i en kompleks ytring, ligesom vi ikke kan rette lige meget opmærksomhed mod alle elementer i et billede. Derfor må vi prioritere så der i ytringer opstår en skelnen mellem elementer der er diskursivt primære (opmærksomhedsmæssig forgrund i forhold til syntagmatisk relaterede elementer), og elementer der er diskursivt

sekundære (opmærksomhedsmæssig baggrund i forhold til syntagmatisk relaterede elementer). Denne prioritering kan konventionaliseres således at der opstår elementer der pr. konvention er diskursivt sekundære og ikke kan anvendes som primære (uden for metasproglige kontekster som *jeg snakker om løsningEN, ikke om løsningER*, hvor konventionen brydes). Sproglige elementer der pr. konvention er opmærksomhedsmæssig baggrund, er grammatiske, og på grund af deres inhærente diskursivt sekundære status kan de ikke udbygges gennem modifikation. Leksikalske elementer derimod kan være diskursivt primære, og derfor kan de modificeres. Et eksempel på et leksikalsk element er *rød*, som kan modificeres af *meget*, mens de ubestemte artikler *en* og *et* på grund af deres grammatiske status ikke kan modificeres. Vi foreslår derfor at bruge modificerbarhed til empirisk at identificere elementer med leksikalsk status (se også Bybee, Perkins & Pagliuca 1994: 7; Keizer 2007; Messerschmidt et al. 2018).

Haspelmath (2023) giver hvad der efter vores mening er det bedste eksisterende bud på en tværsproglig definition af affikser. Ifølge Haspelmath er et affiks et bundet morfem som ikke er en rod, som skal forekomme på en rod (evt. med andre affikser), og som ikke kan forekomme på rødder af forskellige klasser (p. 286). Affiksdefinitionen forudsætter en definition af rod, som Haspelmath også giver os: En rod er et morfem der betegner en handling, en genstand eller en egenskab – det kalder Haspelmath et indholdsrigt ("contentful") morfem – og som kan forekomme som en del af en fri form uden et andet indholdsrigt morfem (p. 287). Når Haspelmath betragter handlings-, genstands- og egenskabsbetydninger som indholdsrige, er det fordi Croft (2001) har argumenteret for at disse betydninger kan bruges til at identificere ordklasserødder på tværs af sprog. Verbumrødder betegner handlinger, substantivrødder betegner genstande, og adjektivrødder betegner egenskaber. Som Haspelmath selv bemærker, er hans definition imidlertid ikke uproblematisk og muligvis ikke synderlig interessant. Den diskussion tager vi op i afsnit 5.

Idet vi kombinerer definitionerne i Boye og Harder (2012) og Haspelmath (2023), foreslår vi følgende tværsproglige definition af leksikalske affikser:

Tværsproglig definition af leksikalske affikser

Leksikalske affikser er bundne, rodklasseafhængige ikke-rodsmorfemer der pr. konvention er potentielt diskursivt primære.

3.2. *Leksikalske affikser fra litteraturen*

Ud fra kriteriet om modificerbarhed kan vi umiddelbart konstatere at de engelske affikser som Hopper og Traugott (2003) foreslog var leksikalske, i virkeligheden er grammatiske: *Un-* og *-ling* kan ikke modificeres uafhængigt af stammen. I *ugly duckling* udbygger *ugly* ikke diminutivbetydningen som *-ling* tilføjer til *duck*.

Harleys (2006) forslag om at de ordklasseændrende affikser i engelsk er leksikalske, er en mere kompliceret sag. Fordi affikserne ændrer ordklasse, tillader de modifikation som ellers ikke tillades af stammen. Når *friendly* kan modificeres af *extremely* er det således muliggjort af *-ly*, jf. at *friend* ikke kan modificeres af *extremely*. Af den grund kunne man fristes til at karakterisere *-ly* som leksikalsk. Men dette ville være udtryk for en syntaktisk analyse og ikke en semantisk. I kognitiv lingvistisk forstand ”profilerer”, dvs. fremhæver, *-ly* et egenskabsaspekt ved konceptet FRIEND (jf. Langacker 1987). I denne proces kommer det afledte ord *friendly* til at tilhøre den distributionelle klasse af adjektiver, og denne klasse kan modificeres af adverbier som fx *extremely*. Men det er ikke egenskabsaspektet i sig selv som *extremely* udbygger. *Extremely friendly* betyder nemlig ikke at noget er egenskabsagtigt i ekstrem grad, men at nogen manifesterer konceptet FRIEND i ekstrem grad.

Det er dog muligt at andre engelske afledningsaffikser kan modificeres semantisk uafhængigt af stammen. Et eksempel er *-able*. *Extremely breakable* betyder at noget i ekstrem grad er i stand til at blive ødelagt, ikke at noget er i stand til at blive ekstremt ødelagt. Her lader det altså til at det er den dynamisk modale betydning af *-able* (‘være i stand til’) og ikke et aspekt ved konceptet BREAK som muliggør modifikationen af *extremely*. Tilsvarende kan man i dansk analysere *-bar* som leksikalsk for så vidt som *ekstremt brugbar* betegner et ekstremt potentiale, ikke en ekstrem brug. Andre danske affikser som vi mistænker kan modificeres, er *-mager*, *-vorn*, *eks-* og *sam-*.

I forhold til de omtalte affikser i wakashanske, salishanske og chimakuanske er det for nuværende uafklaret om de er leksikalske ifølge vores definition. Mithun (1997: 363-364) giver dog eksempler på at man i det salishanske sprog nuxalk bruger frie rødder til at introducere referenter og semantisk relaterede suffikser til at omtale samme referenter efterfølgende. Det at suffikserne ikke bruges til at introducere ny information, kunne tyde på at de er diskursivt sekundære og dermed grammatiske.

4. Identifikation af leksikalske affikser i vestgrønlandsk

4.1. Vestgrønlandsk

For at finde eksempler på leksikalske affikser vil vi se til vestgrønlandsk, som tales på Grønlands vestkyst og betragtes som standardsproget i landet. Sproget har 300-400 produktive affikser (Fortescue 1980) og fremhæves ofte som et prototypisk eksempel på et polysyntetisk sprog (Mattissen 2017; Trudgill 2017). Alle de vestgrønlandske affikser vi kommer til at se nærmere på i det følgende, falder under Haspelmaths (2023) affiksdefinition: De forekommer på rødder af én klasse, og selv om nogle af dem betegner handlinger og derfor er det Haspelmath kalder indholdsrige morfemer, er de ikke selv rødder idet de ikke kan forekomme uden et andet indholdsrigt morfem.

4.2. Fremgangsmåde

Indledningsvist har vi konsulteret Fortescues (1980) oversigt over vestgrønlandske affikser og fundet affikser med betydninger som vi umiddelbart vurderede ville kunne modificeres af andre affikser. En af artiklens forfattere, som taler vestgrønlandsk som modersmål, har derefter på baggrund af sin intuition vurderet om mistanken var korrekt. For at underbygge den intuitive vurdering har vi for flere af affikserne fundet autentiske eksempler med entydige læsninger i et online skriftsprogskorpus hostet af Oqaasileriffik, Grønlands Sprogsekretariat (<https://oqaasileriffik.gl/da/sprogteknologi/korpus/>). Eksemplerne i det følgende afsnit stammer fra dette korpus. Vi arbejder på at finde autentiske eksempler på alle leksikalske affikser i vestgrønlandsk.

Målet med undersøgelsen er at finde ud af om der i det hele taget findes leksikalske affikser i vestgrønlandsk, ikke at lave en kortlægning af leksikalske og grammatiske affikser i sproget. Vi har desuden kun set på verbale affikser.

4.3. Eksempler

Vi har indtil videre identificeret 22 vestgrønlandske affikser som kan modificeres. I (4) ses et eksempel på at *-tuar* 'hele tiden' modificerer *-nerar* 'sige', som dermed kan betragtes som leksikalsk.

- (4) *pinngitsuunerartuarsinnarpog*
pinngitsuu-nerar-tuar-sinnar-voq
 uskyldig-**sige-hele.tiden**-bare-3SG.DECL
 'han siger bare hele tiden at han er uskyldig'

Det er utvetydigt at *-tuar* modificerer *-nerar* og ikke stammen *pinngitsuu* 'uskyldig'. Hvis *-tuar* modificerede stammen, ville sætningen betyde 'han siger at han hele tiden er uskyldig', og det er ikke en mulig læsning af (4). Stammen modificeres altså af *-nerar* og *-tuar* tilsammen. Både *-sinnar* og *-voq* modificerer alt det forudgående.

Det næste eksempel viser at *-pallag* 'hurtigt' kan modificeres af *-ngaar* 'meget':

- (5) *neripallangaarmallu*
neri-pallag-ngaar-mat=lu
 spise-**hurtigt-meget**-3SG.CAUS=og
 'og da hun spiste meget hurtigt'

Det er kun muligt at forstå eksemplet således at hun spiste på en meget hurtig måde, og ikke at hun spiste meget på en hurtig måde. Affikset *-ngaar* udbygger mådesbetydningen 'hurtigt', ikke begivenhedsbetydningen 'spise'.

I både (4) og (5) ser vi altså at affikserne danner komplekse enheder som modificerer stammerne, fuldstændig som når vi er vant til at beskrive den hierarkiske opbygning af fraser som fx *hun fandt en tier*, hvor *en tier* udgør et komplekst led i sætningen. Det er altså ikke sådan at affikser nødvendigvis modificerer hele den stamme de sættes udenpå (se også Fortescue 1980: 261-262). Det vender vi tilbage til i afsnit 5.

4.4. Oversigt

I tabel 1 ses en oversigt over de 20 leksikalske affikser vi har fundet, og deres placering i vestgrønlandske verbaler. Verbalskemaet og affiksernes placering er baseret på Fortescue (1980: 261).

Table 1 Leksikalske affikser og deres placering i vestgrønlandske verbaler.¹

+1 verbalisering	+2 udvidelse	+4 modifikation
-erut 'fratage'	-gunar 'ligne'	-jaar 'tidligt'
-iar 'fjerne'	-lersaar 'agte'	-juannar 'vedvarende'
-ler 'forsyne med'	-nerar 'sige'	-lerlor 'hurtigt (på kort tid)'
-liar 'gå hen til'	-niar 'prøve'	-neru 'mere'
-si 'få, købe, servere'		-pallag 'skynde sig'
-sior 'lede efter'		-qqissaar 'forsigtigt, nøje'
-sior 'fejre'		-riataar 'pludselig'
-tar 'samle'		
-tor 'bruge, spise'		

Det er ikke sådan at alle affikser på de pladser hvor vi har fundet leksikalske affikser, er leksikalske. På plads +1 kan man fx også bruge *-nngor* 'blive' og *-u* 'være', som ikke kan modificeres, og på plads +2 og +4 kan man bruge hhv. det passiverende suffiks *-neqar* og *-kaa* 'i flok', som heller ikke kan modificeres. På den samme plads kan der altså både forekomme grammatiske og leksikalske affikser. Tilsvarende heterogenitet inden for det danske præpositionsparadigme er beskrevet af Messerschmidt et al. (2018), som viser at bl.a. *før* og *over* er leksikalske fordi de kan modificeres (jf. **kort før** ham sang hun og han cyklede **lige over** marken), mens *før* og *af* er grammatiske fordi de ikke kan modificeres (jf. **kort før* ham sang hun og **han blev kysset lige af* hende).

Når vi for mange affikser ikke har fundet belæg for at de kan modificeres, kan det skyldes forskellige ting. Dels har to af artiklens forfattere ikke modernmålsintuition, hvilket kan gøre det svært at finde på måder at modificere affikser på. Dels er det i mange tilfælde svært at komme med eksempler hvor modifikation er pragmatisk meningsfuld. Det er ikke sikkert at de affikser som vi ikke har fundet passende modifikationsmuligheder for, er grammatiske. Bestemmelse af grammatisk status er nemlig mere usikker end bestemmelse af leksikalsk status idet man kan finde positivt belæg for at et givent element er leksikalsk (her: tilstedeværelsen af modifikati-

1 På følgende pladser har vi ikke fundet leksikalske affikser: +3 (negation), +5 (tid), +6 (epistemicitet), +7 (negation), +8 (subjektiv farvning), +9 (konjunkional), +10 (illokution/underordning + kongruens). Pladserne er baseret på Fortescue (1980), som dog ikke bestemmer konjunkionalernes plads entydigt, men fordi konjunkionaler altid skal følges umiddelbart af en underordningsmarkør, har vi placeret dem lige inden underordningsmarkørernes plads.

onsmulighed), mens grammatisk status kun kan bestemmes på et negativt grundlag (her: fraværet af modifikationsmulighed). Denne asymmetri er bygget ind i Boye og Harders (2012) grammatikdefinition: Grammatiske elementer kan ikke være diskursivt primære. For en del affikser som vi ikke har fundet modifikationsmuligheder for, gælder det dog uden tvivl at de er grammatiske, og at det er derfor de ikke kan modificeres.

5. Perspektiver og implikationer

Identifikationen af leksikalske affikser i vestgrønlandsk har forskellige teoretiske og praktiske implikationer. Allererst kan nævnes at vi har set at der på den samme affiksplads i vestgrønlandske verber kan forekomme både grammatiske og leksikalske affikser. Samme paradigmeinterne heterogenitet er, som nævnt i afsnit 4.4 ovenfor, beskrevet for danske præpositioner (Messerschmidt et al. 2018). Heterogeniteten inden for paradigmer er uforenelig med traditionens lighedstegn mellem leksikon og åbne klasser og mellem grammatik og lukkede klasser. Grammatisk og leksikalsk status er paradigmeafhængigt.

Forekomsten af leksikalsk og grammatisk materiale på den samme plads i vestgrønlandske verber må dog tages med det forbehold at pladserne måske ikke er så veldefinerede som de tager sig ud i verbalkemaet i tabel 1. Skemaet er som nævnt afledt af Fortescues (1980: 261) frasestrukturregler, men som Fortescue selv bemærker, er rækkefølgen af visse affikser i vestgrønlandsk ofte bestemt af skopusforhold (p. 259-260). Således betyder (6) og (7) to forskellige ting:

(6) *suleqqilersaarpoq*
suli-qqig-lersaar-voq
 arbejde-igen-agte.at-3SG.DECL
 'hun agter at arbejde igen'

(7) *sulilersaaqqippooq*
suli-lersaar-qqig-voq
 arbejde-agte.at-igen-3SG.DECL
 'hun agter igen at arbejde'

I (6) modificeres *suli* 'arbejde' først af *-qqig* 'igen', og deres samlede betydning modificeres herefter af *-lersaar* 'agte at'. Subjektreferenten har altså arbejdet før og agter nu at gøre det igen. I (7) modificeres *suli* først af

-lersaar, og deres samlede betydning modificeres herefter af *-qqig*. Her har subjektreferenten tidligere agtet at arbejde, men ikke nødvendigvis arbejdet. Eksemplerne er interessante i distributionsojemed fordi *-qqig* ifølge Fortescue hører til på plads +4, og *-lersaar* på plads +2. Det burde altså kun være (7) der er tilladt.

Men hvor *-qqig* og *-lersaar* altså kan forekomme i forskellige indbyrdes rækkefølger, har andre affikser i vestgrønlandsk en fast placering. Tempusaffikser kommer eksempelvis altid efter gradsaffikser, epistemiske affikser efter tempusaffikser, og illokutionære affikser kommer altid til sidst. Disse faste rækkefølgeforhold afspejler imidlertid også skopusrelationer: Ilokutionære affikser opererer på en illokution og har skopus over epistemiske affikser, der operer på en proposition og selv har skopus over tempusaffikser der opererer på et sagforhold (fx Boye 2012: kap. 4).

Tilstedeværelsen af leksikalske affikser har praktiske implikationer for talere af vestgrønlandsk for så vidt som affiksernes opmærksomhedspotentiale må forventes at vise sig i processeringen af dem. Fordi leksikalske affikser i modsætning til grammatiske er potentielt diskursivt primære, vil vi eksempelvis i et ændringsblindhedsforsøg forvente at læsere vil være mere tilbøjelige til at opdage når leksikalske affikser ændres, end når grammatiske affikser gør det, og i en bogstavssøgningsopgave vil vi forvente at læsere vil være dårligere til at opdage bogstaver i grammatiske end i leksikalske affikser. Leksikalsk status kan også have betydning for hvordan talere med grammatiske vanskeligheder som agrammatisme udtrykker sig. Bau og Boye (2023) har fundet at engelsktalere med agrammatisme bruger færre grammatiske end leksikalske præpositioner, og vi forventer at finde noget tilsvarende for fakultative affikser i en igangværende undersøgelse af grønlandsktalere med agrammatisme.

Af implikationer med mere teoretisk end praktisk relevans kan det nævnes at vores forståelse af leksikalske affikser muligvis indebærer en empirisk efterprøvelig test af polysyntese. Mattissen (2017) har foreslået at det afgørende kriterie for polysyntese er tilstedeværelsen af leksikalske affikser. For så vidt som dette forslag er gangbart, kan vores procedure for identifikation af leksikalske affikser tjene som procedure for identifikation af polysyntetiske sprog. Inden forslaget antages, ville det imidlertid være oplagt at undersøge i hvilken grad tilstedeværelsen af leksikalske affikser, som vi har defineret dem, falder sammen med de intuitive klassifikationer af sprog som polysyntetiske som findes i litteraturen.

I det videre arbejde med identifikationen af leksikalske affikser vil det

være interessant at lave en decideret kortlægning af distributionen af leksikalske og grammatiske affikser, ikke bare i vestgrønlandsk, men også i andre sprog. Sådanne kortlægninger ville gøre det muligt at se eventuelle mønstre på tværs af sprog. Ifølge Hengeveld (2017) er grammatikalisering ofte forbundet med skopusudvidelse, og fordi skopus ofte er forbundet med rækkefølge, kunne man forestille sig at grammatiske affikser alt andet lige forekommer længere væk fra roden end leksikalske affikser. Princippet kan illustreres med en sætning som *hun har da virkelig oppet sig*, hvor det grammatiske sætningsadverbial *da* står længere fra fuldverbet end det leksikalske adverbial *virkelig*. Man må se for sig at et led i grammatikaliseringen af et element er at elementet kommer på afstand af kernen (eller at det er elementer der allerede er på afstand af kernen, som grammatikaliseres). Umiddelbart ser tabel 1 ud til at underbygge tesen idet vi indtil videre kun har fundet leksikalske affikser på tre af de fire første pladser, mens vi ikke har fundet nogen på de sidste fem pladser. Men tabellen bygger som nævnt ikke på en gennemtrawling af alle vestgrønlandske affikser. En vurdering af hvor mange leksikalske affikser der findes på de forskellige pladser, må desuden ses i relation til det totale antal affikser der kan forekomme på samme pladser.

Det sidste perspektiv på undersøgelsen som vi vil trække frem her, er spørgsmålet om hvorvidt, og i givet fald i hvilket omfang, det er hensigtsmæssigt at skelne mellem fraser og ord og dermed mellem syntaks og morfologi. Haspelmath (2023) adresserer det samme problem umiddelbart efter at have defineret begreberne ord og affiks i overensstemmelse med den lingvistiske tradition: "it could be that the notions of 'word', 'morphology' and 'syntax' are stereotypes derived from our tradition that have no actual basis in the reality of languages" (p. 295). Fraser og ord antages at adskille sig på en række punkter, men i hvert fald nogle af forskellene opløses i lyset af et teoretisk forankret begreb om leksikalske affikser.

En første sådan forskel har at gøre med bestanddelene i fraser og ord. Simple fraser (dvs. fraser der ikke indeholder andre fraser) anses for at bestå af et hoved omgivet af materiale der kan være både leksikalsk og grammatisk, nemlig ord (jf. sondringen mellem leksikalske ord som fx fuldverber og grammatiske ord som fx hjælpeverber). I modsætning her til anses simple ord (dvs. ord der ikke er sammensatte, og som altså ikke indeholder andre ord) for at bestå af et hoved omgivet af materiale der kun kan være grammatisk, nemlig affikser (jf. afsnit 1 om den udbredte opfattelse af affikser som grammatiske). Denne forskel forsvinder når

leksikalske affikser tages i betragtning: I både fraser og ord kan hovedet ledsages af både grammatisk og leksikalsk materiale.

En anden forskel har at gøre med strukturelt hierarki. Hvor fraser kan have komplekse konstituent, regner man normalt ikke med at ikke-sammensatte ord har det. Hvor der fx er gode argumenter for at regne *en tier* som en kompleks konstituent inden for frasen *hun fandt en tier*, er der ingen argumenter for at regne affikskombinationen *-er-ne* for en kompleks konstituent i *bilerne*. Det skyldes imidlertid at både *-er* og *-ne* er grammatiske. Når man tager leksikalske affikser i betragtning, bliver det klart at affikskombinationer kan være komplekse ordkonstituent. Vi har allerede set at dette gælder for nogle affikskombinationer inden for vestgrønlandske ord, fx i (4), hvor *-nerar* og *-tuar* danner et affikskompleks der til sammen modificerer *pinngitsuu*.

Eksistensen af leksikalske affikser underminerer altså i hvert fald nogle grunde til på tværs af sprog at skelne mellem fraser og ord og mellem syntaks og morfologi, og dermed også til at opfatte affiks som en tværsproglig kategori.

Om forfatterne

Mads Nielsen, ph.d.-stipendiat, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.

Naja Blytmann Trondhjem, ph.d., lektor, Institut for Tværkulturelle og Regionale Studier, Københavns Universitet.

Kasper Boye, ph.d., lektor, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.

Litteratur

Bau, Naja & Kasper Boye (2023): Præpositioner som heterogen ordklasse.

En afasiologisk undersøgelsen af sondringen mellem grammatiske og leksikalske præpositioner i engelsk. *Ny forskning i grammatik* 30, 5-23. <https://doi.org/10.7146/nfg.v1i30.137950>

Boye, Kasper (2012): *Epistemic meaning: A functional-cognitive study*. Berlin/Boston: de Gruyter Mouton.

Boye, Kasper & Peter Harder (2012): A usage-based theory of grammatical status and grammaticalization. *Language* 88(1), 1-44. <https://doi.org/10.1353/lan.2012.0020>

- Bybee, Joan, Revere Perkins & William Pagliuca (1994): *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Croft, William (2001): *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Escalante, Fernando (1990): *Voice and argument structure in Yaqui*. Ph.d.-afhandling, University of Arizona, Tucson.
- Fortescue, Michael (1980): Affix ordering in West Greenlandic derivational processes. *International journal of American linguistics* 46(4), 259-278.
- Gerdts, Donna B. (2003): The morphosyntax og Halkomelem lexical suffixes. *Internaitonal Journal of American Linguistics*, 69(4), 345-356.
- Harley, Heidi (2006): *English words: A linguistic introduction*. New York: John Wiley & Sons.
- Haspelmath, Martin (2001): The European linguistic area: Standard Average European. I: *Language typology and language universals*. Vol. 2. Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (red.). Berlin/Boston: de Gruyter Mouton, 1492-1510. <https://doi.org/10.1515/9783110194265-044>
- Haspelmath, Martin (2023): Defining the word. *WORD* 69(3), 283-297. <https://doi.org/10.1080/00437956.2023.2237272>
- Hengeveld, Kees (2017): A hierarchical approach to grammaticalization. I: *The grammaticalization of tense, aspect, modality and evidentiality*. Kees Hengeveld, Heiko Narrog & Hella Olbertz (red.). Berlin/Boston: de Gruyter Mouton, 13-37.
- Hopper, Paul & Elizabeth Traugott (2003): *Grammaticalization* (2. udg.). Cambridge: Cambridge University Press.
- Keizer, Evelien (2007): The grammatical-lexical dichotomy in Functional Discourse Grammar. *Alfa – Revista de Lingüística* 51(2), 35-56.
- Lahaussais, Aimée (2020): Predicate derivations in Thulung. *Journal of South Asian Languages and Linguistics* 7(1), 57-84. <https://doi.org/10.1515/jsall-2021-2024>
- Langacker, Ronald W. (1987): *Foundations of cognitive grammar. Theoretical prerequisites*. Vol. 1. Stanford, CA: Stanford University Press.
- Martinet, André (1960): *Éléments de linguistique générale*. Paris: Armand Colin.
- Mattissen, Johanna (2017): Sub-types of polysynthesis. I: *The Oxford handbook of polysynthesis*. Micwhael Fortescue, Marianne Mithun & Nicholas Evans (red.). Oxford: Oxford University Press, 70-98.

- Messerschmidt, Maria, Kasper Boye, Mathilde Marie Overmark, Sofie Theilmann Kristensen & Peter Harder (2018): Sondringen mellem grammatiske og leksikalske præpositioner. *Ny forskning i grammatik* 25, 89-106. <https://doi.org/10.7146/nfg.v26i25.109304>
- Mithun, Marianne (1997): Lexical affixes and morphological typology. I: *Essays on language function and language type*. Joan Bybee, Johan Haiman & Sandra A. Thompson (red.). Amsterdam/Philadelphia: John Benjamins, 357-371.
- Nedergaard Thomsen, Ole (1992): Unit accentuation as an expression device for predicate formation. The case of syntactic noun incorporation in Danish. I: *Layered structure and reference in a functional perspective*. Michael Fortescue, Peter Harder & Lars Kristoffersen (red.). Amsterdam/Philadelphia: John Benjamins, 173-229.
- O'Grady, William, John Archibald & Francis Katamba (2011): *Contemporary linguistics* (2. udg.). Harlow: Longman – Pearson Education.
- Segalowitz, Sidney & Korri Lane (2000): Lexical access of function vs content words. *Brain and Language* 75, 376-389.
- Trudgill, Peter (2017): The anthropological setting of polysynthesis. I: *The Oxford handbook of polysynthesis*. Michael Fortescue, Marianne Mithun & Nicholas Evans (red.). Oxford: Oxford University Press, 186-202.
- Whorf, Benjamin Lee (1960) [1941]: The relation of habitual thought and behavior to language. I: *Language, Culture, and Personality: Essays in Memory of Edward Sapir*. Leslie Spier, A. Irving Hallowell & Stanley S. Newman (red.). Salt Lake City, UT: University of Utah Press, 75-13.

Det danske indirekte objekt i tre tekster fra 1500-tallet

Peter Juul Nielsen

Abstract

Since the Middle Ages, the Danish indirect object (IO) has undergone a narrowing of its potential in terms of semantic role and its ability to occur without a valence slot provided by a 3-place verb, i.e. the possibility of free IOs. The study examines a particular stage in this diachronic process: the IOs in 16th c. Danish. The study is based on three religious texts. It finds that about half the IOs designate Recipients, which is a lower share than what has been found in a study of 17th c. Danish IOs. The 16th c. IOs include around 10% free IOs, which is noticeably higher than in the 17th c. data. The results of the study provide empirical underpinning for the theoretical account of the IO's narrowing process.

Nøgleord

indirekte objekt, dansk sproghistorie, semantiske roller, valens

1. Indledning

Igennem historien er mulighederne for konstruktion med indirekte objekt (IO) i dansk blevet indskrænket, og IO's semantiske potentiale har undergået en specialisering. Som bidrag til analysen af denne diakrone proces har jeg i Nielsen (2021) foretaget en synkron undersøgelse af IO i det sene 1600-tal repræsenteret ved Leonora Christinas Jammers Minde. I denne artikel følger jeg samme spor og slår ned på den foregående periode med en undersøgelse af IO i 1500-tallets dansk.

IO er defineret som argument 3, dvs. et tredje nominalt argument i tillæg til subjekt og direkte objekt (DO) således at en IO-konstruktion betegner et sagforhold hvori der indgår tre partcipanter (jf. GDS: 1304-1306; Nielsen & Heltoft 2023; Heltoft 2019: 151). I moderne dansk findes IO næsten udelukkende ved trivalente verber med en valensplads til de tre nominale argumenter subjekt, DO og IO. IO's karakteristiske semantiske rolle er

modtager (1)¹ (jf. Hansen & Heltoft 2011 (herefter GDS): 1305-1306, 1310).

(1) Olga gav **Jokum** en bog

I de ældre perioder af dansk er karakteristikkene af IO anderledes. De træk ved IO der igennem historien adskiller sig fra de moderne træk, skitseres i følgende gennemgang af IO's udvikling.

1.1. Hovedtræk i IO's historiske udvikling

De store linjer i IO-udviklingen fra middeldansk til nutidsdansk kan beskrives som en indsnævring af IO's semantiske potentiale (Heltoft 2014, 2019; Nielsen 2021). I middeldansk har IO den brede rolle interessant, dvs. en participant der står i et interesseforhold til verbalhandlingen.² Rollen modtager er en central variant, men det er karakteristisk at det ikke er nødvendigt for IO-konstruktionen at betegne en overdragelse med IO som modtager. Der er vide muligheder for fri tilføjelse af IO med ren interessentrolle uden valensstyring, jf. (2) hvor IO *henne* 'hende' ikke betegner en modtager af hjulet, men en referent med en anden form for interesse eller involvering, her den involvering at skulle tortureres på hjulet (Heltoft 2019: 161).

(2) Vrbanus (...) lot **henne** eth hiull giøræ
'Urbanus lod gøre et hjul til hende'
(fra Kristina-legenden, ca. 1480, De hellige Kvinder, 43, 35-36)

I 1700-tallets dansk er det ifølge Høysgaard (1752: 108) stadig muligt frit at tilføje IO med interessentrolle uden overdragelsessemantik, fx *at åbne nogen en dør*. I 1800-tallet indsnævres IO-semantikken omkring overdragelsesbetydning (Heltoft 2014). Frit IO med indholdet intenderet modtager ved anskaffelses- og fremstillingsverber er stadig muligt (3), men de rene interessenter er faldet bort som del af den generelle norm (4).

(3) Jeg byggede **hende** en Lykke-Verden (Wilhelm Bergsøe, Fra den gamle Fabrik (1869), 333)³

1 I alle eksempler er IO fremhævet med fed.

2 Rollen interessant kan ses som overbegreb til 'beneficiary' og 'maleficiary', og interesse skal dermed ikke forstås sådan at IO-referenten nødvendigvis er interesseret i handlingen; handlingen kan også være til gene for IO-referenten.

3 Efter gengivelsen hos Arkiv for Dansk Litteratur: <https://tekster.kb.dk/text/adl-texts-bergsoc02-root> (lokaliseret 30.4.24).

(4) *jeg åbnede **hende** døren

I nutidsdansk ses en stadigt tydeligere indsnævring omkring overdragelsessemantik. Fra det 20. århundrede er det frie IO desuden underkastet et krav til den pragmatiske kontekst: Frit IO ved anskaffelses- og fremstillingsverber kræver regulativ sproghandlingskontekst, jf. (5)-(6) (Nielsen & Heltoft 2023; Nielsen 2024).

(5) kunne du ikke lige købe **mig** en øl?

(6) */?han købte **naboen** en bog

1.2. 1500-tallets IO i den historiske proces

Den ovenfor skitserede diakrone udvikling udstikker nogle aspekter af IO-konstruktionen og brugen af den som er af særlig interesse i kortlægningen af de enkelte historiske trin undervejs. De centrale spørgsmål drejer sig om IO's semantik og frihed/binding ift. verbalvalensen. For at opnå sammenlignelighed med Nielsen (2021) foretager jeg ligesom i min undersøgelse af Jammers Minde en semantisk klassifikation af de verber der optræder med IO (afsnit 2.1), og en analyse af IO's semantiske rolle i de fundne IO-konstruktioner med særlig fokus på rollen interessant (afsnit 2.2), og jeg undersøger forekomsten af frie IO'er nærmere (afsnit 3).

1500-talsresultaterne i afsnit 2 og 3 danner grundlag for en sammenligning med Jammers Minde-resultaterne i Nielsen (2021) i afsnit 4, og i afsnit 5 afslutter jeg med en kort konklusion. Som sidste del af dette indledende afsnit præsenteres kilder og metode i afsnit 1.3.

1.3. Kilder og metode

Undersøgelsen er baseret på tre kilder:

- Jesu Barndoms Bog (1508)
- Peder Palladius' Visitatsbog (ca. 1543)
- Christian 3.s danske Bibel (1550): Første Mosebog, kap. 12-25

Jesu Barndoms Bog, herefter JBB, er en apokryf bibelfortælling om Jesu barndom og hans families liv, trykt i 1508 af Gotfred af Ghemen, udgivet i Danske Folkebøger bd. 1. Palladius' Visitatsbog, herefter PV, er en sammenstyknings af notater skrevet af den danske biskop Peder Palladius (1503-1560) i forbindelse med hans kirkevisitationer på Sjælland. Den

undersøgte udgave er Lis Jacobsens (1925-1926) som er baseret på en senere afskrift af Palladius' original. Christian 3.s Bibel, herefter C3B, er en oversættelse af Luthers tyske Bibel (1534), men sproget i oversættelsen anses for at være i overensstemmelse med tidens almindelige danske skriftsprog uden stærkt tysk præg (Nielsen, u.å.). Den undersøgte udgave er online-udgaven af C3B på salmer.dsl.dk/texts. De tre tekster hører alle til den religiøse tekstgenre hvilket kan præge den observerede brug af IO, fx i forekomsten af verber såsom *forlade* 'tilgive' med IO.

Metodisk følger undersøgelsen samme fremgangsmåde som Nielsen (2021). Alle tre tekster er manuelt exciperet, og alle registrerede IO-forekomster er analyseret.⁴ Som ægte IO'er er regnet ikke-refleksive nominaler der fungerer som argument 3. Det præpositionelle modstykke til det nominale IO, IO's antiled (GDS: 1183-1185), er ikke inkluderet, ligesom reflexiver, som ikke skal regnes for egentlige argumenter (jf. GDS: 592-594, Nielsen & Heltoft 2023), også er udeladt.

Verberne i konstruktion med IO er klassificeret efter deres semantik ud fra det samme kategoriseringssystem som det der benyttes i Nielsen (2021) (se her for yderligere referencer), dog med tilføjelse af kategorierne *udelukkelse* (verber som fx *forvise*) og *perception* (fx *høre*) da disse kategorier viste sig nødvendige for klassifikationen af fundene i 1500-talsmaterialet. Kategoriseringssystemet er vist i tabel 1.⁵

Kategorier	Forklaring	Eksempelverber
Overdragelse		
Gennemført overdragelse	Overdragelse af DO til IO i og med handlingen	<i>give, bevilge</i>
Planlagt overdragelse	Forberedelse af senere overdragelse af DO til IO	<i>reservere, efterlade</i>
Fratagelse	Omvendt overdragelse, IO mister DO	<i>fratage, koste</i>
Kompensation	Overdragelse til IO af godtgørelse for DO	<i>godtgøre</i>
Kommunikation		
Meddelelsestype	Overdragelse af budskab (=DO) af en vis type	<i>fortælle, spørge</i>
Kommunikationsform	Specifikation af måden eller redskabet til overdragelse af budskab (=DO)	<i>råbe, skrive</i>

4 Den manuelle excipering af JBB er foretaget af Lars Heltoft som jeg skylder en stor tak for dette bidrag til min undersøgelse samt for bistand i analysen af vanskelige eksempler fra mine kilder.

5 Eksempelverberne illustrerer de semantiske kategorier uden nødvendigvis at kunne optræde med IO i (moderne eller ældre) dansk og uden nødvendigvis at optræde i 1500-talsmaterialet.

Regulativ			
	Handlingskommissiv	S forpligter sig til at udføre handling (=DO)	<i>tilbyde, love</i>
	Overdragelseskommisiv	S forpligter sig til at overdrage DO	<i>tilbyde, love⁶</i>
	Udførelsesdirektiv	S søger at få IO til at udføre handling (=DO)	<i>tillade, pålægge</i>
	Undladelsesdirektiv	S søger at få IO til at afstå fra handling (=DO)	<i>fravåde, forbyde</i>
Muliggørelse og facilitering		S gør det muligt eller lettere for IO at besidde eller håndtere DO	<i>lette, forunde</i>
Forårsagelse af intern proces			
	Forårsagelse af perception	S forårsager en perceptionsoplevelse i IO	<i>vise</i>
	Forårsagelse af kognition	S forårsager en kognitionsproces for IO	<i>lære, forklare</i>
Hindring			
	Hindring af overdragelse	S hindrer eller vanskeliggør overdragelsen af DO til IO	<i>forholde, nægte</i>
	Hindring af handling	S hindrer eller vanskeliggør IO's udførelse af handling (=DO) (ikke sprogligt, dvs. m. direktiv)	<i>forhindre</i>
Udelukkelse		S udelukker IO fra et sted betegnet m. DO	<i>forvise</i>
Fremstilling og klargøring		DO skabes eller gøres klar til et/dets formål af S	<i>bage, bygge</i>
Anskaffelse		DO tilvejebringes af S	<i>købe, finde</i>
Flytning		DO flyttes fra ét sted til et andet (evt. til deiktisk centrum)	<i>sende, bære</i>
Manipulation af genstand		Ikke-tilvejebringende handling rettet mod DO	<i>åbne, vaske</i>
Perception		S perciperer DO	<i>se, lytte</i>
Påvirkning		S udsætter IO for handling el. omstændighed (=DO)	<i>gøre (nogen noget), forårsage</i>
Attituder og følelser		S's indstilling til relationen mellem IO og DO	<i>misunde, tilgive</i>

Tabel 1: Kategorisering af IO-konstruktionernes verbers semantik

2. IO-fund i de tre 1500-talskilder

Den manuelle excerpering gav i første omgang 436 fund. Fra disse er der frasorteret 16 som er vurderet irrelevante for IO-analysen, først og fremmest konstruktioner hvor et betydningslet verbum danner forbindelse med et substantiv således at dette ikke betegner et reelt DO-argument, fx *give hånd(en)* som betegner en hilsehandling, ikke en 'give'-handling med *hånd(en)* som udtryk for en genstand der overdrages (jf. at det ikke følger

6 Ved handlingskommissiv vil DO for fx tilbyde og love være neksualt, ved overdragelseskommisiv innekksualt.

af *hun gav Bo hånd(en)* at Bo efterfølgende har denne hånd).⁷ Analyserne er baseret på de resterende 420 fund som fordeler sig således på de tre kilder:

- Jesu Barndoms Bog: 140 IO'er
- Palladius' Visitatsbog: 164 IO'er
- Christian 3.s Bibel: 116 IO'er

2.1. Verber i IO-konstruktioner

Verberne i IO-konstruktionerne er klassificeret sådan at et verbum med flere betydningsfunktioner – fx *skænke* med betydningerne 'overdrage, forære' og 'tilgive' – regnes for ét verbum, men henregnes til forskellige semantiske kategorier alt efter betydningsvariant i eksemplet. Det betyder at nogle verber optræder i flere semantiske kategorier i tabel 3 med resultaterne. Mere eller mindre faste forbindelser af verbum + andet materiale regnes for særskilte leksemer, fx *holde til gode* 'bære over med' og *give til tilgive*. Derfor finder man i tabel 3 fx både *holde* (om løfter) og *holde til gode*.

Ved klassifikation efter disse principper findes der i de 420 IO-konstruktioner 78 forskellige verber. Deres inddeling efter semantiske kategorier er vist i tabel 3. En lille gruppe verber står for det store flertal af IO-forekomsterne, jf. tabel 2 med de 8 hyppigste verber⁸ som tilsammen findes i 252 forekomster der udgør 60,00% af alle IO-fund.

Verbum	Antal	% af samlet (420)
<i>give</i>	98	23,33%
<i>sige</i>	53	12,62%
<i>gøre noget/ tjeneste/ skade/ etc.</i>	32	7,62%
<i>føde</i>	15	3,57%
<i>lære</i>	15	3,57%
<i>bede</i>	14	3,33%
<i>unde</i>	13	3,10%
<i>befale</i>	12	2,86%
I alt	252	60,00%

Tabel 2: De 8 hyppigste verber i IO-konstruktionerne

7 Typen *give hånd* med noget substantiv kan muligvis opfattes som en inkorporationsforbindelse. Jeg vil dog ikke i denne artikel forpligte mig på en sådan analyse.

8 Fremfor det pæne runde 10 er 8 oplistet fordi kun disse har tocifrede forekomster, og fordi der herefter er et lidt større spring ned til nr. 9 (*love*, 8 forekomster) og 10 (*forbyde*, 7 forekomster).

De to hyppigste verber alene står for 151 forekomster og udgør 35,95%. 47 af verberne forekommer kun i IO-konstruktion en eller to gange.

I tabel 3 vises IO-forekomsterne fordelt efter verballeksemernes semantiske kategorier.⁹ For de kategorier der har underkategorier, vises antal og procentdel dels for overkategorien, dels for underkategorierne.

Kategori	Verber	Antal	% af samlet (420)
Overdragelse		135	32,14%
Gennemført overdragelse	<i>antworde, befale, få, give, gøre redsel/rettighed</i> <i>'betale',¹⁰ låne, ofre, overgive, række, sende, skikke</i> <i>'sende, tildele', skænke, sælge, sætte, vede 'yde,</i> <i>give'</i>	131	31,19%
Planlagt overdragelse	(ingen)	0	0,00%
Fratagelse	<i>koste, løbe 'beløbe, koste', rove</i>	4	0,95%
Kompensation	(ingen)	0	0,00%
Kommunikation		68	16,19%
Meddelelsestype	<i>drage for 'foredrage', fortælle, kundgøre, kynde</i> <i>'forkynde, bekendtgøre', prædike, ride,¹¹ sige,</i> <i>spørge, svare</i>	66	15,71%
Kommunikationsform	<i>læse, skrive</i>	2	0,48%
Regulativ		57	13,57%
Handlingskommissiv	<i>byde til 'tilbyde', jette 'love', love, sværge, tilsige</i> <i>'love'</i>	15	3,57%
Overdragelseskommissiv	<i>tilsige 'love'</i>	2	0,48%
Udførelsesdirektiv	<i>bede, befale, byde, lægge, råde, sige til, tilstede</i> <i>'tillade'</i>	34	8,10%
Undladellesdirektiv	<i>forbyde, sjønne 'nægte'</i>	6	1,43%
Muliggørelse og facilitering	<i>give, levne, stå til 'tilstå, erkende',¹² unde</i>	17	4,05%
Forårsagelse af intern proces		39	9,29%
Forårsagelse af perception	<i>tegne 'vise ved tegn', vise, åbenbare</i>	7	1,67%
Forårsagelse af kognition	<i>bevise, give til kende, kende 'lære', lære, påminde,</i> <i>skyde i hu</i>	32	7,62%

9 Uden for citerede eksempler benyttes moderne ortografi for alle verber undtagen *sjønne* 'nægte' som ikke har en moderne form; her anvendes Kalkars (1976) stavning.

10 Både *redsel* og *rettighed* kan oversættes til 'afgift', og i kombination med disse har *gøre* betydningen *betale* (afgift).

11 I forbindelsen *ride i næsen* 'foreholde, anføre mod'

12 Dvs. '(for)unde, lade have'

Hindring			8	1,90%
	Hindring af overdragelse	<i>forhindre, formene, forsømme, holde inde</i> 'forholde, tilbageholde'	8	1,90%
	Hindring af handling	(ingen)	0	0,00%
Udelukkelse		<i>forbyde</i>	2	0,48%
Fremstilling og klargøring		<i>bygge, føde, gøre, rede</i> 'berede, indrette'	22	5,24%
Anskaffelse		<i>hente, skikke</i> 'ordne, arrangere', <i>tage</i>	12	2,86%
Flytning		<i>bære, føre, lægge</i>	4	0,95%
Manipulation af genstand		<i>bedre, beskikke, rette op, veje (til)</i>	4	0,95%
Perception		<i>høre</i>	1	0,24%
Påvirkning		<i>bevise, fly, give, gøre noget/tjeneste/skade</i> m.fl., <i>holde</i> , ¹³ <i>slå</i>	40	9,52%
Attituder og følelser		<i>byde, forlade, give til</i> 'tilgive', <i>holde til gode</i> , 'bære over med' <i>regne, skyld</i> 'beskyld for', <i>skænke</i> 'tilgive'	11	2,62%

Tabel 3: Fordeling af verber i IO-konstruktion efter verbernes semantiske kategorier med antallet af IO-forekomster

Som vist i tabel 3 udgør overdragelse, kommunikation og regulativer størstedelen af verberne og af IO-forekomsterne, i alt 260 af de 420 forekomster hvilket udgør 61,90%.

2.2. IO's semantiske rolle

Som beskrevet i afsnit 1 har IO i nutidigt dansk prototypisk den semantiske rolle modtager, og denne rolle er også rigt repræsenteret i 1500-talsmaterialet med undertyperne faktisk modtager (7), betinget modtager (fx ved løfter om overdragelse) (8), non-modtager (ved forhindret overdragelse) (9) og invers modtager (ved fratagelse) (10). Dertil kommer rollen som modtager ved kommunikationshandlinger ved verber som fx *kynde* 'forkynde, bekendtgøre' (11).

- (7) hon gaff **meg** mad (JBB: 96, 1)
- (8) **Abram** tilsigis en søn (C3B: margennote til kap. 15)
- (9) forhindrer du **dem** deeris salighed, da schalt du dybere neder i helffuede (PV: 50, 25)
- (10) paa det hand iche skal røffue **eder** eders retfongen arbeed fraa (PV: 122, 19-20)
- (11) ieg kynder **dig** et got raad (PV: 111, 15-15)

13 Om løfter.

I flere sammenhænge betegner IO dog noget andet end en modtager af en overdraget konkret eller abstrakt genstand. Det gælder bl.a. for konstruktioner der udtrykker en direktiv sproghandling rettet mod IO-referenten, hvor IO er et argument i en af DO betegnet prædikation,¹⁴ nærmere betegnet agent for DO-handlingen, som i (12). En lignende argumentrolle i en DO-prædikation har IO i sætninger der udtrykker at IO-referenten hindres i at udføre DO-handlingen (13). Ved verber der betegner en indre oplevelse i IO-referenten, betegner IO en eksperient (14).

- (12) tha bøth han **sith folck** ath ride effter them oc gripe them oc sla them i hiel (JBB: 52, 29-53, 2)
- (13) ... der hand siger, lader smaa bør [= børn]¹⁵ komme til mig, och forbiuder **dem** det icke (PV: 50, 1-3)
- (14) tha wisthe hon **them** barnet (JBB: 46, 6-7)

Endelig kan IO betegne en interessent. Der skelnes mellem tre undertyper: interessenter hvis interesse i verbalhandlingen udmønter sig i en (planlagt) overdragelse af DO-referenten til IO-referenten (som modtager-interessent) (15), de rene interessenter hvor der ikke er involveret nogen form for overdragelse til IO-referenten (16), og interessenter der er uafgørlige mht. modtagelse, som i (17) hvor det ikke kan afgøres om IO-referenten (Abraham) skal anses for modtager af sønnen, eller blot som den der nyder godt af sønnens tilvejebringelse.

- (15) her haffuer hand skicket **eder** eders spisning inden diße dørre (PV: 131, 13-14)
 'her har han ordnet/arrangeret mad inden for disse døre for og til jer'
- (16) och dersom du gjør **hannem** megit fortræd (PV: 41, 6)
- (17) Sara din hustru skal føde **dig** en søn (C3B: 17)

Analyseret efter dette sæt af rollebeskrivelse fordeler IO-forekomsterne sig som vist i tabel 4.

14 Ud over sætninger og infinitivhelheder tæller også andre neksuale DO'er med som prædikationelle, fx verbalsubstantiver.

15 *Bør* er ifølge udgavens noteapparat en dialektal form af *børn*.

IO-rolle		Eksempelverber	Antal	Procentdel
Modtager			159	37,86%
	Faktisk modtager	<i>give, sende</i>	144	34,29%
	Betinget modtager	<i>tilsige</i>	2	0,48%
	Non-modtager	<i>forhindre, holde inde</i>	9	2,14%
	Invers modtager	<i>koste, røve</i>	4	0,95%
Modtager af kommunikation		<i>fortælle, spørge</i>	68	16,19%
Argument for DO-prædikation			39	9,29%
	Argument for DO i direktiv	<i>befale, råde</i>	35	8,33%
	Argument for hindret handling	<i>forbyde, sjonne</i>	4	0,95%
Eksperient		<i>lære, åbenbare</i>	39	9,29%
Interessant			115	27,38%
	Uafgørlig mht. modtagelse	<i>føde, gøre</i>	18	4,29%
	Modtager-interessant	<i>hente, tage</i>	16	3,81%
	Ren interessant	<i>love,¹⁶ gøre noget/tjeneste/skade m.fl.</i>	81	19,29%
I alt			420	100,00%

Tabel 4: Forekomst af forskellige IO-roller

Som man kan se af tabel 4, udgør de forskellige varianter af modtagerrollen en stor andel. Tilsammen udgør de egentlige modtagere og kommunikationsmodtagerne 54,05% (N = 227).

Rollen interessant udgør en betydelig andel, de rene interessenter alene næsten en femtedel af alle IO'er (19,29%). Inden for de 81 rene interessenter er den største enkeltgruppe de 32 konstruktioner med det lette verbum *gøre* og et DO der betegner noget som IO-referenten udsættes for, bl.a. *noget, tjeneste, skade, løgn, troskab og ondt*, fx (18).¹⁷ Den næststørste gruppe er de 15 IO'er der betegner den der nyder godt af en handling som subjektsreferenten forpligter sig til med verber som fx *love* (19). Også udtryk for at nogen holder et afgivet løfte, findes blandt IO-konstruktioner med ren interessant (20), af disse er der fire. IO'er for verber der udtrykker attituder og følelser, fx *skænke* i betydningen 'tilgive' og *holde til gode* 'bære over med' (21), udgør 11 af de 81.

16 IO-rolle ren interessant forekommer ved *love* med neksualt DO som i (21).

17 Forbindelser af typen *gøre noget/tjeneste/ ondt* kunne ligne den frasorterede *give hånd(en)*-type (jf. afsnit 2), men den omstændighed eller det sagforhold der betegnes af DO'er som *noget, tjeneste* etc., analyseres her som er et faktisk (neksualt) argument således at IO-konstruktionen betegner en relation mellem den neksuale DO-referent og IO-referenten.

- (18) huat tycker teg wi wille gøre om thenne løyen the acthe ath gøre **oss** vether willie (JBB: 85, 29-86, 1)
- (19) loffue j **mig** det at j vil gjerne søge til eders sogendegen? (PV: 87, 14-15)
- (20) men at i schulle tencke til udj Iesu naffn at holde **gud och eders neste**, huad i der loffue och tilsige (PV: 35, 23-25)
- (21) sandelig det bliffuer **os** iche schenckt, eller til gode holden. (PV: 63, 32-33)

Disse fire nævnte grupper udgør 62 af de rene interessenter som ved disse verber kan betragtes som valensbundne IO'er, altså IO'er som indgår i det leksikalsk bestemte sæt af argumenter; dette gælder også for konstruktionerne af typen *gøre noget/tjeneste/ondt*/etc. som udtrykker en påvirkning af IO-referenten og dermed forudsætter et IO-led. Blandt de resterende 19 finder man bl.a. IO for udelukkelse (fra et sted) udtrykt med *forbyde* (22) og IO ved *unde* 'forunde' med et DO der betegner et vilkår for IO-referenten (23), dvs. verber der ligesom dem i de fire ovennævnte grupper har IO som valensled.

- (22) ... och hand da forbiuder **dem** sognen och hærredt... (PV: 139, 17-18)
- (23) gud haffue loff heder oc ære ther **oss** haffuer vnt thet ath wi ære fundnæ i liffue sammen (JBB: 82, 15-17)

Ud over dette store flertal af valensbundne interessent-IO'er findes 11 frit tilføjede rene interessenter som behandles i næste afsnit.

3. Frie IO'er i 1500-talskilderne

Blandt de 420 IO'er kan 45 beskrives som frie, dvs. frit tilføjet uden valensstyring, altså 10,71%. Omkring halvdelen er tilføjelser ved fremstillings- og klargøringsverber, fx *bygge*. Cirka en fjerdedel optræder ved anskaffelsesverber såsom *hente*, og den sidste fjerdedel er tilføjelser ved flytningsverber (fx *bære*), manipulation af en genstand (fx *veje*), perception (verbet *høre*) og påvirkning (fx *slå*), jf. kategorierne i tabel 1. De frie IO'ers semantiske roller fordeler sig som vist i tabel 5:

IO-rolle	Antal	% af alle frie IO'er
Faktisk modtager	3	6,67%
Interessant, uafgørlig mht. modtagelse	16	35,56%
Modtager-interessant	15	33,33%
Ren interessant	11	24,44%
I alt	45	100,00%

Tabel 5: Frie IO'ers semantiske rolle

Den største enkelgruppe, interessenter uspecificeret for modtagelse, udgøres af 16 forekomster af verbet *føde*, alle fra C3B, jf. (17). Tekstgenren må betragtes som afgørende for dette fund; skildringen af slægtskabsforhold og videreførelsen af slægterne spiller en væsentlig rolle i Første Mosebog. Som faktiske modtagere er regnet IO'er i konstruktion med flytningsverber som helt entydigt betegner overdragelse af DO-referenten, fx (24). I sætninger hvor et frit tilføjet IO betegner en der står i et interesseforhold til tilvejebringelsen af DO-referenten som en tiltænkt modtager, fx (25), er IO klassificeret som modtager-interessant.

- (24) han ... dreff mijne fjske borth the ieg haffde achtheth at bære **eder** hiem (JBB: 90, 16-18)
- (25) Tha begynte iomfru maria ... ath gøre **sijn welsinedhe søn** een kiortel (JBB: 70, 27-29)

De 11 frie rene interessenter er særligt interessante (jf. afsnit 1.2). De viser at det i 1500-tallet var muligt at lade en sætning udbygge med et frit IO-led uden dermed at udtrykke overdragelsesbetydning, og dermed kaster fundene yderligere empirisk lys på de primært teoretiske fremstillinger af denne mulighed (jf. afsnit 1.1-1.2). Omhyggelig tolkning af den intenderede betydning i teksten er nødvendigt for at fastslå at der er tale om rene interessenter. Et godt eksempel er (26) med verbet *beskikke*.

- (26) derfor ehr **hannem** beschicket it kar her inden døren (PV: 34, 9)

Beskikke kan have en række betydninger (jf. Kalkar 1976), herunder 'ordne', 'bringe i en vis tilstand', 'sende' og 'forkynde', og flere af disse giver mulighed for at tolke IO-referenten som en modtager og eventuelt et styret IO (fx ved 'sende'-betydningen). Men (26) er fra en passage der handler om kirkens døbefont, og om hvordan dette kar er placeret (*beskikket*)

inde i kirken således at præsten slipper for at gå til nærmeste vandløb når børn skal døbes. IO'et *hannem* (= præsten) er således en ren interessent.

Også ved verber der entydigt betegner tilvejebringelse af en genstand som ville kunne overdrages, finder man IO-konstruktioner med rene interessenter. Et eksempel på dette er (27) med *bygge*.

- (27) Saa brød Abram sit paulun op / kom oc bode i Mamre lund / som er i Hebron / oc bygde **HERREN** der it Altere (C3B: 13)

Pointen med byggeriet er ikke at Abram skal overdrage det tilvejebragte alter til Herren; alteret er bygget for at Herren kan nyde godt af Abrams tilbedelse, jf. den moderne bibeloversættelse hvor relationen til Herren udtrykkes med *for*, ikke *til* (28).

- (28) Dér slog han sig ned og byggede et alter for Herren¹⁸

Et af de interessanteste eksempler på frit tilføjet ren interessent er forekomsten med perceptionsverbet *høre* i (29).

- (29) kiere herre vil i høre **mig** min skriffDMAALL (PV: 65, 13)

Her foregår der en form for overdragelse, nemlig af (indholdet af) IO-referentens skriftemål til subjektsreferenten, og i tolkningen bliver IO ift. denne overdragelse faktisk afsender. ”Uden på” denne kommunikative overdragelse lægger sig så den markering af interesseforhold som IO-referenten står i: Det er til fordel for sognebarnet (= IO) at præsten (= subjekt) skal lytte til skriftemålet.

4. Sammenligning med Jammers Minde-undersøgelse

På baggrund af resultaterne præsenteret ovenfor kan undersøgelsen af IO i et lille udvalg af 1500-talstekster nu sammenlignes med undersøgelsen af IO i Leonora Christinas Jammers Minde (herefter JM) fra slutningen af 1600-tallet i Nielsen (2021). JM-undersøgelsen beskæftiger sig kun med forhold omkring IO i en tekst af én bestemt forfatter, hvilket selvsagt har

¹⁸ Bibelen online: https://www.bibelselskabet.dk/brugbibelen/bibelenonline/1_Mos/13 (lokaliseret 30.4.24).

begrænset gyldighed som billede på situationen i det sene 1600-tal som sådan, og på samme måde hvad angår de tre 1500-talstekster, som alle er religiøse tekster, ift. 1500-tallets dansk som sådan. Ikke desto mindre kan sammenligningen kaste lys på den del af den lange historiske proces fra middeldansk til nutidsdansk som finder sted i 1500- og 1600-tallet, ved at fundene i denne artikel og i Nielsen (2021) jævnføres med de overordnede udviklingslinjer beskrevet i afsnit 1.1.

4.1. IO-verbernes semantik

En sammenligning af fordelingen af de fundne IO'er på verbalklasser i tabel 2 ovenfor og i Nielsen (2021: 107) viser en række forskelle hvoraf tre skal omtales her. For det første er der en bemærkelsesværdig forskel mht. hvor stor en andel der udgøres af verber med en leksikalsk bestemt modtagerrolle til IO'et. Lægger man forekomster inden for verbalkategorierne overdragelse, kommunikation, overdragelseskommissiv og hindring af overdragelse sammen, udgør denne samlede modtagergruppe godt halvdelen af IO'erne i 1500-talsmaterialet (N = 213; 50,71% af de 467 IO'er), men to tredjedele af JM-IO'erne (N = 309; 66,17% af 467). I den overordnede bevægelse i retning af indsnævring af IO's betydningspotentiale omkring modtagerrollen kan denne forskel betragtes som en afspejling af udviklingen i *parole*-tendenserne i hhv. 1500- og (slut-)1600-tallet.

For det andet kan man bemærke at der er langt flere forekomster af IO ved fremstillings- og klargøringsverber i 1500-talsmaterialet (N = 22; 5,24%) end blandt JM-IO'erne (N = 5; 1,07%). Det må tages med i betragtning at 15 af 1500-talsforekomsterne udgøres af IO ved *føde* i C3B (jf. (18)) der som nævnt i afsnit 3 må betragtes som særligt motiveret af teksttypen. Dog stemmer forskellen med en overordnet diakron udvikling i retning af en ringere tilbøjelighed til at føje et frit IO til fremstillings- og klargøringsverber.

For det tredje er der kun fire forekomster med flytningsverber i 1500-talsmaterialet (0,95%), men 27 forekomster i JM (5,78%). Forskellen kan formentlig tilskrives et hyppigt anslået tema i JM: hvordan ting bringes ind og ud af Leonora Christinas fængselscelle, hvilket medfører mange sætninger med fx *bære*, *sende* og *skaffe*. Forskellen er derfor næppe en afspejling af en systematisk udvikling.

4.2. Frie IO'er

Som omtalt i afsnit 1.2 har de frie IO'er særlig interesse. Sammenligningen af 1500-talsmaterialet og JM viser en bemærkelsesværdig forskel.

Nielsen (2021: 114) finder 23 frie IO'er blandt JM-fundene, og de udgør beskedne 4,93% af samtlige IO'er.¹⁹ I 1500-talsmaterialet er der fundet 45 frie IO'er (jf. afsnit 3), og de udgør 10,71% af samtlige IO'er, altså en andel der er over dobbelt så stor som i JM.

Datamaterialet er ikke stort og bredt nok til at komme med skråsikre konklusioner om signifikante forskelle mellem århundrederne, men sammenligningen af nærværende undersøgelse med Nielsen (2021) tegner dog et billede af en gradvis indskrænkning af brugen af frit IO som stemmer godt overens med de store linjer beskrevet i afsnit 1.1.

4.3. Sproghandlingsrammen for frie IO'er

Som beskrevet i afsnit 1.1 er forholdene i det 20. og 21. århundrede kendetegnet ved at der kræves en regulativ sproghandlingskontekst ved tilføjelse af frit IO. Nielsen (2024: 298-299) peger på at der i materiale fra 1800-tallet kan spores en tendens i sprogbrugen til at favorisere regulative kontekster også forud for en egentlig grammatisk regel om binding til regulativer. Der viser sig en bemærkelsesværdig forskel mellem 1500-talsmaterialet og JM i koblingen mellem forekomst af frit IO og sproghandlingsramme.

Nielsen (2021: 114) diskuterer hvordan man skal tolke forholdene omkring sproghandlingsrammen for de frie IO'er i JM. Det bemærkes bl.a. at tre af de fem forekomster af frit IO med fremstillingsverber findes i regulativ kontekst, fx (30), og det konkluderes at de frie IO'er i påfaldende grad synes at tendere mod regulative kontekster (Nielsen 2021: 118).

(30) giør **mig** deraff en Rem i miine Boxer (JM: 113, 4-5)

I 1500-talskilderne er der også eksempler på frit IO ved fremstillings- og anskaffelsesverber i regulativ ramme (31), men langt de fleste frie IO'er ved disse verbaltyper findes i konstativ kontekst, jf. (25) og (27) ovenfor og (32) og (33) herunder.

(31) gack bort oc hænt **meg** skarderne [=skårene] (JBB: 87, 14)

(32) hon ... redde **henne** eth bat oc bad henne bade sit barn (JBB: 59, 7-11)

(33) Oc hans moder tog **hannem** en hustru aff Egipiti land (C3B)

19 I Nielsen (2021: 114) angives de 23 fejlagtigt som 2,65% af de 467 IO'er.

Det kunne altså se ud til at koblingen til den regulative ramme gradvist udvikler sig i sprogbrugen hen over århundrederne; i 1500-tallet er der ikke en klar kobling, i det sene 1600-tal træder en tendens frem som udvikler sig videre frem mod situationen i dag hvor normen – forenklet sagt (jf. Nielsen & Heltoft 2023; Nielsen 2024) – er at frit IO forudsætter regulativ kontekst.

5. Konklusion

Undersøgelsen af 1500-tallets IO som det fremtræder i data fra de tre kilder, udfolder detaljerne i et afsnit af det danske IO's historie og leverer både kvalitativ og kvantitativ empirisk underbygning af de teoretisk orienterede beskrivelser af IO's udvikling som fx Heltoft (2014; 2019).

Studiet af det undersøgte materiale fra 1500-tallet viser at den semantiske rolle modtager med sine 50,71% står stærkt og centralt i brugen af IO, men også at andelen af modtager-IO'er er klart mindre end de 66,17% i Jammers Minde-materialet undersøgt i Nielsen (2021). Rene interessenter finder man ganske mange af i 1500-talsmaterialet, men antallet af frit tilføjede interessenter er ganske lavt, og det gælder for frie IO'er i det hele taget. Det betyder ikke at muligheden for frit IO i 1500-tallet har været strukturelt begrænset ift. mulighederne i ældre tider, men det kunne se ud til at muligheden kun sjældent udnyttes i 1500-tallets sprogbrug – dog betydeligt hyppigere i de tre undersøgte 1500-talstekster end i Leonora Christinas 1600-talsdansk.

Om forfatteren

Peter Juul Nielsen, ph.d., Institut for Kultur- og Sprogvidenskaber ved Syddansk Universitet.

Litteratur

Christian 3.s Bibel (1550): Onlineudgave <https://salmer.dsl.dk/christian-3-bibel/titelblad> (lokaliseret 30.4.24).

GDS: se Hansen & Heltoft (2011).

De hellige Kvinder, en Legende-Samling (1859): Udg. af C. J. Brandt. København: Gad.

Hansen, Erik & Lars Heltoft (2011): *Grammatik over det Danske Sprog* (GDS). København: Det Danske Sprog- og Litteraturselskab.

- Heltoft, Lars (2014): Constructional change, paradigmatic structure and the orientation of usage processes. I *Usage-Based Approaches to Language Change*. Evie Coussé & Ferdinand von Mengden (red.). Amsterdam: John Benjamins, 203-241. [10.1075/sfsl.69.08hel](https://doi.org/10.1075/sfsl.69.08hel)
- Heltoft, Lars (2019): Fra gammeldansk til nutidsdansk. I *Dansk Sproghistorie* bd. III. Ebba Hjorth et al. (red.). København: Det Danske Sprog- og Litteraturselskab, 129-225.
- Høysgaard, Jens Pedersen (1752): *Methodisk Forsøg til en fuldstændig dansk Syntax*. København. (= Bertelsen, Henrik (red.), Danske Grammatikere V. Det danske Sprog- og Litteraturselskab [1919-1920]).
- Jesu Barndoms Bog* (1915): Danske Folkebøger bd. 1., udg. af J. P. Jacobsen. København: Gyldendal.
- Kalkar, Otto (1976): *Ordbog til det ældre danske sprog (1300-1700)*, 2. udgave. København: Akademisk Forlag.
- Leonora Christinas Jammers Minde* (1998): Udg. af Poul Lindegård Hjorth og Marita Akhøj Nielsen. København: Det danske Sprog- og Litteraturselskab.
- Nielsen, Marita Akhøj (u.å.): Om teksten. Præsentation af Christian 3.s Bibel på hjemmesiden Danske Reformationssalmer. <https://salmer.dsl.dk/christian-3-bibel/front/3?meta=introductions> (lokaliseret 30.4.24).
- Nielsen, Peter Juul (2021): Det indirekte objekt i *Jammers Minde*: Et nedslag i en diakron proces. *Ny forskning i grammatik* 28: 101-120. [10.7146/nfg.vi28.128786](https://doi.org/10.7146/nfg.vi28.128786)
- Nielsen, Peter Juul (2024): From symbol to index: The semiotic harmonisation of the Danish free indirect object. I *Indexicality: The role of indexing in language structure and language change*. Peter Juul Nielsen & María Sol Sansiñena (red.). Berlin: De Gruyter, 283-328. DOI: [10.1515/9783110791433-009](https://doi.org/10.1515/9783110791433-009)
- Nielsen, Peter Juul & Lars Heltoft (2023): Indexicality across the boundaries of syntax, semantics and pragmatics: The constructional content of the Danish free indirect object. I: *Ditransitives in Germanic languages: Synchronic and diachronic aspects*. Eva Zehentner, Melanie Rothlisberger & Timothy Colleman (red.). Amsterdam: John Benjamins, 150-194. [10.1075/sigl.7.05nic](https://doi.org/10.1075/sigl.7.05nic)
- Palladius, Peder (c. 1543): *Visitatsbogen*. Udg. af Lis Jacobsen, *Peder Palladius' Danske Skrifter*, bd. 5, 1926. København: S.S. Thiele.

Når en ø er fremmed land: Danskernes bedømmelse af ekstraktion fra relativsætninger i engelsk som andetsprog

Anne Mette Nyvad &
Ken Ramshøj Christensen

Abstract

In the generative literature, it is generally assumed that relative clauses are so-called ‘islands’ (i.e. syntactic domains from which extraction is impossible), originally based on data from English. However, counterexamples have been widely attested in the Mainland Scandinavian languages. Previous research has found that relative clauses are not impermeable in Danish and that the acceptability of extraction appears to be reduced due to structural complexity, potentially interacting with non-structural factors. If Danish does not adhere to the proposed universal constraint on extraction from relative clauses, how do native speakers of Danish judge this type of sentence in English as a second language? We present findings from an acceptability study which suggest syntactic transfer from Danish to English.

Nøgleord

andetsprogstilegnelse, ekstraktion, relativsætninger, acceptabilitet, tværsproglig variation

1. Indledning

Denne artikel udforsker forskellen mellem engelsk og dansk i hvor acceptabelt det er at ekstrahere fra indlejrede sætninger. Ved ekstraktion menes fremflytning ud af sætning, som i (1b) og (1c):

(1a) **Hvem** så Peter __ i går?

- (1b) **Hvem** tror du [at Peter så __ i går]?
- (1c) **Hvem** tror du [at hun sagde [at jeg havde hørt [at Peter så __ i går]]]?

I eksemplerne i (1) etableres der således en såkaldt *filler-gap*-dependens over kortere eller længere afstand (Sprouse m.fl. 2016): *hvem* er 'filler', og den understregede basisposition er 'gap'. Vi fokuserer her udelukkende på sætninger, hvor der er ekstraktion fra en indlejret sætning – det der i dansk traditionel grammatik kaldes en 'sætningsknude'. Eksemplerne ovenfor illustrerer spørgekonstruktioner med varierende grader af dependenslængde. Hvor *filler-gap*-dependensen i (1a) er kort, krydser den derimod en sætningsgrænse (angivet med firkantede parenteser) i (1b). Indenfor generativ grammatik antages det, at der teoretisk set ikke er nogen grænse for den mulige afstand mellem *filler* og *gap* (Chaves & Putnam 2020). Eksemplet i (1c) bekræfter dette, idet det fremrykkede element her bevæger sig på tværs af ikke mindre end tre sætningsgrænser på vej til sit endelige landingssted i Fundamentfeltet.

Visse strukturer med syntaktiske dependenser ser dog ud til at være underlagt restriktioner, i den forstand at modersmålstalende har en tendens til at finde dem uacceptable. Mere specifikt synes der at opstå problemer i specifikke konstruktioner, som Ross (1967) metaforisk døbte '**øer**' (se også Christensen & Nyvad 2023), fordi de er svære eller umulige at flygte fra (med henvisning til historien om Robinson Crusoe). Blandt de konfigurationer, der menes at være ø-strukturer, er relativsætninger, (2), og ekstraktion fra sådanne indlejrede sætninger antages således at føre til ugrammatikalitet (signaleret med '*'):

- (2) ***Who** does Phineas know a girl [**who** __ is working with __]?
(Ross 1967: 124)

For at redegøre for den manglende mulighed for ekstraktion fra relativsætninger mere specifikt, er der i den generative litteratur fremsat en række syntaktiske begrænsninger, herunder *Complex NP Constraint* (CNPC, Ross 1967), samt de mere generelle principper i *Subjacency Condition* (Chomsky 1973) og *Minimal Link Condition* (Chomsky 1995). CNPC ("no element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase", Ross 1967: 127) blev oprindeligt foreslået med henvisning til data fra engelsk, såsom i (2), hvor der

er en uacceptabel ekstraktion fra den komplekse NP, der begynder med *a girl* og indeholder en relativsætning.

Der ser dog ud til at være en masse variation både på tværs af tale-re og på tværs af sprog. Allerede for mere end 50 år siden var der en række teoretikere, der pointerede, at de fastlandsskandinaviske sprog i usædvanlig vid udstrækning tillader ekstraktion fra relativsætninger. Flere teoretikere har således argumenteret for, at disse sprog er mere tolerante mht. ekstraktion, da ekstraktion fra forskellige øer ofte er blevet attestet i naturligt forekommende tale (Allwood 1976, Erteschik-Shir 1973, Lindahl 2017. Se også Engdahl & Ejerhed 1982). De påståede universelle begrænsninger er således konsekvent blevet mødt med modeksempler, i særdeleshed på de fastlandsskandinaviske sprog. Her er et lille udpluk af danske modeksempler med ekstraktion fra relativsætninger fra den grammatiske litteratur, der går helt tilbage til 1800-tallet og frem til i dag (se også Christensen & Nyvad 2023):

- (3) Han er **en Person**, som du næppe kan nævne nogen [der kan maale sig med ____].
(Mikkelsen 1894: 422)
- (4a) **Her** har jeg en god ven [der bor ____]
(Hansen 1984: 75)
- (4b) **Det** kender jeg da en fyr [der gør ____ hver dag]
(Hansen 1984: 75)
- (4c) **Suppe** kender jeg mange [der kan lide ____]
(Erteschik-Shir 1973: 67).
- (4d) **Det nummer** har jeg engang set en artist [der udførte ____]
(Hansen & Heltoft 2011: 1815)

Forventningen fra teorien er ellers at sådanne ekstraktioner ville blive bedømt som mindre acceptable eller være decideret ugrammatiske, så hvad karakteriserer mulige ekstraktioner? Det er blevet observeret et væld af gange, at acceptabiliteten af ekstraktion fra relativsætninger bl.a. afhænger af matrixverbet (Engdahl 1997, Lindahl 2017). For hovedparten af de naturligt forekommende eksempler i Lindahls (2017) svenske dataindsamling gælder det, at de optræder i eksistentielle eller såkaldt 'præsentationelle' konstruktioner, hvis funktion det er at introducere nye referenter i diskursen (fx *Chokoladeis er der mange der kan lide*). Hvad angår dansk, fandt Müller & Eggers (2022) 940 eksempler med ekstraktion fra en relativsæt-

ning i *KorpusDK* (der består af 56 mio. ord), hvoraf der var 910 eksempler med topikalisering (som i (4)), 30 med relativisering (som i (3)) og nul eksempler med *hw*-ekstraktion (spørgekonstruktion). I 933 af de 940 eksempler (= 99,3%) var matrixverbet en form af *være*. Sidst men ikke mindst var subjektet i matrixsætningen som oftest pronominent (Müller & Eggers 2022: 14-15):

- (5a) **Det** er der allerede mange [der gør ___].
- (5b) Der var **noget damegymnastik på skærmen**, som der ikke var nogen [der så på ___].
- (5c) Men **den slags spekulationer** kender jeg ikke nogen [der tør fremsætte ___ offentligt].

Det er bemærkelsesværdigt, at ekstraktion fra relativsætninger optræder næsten lige så ofte som ekstraktion fra almindelige deklarative *at*-sætninger i *KorpusDK* (hhv. 910 og 1250 eksempler, Müller og Eggers 2022: 11). Overordnet set er mønstret således, at ekstraktion primært foregår ved topikalisering, mens relativisering ud af en relativsætning forekommer markant sjældnere, og at fremrykning ved *hw*-ekstraktion (spørgekonstruktion) ud af en relativsætning nærmest ikke forekommer (Lindahl 2017, Müller & Eggers 2022, Jensen 1998).

Men hvordan forholder det sig på engelsk? Ifølge Chaves & Putnam (2020) er visse ekstraktioner fra relativsætninger acceptable på engelsk, og man kan også finde modeksempler i den generative syntakslitteratur:

- (6a) This is **a paper** that we really need to find someone [who understands ___]
(Chung & McCloskey 1987: 708).
- (6b) Then you look at what happens in languages that you know and **languages** that you have a friend [who knows ___]
(Kuno 1976: 423)

Müller & Eggers (2022) fandt **ingen** eksempler på ekstraktion i *British National Corpus* eller i *Corpus of Contemporary American English*, på trods af at de to korpusser er mange gange større end *KorpusDK* (henholdsvis 100 millioner og 1 milliard ord), mens der blev fundet 18 eksempler i *GloWbE*-korpusset på knap 2 milliarder ord. Det tyder på, at der en tværsproglig forskel, men hvor stor er denne forskel så mellem de fastlandsskandinaviske

sprog og engelsk? (Müller 2019, Nyvad m.fl. 2022, Nyvad m.fl. 2024).

Denne artikel udforsker, hvordan grammatikken i vores modersmål (L1) påvirker processeringen (dvs. forståelsen) af disse langdistancedependenser på et andetsprog (L2). Nærmere bestemt ser artiklen på danske modersmålstalendes bedømmelse af ekstraktion fra relativsætninger på engelsk som andetsprog. Vi vil præsentere resultatet fra undersøgelser, der peger i retning af, at dele af L1-grammatikken overføres til L2, når det drejer sig specifikt om ø-konstruktioner som dem i (2)-(6). For at sætte L2-studiet i perspektiv introducerer vi i afsnit 2 to af vores tidligere studier, hvor vi har undersøgt acceptabiliteten af ekstraktion fra relativsætninger på henholdsvis dansk L1 (Christensen & Nyvad 2014) og engelsk L1 (Christensen & Nyvad 2022). I afsnit 3 præsenterer vi ganske kort den teoretiske baggrund inden for forskning i andetsprogstilegnelse og vores forudsigelser for L2-studiet, samt vores studies overordnede resultater. I det afsluttende afsnit 4 diskuterer vi, hvordan L2-studiets fund kan tolkes.

2. Eksperimentelle studier i Christensen & Nyvad (2014, 2022)

En række syntaktikere har argumenteret for, at modeksemplerne i de fastlandsskandinaviske sprog egentlig ikke er modbeviser, fordi den underliggende struktur af de pågældende konstruktioner angiveligt er forskellig fra deres overfladestruktur (fx Chomsky 1982, Kush & Lindahl 2011, Kush m.fl. 2013, Phillips 2013a,b. Se Engdahl 1997), og at de foreslåede restriktioner på langdistancedependenser således stadig kan betragtes som universelle. Mere konkret antager Kush & Lindahl (2011) for svensk, at det kun er bestemte matrixverber, der tillader ekstraktion fra relativsætninger, og at disse verber i virkeligheden ikke har en NP (eller DP) som objekt, men derimod en særlig slags sætning (en *Predicate Phrase*, PredP), der på overfladen er identisk med en NP med en relativsætning, og dermed overtræder de pågældende eksempler ikke CNPC. De angiver følgende kontrast, hvor *träffade* og *kysste* ('mødte' og 'kyskede') har en NP med 'rigtig' relativsætning som objekt, mens *ser* og *känner* ('ser' og 'kender') efter sigende har en PredP som objekt:

- (7a) ***De blommorna** träffade/kysste jag en man [som sålde__].
 (7b) **De blommorna** ser/känner jag en man [som sålde __]

Kush & Lindahl (2011) testede denne tese i et eksperiment, der viste en

signifikant forskel i bedømmelse af ekstraktion fra relativsætninger med henholdsvis matrixverberne *se* og *træffa*. For at se om Kush & Lindahls (2011) teori også holder på dansk, designede vi et eksperiment (Christensen & Nyvad 2014), hvor vi undersøgte effekten af hele 16 forskellige matrixverber. Et forudgående pilotstudie viste, at halvdelen af disse verber var i stand til at have en PredP som objekt (*behøve, foretrække, beholde, mistænke, anmelde, udnytte, kende, se, modarbejde, tilgive, kysse, drille, genere, myrde, træffe, møde*), fx *Vi kendte Troels som barn*, hvor *Troels som barn* er PredP. Acceptabilitetsundersøgelsen bestod af stimulussætninger af følgende type:

- (8a) Pia har engang mødt en pensionist [som havde **sådan en hund**].
 (8b) **Sådan en hund** har Pia engang mødt en pensionist [som havde
 —].

Acceptabilitetsstudiet viste, i modstrid med Kush & Lindahls teori om en PredP, at bedømmelserne var helt uafhængige af, om matrixverbet i andre kontekster er kompatibelt eller ej med en PredP som objekt, men i stedet så ud til at være påvirkelige af dets frekvens. Den kendsgerning, at acceptabiliteten så ud til at variere som funktion af manipulation af en ikke-syntaktisk faktor (frekvens), så vi som et tegn på, at ekstraktion fra relativsætninger faktisk er mulig på dansk, men nedsat i acceptabilitet pga. det processeringsbesvær, som konstruktionens kompleksitet indebærer (Christensen & Nyvad 2023, 2024). Dertil kommer, at forsøgspersonerne syntes signifikant bedre om ekstraktion fra relativsætninger, jo mere de blev eksponeret for dem, og denne form for *priming* kan kun forekomme med sætninger, der grundlæggende er grammatiske (Sprouse 2007: 123-124).

På baggrund af disse resultater lavede vi en tilsvarende undersøgelse af ekstraktion fra relativsætninger på engelsk (Christensen & Nyvad 2022), hvor vi fandt, at der ikke var den samme effekt af leksikalsk frekvens: Hvor stimulussætningerne med de tre mest frekvente matrixverber lå over middel i acceptabilitet i det danske studie, lå alle matrixverberne noget under middel i det engelske studie. Derudover fandt vi, at forsøgspersonernes bedømmelse af denne type sætning var på niveau med utvetydigt ugrammatiske sætninger, såsom (9d) (*What type of man did Peter once kiss a girl who preferred men?*), og der var ingen effekt af priming (altså af om de havde set samme sætningstype før i forsøget), sådan som vi havde set det i det danske studie. Disse resultater peger på, at standardantagelsen i den generative syntakslitteratur vedrørende relativsætningers status som 'stær-

ke' øer på engelsk har hold i virkeligheden. Denne sætningstype ser ud til kategorisk at blokere for ekstraktion på engelsk, uafhængigt af leksikalske effekter. Med andre ord fører ekstraktion uvægerligt til uacceptabilitet på engelsk, mens vi tidligere har konkluderet (Christensen & Nyvad 2022), at relativsætninger er 'svage' øer på dansk (hvad man måske kunne kalde 'halvøer'), fordi de udviser variation som funktion af manipulation af ikke-syntaktiske faktorer, såsom frekvens. Med andre ord peger vores tidligere studier på, at dansk og engelsk har forskellige begrænsninger for ekstraktioner fra ø-konfigurationer. I det følgende vil vi undersøge hvordan dette spiller ind i acceptabiliteten af ekstraktion fra relativsætninger, når engelsk er L2 og dansk L1.

3. Ø-konstruktioner på L2 engelsk med dansk som modersmål

Et centralt spørgsmål i litteraturen om andetsprogstilegnelse er, hvordan modersmålet (L1) påvirker det andet, tredje eller fjerde sprog, man tilegner sig. Særligt relationen mellem L1 og andetsproget (L2) er blevet studeret. Ifølge den fremtrædende teori, *Full Transfer Full Access*-hypotesen (White 2003: 61), er L2-grammatikken i de tidlige stadier af andetsprogstilegnelsen lig med L1-grammatikken ('Full Transfer'), og læreren har fuld adgang til universalgrammatikken ('Full Access'), der antages at begrænse 'intersproget' (dvs. stadierne i den glidende overgang mellem L1 og L2). Disse begrænsninger inkluderer ø-restriktioner, såsom CNPC (se indledningen). En ændring eller justering i den engelske L2-grammatik kræver formentlig positiv evidens i inputtet (dvs. det sprog, læreren eksponeres for) (Pullum & Scholz 2002, Berwick m.fl. 2011), der viser, at der er en uoverensstemmelse mellem dansk L1 og engelsk L2. De danske L1-talere lærer ikke, at man ikke kan ekstrahere som på dansk, når man skal tale engelsk. Det skyldes nok, at mængden af observerbare langdistance-*filler-gap*-dependenser på engelsk sandsynligvis udgør en brøkdel af dem, der produceres på de fastlandsskandinaviske sprog (der er færre muligheder på engelsk). Det er således sandsynligt, at talere med dansk L1 aldrig rigtigt justerer deres grammatik for engelsk L2, hvad angår netop ø-konstruktionerne (se Kush m.fl. 2023).

Baseret på vores tidligere resultater (se diskussion ovenfor) og den grundlæggende antagelse af, at (dele af) grammatikken på modersmålet overføres til andetsproget, havde vi to overordnede forudsigelser om acceptabiliteten af ekstraktion fra relativsætninger på L2 engelsk:

- **Forudsigelse 1:** Acceptabiliteten for ekstraktion ud af en relativsætning er positivt korreleret med matrixverbets frekvens.
- **Forudsigelse 2:** Ekstraktion af både topic-elementer og spørgeelementer (*hw*-led) fra relativsætninger er reduceret i acceptabilitet (pga. processeringsbesvær/byrde på arbejdshukommelsen), men stadig acceptable.

Den første forudsigelse var baseret på det mønster, vi fandt på L1-dansk (Christensen & Nyvad 2014) samt antagelsen af, at der er transfer fra L1 til L2. Hvad angår den anden forudsigelse, antog vi, at en langdistance-dependens mellem en *filler* og et *gap* i sig selv involverer en så stor byrde på arbejdshukommelsen, at acceptabiliteten påvirkes af det (Fanselow & Frisch 2006, Hofmeister & Sag 2010). Som pointeret i Christensen & Nyvad (2022) er topikalisering og *hw*-flytning rettet mod den samme position (CP-Spec/Fundamentfeltet), og da de begge forbinder sætningen til den forudgående diskurs, kunne man forvente, at de udviser parallelle acceptabilitetsmønstre (omend de to typer flytning har forskellig informationsstruktur og funktion). Da vores grundantagelse ydermere var, at relativsætninger tillader ekstraktion på dansk men ikke engelsk, uafhængigt af dependensstypen, forudså vi, at der *ikke* ville være en forskel i danske modersmålstalendes bedømmelse af L2-engelsk, hvis der er transfer fra L1.

Forsøgsdeltagerne (41 kvinder og 35 mænd, 76 i alt) havde dansk som modersmål og engelsk som andetsprog. Deres gennemsnitsalder var 36,7 (yngste 22 og ældste 65). Forsøgspersonerne blev bedt om at angive, om de var lingvister, hvilket 12 svarede ja til, og 28 af de deltagende var engelskstuderende (hvis forsøgspersonerne sagde ja til at være 'lingvist' eller 'engelskstuderende', kunne de potentielt være primet af teoretiske diskussioner om (u)muligheden af ekstraktion fra relativsætninger, og denne effekt ville vi gerne kontrollere for). Spørgeskemaet blev udformet på Google Forms og delt på diverse SoMe-platforme. Opgavebeskrivelsen til spørgeskemaet specificerede, at forsøgsdeltagere skulle have dansk som modersmål og engelsk som andetsprog, og dette krav blev gentaget i et kontrolspørgsmål i spørgeskemaet.

Stimuli var identiske med dem, der blev brugt i Christensen & Nyvad (2022), og bestod af 140 sætninger, hvoraf 48 var målsætninger, 10 utvetydigt grammatiske sætninger og 10 utvetydigt ugrammatiske sætninger

samt 72 sætninger fra et andet eksperiment (Christensen & Nyvad 2019). For at få et indblik i effekten af dependenstype på *ø*-konstruktionernes acceptabilitet på L2-engelsk anvendte vi testbatteriet fra Christensen & Nyvad (2022), hvor vi havde konstrueret minimalt forskellige sætningstyper med henholdsvis topikalisering og *hw*-flytning ud af en relativsætning:

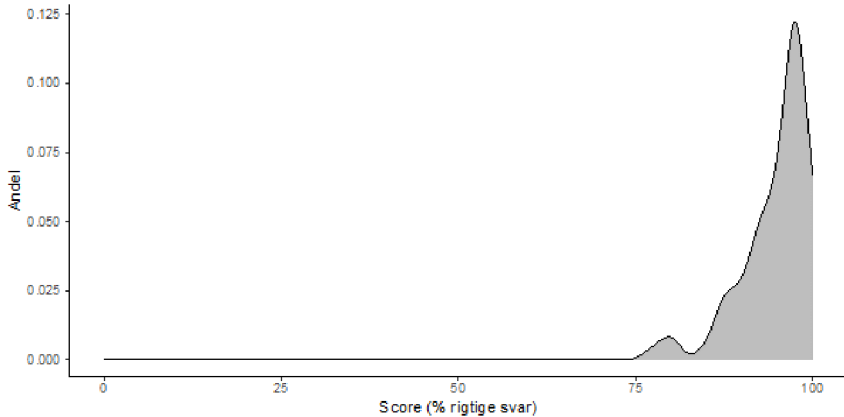
- (9a) Peter once kissed a girl [who preferred **that type of man**]. (BASIS)
- (9b) **What type of man** did Peter once kiss a girl [who preferred __]? (HV)
- (9c) **That type of man Peter** once kissed a girl [who preferred __]. (TOPIC)
- (9d) ***What type of man** did Peter once kiss a girl [who preferred **men**]? (ANOMALI)

Disse sætninger blev fordelt på 6 lister med hver 40 testsætninger i randomiseret rækkefølge i et såkaldt romerkvadrat-design, så forsøgspersonerne kun så hvert matrixverbum (*have, know, love, meet, visit, hate, bother, respect, kiss, murder, dislike, bully*) én gang i løbet af eksperimentet. Hver forsøgsdeltager skulle bedømme stimulussætningerne på en 5-punktsskala fra 1 ('helt uacceptabel') til 5 ('helt acceptabel') og blev instrueret i udelukkende at basere deres bedømmelse på egen sproglig intuition.

De engelske stimuli var alle i simpel datid, og for at undgå faktorer, der har vist sig at nedsætte acceptabilitet uafhængigt, var alle matrixverberne ikke-faktive, dvs. sandhedsværdien af propositionen i den indlejrede sætning ikke er givet, fx *Hun troede, at han var spion (og det var han også / men det var han ikke)* (jf. Kiparsky & Kiparsky 1970, Christensen, Kizach & Nyvad 2013a,b). Af samme grund modificerede relativsætningerne en ikke-definit NP (Engdahl 1982, Hofmeister & Sag 2010). Med undtagelse af *do*-support og den overflødige NP i anomali-sætningerne var antallet af ord konstant på tværs af typerne (9a)-(9d). Derudover søgte vi at facilitere ekstraktion ved at sørge for kohæsion mellem matrixsætningen og relativsætningen i Allwoods (1976) forstand og ved at bruge komplekse *hw*-nominaler i stedet for pronominer (jf. Cinque 1990, Hofmeister & Sag 2010).

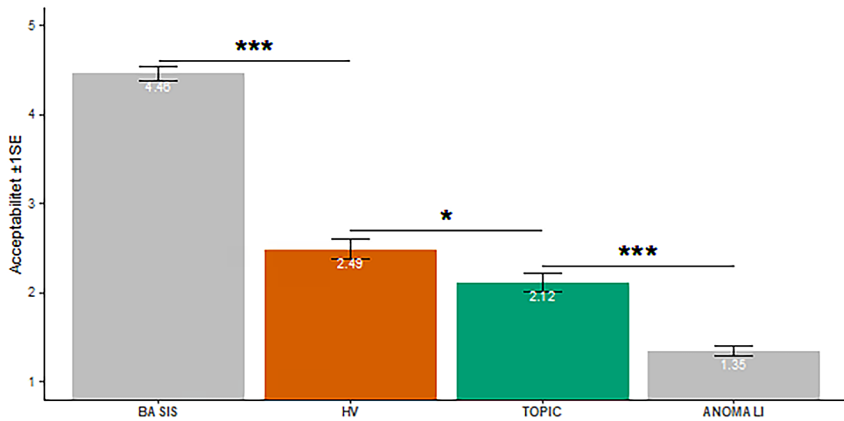
Inden selve forsøget skulle deltagerne udføre en cloze-test, udviklet af O'Neill, Cornelius & Washburn (1981), der ifølge Ionin & Montrul (2010) er en pålidelig sprogfærdighedstest. Testen bestod af en tekst med 40 tomme huller, der skulle fyldes af det ord (ud af tre valgmuligheder),

der gav pragmatisk, semantisk og grammatisk mening. Fx lyder tekstens start således: *Joe came home from work on Friday. It was payday, but he wasn't (1) even / more / ever excited about it.* Denne test blev brugt som en mål for sprogfærdighed for at kontrollere for, hvilken rolle graden af sprogfærdighed spiller for bedømmelsen af ekstraktion fra relativsætninger på L2-engelsk. I figur 1 ses resultatet af clozetesten. Alle forsøgsdeltagerne havde en sprogfærdighed på > 75 % (og langt de fleste > 90 %).



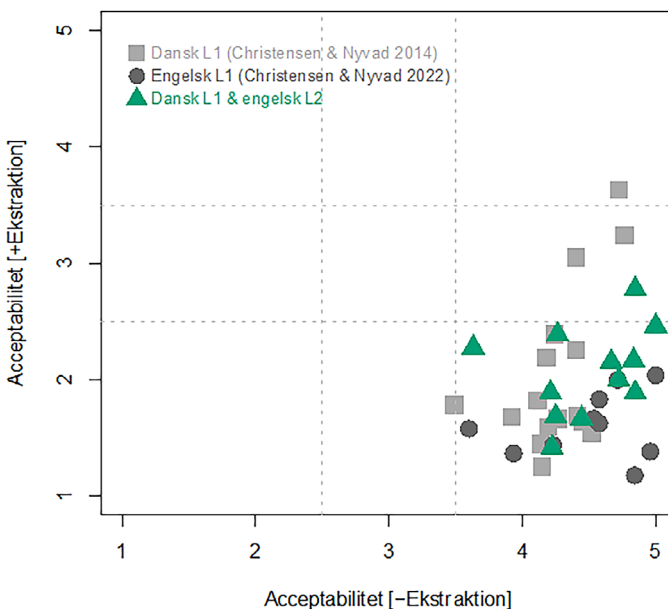
Figur 1: Resultatet af cloze-testen. Y-aksen angiver andelen af deltagere fra 0-1, svarende til 0-100 %. (Det er et såkaldt density-plot, hvor arealet af det grå område er 1, hvilket svarer til 100 %.)

Resultatet af acceptabilitetstesten viste, at der ikke var nogen signifikant overordnet effekt af sprogfærdighed ($p = 0,28$), status som engelskstuderende ($p = 0,51$) eller lingvist ($p = 0,82$). Interessant nok var der, i modsætning til hvad vi fandt i det danske studie (Christensen & Nyvad 2014), ingen effekt af eksponering over tid, dvs. der var ingen primingeffekt ($p = 0,33$), men generelt responderede forsøgspersonerne meget lig det danske L1-mønster: Som forventet medfører ekstraktion i sig selv en reduktion i acceptabilitet:



Figur 2: De gennemsnitlige acceptabilitetsbedømmelser (± 1 standardfejl) på en skala fra 1 til 5 for de fire sætningstyper i (9). * $p < 0,05$, *** $p < 0,001$.

Forsøgspersonerne med dansk som modersmål ser således ud til generelt at bedømme de samme engelske ekstraktionssætninger lidt højere end deltagerne med engelsk som modersmål (Christensen & Nyvad 2022), og der er desuden en positiv korrelation mellem matrixverbets frekvens og acceptabilitet ($p < 0,001$), og vores første forudsigelse kan således betragtes som bekræftet.



Figur 3: De gennemsnitlige acceptabilitetsbedømmelser på tværs af vores studier af ekstraktion fra relativsætninger.

Derudover bedømmer forsøgspersonerne ekstraktion fra relativsætninger signifikant bedre end den syntaktiske anomali, (9d), hvilket ikke var tilfældet for forsøgspersonerne med engelsk som modersmål (Christensen & Nyvad 2022), så der er basis for at formode, at der er overførsel af noget af L1-grammatikken på dansk i processeringen af engelsk som andetsprog. Det faktum at ekstraktion fra relativsætninger blev bedømt lavt, men signifikant bedre end den syntaktiske anomali, er også kompatibelt med vores forudsigelse 2. Sidst men ikke mindst var der signifikant forskel på bedømmelserne af sætningerne med henholdsvis topikaliserings- og *hw*-flytning ($p = 0,006$): *Hw*-flytning bedømmes 0,37 point bedre end topikaliserings-. Dette fund er bemærkelsesværdigt i den forstand, at topikaliserings- ud af relativsætninger tilsyneladende er mange gange mere frekvent end *hw*-flytning (Müller & Eggers 2022, se indledning), mens topikaliserings- er markeret på engelsk (Engdahl 1997, Poole 2011). De få eksempler på acceptabel ekstraktion fra en relativsætning på engelsk, der findes i litteraturen, er ofte karakteriseret ved at fremrykningen sker vha. relativisering (Sprouse et al. 2016, Nyvad m.fl. 2022. Se også eksemplerne i (6), hvor fillerens fremrykkes vha. relativisering hele to gange, men det er kun den sidste relativisering, der rykker på tværs af en 'ø-grænse', nemlig ud af den dybest indlejrede relativsætning). Dette resultat vidner sandsynligvis om, at forsøgspersonerne har et højt niveau af sprogfærdighed på engelsk, siden deres L2-mønster her overordnet er parallelt med den tendens, vi fandt på L1-engelsk (Christensen & Nyvad 2022). Med andre ord ser forsøgspersonerne ud til allerede at have tilegnet sig den del af den engelske grammatik, der er ansvarlig for, at topikaliserings- er markeret på dette sprog. På den anden side bedømmer vores deltagere her (L2-engelsk) ekstraktion fra en relativsætning signifikant højere end anomalien (0,77 point, $p < 0,001$), hvilket ikke var tilfældet for L1-engelsk.

4. Diskussion

I litteraturen er der blevet konstateret en slående kontrast mellem bedømmelser i forsøg og hvad der produceres i naturligt sprogbrug (fx Müller 2019, Müller & Eggers 2022). Ekstraktion fra \emptyset -konstruktioner findes helt klart i naturligt forekommende tale, angiveligt særligt på de fastlandsskandinaviske sprog, men også i nogen udstrækning på engelsk (jf. Chaves & Putnam 2020, Jensen 1998, Lindahl 2017). I eksperimentelle acceptabilitetsstudier modtager de ofte overraskende lave bedøm-

melser (fx Poulsen 2008, Tutunjian m.fl. 2017, Kush m.fl. 2018), hvilket evt. kunne tolkes som et udtryk for at de slet ikke er mulige. Men en række psyko- og neurolingvistiske eksperimenter har, som tidligere nævnt, fundet evidens for, at en del ekstraktioner fra \emptyset -strukturer faktisk ser ud til at være grammatiske på dansk (Christensen, Kizach & Nyvad 2013a,b, Christensen & Nyvad 2014, Nyvad, Christensen & Vikner 2017, Nyvad & Christensen 2023). Desuden kunne man måske endda forvente, at acceptabiliteten kunne forbedres både på dansk og engelsk (samt L2-varianter af disse sprog, jf. de forholdsvis lave bedømmelser i nærværende studie), hvis man anvendte formuleringer, der var tættere på talesprog, fx med *at være* som matrixverbum og pronominer i stedet for substantiver eller fulde NP'er: *Det er der jo mange/ingen der tror på* (se diskussionen i Müller & Eggers 2022: 14-15 og Nyvad & Christensen 2022).

Som nævnt i indledningen, er relativsætninger traditionelt blevet anset som syntaktiske øer, der på tværs af sprog blokerer for ekstraktion, men denne antagelse udfordres af variationen i bedømmelserne – både på tværs af konstruktioner og sprog. Noget af denne variation skyldes sandsynligvis egenskaber ved matrixverbet, hvilket potentielt kan forklares med henvisning til Chaves & Putnams (2020) begreb *Relevance Islands*, der bygger på en lang række informationsstrukturelle tilgange (Erteschik-Shir 1982, Deane 1991, Goldberg 2005. Se også Abeillé m.fl. 2020). De forsøger at give en pragmatisk forklaring på, hvorfor \emptyset -konstruktioner ofte opfattes som mere eller mindre uacceptable: Ifølge Chaves & Putnams (2020) er de konstruktioner, som traditionelt er blevet betragtet som værende underlagt syntaktiske \emptyset -restriktioner (fx indlejrede spørgsmål, adverbielle ledsætninger og relativsætninger), i virkeligheden begrænset af pragmatikken: Når et element bliver fremrykket (ved topikalisering, relativering eller *hw*-ekstraktion), bliver det også fremhævet, og denne fremhævelse skal være kompatibel med pragmatikken, således at det fremrykkede element også er tilstrækkeligt relevant for ytringens hovedbegrivenhed. Derudover bør det fremrykkede elements *gap* være i en konstituent, der er 'prominent' eller 'relevant' (Chaves & Putnam 2020: 206), og eftersom relativsætninger ofte er karakteriseret ved at indeholde præsupperet eller givet information, er et element ekstrahe- ret herfra er prominent men typisk ikke relevant 'nok' til at sætningen er pragmatisk naturlig, dvs. overholder et pragmatisk princip om relevans. Det er dog muligt, at semantisk 'transparente' matrixverber som *være* i

højere grad end semantisk rige verber tillader, at den indlejrede sætning er i forgrunden eller er *main point of utterance* (jf. Simons 2007, Wiklund m.fl. 2009, Kush m.fl. 2021). Kort sagt er det ikke umuligt, at nogle af forskellene i acceptabilitetsbedømmelserne på tværs af konstruktioner og forsøg kan skyldes variationer i semantiske og pragmatiske faktorer, som vi ikke har kontrolleret for. I Christensen & Nyvad (2022), der havde fokus på L1-engelsk, og i nærværende L2-forsøg med fokus på L2-engelsk har vi dog brugt de præcis samme stimuli, så sammenligningen mellem de to studier lider ikke under den manglende kontrol af disse potentielle (i høj grad ubekendte) faktorer.

Overordnet set er vores resultater kompatible med antagelsen om, at der er transfer fra L1 til L2. For forsøgspersoner med engelsk som modersmål gælder det tilsyneladende, at ekstraktion fra relativsætninger er (så godt som) umulig (Christensen & Nyvad 2022). På dansk, derimod, tyder signifikant variation på tværs af matrixverber på, at denne type konstruktion er mulig, selvom acceptabiliteten er nedsat pga. processeringsbesvær eller pres på arbejdshukommelsen (Christensen & Nyvad 2014). Forsøgspersonerne ser ud til at overføre den mere liberale tilgang til ekstraktion i dansk (L1) til engelsk (L2) (dvs. at visse dele af den danske grammatik muligvis overføres til engelsk), og det medfører, at ekstraktionerne bedømmes højere. På den anden side ser deres høje sprogfærdighed ud til at øge den engelske forskel mellem *hv*-flytning og topikalisering. De har med andre ord styr på, at topikalisering er mere markeret på engelsk.

Eftersom matrixverbets frekvens interagerer med acceptabiliteten på L2 engelsk ligesom på L1 dansk, og modsat L1 engelsk, ser det ud til, at der er transfer. Det fundne mønster skyldes ikke baggrund som engelskstuderende eller lingvist, eller eksponering over tid. Ingen af disse var statistisk signifikante. Selvom vi ikke fandt en statistisk effekt af sprogfærdighed på acceptabiliteten, er det dog ikke utænkeligt at den meget høje sprogfærdighed har medført, at deltagerne bedømmer topikalisering ud af en relativsætning som mindre acceptabelt end *hv*-flytning (modsat af hvad vi ville forvente, hvis der var 'full transfer' fra dansk). Det er muligt, at den overordnede tilbøjelighed på dansk til at fremflytte til Fundamentfeltet i sig selv spiller en ikke ubetydelig rolle for vores villighed til at acceptere den type sætning på engelsk, men det er et spørgsmål, som må adresseres i fremtidig forskning.

Om forfatterne

Anne Mette Nyvad, ph.d., lektor, Institut for Kommunikation og Kultur, Afdeling for Engelsk, Aarhus Universitet.

Ken Ramshøj Christensen, ph.d., lektor, Institut for Kommunikation og Kultur, Afdeling for Engelsk, Aarhus Universitet.

Litteratur

- Abeillé, A., B. Hemforth, E. Winckel & E. Gibson (2020): Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204: 104293. <https://doi.org/10.1016/j.cognition.2020.104293>.
- Allwood, J. (1976): The Complex NP Constraint as a non-universal rule and some semantic factors influencing the acceptability of Swedish sentences which violate the CNPC. *University of Massachusetts Occasional Papers in Linguistics* 2.
- Berwick, R.C., P.P. Pietroski, B. Yankama & N. Chomsky (2011): Poverty of the stimulus revisited. *Cognitive Science* 35, 1207–1242. <https://doi.org/10.1111/j.1551-6709.2011.01189.x>.
- Chaves, R.P. & M.T. Putnam (2020). *Unbounded dependency constructions: Theoretical and experimental perspectives*. Oxford: Oxford University Press.
- Chomsky, N. (1973): Conditions on transformation. I: *A festschrift for Morris Halle*, S.R. Anderson & P. Kiparsky (red.). New York, NY: Holt, Rinehart & Winston. 232–286.
- Chomsky, N. (1982): *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995): *The minimalist program*. Cambridge, MA: MIT Press.
- Christensen, K.R., J. Kizach & A.M. Nyvad (2013a): Escape from the island: Grammaticality and (reduced) acceptability of *wh*-island violations in Danish. *Journal of Psycholinguistic Research* 42, 51–70. <https://doi.org/10.1007/s10936-012-9210-x>.
- Christensen, K.R., J. Kizach & A.M. Nyvad (2013b): The processing of syntactic islands – An fMRI study. *Journal of Neurolinguistics* 26, 239–251. <https://doi.org/10.1016/j.jneuroling.2012.08.002>.
- Christensen, K.R. & A.M. Nyvad (2014): On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37, 29–45. <https://doi.org/10.1017/S0332586514000055>.

- Christensen, K.R. & A.M. Nyvad (2019): No escape from the island: On extraction from complement wh-clauses in English. I: *The sign of the V – Papers in honour of Sten Vikner*. K.R. Christensen, H. Jørgensen & J.L. Wood (red.). Aarhus: Dept. of English, School of Communication and Culture, Aarhus University. 95–112. <https://doi.org/doi:10.7146/aul.348.91>.
- Christensen, K.R. & A.M. Nyvad (2022): The island is still there: Experimental evidence for the inescapability of relative clauses in English. *Studia Linguistica* 76, 684–708. <https://doi.org/10.1017/S0332586514000055>.
- Christensen, K.R. & A.M. Nyvad (2023): Økuller? Det er ikke altid lige (u) acceptabelt at forlade en indlejring. I: *MUDS 19*. K. Boas, I.S. Hansen, T.T. Hougaard & E.L. Overgaard. Aarhus: Aarhus Universitet. 77–92.
- Christensen, K.R. & A.M. Nyvad (2024): Complexity, frequency, and acceptability. *Glossa* 9, 1–44. <https://doi.org/10.16995/glossa.10618>.
- Chung, S. & J. McCloskey (1987): Government, barriers and small clauses in Modern Irish. *Linguistic Inquiry* 18, 173–237.
- Cinque, G. (1990): *Types of A²-Dependencies*. Cambridge, MA: MIT Press.
- Deane, P. (1991): Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2, 1–64. <https://doi.org/10.1515/cogl.1991.2.1.1>.
- Engdahl, E. (1982): Restrictions on unbounded dependencies in Swedish. I: E. Engdahl & E. Ejerhed (red.). 151–174.
- Engdahl, E. (1997): Relative clause extraction in context. *Working Papers in Scandinavian Syntax* 60, 51–79.
- Engdahl, E. & E. Ejerhed (red.) (1982): *Readings on unbounded dependencies in Scandinavian languages*. Umeå: Universitetet i Umeå.
- Erteschik-Shir, N. (1973): *On the nature of island constraints*. Ph.d.-afhandling, MIT.
- Erteschik-Shir, N. (1982): Extractability in Danish and the pragmatic principle of dominance. I: E. Engdahl & E. Ejerhed (red.). 175–192.
- Fanselow, G. & S. Frisch (2006): Effects of processing difficulty on judgments of acceptability. I: *Gradience in Grammar: Generative Perspectives*. G. Fanselow, C. Fery & M. Schlesewsky (red.). Oxford: Oxford University Press. 291–316.
- Goldberg, A.E. (2005): *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

- Hansen, E. (1984). *Dæmonernes port: Støttmateriale til undervisningen i nydansk grammatik*. 3. udg. København: Reitzel.
- Hansen, E. & L. Heltoft (2011): *Grammatik over det danske sprog*. København: Det danske sprog- og litteraturselskab.
- Hofmeister, P. & I.A. Sag (2010): Cognitive constraints and island effects. *Language* 86, 366–415. <https://doi.org/10.1353/lan.0.0223>.
- Ionin, T., & S. Montrul (2010): The role of L1-transfer in the interpretation of articles with definite plurals in L2-English. *Language Learning* 60, 877–925. <https://doi.org/10.1111/j.1467-9922.2010.00577.x>.
- Jensen, A. (1998): *Knudekonstruktioner – en syntaktisk, semantisk og pragmatisk analyse af sætningsknuder i dansk*. Speciale, Københavns Universitet.
- Kiparsky, P. & C. Kiparsky (1970). Fact. I: *Progress in linguistics: A collection of papers*. M. Bierwisch & K.E. Heidolph (red.). Berlin: De Gruyter. 143–173.
- Kuno, S. (1976): Subject, theme, and the speaker's empathy: A re-examination of relativization phenomena. I: *Subject and Topic*. Charles N. Li (red.). New York, NY: Academic Press. 417–444.
- Kush, D. & F. Lindahl (2011): On the escapability of islands in Scandinavian. Conferencepresentation, 85th Annual Meeting of the Linguistic Society of America. Pittsburgh, PA.
- Kush, D., A. Omaki & N. Hornstein (2013): Micro-variation in Islands? I: *Experimental syntax and island effects*. J. Sprouse & N. Hornstein (red.). Cambridge, MA: Cambridge University Press. 239–265.
- Kush, D., T. Lohndal & J. Sprouse (2018): Investigating variation in island effects: A case study of Norwegian *wh*-extraction. *Natural Language & Linguistic Theory* 36, 743–779. <https://doi.org/10.1007/s11049-017-9390-z>.
- Kush D., C. Sant & S.B. Strætkvern (2021): Learning island-insensitivity from the input: A corpus analysis of child- and youth-directed text in Norwegian. *Glossa* 6, 1–50. <https://doi.org/10.16995/glossa.5774>.
- Kush, D., A. Dahl & F. Lindahl (2023). Filler-gap dependencies and islands in L2 English production: Comparing transfer from L1 Norwegian and L1 Swedish. *Second Language Research*. Online udg. før tryk. <https://doi.org/10.1177/02676583231172918>.
- Lindahl, F. (2017): *Extraction from relative clauses in Swedish*. Ph.d.-afhandling, Göteborg Universitet.
- Mikkelsen, K. (1894): *Dansk sproglære, med sproghistoriske tillæg: Haandbog for lærere og viderekomne*. København: Lehman & Stages Forlag.

- Müller, C. (2019). *Permeable islands. A contrastive study of Swedish and English adjunct clause extractions*. Ph.d.-afhandling, Lunds Universitet.
- Müller, C. & C.U. Eggers (2022): Island extractions in the wild: A corpus study of adjunct and relative clause islands in Danish and English. *Languages* 7, 125. <https://doi.org/10.3390/languages7020125>.
- Nyvad, A.M. & K.R. Christensen (2023): *Hvis, da og fordi* ekstraktion fra adverbielle ledsætninger var mulig. *Ny forskning i grammatik* 30, 185–200. <https://doi.org/10.7146/nfg.v1i30.137960>.
- Nyvad, A.M., C. Müller & K.R. Christensen (2022): Too true to be good? The non-uniformity of extraction from adjunct clauses in English. *Languages* 7, 244. <https://doi.org/10.3390/languages7040244>.
- Nyvad, A.M., K.R. Christensen & S. Vikner (2017): CP-recursion in Danish: A cP/CP-analysis. *The Linguistic Review* 34, 449–477. <https://doi.org/10.1515/tlr-2017-0008>.
- O’Neill, R., E. Cornelius & G. Washburn (1981): *American kernel lessons: Advanced students’ book*. New York, NY: Longman Publishing Group.
- Phillips, C. (2013a): On the nature of island constraints I: Language processing and reductionist accounts. I: *Experimental syntax and island effects*. J. Sprouse & N. Hornstein (red.). Cambridge: Cambridge University Press. 64–108.
- Phillips, C. (2013b): On the nature of island constraints II: Language learning and innateness. I: *Experimental syntax and island effects*. J. Sprouse & N. Hornstein (red.). Cambridge: Cambridge University Press. 132–158.
- Poole, G. (2011): *Syntactic theory*. Houndsmills: Palgrave Macmillan.
- Poulsen, M. (2008): Acceptability and processing of long-distance dependencies in Danish. *Nordic Journal of Linguistics* 31, 73–107. <https://doi.org/10.1017/S0332586508001832>.
- Pullum, G.K. & B.C. Scholz (2002): Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 18, 9–50. <https://doi.org/10.1515/tlir.19.1-2.9>.
- Ross, J.R. (1967): *Constraints on variables in syntax*. Ph.d.-afhandling, MIT.
- Simons, M. (2007): *Observations on embedding verbs, evidentiality, and presupposition*. *Lingua* 117, 1034–1056. <https://doi.org/10.1016/j.lingua.2006.05.006>.
- Sprouse, J. (2007): Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1, 123–134. <https://doi.org/10.5964/bioling.8597>.

- Sprouse, J., I. Caponigro, C. Greco & C. Cecchetto (2016): Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34, 307–344. <https://doi.org/10.1007/s11049-015-9286-8>
- Tutunjian, D., F. Heintz, E. Klingvall & A.-L. Wiklund (2017): Processing relative clause extractions in Swedish. *Frontiers in Psychology* 8, 2118. <https://doi.org/10.3389/fpsyg.2017.02118>.
- Vikner S., K.R. Christensen & A.M. Nyvad (2017): V2 and cP/CP. I: *Order and structure in syntax I: Word order and syntactic structure*. L.R. Bailey & M. Sheehan (red.). Berlin: Language Science Press. 313–324.
- White, L. (2003): *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Wiklund, A.-L., K. Bentzen, G.H. Hrafnbjargarson & Þ. Hróarsdóttir (2009): On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119, 1914–1938. <https://doi.org/10.1016/j.lingua.2009.03.006>.

*Ved Lejre græsse nu Faar paa Vold, hvor
fordum Kæmperne drukke.*

Udviklingen af numeruskongruens mellem subjekt og finit verbum i dansk fra ca. 1500 til ca. 1900

Kathrine Thisted Petersen & Steffen Krogh

Abstract

In this paper, we explore the development of number agreement between the subject and the finite verb in written Early and Late Modern Danish. Previous research has shown that the plural could occasionally be substituted for the singular already in Late Middle Danish. Between 1500 and 1900, the use of the plural gradually decreases, but it varies greatly from text to text. In order to explain this variation systematically, we base our study on seven translations of The Gospel of Luke from 1550 until 1907. What we find is a unidirectional and systematic development from consistent number agreement in all the texts up to 1746, where the plural form winds up in the preterite of strong verbs with stem change between the singular and the plural.

Nøgleord

verbalbøjning, pluralis, numeruskongruens, sproghistorie, bibeloversættelser

1. Indledning¹

Udforskningen af det danske sprogs historie har hidtil haft fokus på gammeldansk, mens forholdene i ældre og yngre nydansk fra ca. 1500 til ca. 1700 resp. ca. 1700 til ca. 1900 er blevet mere stedmoderligt behandlet. Som Ruus (2019: 73-74) skriver, findes der ”ikke nogen større systematiske

¹ Citatet i titlen stammer fra B.S. Ingemanns sang *Ved Lejre græsse nu Faar paa Vold* fra 1837.

undersøgelser af dansk ordbøjning i perioden 1500-1700". Særligt inden for verbalbøjningen er der i ældre nydansk en del afvigelser i forhold til nutidsdansk. Det gælder for kategorierne tempus, modus, diatese, numerus og person. Et af de mest fremtrædende træk er pluralisbøjningen i præsens indikativ og imperativ og for de stærke verbers vedkommende også i præteritum indikativ. I forskningslitteraturen bemærkes det at pluralisbøjningen gradvis afvikles i løbet af ældre og yngre nydansk, og at singularisformer ved subjekter i pluralis allerede kan belægges i begrænset omfang i gammeldansk. Det er en kendsgerning at pluralisbøjning som skriftsproglig norm først blev officielt afskaffet i 1900, men hvordan den lange udvikling henimod afskaffelsen mere præcist fandt sted, er stadig uvist. En nærmere undersøgelse heraf vil kræve en grundig gennemgang af et omfattende kildemateriale som man i første omgang kun kan se på en lille del af. Her præsenteres en undersøgelse af forekomsten af verbal pluralisbøjning i forskellige udgaver af den samme tekst, nemlig Lukas-evangeliet. Denne tekst foreligger i forskellige oversættelser og revisioner udgivet løbende siden 1500-tallet. På dette grundlag kan man tegne et billede af en udvikling inden for en enkelt, men meget markant tekstgenre. Forlods giver vi et overblik over den hidtidige litteratur om emnet i et normativt og et deskriptivt perspektiv.

2. Kongruens mellem subjekt og finit verbum i rune- og gammeldansk

Ligesom i mange andre indoeuropæiske sprog herskede der i det ældste danske skriftsprog streng kongruens mellem subjekt og finit verbum inden for person- og numeruskategorien. Pluralisbøjningen af verbet kommer i stand ved kongruens i numerus mellem sætningens finitte verbum og subjektet der kan være eksplicit eller implicit. I førstnævnte tilfælde er udgangspunktet et nominalt eller pronominalt subjekt i pluralisform eller flere sideordnede led i singularis. Kollektiver (fx *kvæg*), der formelt er singularis, fremkalder som regel ikke verbal pluralisbøjning (for undtagelser med *constructio ad sensum* se eksempel (5) og (6) nedenfor). I tilfælde med implicit subjekt er det derimod en uudtalt forestilling om en flerhed der udløser en tilsvarende formforandring af finittet. Fra ca. 1500 til ca. 1900 ses dette kun ved imperativ, men det var langt mere gængs på tidligere sprogrtrin hvor pro-drop endnu var fremherskende.

Person- og numeruskongruens kommer til udtryk i nedenstående runeindskrifter på to benstykker fra Lund fra 1000-1300 resp. 1050-1300:

- (1) ben æs þætta, ben æs þ[æ]tta (LB 3)
 ben er.PR.ÆS.3.SG dette.N.NOM.3.SG ben er.PR.ÆS.3.SG dette.N.NOM.3.SG
 Dette er ben. Dette er ben.
- (2) ar-ar ara ær-u fiaprar (LB 4 side B)
 åre.F-NOM.3.PL.F ørns er-PR.ÆS.3.PL fjer'
 Ørnens årer er fjere (vingerne).

Her kongruerer verbet *æs* med pronomenet *þætta* i 3. person singularis og verbet *æru* med substantivet *arar* i 3. person pluralis. Uden for skånsk aftager personbøjningen kraftigt i løbet af middelalderen. Således erstattes i præsens indikativ endelsen i 1. person singularis med fællesformen i 2. og 3. person singularis, mens 1. og 2. person pluralis erstattes med endelsen i 3. person pluralis (Skautrup 1944: 274; Frederiksen 2019: 46). I præteritum indikativ hos de stærke verber og præsens indikativ hos de præterito-præsentiske verber samt *vilja* og *væra* fortsætter endelsen i 2. person singularis i længere tid, ”men efter ca. 1400 ses 1./3. singularisendelsen *-θ* ofte i konkurrence med 2. singularis’ *-t* eller den stadig almindeligere endelse *-st*” (Frederiksen 2019: 46). Ved indgangen til ældre nydansk er dermed stort set kun numerusbøjningen bevaret. Som udgangspunkt for korpusundersøgelsen i afsnit 4 gives nedenfor et overblik over verbalbøjningen i aktiv omkring år 1500. Paradigmerne i figur 1-3 skal vise skriftsprogsnormen på dette tidspunkt. Grundet den store variation i 1500-tallets retskrivning har vi anvendt de moderne retskrivningsprincipper.

	Indikativ		Konjunktiv		Imperativ
	Præsens	Præteritum	Præsens	Præteritum	
Singularis					
1. person	<i>finder</i>	<i>fandt</i>	<i>finde</i>	<i>finde</i>	
2. person	<i>finder</i>	<i>fandt</i>	<i>finde</i>	<i>finde</i>	<i>find</i>
3. person	<i>finder</i>	<i>fandt</i>	<i>finde</i>	<i>finde</i>	
Pluralis					
1. person	<i>finde</i>	<i>finde</i>	<i>finde</i>	<i>finde</i>	
2. person	<i>finde</i>	<i>finde</i>	<i>finde</i>	<i>finde</i>	<i>finder</i>
3. person	<i>finde</i>	<i>finde</i>	<i>finde</i>	<i>finde</i>	

Figur 1. Prototypisk bøjningssystem for stærke verber i aktiv i skriftsproget ca. 1500 eksemplificeret med *finde*

Singularis	Indikativ		Konjunktiv		Imperativ
	Præsens	Præteritum	Præsens	Præteritum	
1. person	<i>elsker</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	
2. person	<i>elsker</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	<i>elsk</i>
3. person	<i>elsker</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	
Pluralis					
1. person	<i>elske</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	
2. person	<i>elske</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	<i>elsker</i>
3. person	<i>elske</i>	<i>elskede</i>	<i>elske</i>	<i>elskede</i>	

Figur 2. Prototypisk bøjningssystem for svage verber i aktiv i skriftsproget ca. 1500 eksemplificeret med *elske*

Singularis	Indikativ		Konjunktiv	
	Præsens	Præteritum	Præsens	Præteritum
1. person	<i>kan</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>
2. person	<i>kant</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>
3. person	<i>kan</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>
Pluralis				
1. person	<i>kunne</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>
2. person	<i>kunne</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>
3. person	<i>kunne</i>	<i>kunde</i>	<i>kunne</i>	<i>kunde</i>

Figur 3. Prototypisk bøjningssystem for præterito-præsensiske verber (inkl. ville) i aktiv i skriftsproget ca. 1500 eksemplificeret med *kunne*

3. Hidtidig litteratur om udviklingen af numeruskongruens mellem subjekt og finit verbum fra gammel- til nutidsdansk

I forskningslitteraturen samt i grammatikker og andre sprogrigtigheds skrifter fra 1600-tallet og frem findes bemærkninger om i hvilket omfang de verbale pluralisformer anvendtes eller burde anvendes. I det følgende begynder vi fra et normativt perspektiv ved først at opsummere reglerne i et udvalg af ældre danske grammatikker fra 1600-1800-tallet. Ud fra et deskriptivt perspektiv præsenteres derefter den hidtidige forskningslitteratur om emnet.

3.1. Det normative perspektiv

Fire af de kendteste grammatikere fra 1600-tallet, Lavrids Kock (1660:

59-62), Erik Pontoppidan (1668: 64, 211-212, 214), Erik Olsen Marsløv (1678: 10) og Peder Syv (1685: 178), anser verbal numerusbøjning for vigtig i dansk grammatik. Pontoppidan og Marsløv pointerer at numeruskongruens gælder i både skrift og tale. I overensstemmelse med reglen skelner Marsløv og Syv, der i modsætning til Kock og Pontoppidan skriver på dansk frem for latin, konsekvent mellem singularis og pluralis i deres eget skriftsprog. Derudover fremstiller de begge udeblivende kongruens som en uheldig tendens blandt datidens sprogbrugere (Marsløv 1678: 10; Syv 1685: 237).

Omtrent samtidig giver Henrich Gerner (1678-1679: 128) udtryk for at pluralisformen er forældet, men at den stadig findes i skriftsproget. Gerner foretrækker det forenkede system med kun én endelse, men bruger i nogle tilfælde stadig numerusbøjning. Kun 11 år senere skriver han imidlertid at "[n]umeri pluralis endelse in verbis er lige med Singularis numeri endelse i alle temporibus" (Gerner 1690: 283). Samtidig advarer Gerner (1690: 273) læseren mod at lade sig påvirke af den skånske dialekt hvor pluralisbøjningen stadig lever i bedste velgående.

At der i 1700-tallet har hersket stor variation i brugen af og holdningerne til numeruskongruens, bekræftes af bl.a. Peder Schulz. Først præsenterer han formen *elsker* som et eksempel på præsens indikativ i både singularis og pluralis (Schulz 1724: 81). Efterfølgende bemærker Schulz (1724: 84) dog at pluralisformen uden *-r* også findes, men fraråder alligevel brugen af denne form. Ironisk nok skelner han i sit eget skriftsprog ved en del verber konsekvent mellem singularis og pluralis.

Jens Høysgaard (1743: 200-215) blander sig livligt i debatten mellem Syv og Gerner ved at afprøve deres argumenter. Ligesom Gerner ender Høysgaard (1747: 323-324) med at opfordre til at lade skriftsproget følge talesproget hvor numerusbøjningen er helt forsvundet. Også i sit eget skriftsprog følger han normalt denne praksis.

I 1800-tallet bliver billedet kun mere broget. At forvirringen nok har været stor i den almindelige befolkning, undrer ikke når én og samme person kan give udtryk for to forskellige normer sådan som Edwin Jessen gør det i to skrifter fra henholdsvis 1891 og 1864. I 1891 formuleres reglen om numeruskongruens klart (Jessen 1891: 122), og i samme tekst holder Jessen sig også selv strikt til denne regel. I 1864 går Jessen derimod mere deskriptivt til værks samtidig med at han giver sin egen holdning til kende: "Men alt i de ældste hss o. 1300 kan i handlef. ental findes for flertal, i lidef. omvendt; det trænger igennem för 1600, næppe för 1500; skriften

skjuler, især efter 1500, virkeligheden, der dog ikke sjældent bryder det kunstige bånd, forfatterne pålægger sig” (Jessen 1864: 4). Også implicit udtrykker han sin mening ved konsekvent brug af singularis.

Som nævnt i indledningen blev obligatorisk numeruskongruens mellem subjekt og finit verbum i dansk skriftsprog først officielt afskaffet i 1900 ved en ministeriel bekendtgørelse (BEK nr. 184 af 15.12.1900).

3.2. Det deskriptive perspektiv

Allerede i gammeldansk findes der adskillige belæg med singularisformer i indikativ (Jessen 1864: 5-8; Bjerrum 1949; Brøndum-Nielsen 1918: 127, 129-131; 1959: 54-55), men som hovedregel kun hvis subjektet står efter verbet, hvis det udgøres af en relativpartikel, eller hvis det består af flere sideordnede led i singularis (Jessen 1864: 5; Bjerrum 1949: 162-166, 173; Brøndum-Nielsen 1959: 54-55). Generelt finder man flere belæg i jyske end i sjællandske og skånske kilder, men fordi de samme betingelser for brugen af singularis gør sig gældende i alle tre dialekter (Bjerrum 1949: 174-176), kan man ikke uden videre slutte at pluralisbøjningen i jysk talesprog afvikledes allerede i 1300-tallet sådan som fx Falk & Torp (1900: 13) og Skautrup (1944: 273) gør det. Eksempel (3) stammer fra Valdemars Sjællandske Lov i håndskriftet Stockholm C 69,8^o fra ca. 1400; her figurerer subjektet netop som to sideordnede led i singularis:

- (3) Fathær oc mothær **ær** syn oc dothær næst at ærwæ (VL 3)

Den egentlige udvikling hen imod pluralisformernes undergang i skriftsproget ses først i ældre nydansk fra 1500 til 1700 hvor der ifølge Ruus (2019: 78) sker en gradvis udtynding af verbal pluralisbøjning. Ruus (2019: 78-80) giver flere eksempler på opretholdt numeruskongruens, men nævner kun enkelte undtagelser fra 1600-tallet. I et belæg fra 1641 omtaler Christian 4. sig selv med et pronomen i pluralis majestatis tilknyttet to verber i singularis:

- (4) At indbilde børnen, som uy ved hinders datter Aulid **haffuer**, At uy derris moder **gør** stor wrett (Chr. 4. Br 140)

Da subjektet *uy* her kun refererer til ét individ, opviser de verbale singularisformer i (4) pragmatisk set numeruskongruens, og eksemplet siger dermed ikke meget om pluralisformens status på dette tidspunkt. Tilsvarende

findes der fx i Peder Palladius' visitatsbog fra ca. 1543 og Resens Bibel fra 1607 subjekter i singularis med verber i pluralis:

- (5) om een christne almue **faa** deeris rette føde till deris siæls salighed (Visit 25)
- (6) Folcket lete effter hannem / oc **komme** hen til hannem (Res Bib 32^v)²

Subjekterne *een christne almue* og *Folcket* opfattes som en gruppe af individer der dermed naturligt udløser pluralisformerne *faa* og *komme*.

I forlængelse af Ruus skriver Jacobsen om situationen i yngre nydansk fra 1700 til 1900 at ”kategorierne *tal* og *person* ikke længere er relevante for verbalbøjningen, men kun forekommer som stivnede former i faste udtryk og talemåder, i ældre salme- og sangtekster og i arkaiserende eller spøgende sprogbrug” (Jacobsen 2019: 93). Derudover kan man også finde pluralisformer i prædiken- og foredragssprog (Jacobsen 2019: 102). Jacobsen (2019: 101) tilføjer at singularisformen længe havde været det almindelige i talesproget, men en nærmere datering fremgår ikke.

Jessen (1864) og Sørensen (1900) giver en lang liste over spredte be-læg på singularisformer i ca. 300 forskellige forfatters tekster inden for prosa og poesi frem til ca. 1900, men nogen egentlig sammenhængende udvikling finder de ikke. Alligevel vurderer Sørensen at afviklingen af numerusbøjning i talesproget ”er begyndt i det 13de århundrede og maa betragtes som afsluttet i det 16de, selv om skriftsproget stadig har slæbt den døde flertalsform med sig” (Sørensen 1902: 28).

4. Korpusundersøgelse af verbale pluralisformer i Lukasevangeliets 1550-1907

4.1. Valg af kilder

Brugen af numeruskongruens mellem subjekt og finit verbum kan afhænge af forskellige faktorer som fx tekstgenre, stilleje, målgruppe og skribent, og det er ikke nødvendigvis sådan at én og samme faktor altid udløser den samme numerusform. For at kunne tegne et billede af en sammenhængende udvikling tager vi udgangspunkt i Lukasevangeliets, en tekst der løbende er blevet udgivet i forskellige oversættelser og reviderede versioner

2 I alle de anvendte bibeludgaver har Det Nye Testamente særskilt paginering.

fra begyndelsen af 1500-tallet frem til i dag. Selv om bibelsproget er en stilart der i løbet af denne periode fremtræder stadig mere konservativ, står det ikke stille, og den sproglige udvikling man kan observere fra de ældre til de yngre bibeltekster, peger bl.a. på i hvilken rækkefølge de forskellige verbalklasser opgav numerusbøjningen.

Ifølge Ejrnæs (2020) kan dansk bibeloversættelse inddeles i tre faser. I den første fase bygger oversættelserne på Luthers tyske oversættelse fra 1545. Denne fase indledes med udgivelsen af Christian 3.s Bibel i 1550 der på Frederik 2.s foranledning revideres og genudgives i 1589, hvorefter Christian 4. i 1633 atter giver teksten nyt liv i en revideret udgave. I disse oversættelser er sproget generelt mundret og utvungent i overensstemmelse med Luthers eget oversættelsesideal (Ejrnæs 2020: 72).

I 1600-tallet begynder anden fase der strækker sig frem til begyndelsen af 1900-tallet. I modsætning til tidligere går man nu ud fra henholdsvis den hebraiske og den græske originaltekst. Det er Hans Poulsen Resen der indleder denne tradition med sin oversættelse fra 1607. Denne udgave blev efterfølgende revideret adskillige gange – først af Hans Svane i 1647. Hermed var grunden lagt til den resen-svaningske oversættelsestradition der videreførtes med nye oplag og revisioner helt frem til revisionen af Det Nye Testamente fra 1907. Af vigtige udgivelser inden for dette tidsrum nævner Ejrnæs (2020: 62) Christian 6.s Bibel fra 1740 samt revisionerne af Det Nye Testamente og Det Gamle Testamente fra 1819 resp. 1871. Sproget i de resen-svaningske bibler adskiller sig fra Luthertraditionens ved at være ”tungt og indimellem knudret og vanskeligt at læse” (Ejrnæs 2020: 61).

Den tredje fase, der strækker sig over hele 1900-tallet, er kendetegnet ved mange forskellige nyoversættelser af både officiel og privat art. Her tilstræbes et mere tidssvarende sprog der i løbet af århundredet fjerner sig stadig mere fra den tradition der har hersket i flere århundreder.

Da vi i nærværende undersøgelse beskæftiger os med verbalbøjningen indtil 1900, medtages kun oversættelser fra de to første faser. Nedenfor følger en liste over de syv udvalgte udgaver i kronologisk rækkefølge med forkortelser der henviser til kildelisten. Da det ikke har været muligt at få adgang til de første oplag af Chr. 6.s Bibel fra 1740 og 1819-revisionen af Den Resen-Svaningske Bibel, har vi medtaget to senere oplag fra henholdsvis 1746 og 1824.

1500-tallet:

- Christian 3.s Bibel 1550 (Chr. 3. Bib): oversættelse af Luthers tekst

1600-tallet:

- Resens Bibel 1607 (Res Bib): Hans Poulsen Resens oversættelse af den hebraiske og den græske originaltekst
- Christian 4.s Bibel 1633 (Chr. 4. Bib): revision af Frederik 2.s Bibel fra 1589 der igen er en revision af Chr. 3.s Bibel fra 1550

1700-tallet:

- Christian 6.s Bibel 1746 (Chr. 6. Bib): revision af Hans Svanes reviderede udgave af Hans Poulsen Resens Bibel 1647, også kaldet Den Resen-Svaningske Bibel

1800-tallet:

- Den Resen-Svaningske Bibel 1824 (Bib 1824): revision af Den Resen-Svaningske Bibel
- Den Resen-Svaningske Bibel 1871 (Bib 1871): revision af Den Resen-Svaningske Bibel

1900-tallet:

- Den Resen-Svaningske Bibel 1907 (Bib 1907): revision af Den Resen-Svaningske Bibel

4.2. Annotering, resultater og analyse af en udvikling

Det samlede korpus består af Lukasevangeliets kapitel 1-6 i alle de syv ovennævnte udgaver, hvilket svarer til ca. 10 sider pr. tekst. Vi har markeret alle forekomster af aktive verber med subjekter i pluralis og inddelt dem i grupper efter verbalklasse, modus og tempus. Ud fra disse kriterier har vi fundet eksempler på stærke, svage og præterito-præsensiske verber (inkl. *ville*) samt verbet *være* i præsens og præteritum indikativ samt imperativ. Verber i konjunktiv med subjekter i pluralis forefindes ikke i materialet. Undervejs viste det sig at kongruens i 2. person singularis ved verberne *kunne*, *skulle*, *ville* og *være* har gennemgået en parallel udvikling. Disse former er derfor også medtaget i undersøgelsen.

Gennem hele forløbet er bøjningsforholdene konsistente inden for én og samme tekst. Ligeledes kan det konstateres at udviklingen henimod mindre grad af bøjning er ensrettet gennem hele perioden, dog med en-

kelte undtagelser i Chr. 6.s Bibel fra 1746 som er den eneste tekst hvor præsensformerne *kand*, *skal* og *vil* anvendes frem for *kunne*, *skulle* og *ville*. Udviklingen forløber sådan at der til og med Chr. 4.s Bibel fra 1633 hersker konsekvent numeruskongruens inden for alle verbalklasser, modi og tempora, hvorefter pluralisbøjningen reduceres delvis. Af eksempel (7)-(12) fra Chr. 3.s Bibel fremgår det hvordan stærke verber som *fare*, *finde*, *faa*, *komme*, *see* og *tage*, svage verber som *frycte*, *gøre* og *haffue* samt verbet *være* opviser entydig pluralisbøjning:

- (7) Oc der **vaare** Hyrder i den samme Egn (E 2^r)
- (8) Oc Engelen sagde til dem: **Frycter** eder icke (E 2^r)
- (9) Oc der Englene **fore** fra dem til Himmelen (E 2^v)
- (10) Oc de **komme** hastelige / oc **funde** baade Mariam oc Joseph (E 2^v)
- (11) **Seer** til / **gører** ret Penitentzis fruct / Oc **tager** eder icke faare at sige / Wi **haffue** Abraham til fader (E 3^r)
- (12) Da **finge** de hannem Esaie Prophetis Bog (E 3^v)

Den eneste tilsyneladende afvigelse fra streng numeruskongruens findes ved verbet *see* der i imperativ i de fleste tilfælde forekommer uden pluralisendelsen *-r*:

- (13) **See** / ieg forkynder eder stor Glæde (E 2^r)

Denne afvigelse må ultimativt skyldes at der i den græske originaltekst på dette og tilsvarende steder ikke står en 2. person pluralis imperativ af det græske verbum for ‘se’, men derimod *ἰδοὺ*, en demonstrativ partikel med betydningen ‘(se) dér’. Denne partikel videreføres i den latinske Vulgata af *ecce*, ligeledes en demonstrativ partikel, og i Luthers tyske oversættelse fra 1546³ af *sihe* (DDB 217), oprindeligt en 2. person singularis imperativ af verbet *sehen* ‘se’ hvis fossilerede karakter fremgår af den ellers entydigt pluraliske kontekst som kommer til udtryk i verbet *fürchtet*:

3 Vi anvender den autoritative udgave af Luthers oversættelse fra 1546.

- (14) **Fürchtet** euch nicht, **sihe**, Ich verkündige euch grosse
 frygt REFL.3.PL. ikke se jeg forkynder jer stor
 Freude (DDB 217)
 glæde
Frygt ikke, se, jeg forkynder jer stor glæde

Sihe svarende til *See* i (13) står i modsætning til den ægte 2. person pluralis imperativ *Sehet zu* (DDB 222) svarende til *Seer* i (11).⁴

Afviklingen af personkongruens i 2. person singularis ved *kunne*, *skulle*, *ville* og *være* følger parallelt med reduktionen af numeruskongruens. I Chr. 3.s Bibel kongruerer verberne endnu konsekvent med subjekter i 2. person singularis:

- (15) Paa det du **kant** haffue viss grund paa den Lærdom / i huilcken du
est vnderuist (E 1^r)

Reduktionen i verbalbøjningen viser sig første gang i Chr. 6.s Bibel fra 1746 hvor pluralisbøjningen inden for en afgrænset gruppe af stærke verber i præteritum konsekvent afvikles. Det drejer sig om verber med vokalisk eller konsonantisk stammeforandring mellem singularis og pluralis såsom *finde*, *faa* og *gaa* hvor de oprindelige pluralisformer *fundede*, *finge* og *ginge* erstattes af singularisformerne *fandt*, *fik* og *gik*. Også bøjningen af *komme* reduceres idet singularisformen *kom* overtager pluralisdomænet, måske fordi den oprindelige præteritumsform *komme* falder sammen med præsensformen. Bøjningen af stærke verber uden stammeforandring ændres derimod ikke, og de præterito-præsensiske og de svage verber opviser som bekendt i forvejen ingen numerusforskel i præteritum. Følgende eksempler fra Chr. 6.s Bibel svarer til eksempel (7)-(13) fra Chr. 3.s Bibel. De understregede former viser hvordan singularisformen har overtaget pluralisdomænet ved *finde*, *faa* og *komme*:

- (16) Og der **vare** hyrder i den samme egn (60)

- (17) Og Engelen sagde til dem: **frygter** ikke (60)

4 Formerne i nytestamentligt græsk, latin og tysk fungerer på synkront plan som partikler uagtet deres forskelligartede opståen, jf. Bauer (1988: 753-754 s.v. *idōi*), Ernout & Meillet (1985: 190-191 s.v. *ecce*) og Grimm & Grimm (1905: 144-145 s.v. *sehen*, specielt om imperativ).

- (18) Og det skede, der Englene **fore** fra dem til himmelen (60)
- (19) Og de **kom** hastelig, og **fandt** baade Maria og Joseph (61)
- (20) **Gjører** derfor omvendelsens værdige frugter, og **begynder** ikke at sige ved eder selv: vi **have** Abraham til fader (62)
- (21) Da **fik** de ham Esaias prophetes bog (63)

En konsekvent afvikling af personkongruens ved *kunne*, *skulle*, *ville* og *være* gennemføres på samme tid:

- (22) Paa det du **kand** kiende vis grund paa den lærdom, i hvilken du **er** bleven mundtlig underviist (58)

Som nævnt ovenfor er pluralisbøjningen i Chr. 6.s Bibel også fraværende i præsens indikativ ved *kunne*, *skulle* og *ville*, men det er en undtagelse der ikke gentages i nogen af de tre efterfølgende revisioner frem til 1907. Inden for de øvrige verbalklasser er der på intet tidspunkt vaklen i brugen af numeruskongruens i præsens indikativ og imperativ.

Figur 4 viser alle stærke verber med pluralissubjekt i Lukasevangeliets kapitel 1-6 i henholdsvis Chr. 3.s Bibel fra 1550 og Chr. 6.s Bibel fra 1746. Antallet af forekomster er angivet i parentes. Former uden numeruskongruens i den senere udgave er markeret med fed skrift.

Infinitiv	Chr. 3.s Bibel 1550			Chr. 6.s Bibel 1746		
	Indikativ		Imperativ	Indikativ		Imperativ
	Præs.pl.	Præt.pl.		Præs.pl.	Præt.pl.	
<i>bede</i>	<i>bede</i> (1)	<i>bade</i> (1)	<i>beder</i> (1)	<i>bede</i> (1)	<i>bade</i> (1)	<i>beder</i> (1)
<i>blijfue/blive</i>	<i>blijfue</i> (1)	<i>bleffue</i> (6)		<i>blive</i> (1)	<i>bleve</i> (6)	
<i>bære</i>		<i>baare</i> (1)			<i>bare</i> (1)	
<i>drage</i>	<i>drage</i> (1)					
<i>dricke/drikke</i>	<i>dricke</i> (2)			<i>drikke</i> (2)		
<i>falde</i>	<i>falde</i> (1)					
<i>fåre</i>	<i>fare</i> (1)	<i>fore</i> (3)		<i>fåre</i> (1)	<i>fore</i> (3)	
<i>finde</i>		<i>funde</i> (3)			<i>fandt</i> (2)	
<i>forlade</i>		<i>forlode</i> (1)	<i>forlader</i> (1)		<i>forlode</i> (1)	<i>forlader</i> (1)
<i>forskiude</i>	<i>forskiude</i> (1)			<i>forskiude</i> (1)		
<i>forstaa</i>		<i>forstode</i> (1)			<i>forstode</i> (1)	

	Chr. 3.s Bibel 1550			Chr. 6.s Bibel 1746		
<i>fāa</i>		<i>finge</i> (3)			<i>fik</i> (2)	
<i>giffue/give</i>		<i>gaffue</i> (1)	<i>giffuer</i> (1)		<i>gave</i> (2)	
<i>græde</i>	<i>græde</i> (1)			<i>græde</i> (1)		
<i>gaa</i>		<i>ginge</i> (6)			<i>gik</i> (5)	
<i>holde</i>		<i>hulde</i> (1)				
<i>komme</i>		<i>komme</i> (7)			<i>kom</i> (7)	
<i>lade</i>			<i>lader</i> (2)		<i>lode</i> (2)	<i>lader</i> (2)
<i>lee</i>	<i>lee</i> (1)			<i>lee</i> (1)		
<i>see</i>		<i>saae</i> (1)	<i>see</i> ⁵ (2), <i>seer</i> (1)		<i>saae</i> (1)	<i>see</i> ⁶ (2)
<i>side/sidde</i>	<i>side</i> (1)	<i>sade</i> (2)		<i>sidde</i> (1)	<i>sadde</i> (2)	
<i>springe</i>			<i>springer</i> (1)			<i>springer</i> (1)
<i>staa</i>		<i>stode</i> (1)			<i>stode</i> (1)	
<i>syncke/synke</i>		<i>suncke</i> (1)			<i>sank</i> (1)	
<i>tage</i>		<i>toge</i> (1)	<i>tager</i> (1)		<i>toge</i> (1)	
<i>udgaa</i>					<i>udgik</i> (1)	
<i>vinde</i>		<i>vunde</i> (1)				
<i>vride</i>		<i>vrede</i> (1)			<i>vredde</i> (1)	
<i>æde</i>	<i>æde</i> (2)	<i>ode</i> (1)		<i>æde</i> (2)	<i>aade</i> (1)	

Figur 4. De stærke verbers bøjning i indikativ og imperativ pluralis i Chr. 3.s Bibel fra 1550 og Chr. 6.s Bibel fra 1746

Som det ses af figur 4, er der i Chr. 6.s Bibel sket en systematisk afvikling af pluralisbøjningen i præteritum ved stærke verber med vokalisk eller konsonantisk stammeforandring mellem singularis og pluralis sådan at fx *suncke* og *finge* er vejet for *sank* og *fik*. Denne udvikling svarer til det overordnede indtryk som Mikkelsen (1975 [1911]: 463) har af overleveringen. Bøjningsmønstret i figur 4 fortsætter uændret i alle de undersøgte tekster helt frem til 1907.

I modsætning til alle hittidige undersøgelser kan vi således påvise en egentlig udvikling i verbal numeruskongruens i dansk skriftsprog. Mens Jessen (1864) og Sørensen (1900) slår ned forskellige steder i litteraturen for at eksemplificere den store variation i numerusbøjningen der har hersket i flere århundreder, anskueliggøres i nærværende studie en ensrettet om end langtrukken udvikling henimod mindre pluralisbøjning. Selv om de udvalgte tekster tilhører blot en enkelt genre, er de på flere punkter interessante. For det første kan de være med til at sandsynliggøre at nu-

5 Om denne form se forklaringen til eksempel (13) ovenfor.

6 Se fodnote 5.

meruskongruens i talesproget løbende afvikledes inden for det undersøgte tidsrum sådan som det også er blevet foreslået af bl.a. Sørensen (1902: 28-29). Man kunne forestille sig at pluralisformerne omkring udgivelsen af Chr. 6.s Bibel i 1740 stort set var forsvundet fra talesproget med det resultat at den almindelige dansker ikke længere var fortrolig med de pluralisformer der i præteritum opviser stammeforandring mellem singularis og pluralis. De øvrige pluralisformer har været lettere at afkode og oprettholde i skriften fordi de blot fik tilføjet *-e* til den velkendte singularisform. Dermed kan vi også give et bud på i hvilken rækkefølge formerne afvikledes. Endvidere indicerer resultatet at genre og stilleje må have påvirket afviklingens omfang og hastighed. Bibelgenren har udviklet sig fra at være præget af Luthers ideal om en så mundret og hverdags sproglig stil som muligt til at blive stadig mere traditionsbundet og konservativ. Det forklarer hvordan man så sent som i 1907 efter afskaffelsen af obligatorisk pluralisbøjning kunne udgive en bibel der med hensyn til numeruskongruens var stort set identisk med de foregående udgaver helt tilbage til 1700-tallet.

5. Konklusion

Med nærværende artikel har vi taget første skridt på vejen til en nærmere udforskning af verbal numeruskongruens i dansk fra ca. 1500 til ca. 1900. På baggrund af syv forskellige bibeloversættelser fra tidsrummet 1550-1907 kan vi i modsætning til Jessen (1864) og Sørensen (1900) påvise en ensrettet og systematisk udvikling inden for bibelsproget fra konsekvent numeruskongruens i alle de undersøgte bibler frem til 1746 hvor pluralisbøjningen afvikles i præteritum ved stærke verber med vokalisk eller konsonantisk stammeforandring mellem singularis og pluralis. Dette indicerer at der inden for visse genrer og stilarter kan findes systematiske udviklinger som bidrager til at tegne et mere nuanceret billede af forløbet frem til den endelige afvikling af verbal pluralisbøjning. Dermed er grunden lagt til videre undersøgelser.

Om forfatterne

Kathrine Thisted Petersen, lektor, ph.d., Institut for Kommunikation og Kultur, Aarhus Universitet.

Steffen Krogh, lektor, ph.d., Institut for Kommunikation og Kultur, Aarhus Universitet.

Kilder

- Bib 1824 = Revision af Den Resen-Svaningske Bibel 1824. *BIBLIA, det er: den ganske Hellige Skrifts Bøger, med Flid efterseete og rettede efter Grundtexten, saa og med mange Parallelsteder og udførlige Indholdsfortegnelser forsynede.* Kjøbenhavn: det Kongelige Vaisenhus.
- Bib 1871 = Revision af Den Resen-Svaningske Bibel 1871. *Bibelen, det er: den ganske Hellige Skrifts Bøger, med Flid eftersete og rettede efter Grundtexten, saa og med mange Parallelsteder og udførlige Indholdsfortegnelser forsynede.* Kjøbenhavn: det Kongelige Vaisenhus.
- Bib 1907 = Revision af Den Resen-Svaningske Bibel 1907. *Bibelen, det er: den ganske Hellige Skrifts Bøger, med Flid eftersete og rettede efter Grundtexten, saa og med mange Parallelsteder og udførlige Indholdsfortegnelser forsynede.* Kjøbenhavn: Bibelselskabet for Danmark.
- Chr. 3. Bib = Christian 3.s Bibel 1550. *Biblia/ Det er den ganske Hellige Scrifft/ vdsæt paa Danske.* Københaffn: Ludowich Dietz.
- Chr. 4. Bib = Christian 4.s Bibel 1633. *Biblia Det er/ Den ganske Hellige Scrifft/ paa Danske/ igien offuerseet oc Prentet effter vor allernaadigste Herris oc Kongis K. Christian den IV. Befaling. Mett Register/ alle D. Lutheri Fortaler/ hans Vdlegning i Brædden/ oc Viti Theodori Summarier.* Kiøbenhaffn: forlag ikke angivet.
- Chr. 6. Bib = Christian 6.s Bibel (revision af Den Resen-Svaningske Bibel) 1746. *BIBLIA, Det er, Den ganske Hell. Skriftes Bøger, Ved Hans Kongel. Måjests. Vor Allernaadigste Arve-Herres, Kong Christian Den Siettes Christelige Omsorg, Med Fliid og efter Grund-Texten efterseete og rettede, Saa og med Paralleler forsynede.* 4. oplag. Kjøbenhavn: det Kongel. Waysenus.
- Chr. 4. Br = Christian 4.s Breve 1584-1648. J. Skovgaard (red.) (1947): *Kong Christian den Fjerdes egenhændige Breve.* Bd. 8. 1584-1648. Supplement til C.F. Brickas & J.A. Fridericias udgave. København: Gad.
- DDB = Die Deutsche Bibel 1546. *D. Martin Luthers Werke. Kritische Gesamtausgabe.* Bd. 6. *Die Deutsche Bibel.* Weimar: Hermann Böhlhaus Nachfolger.
- LB 3 = Lund-benstykke 3. <https://runer.ku.dk/q.php?p=runer/genstand/genstand/298> (lokaliseret 2.1.2024).
- LB 4 = Lund-benstykke 4. <https://runer.ku.dk/q.php?p=runer/genstand/genstand/602> (lokaliseret 2.1.2024).
- Res Bib = Hans Poulsen Resens Bibel 1607. *BIBLIA, Paa Danske/ Det er/ Den ganske hellige scriftis Bøgger igennem seete med flijd effter den Ebræiske oc Grækske text etc. det negste mueligt vaar/ Oc effter som de paa andre atskillige Tungemaal vaare best vdsatte.* Kiøbinghaffn: Niels Michelson.

- Visit = Peder Palladius' visitatsbog ca. 1543. Lis Jacobsen (red.) (1925-1926): *Peder Palladius' Danske Skrifter*. Bd. 5. København: H.H. Thiele.
- VL = Valdemars Sjællandske Lov i Stockholm C 69,8° ca. 1400. I: *Danmarks gamle landskabslove med kirkelovene*. Bd. 8. *Valdemars sjællandske lov. Ældre og yngre redaktion samt Sjællandske kirkelov*. Erik Kroman (red.) (1941). København: Gyldendal, 3-104.

Litteratur

- Bauer, Walter (1988): *Griechisch-deutsches Wörterbuch zu den Schriften des Neuen Testaments und der frühchristlichen Literatur*. 6. Auflage herausgegeben von Kurt Aland und Barbara Aland. Berlin/New York: Walter de Gruyter.
- BEK nr. 184 af 15.12.1900 = *Bekendtgørelse angaaende Brugen af Verbernes Flertalsform ved Undervisningen i Modersmaalet*. <https://www.retsinformati-on.dk/eli/lta/1900/184> (lokaliseret 2.1.2024).
- Bertelsen, Henrik (1979 [1915-1929]): *Danske Grammatikere fra Midten af det syttende til Midten af det attende Aarhundrede*. Bd. 1-6. København: C.A. Reitzel.
- Bjerrum, Anders (1949): Verbal Number in the Jutlandic Law. *Travaux du Cercle Linguistique de Copenhague* 5. *Recherches structurales*, 156-176.
- Brøndum-Nielsen, Johs. (1918): Danske lovhåndskrifter og dansk lovsprog i den ældre middelalder. *Arkiv för nordisk filologi* 34, 105-137.
- Brøndum-Nielsen, Johs. (1959): *Festskrift udgivet af Københavns Universitet i Anledning af Hans Majestæt Kong Frederik IX's 60 Aars Fødselsdag den 11. Marts 1959. Fra Skaanes Senmiddelalder*. København: Bianco Lunos Bogtrykkeri.
- Ejrnæs, Bodil (2020): Bibelsprog. I: *Dansk Sproghistorie*. Bd. 4. *Dansk i brug*. Ebba Hjorth (hovedredaktør), Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). Aarhus: Aarhus Universitetsforlag, 59-98.
- Ernout, Alfred & Antoine Meillet (1985): *Dictionnaire étymologique de la langue latine. Histoire des mots*. Quatrième édition par Jacques André. Paris: Éditions Klincksieck.
- Falk, Hjalmar & Alf Torp (1900): *Dansk-norskens syntax i historisk fremstilling*. Kristiania: H. Aschehoug & Co.
- Frederiksen, Britta Orlík (2019): Gammeldansk. I: *Dansk Sproghistorie*. Bd. 3. *Bøjning og bygning*. Ebba Hjorth (hovedredaktør), Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). Aarhus: Aarhus Universitetsforlag, 41-72.

- Gerner, Henrich (1678-1679): *Orthographia Danica*. I: Bertelsen (1979 [1915-1929]) Bd. 3, 11-146.
- Gerner, Henrich (1690): *Epitome Philologiæ Danicæ*. I: Bertelsen (1979 [1915-1929]) Bd. 3, 251-317.
- Grimm, Jacob & Wilhelm Grimm (1905): *Deutsches Wörterbuch*. Bd. 16. Leipzig: S. Hirzel.
- Høysgaard, Jens (1743): *Tres faciunt collegium*. I: Bertelsen (1979 [1915-1929]) Bd. 4, 185-215.
- Høysgaard, Jens (1747): *Accentuered og Raisonnered Grammatica*. I: Bertelsen (1979 [1915-1929]) Bd. 4, 249-488.
- Jacobsen, Henrik Galberg (2019): Yngre Nydansk. I: *Dansk Sproghistorie*. Bd. 3. *Bøjning og bygning*. Ebba Hjorth (hovedredaktør), Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). Aarhus: Aarhus Universitetsforlag, 93-111.
- Jessen, Edwin (1864): Danske bøjningsformers historie. *Særskilt Aftryk af Tidsskrift for Philologi og Pædagogik* 5, 1-22.
- Jessen, Edwin (1891): *Dansk Grammatik*. Kjøbenhavn: Gyldendal.
- Kock, Lavrids (ca. 1660): *Introductio ad linguam Danicam*. I: Bertelsen (1979 [1915-1929]) Bd. 1, 35-74.
- Marsløv, Erik Olsen (1678): *Nogle Betænkninger og Regle om Skrive-Rictigheden udi ded Danske Sprog* I: Bertelsen (1979 [1915-1929]) Bd. 3, 1-10.
- Mikkelsen, Kristian (1975 [1911]): *Dansk ordføjningslære med sproghistoriske tillæg. Håndbog for viderekomne og lærere*. København: Hans Reitzel.
- Pontoppidan, Erik (1668): *Grammatica Danica*. I: Bertelsen (1979 [1915-1929]) Bd. 2, 1-426.
- Ruus, Hanne (2019): Ældre Nydansk. I: *Dansk Sproghistorie*. Bd. 3. *Bøjning og bygning*. Ebba Hjorth (hovedredaktør), Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). Aarhus: Aarhus Universitetsforlag, 73-92.
- Schulz, P. (1724): *Danskens Skriverigtighed*. I: Bertelsen (1979 [1915-1929]) Bd. 4, 1-153.
- Skautrup, Peter (1944): *Det danske Sprogs Historie*. Bd. 1. *Fra guldhornene til Jyske lov*. København: Gyldendal.
- Syv, Peder (1685): *Den Danske Sprog-Kunst*. I: Bertelsen (1979 [1915-1929]) Bd. 3, 147-250.
- Sörensen, Axel (1900): Udsagnsordenes flertalsformer. En redegørelse. *Dania* 7, 145-209.

Sørensen, Axel (1902): *En ariadnetråd gennem en sproglig labyrinth i spørgsmålet om förtids-tillægsmåderne til velvillig overvejelse for vakkende*. København: Lehmann & Stage.

Markering af bestemthed i litauiske studerendes indlærerdansk

Birutė Spraunienė & Emil Jonas Slott

Abstract

The paper presents a study of how L1 Lithuanian university students acquire the category of definiteness in their L2 Danish. The purpose of the study is to establish the sequence of acquiring the system of definiteness marking in L2 Danish. The research has showed that the informants' acquisition of the types of Danish nominals with (lexical as well as grammatical) (in)definiteness markers depends upon two main factors: the morphosyntactic complexity of the nominals and transfer from the informants' native language, which lacks articles. The latter manifests itself in the overall tendency to overuse bare nouns. The results of the study essentially confirm the findings of previous research of L2 Swedish (Andersson 1994, Nyqvist 2013), albeit with slight deviations.

Nøgleord

andet- og fremmedsprogsindlæring, L2 dansk, grammatikafvigelser, bestemthed

1. Indledning

Denne artikel omhandler indlæring af bestemthedskategori i dansk, som er særligt vanskelig for litauere, der ikke har denne kategori i deres modersmål¹. Os bekendt foreligger der ingen forskning af tilegnelse af bestemthed i voksnes L2 dansk. Til gengæld er der lavet en række undersøgelser af bestemthed og artikelbrug i L2 svensk, både som andet- og fremmedsprog (se afsnit 2), som har dannet baggrund for vores studie. Vi stiller følgende forskningsspørgsmål:

1 Personlig kommunikation med lærere og indlærere.

1. Hvordan tilegner litauiske danskstuderende sig bestemthedsmarkering i dansk?
2. Hvilken udvikling i brug af bestemthedsmarkører sker der i takt med at informanternes L2-niveau stiger?

Vi begrænser os ikke til de grammatiske markører af bestemthed (artikler), men inddrager også de leksikalske markører såsom possessive pronomener (*hans, sin*), demonstrative pronomener (*denne, den*), universelle pronomener (*al, hver*), ubestemte pronomener (*nogen*) o.l. Formålet med undersøgelsen er at etablere en indlæringsgang af bestemthedssystemet i dansk, dvs. at afdække i hvilke tilfælde målsproglignende brug af bestemthedsmarkører indlæres tidligst, hvilke tilfælde indlæres senere, og hvad der især udfordrer de litauiske danskstuderende.

Vores data stammer fra 12 informanter der studerer dansk på Vilnius universitet, fra 1. til 4. studieår. Det består af 24 mundtlige fortællinger baseret på to billedbøger (ca. 4 timers lydoptagelse i alt). Fortællingerne transskriberedes, og en manuel søgning af nominaler med substantivisk kerne blev udført. I alt blev 1797 eksempler excerperet og analyseret for bestemthedsmarkering (se afsnit 3.1.).

Metodisk bygger vores undersøgelse på performansanalyse, som indebærer at man sætter lige så stor fokus på det som er korrekt, som på det som er afvigende fra målsproget, i L2 indlærerens sprogproduktion (Ellis & Barkhuizen 2006: 73, jf. Nyqvist 2013: 20). I lighed med Axelsson (1994) og Nyqvist (2013) udførte vi analysen af nominaler i forhold til obligatoriske kontekster, dvs. vi ser hvilken nominalform der er obligatorisk efter målsprogsnormen (se afsnit 4). I lighed med Nyqvist (2013: 62) antager vi at en høj korrekthedsprocent indikerer tidlig indlæring.

2. Tilegnelse af bestemthed i andet- og fremmedsprogindlæringssammenhæng

Tilegnelsen af bestemtheds kategorien hos L2 danskindlærere er et emne som har savnet fokus i litteraturen. Der er dog foretaget et antal undersøgelser som tilsigter at afdække tilegnelsesprocessen for bestemthed i L2 svensk, hvori markeringen af bestemthed ligger meget tæt op ad den i dansk. I dette afsnit gennemgår vi et udvalg af den forskning der er mest relevant for vores projekt.

Anette Eriksson & Elsie Wijk-Andersson (1988) undersøgte artikelbrug i skriftlig produktion af 10 polske og 10 tyske gæstestuderendes indlærere-

svensk i forhold til en kontrolgruppe af 10 svenske gymnasieelever. Deres undersøgelse har vist at andelen af fejl er betydelig højere i polske end i tyske informanternes materiale når det gælder de såkaldte *speciesfel* dvs. de tilfælde hvor informanten bruger et ubestemt nominal der hvor konteksten kræver et bestemt nominal, eller omvendt (ibid.: 8). Forfatterne påpeger at polakkerne overanvender nøgne substantiver: 61,5 % af alle *speciesfel* består af bortfald af enten ubestemt artikel eller enklitisk bestemt artikel (ibid.: 11). Polakkerne anvender ofte pronomener hvor man i svensk kan eller bør anvende bestemte eller ubestemte artikler. Sammenligning med kontrolgruppen har vist at de polske informanter anvender pronominer og genitiver som bestemmere langt hyppigere end de svenske informanter. Eriksson & Wijk-Anderssons studie tyder på at indlæring af bestemthed i målsproget afhænger af om der findes en grammatikaliseret bestemtheds-kategori i indlæreres modersmål.

Monica Axelssons (1994) doktorafhandling er baseret på uformelle interviews med 60 informanter med hhv. polsk, spansk og finsk som modersmål. Informanterne er inddelt i grupper efter deres L1 og derefter i to grupper efter deres L2's færdighedsniveau (fortsætter- og avanceret niveau). Hendes undersøgelse har vist at der findes tre niveauer i nominalens udvikling som ikke afhænger af indlærernes modersmål (se tabel 1).

Tabel 1. Nominalets udviklingsniveauer (jf. Axelsson 1994: 99-101)

Nominalets udviklingsniveauer	Nominaltype	Eksempler
Niveau 1 98.5 – 92.2 % ²	proprier	<i>Rom, Nederländerna</i>
	ubestemte nominaler med adjektiviske eller pronominale adled	<i>tre flickor, många timar, inget problem</i>
	nøgne substantiver	<i>flicka, flickor</i>
	bestemte nominaler med pronominale adled	<i>min syster, varje dag</i>
Niveau 2 69.2 – 60.6 %	nominaler med sammenfattere/påpegende pronominer/bestemte adjektiver og enklitiske artikler	<i>hela boken, den här dagen, svenska ambassaden</i>
	bestemt form med enklitisk artikel	<i>boken, rosorna</i>
	ubestemt form ental	<i>en bok, ett hus</i>
Niveau 3 33.5 – 21.7 %	ubestemt form ental med adjektivisk attribut	<i>en röd bok</i>
	bestemt form med adjektivisk attribut	<i>den röda boken</i>

2 Tallene angiver korrekthedsprocenter for de nominaltyper der tilhører det pågældende udviklingsniveau.

Axelssons studie har vist at nærmest alle indlærere opnår niveau 1, og niveau 2 opnås af mere avancerede indlærere, mens niveau 3 viser sig at blive opnået af ganske få indlærere (lav hyppighed i anvendelse og lav korrekthed).

Axelsson påpeger dog at der er nogle forskelle som kan skyldes transfer fra modersmålet: De spanske informanter, som har ubestemt artikel i deres L1, anvender nominaler med ubestemt artikel oftere og mere korrekt end de andre informantgrupper. De finske informanter demonstrerer mere frekvent og korrekt brug af nominaler med enklitisk bestemt artikel, sikkert fordi den minder dem om modersmålets fleksiver. De polske informanter har en tendens til forenkling ved at udelade bestemthedsmarkører (jf. Eriksson & Wijk-Anderssons (1988) resultat for skriftsproget).

Eeva Liisa Nyqvist (2013) har undersøgt brug af bestemthed i finske folkeskoleelevers indlærersvensk og har anvendt både skriftligt og mundligt materiale. Hun viser at eleverne har sværest ved indirekte anaforer (*en cykel ... forhjulet*) og tilfælde af dobbelt bestemthed. Nyqvists informanter lærer svensk i Finland, dvs. som fremmedsprog. Hun har lavet en longitudinal undersøgelse hvor hun har fulgt samme informanter i tre år (7., 8. og 9. klasse).

Nyqvists resultater stemmer overens med nominalets indlæringsgang etableret af Axelsson (1994): Nominaler uden artikler indlæres først, bestemt form ental er lettere at indlære end ubestemt form ental og nominaler med adjektiviske attributter og foranstillede artikler er sværest, især dobbeltbestemte nominaler.

Nyqvist har behandlet sine data statistisk for at finde ud af om de forskelle hun finder, er statistisk signifikante. Analysen har påvist at nøgne former er signifikant lettere at indlære end de øvrige former, og at possessive nominaler er signifikant lettere at indlære end alle de andre former med lavere korrekthedsprocent. Bestemt form ental er signifikant lettere end ubestemt form ental og bestemt form flertal.

Nyqvist påpeger at hendes resultater ikke stemmer overens med Procesbarhedsteoriens forudsigelser (for en præsentation af teorien se fx Piemann 1998) som fremgår af tabel 2:

Tabel 2. Applicering af Procesbarhedsteoriens fem stadier på bestemhedsindlæring i svensk intersprog baseret på Nyqvist (2013: 61-62), citeret fra Kołaczek (2020: 137).

Procedur	Betydelse/-frastyt	Exempel
1. Lexikon	Substantivets nakna former, helsekvenser ³	<i>bok, hund, hus, i skolan</i>
2. Kategoriprocedur	Bestämd form singularis och pluralis, ubestämd form pluralis	<i>hunden, hundar, hundarna</i>
3. Frasprocedur	Nominalfraser som innehåller mer än ett ord d.v.s. nominalfraser med attribut, substantivets ubestämda form singularis	<i>min bok, en bok, en bra bok, den bra boken</i>
4. Satsprocedur ⁴	-	-
5. Bisatsprocedur ⁵	Direkta och indirekta anaforer (val av species innebär att information överförs över sats- och meningsgränserna)	<i>motorn, vinterdäcken</i> när det har talats om en bil flera meningar tidigare

Ifølge Nyqvist indlæres ubestemt form flertal og bestemt form ental tidligt selvom de befinder sig på det andet stadiet i Procesbarhedsteoriens model. Når det gælder det tredje stadiet, som indeholder komplekse nominaler med adled, viser Nyqvists resultater at nominaler med possessive attributter samt ubestemte attributter (*min bok, tre dagar, goda böcker*) beherskes godt allerede i 7. klasse (bedre end nogle nominaltyper på andet stadiet) fordi de er formelt enkle (ikke indeholder artikler) og har semantiske ækvivalenter i finsk (Nyqvist 2013: 187-188).

Kołaczek (2020) har udarbejdet et longitudinelt studie af hvordan polske svenskstuderende tilegner sig bestemthed når det gælder indirekte anaforer (*en taxa ... chauffören*). Hun bemærker at de mindre erfarne indlærere har en tendens til at gøre langt mere brug af leksikalske markører (possessive pronominer) end bestemte suffikser når begge konstruktioner er mulige, hvilket er en bekræftelse af de tidligere nævnte resultater af Axelsson (1994) og Nyqvist (2013). Denne tendens lader til at aftage i takt med at indlærerens niveau stiger (Kołaczek 2020: 150).

3 Et svensk begreb for uanalyserbare helheder.

4 *Satsprocedur* (eng. *S-procedure*), som består i at indlæreren bliver i stand til at overføre informationer på tværs af syntagmer og fx tilegner sig kongruens af prædikative adjektiver (jf. Pienemann & Håkansson 1999: 398), har ikke nogen tilsvarende struktur når det gælder bestemhedsindlæring.

5 Bisatsprocedur (eng. *Subordinate clause procedure*) består i at indlæreren bliver i stand til at overføre informationer på tværs af sætningsgrænser, dvs. lærer at skelne mellem hel- og ledsætninger, tilegner sig ordstillingsregler mm. Ifølge Nyquist kan visse direkte og indirekte anaforer (jf. begrebet 'inferentiel bestemthed' hos Hansen & Heltoft 2011: 470) hvor der er flere sætningers afstand mellem korrelatet og det anaforske nominal, regnes til dette højeste stadiet i modellen (Nyquist 2013: 61-62).

Agebjörn (2021) har i sin afhandling undersøgt bestemthedsindlæring i hviderussiske studerendes indlærersvensk. Hans resultater indikerer at inklusion af en adjektivmodifikation øver en negativ indflydelse på brugen af ubestemte artikler i højere grad end det gør sig gældende for bestemte suffikser, hvilket bekræfter tidligere forskning af Axelsson (1994) (Agebjörn 2021: 113). Afhandlingen konkluderer at informanterne var mere tilbøjelige til at anvende hyppigt forekommende bestemtheds morfemer korrekt, og at det i mindre grad var tilfældet med morfemer som bruges sjældnere.

3. Undersøgelsens design

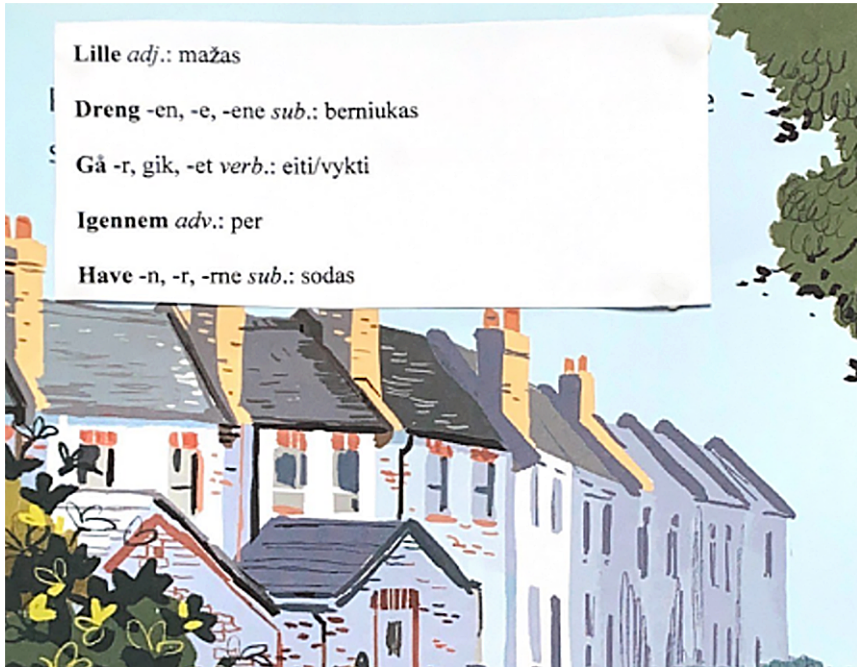
3.1. Dataindsamling

Projektet er baseret på mundtligt produceret L2⁶ dansk fra 12 informanter i alderen 19-23 år. De studerer alle dansk på BA-niveau – tre pr. årgang, 1.-4. år – på fuld tid, og studiet kræver intet forudgående kendskab til dansk sprog. Dataene består dels af interviews og dels af mundtlige fortællinger baseret på billedbøger. Interviewene med informanterne udførtes på litauisk med formålet at indsamle baggrundsdata samt information om hver informants indlæring og brug af dansk. De udgør altså en kontekst for undersøgelsens egentlige analyseobjekt: de mundtlige fortællinger af billedbøger på L2 dansk.

Fortællingerne er baseret på bøgerne *Where the Wild Things Are* af M. Sendak og *Granddad's Island* af B. Davies. Bøgerne er ikke produceret som rene billedbøger, men i stedet har vi fjernet teksten og indført gloser (kun indholdsord: substantiver, verber, adjektiver og enkelte adverbier). Vi forsynede også de danske gloser med bøjningsformerne, fx:

6 Rettere sagt er dansk det tredje eller fjerde fremmedsprog for vores informanter da alle kan engelsk og nogle derudover behersker andre sprog på forskelligt niveau (fx russisk, tysk eller fransk) inden de begynder at studere dansk.

Illustration 1. Et billede fra Granddad's Island med gloseliste



Informanterne var klar over at de deltog som informanter i et forskningsprojekt, men blev ikke fortalt at projektet havde at gøre med bestemt-
 hed specifikt. Både interviews og billedbogsfortællinger er indsamlet af
 en 4.-årsstuderende som var tilknyttet projektet i dets indledende fase⁷.
 Samtlige interviews og fortællinger blev optaget med lyd, men uden video.

3.2. Databehandling

Alle fortællinger er transskriberet, og samtlige nominaler med substanti-
 visk kerne (undtagen *proprier*) – i alt 1797 eksempler – annoteret og kate-
 goreret af to modersmålstalere. Udover nominaler med grammatiske (u)
 bestemthedsmarkører (artikler) har vi inkluderet nominaler med leksikal-
 ske (u)bestemthedsmarkører såsom pronominer (*hans, sin, denne, den, nogen,*
hver) og sammenfattere (*al, hele*) i materialet. Modersmålstalernes annota-
 tioner indikerer hvor informanternes produktion af nominaler afviger fra
 målsprogets normer, sådan at hvis begge modersmålstalere har indikeret

⁷ Projektet modtog økonomisk støtte fra Det Litauiske Forskningsråd til at finansiere Austė Vagoraitės deltagelse som forskningsassistent jf. projektansøgning Nr. P-ST-22-218.

en afvigelse, anses det for bekræftet mens et nominal som kun er markeret af én modersmålstaler, anses for tvetydigt.

De identificerede afvigelser er efterfølgende inddelt i to overordnede kategorier, nemlig 1) **afvigelse i form**, og 2) **afvigelse i brug**. Afvigelser i form er tilfælde hvor den anvendte nominalform er passende for konteksten når det gælder (u)bestemthed, men den konkrete form er afvigende. Vi har valgt at udelade fejl i genus og numerus (fx *en stor træ*, K2C⁸, *hans armer* og *bener*, K2A) samt valg af afvigende possessivpronomen (fx *så komme han på besøg hans bedstefar*, K2C) fra kategorien afvigelse i form. Nedenfor gives eksempler på afvigelser i form:

1. Efterhængt artikel bruges i stedet for foranstillet artikel: *lille drengen* (→ **den** lille dreng), K2C.
2. Bestemt form bruges efter genitiver eller possessiver: *bedstefars loftet* (→ loft) K3C, *hans navnet* (→ navn), K2A.
3. Påpegende pronomen (foranstillet bestemt artikel) bruges i stedet for efterhængt artikel: *det ø* (→ **øen**), K2C.
4. Ubestemt artikel er redundant eller udebliver⁹: *lille dreng* (→ **en** lille dreng) *gik igennem en have til hendes bedstefars hus*, M1B.
5. Adjektivets bestemthedsflektiv mangler: *det størrest skib* (→ det største skib), K2C.

Afvigelse i brug indikerer modsat at bestemthedsmarkørerne er konstrueret korrekt, men er fejlagtigt anvendt. Et nominal blev kategoriseret som afvigelse i brug i fire tilfælde:

1. Nøgen form bruges i stedet for bestemt form: *lille dreng* (→ **den** lille dreng), K3B.
2. Bestemt form bruges i stedet for nøgen form: *i jungleøen der er meget dyrelivet* (→ dyreliv), K4A.

8 Forkortelserne henviser til informanterne: K og M står for henholdsvis kvinde og mand, tallet står for studieår, og så blev informanterne på hvert studieår tildelt et bogstav: A, B eller C.

9 Denne afvigelse kan i princippet kategoriseres som en afvigelse i brug eftersom nøgen form og ubestemt form af substantivet (*avis* vs. *en avis*) har forskellige betydninger og funktioner. Nyquist regner dog bortfald eller overanvendelse af ubestemt artikel som en afvigelse i form. Hendes argument er at denne fejl ikke medfører en ændring i ”speciesbetydelsen” eftersom både nøgne og ubestemte former angiver ubestemt betydning (Nyquist 2013: 137). Selvom vi ikke helt er enige med Nyquist, har vi valgt at følge hendes kategorisering for at kunne sammenligne vores resultater med hendes.

3. Bestemt form bruges i stedet for ubestemt form: *og en dag bedstefar sad i **stolen*** (→ **en stol**) *og i dækken og læste en avis, K3C.*
4. Ubestemt form bruges i stedet for bestemt form: *de dansede og hylede foran **en fuldmåne*** (→ **fuldmånen**) *ligesom rigtig ulv, K2B.*

4. Resultater

4.1. De overordnede resultater og nominalers indlæringsgang

Tabel 3 præsenterer analysens resultater for hver informant.

Tabel 3. Korrekte og afvigende nominaler pr. informant

Infor- mant	Kor- rekt	%	Afvi- gelse i brug	%	Afvi- gelse i form	%	Tve- tydig	%	I alt	%
K4A	105	77 %	15	11 %	13	10 %	3	2 %	136	100 %
K4B	131	90 %	6	4 %	6	4 %	3	2 %	146	100 %
M4C	152	86 %	10	6 %	7	4 %	8	5 %	177	100 %
M3A	141	96 %	3	2 %	1	1 %	2	1 %	147	100 %
K3B	80	59 %	37	27 %	16	12 %	3	2 %	136	100 %
K3C	125	88 %	11	8 %	4	3 %	2	1 %	142	100 %
K2C	126	70 %	19	11 %	29	16 %	5	3 %	179	100 %
K2A	145	81 %	13	7 %	15	8 %	6	3 %	179	100 %
K2B	123	85 %	10	7 %	9	6 %	2	1 %	144	100 %
K1C	79	64 %	29	24 %	11	9 %	4	3 %	123	100 %
K1A	108	65 %	35	21 %	18	11 %	5	3 %	166	100 %
M1B	81	66 %	26	21 %	10	8 %	5	4 %	122	100 %
	1397	78 %	214	12 %	138	8 %	48	3 %	1797	100 %

Som det fremgår af tabel 3, har hver informant produceret fra 122 til 179 nominaler, dvs., 150 i gennemsnit. Førsteårsstuderende har nogenlunde samme korrekthedsprocent, men de øvrige informanternes resultater flugter ikke med hvilket studietrin de er på. Fx har K3B præsteret dårligst, og M3A bedst selv om de begge er på tredje studieår, og K2A og K2B har scoret højere end K4A. Af den grund har vi inddelt informanterne i to grupper efter korrekthed: De informanter hvis korrekthedsprocent ligger på 70 % og lavere, kom i gruppe 2 (lavere niveau), og resten i gruppe 1 (højere niveau). Disse grupper har dannet basis for videre analyse af udviklingen i indlærernes brug af bestemthedsmarkering.

Alle nominaler blev inddelt i 9 typer efter morfosyntaktisk kompleksitet, og der blev beregnet korrekthedsprocent for hver type. Nominalerne blev inddelt i typer på baggrund af de obligatoriske kontekster ifølge målsprogsnormen (jf. Nyqvist 2013: 20). F. eks. blev (1a) - (1c) kategoriseret som **bestemt form ental** da det er den form som konteksten kræver:

- (1a) *de var begge to på **dæk*** (→ dækket), K2C
 (1b) *han styrt **det skib*** (→ skibet), K2C
 (1c) *og han løb hen til **et vindue*** (→ vinduet), K2B

Tablet 4 viser hvordan de to informantgrupper anvender de forskellige nominaltyper:

Tablet 4. Korrekthedsprocent i forhold til nominaltyper for hver informantgruppe

Nominaltyper	Eksempel	Gruppe 2 (lavere niveau)			Gruppe 1 (højere niveau)		
		I alt	Kor- rekt	%	I alt	Kor- rekt	%
1. Nogen form sg og pl	<i>bog, bøger</i>	82	70	85,4 %	84	72	85,7 %
2. Def bestemmer + N	<i>hans/denne bog</i>	144	121	84,0 %	190	182	95,8 %
3. Indf bestemmer/attr + N	<i>mange (gode) bøger</i>	38	29	76,3 %	63	60	95,2 %
4. Bestemt form ental	<i>bogen</i>	172	122	70,9 %	331	295	89,1 %
5. Bestemt form flertal	<i>bøgerne</i>	57	39	68,4 %	56	34	60,7 %
6. Ubestemt form ental	<i>en bog</i>	82	51	62,2 %	158	141	89,2 %
7. Ubestemt form med attr	<i>en god bog</i>	53	26	49,1 %	74	70	94,6 %
8. Sammenfatter + bestemt form	<i>hele dagen, al maden</i>	17	2	11,8 %	22	11	50,0 %
9. Bestemt form med attr	<i>den lange dag</i>	33	2	6,1 %	40	38	95,0 %
		678	463		1018	903	

Inddeling i nominaltyper inspireredes af Axelsson (1994) og Nyqvist (2013). Vi fravalgte dog *proprier*, og nogle typer er strukturelt forskellige fra svensk (fx blev nominaler med påpegende pronominer henført til samme kategori som nominaler med possessive pronominer).

Ser man på gruppe 2 (indlærerne på lavere niveau), er det påfaldende at analyseresultaterne på mange punkter stemmer overens med de tidligere undersøgelser for indlærersvensk (Axelsson 1994, Nyqvist 2013): Nominaler uden de grammatiske bestemthedsmarkører (*bog, bøger, hans/denne*

bog, mange/nye bøger) indlæres først, så følger artikelformerne uden adjektiver (*bogen, bøgerne, en bog*), og nominaler med adjektiviske attributter (*en god bog, den lange dag*) indlæres sidst. Der er dog nogle forskelle. Nyqvists¹⁰ 7. klasse-elever anvender bestemt form flertal med betydeligt lavere korrekthedsgrad end bestemt form ental (hhv. 44 % og 66 %). I 9. klasse bliver forskellen sågar større: 43 % (flertal) vs. 79 % (ental) (Nyqvist 2013: 186-187). Hun konkluderer at bestemt form flertal er betydeligt sværere end bestemt form ental for finske indlærere (ibid.: 118). Dette er forventeligt fordi flertalsformen er mere kompleks med to grammatiske morfemer. I vores data ligger disse former således ifølge korrekthedsprocenten hos gruppe 2: bestemt form ental (70,9 %) og bestemt form flertal (68,4 %). Hos gruppe 1 ligner rækkefølgen dog de finske elevers: bestemt form ental (89,1 %) og bestemt form flertal (60,7 %). Bestemt form flertal er faktisk den eneste nominaltype hvor korrekthedsprocenten falder fra gruppe 2 til gruppe 1, hvilket er bemærkelsesværdigt. Antallet af belæg er dog for lille til at sige noget om litauernes indlæring af den pågældende nominaltype.

En anden forskel er at nominaler med sammenfattere (*hele natten*) anvendes med meget højere korrekthedsprocent hos Axelssons informanter end hos begge vores grupper (se tabel 1). Dette er den næstsværeste (gruppe 2) eller den sværeste (gruppe 1) nominaltype i vores materiale. Specielt volder ordet *hel* problemer da det kan bruges som et adjektiv (*en hel nat*), men den svage form *hele* anvendes som en sammenfatter efterfulgt af et substantiv med enklitisk artikel (*hele natten*). Informanterne på lavere niveau bruger forbindelser uden grammatiske morfemer (*hel dag*) mens alle informanterne udelader enklitisk artikel (*hele dag, alle ting* i stedet for *alle tingene* eller *alle de ting som ...*). De avancerede studerende bøjer nogle gange *hel* som et adjektiv: *den hele nat*. Vi er enige med Nyqvist som antager at den høje korrekthedsprocent for denne nominaltype hos Axelsson kan skyldes at Axelssons informanter, som bor i Sverige, har fået mere sprogligt input end fremmedsprogs elever og kan have indlært disse nominaler som uanalyserbare helheder (sv. *helsekvenser*) (Nyqvist 2013: 148-149).

Sammenlignes vores gruppe 1's præstation med Axelssons og Nyqvists resultater, er det påfaldende at de nøgne substantiver, som beholder et højt korrekthedsniveau (85,7 %), end ikke er den nemmeste nominaltype: De er overgået af 6 andre nominaltyper. Muligvis er en del af de korrekte anvendelser af nøgne substantiver hos gruppe 2 pseudokorrekte da

10 Axelsson henførte bestemt form ental og bestemt form flertal til samme kategori.

indlærere på begynderniveau ofte bruger dem som en slags default-form (jf. Pienemann & Håkansson 1999: 392, 398; Nyqvist 2013: 122). Indlærerne på højere niveau (gruppe 1) begynder at applicere de regler om artikelbrug som de kender, og det giver afvigelser da nøgen form ental er en slags restkategori og bruges i mange forskellige kontekster: med utællelige substantiver (*han savnede et sted hvor der var **kærlighed**, K3C*), visse prædikative nominaler (*men Max blev træt af at være **konge**, K3C*), tryktabsforbindelser (*gå i **seng**, lave **ballade***) og faste udtryk (*bedstefar har en stor erfar erfar som **sømand**, K4A*). Nogle gange bortfalder den ubestemte artikel ved tællelige substantiver (*det er **brev** fra hans bedstefar, K4A*) trods hovedreglen. Nøgen form af slægtskabsbetegnelser kan endda anvendes for at henvise til identificerbare referenter (*drengen spurgte **bedstefar** ville han ikke være ensom, K2B*). Nøgen form ental har altså ingen entydig semantik og kan derfor give besvær.

Den sidste større forskel fra tidligere studier i gruppe 1's data er den høje korrekthed for nominaler med adjektiviske attributter, både ubestemte (94,6 %) og bestemte (95 %). Disse nominaltyper ses både som lavfrekvente og vanskelige at indlære (Axelsson 1994: 99). Det lille antal belæg og manglen af statistiske beregninger tillader næppe entydige konklusioner. Alligevel synes vi at korrekt brug af de pågældende nominaltyper kan bero på at vores informanter er filologistuderende som har fået mange timers systematisk undervisning i dansk med langt større fokus på sproglig korrekthed end indlærere der begår sig i et målsprogligt miljø og som prioriterer at tilfredsstille deres kommunikative behov.

Opsummerende antyder resultaterne følgende indlæringsgang/sværhedshierarki (Nyqvist (2013: 62) bruger disse begreber synonymt) for indlæring af bestemthedsmarkering af de to informantgrupper:

Tabel 5. Indlæringsgang/sværhedshierarki for indlæring af bestemthedsmarkering

Indlæringsstadiet	Gruppe 2 (lavere niveau)	Gruppe 1 (højere niveau)
1. stadiet	<i>bog, bøger (85,4 %)</i> <i>hans/denne bog (84 %)</i> <i>mange (gode) bøger (76,3 %)</i>	<i>hans/denne bog (95,8 %)</i> <i>mange (gode) bøger (95,2 %)</i> <i>den lange dag (95 %)</i> <i>en lang dag (94,6 %)</i>
2. stadiet	<i>bogen (70,9 %)</i> <i>bøgerne (68,4 %)</i> <i>en bog (62,2 %)</i>	<i>en bog (89,2 %)</i> <i>bogen (89,1 %)</i> <i>bog, bøger (85,7 %)</i>
3. stadiet	<i>en lang dag (49,1 %)</i> <i>hele dagen (11,8 %)</i> <i>den lange dag (6,1 %)</i>	<i>bøgerne (60,7 %)</i> <i>hele dagen (50 %)</i>

Som det fremgår af tabel 5 tilegner litauiske indlærere sig først nominaler uden grammatiske bestemthedsmarkører. Dog forbliver nøgne substantiver lige så vanskelige som bestemt og ubestemt form ental for indlærerne på højere niveau. Nominaler med artikler og adjektivattributter er sværest (3. stadié) at indlære for informanterne på lavere niveau, men de topper (1. stadié) sværhedshierarkiet hos informanterne på højere niveau. Vores data bekræfter at bestemt form ental indlæres tidligere end ubestemt form ental; dog volder ubestemt form ental mindre besvær for informanterne på højere niveau og beherskes lige så godt som bestemt form ental. Nominaler med sammenfattere er den sværeste nominaltype for litauerne.

4.2. *Analyse af afvigelser*

Analysen har vist at de litauiske informanter på lavere niveau producerer langt flere afvigelser i brug end afvigelser i form, jf. tabel 6.

Tabel 6. *Afvigelser i brug og afvigelser i form pr. informantgruppe*

Informantgruppe	Afvigelser i brug		Afvigelser i form		Alle nominaler	
	n	%	n	%	n	%
Gruppe 1 (højere niveau)	68	6,3 %	55	5,1 %	1071	100 %
Gruppe 2 (lavere niveau)	146	20,1 %	84	11,6 %	726	100 %

Grunden kan være L1 transfer: Litauisk har ikke artikler, og det er derfor svært for litauiske indlærere at konceptualisere forskellen mellem 'identificerbart' og 'ikke-identificerbart'. Finsk er også artikelløst, men finske elever laver færre fejl i brug end fejl i form hvilket ifølge Nyqvist (2013: 190) kan skyldes at de kan overføre viden om valg af bestemthed fra engelsk. Til gengæld er andelen af afvigelser i brug og afvigelser i form hos informanterne på højere niveau næsten ens. I tabel 7 gives en oversigt over hvilke afvigelsestyper der forekommer i materialet.

Tabel 7. Afvigelsestyper pr. informantgruppe

Afvigelsestype	Gruppe 1		Gruppe 2	
	n	%	n	%
Ingen bestemthedsmarkør	44	35,8 %	86	37,6 %
Bestemthedsmarkør i stedet for nøgen form	9	7,3 %	12	5,2 %
Afvigende bestemthedsmarkør i et bestemt nominal	23	18,7 %	25	10,9 %
Bestemthedsmarkør i stedet for ubestemthedsmarkør	8	6,5 %	26	11,4 %
Ingen ubestemthedsmarkør	13	10,6 %	31	13,5 %
Overflødig bestemthedsmarkør i et bestemt nominal	11	8,9 %	22	9,6 %
Ubestemthedsmarkør i stedet for bestemthedsmarkør	5	4,1 %	4	1,7 %
Overflødig ubestemthedsmarkør i et ubestemt nominal	5	4,1 %	3	1,3 %
Hybride former	0	0,0 %	5	2,2 %
Andre afvigelser	5	4,1 %	16	7,0 %
I alt	123	100 %	230	100 %

Som det ses i tabel 7 er den hyppigste afvigelse hos begge grupper bortfald af bestemthedsmarkør i bestemte nominaler (fx og *pludselig fik drengen øje på en ø på **horisont*** (→ horisonten), K2A). Nyqvist (2013: 140) konstaterer også at overforbrug af nøgen form i stedet for bestemt form er signifikant hyppigere hos de finske elever end alle andre fejl i brug. Vi ser det som udtryk for L1 transfer.

En anden særligt hyppig afvigelsestype hos gruppe 1 er anvendelse af en afvigende bestemthedsmarkør i et bestemt nominal. Normalt bruges demonstrativt pronomener/foranstillet bestemt artikel i stedet for efterhængt artikel (fx *livet var meget dejlig i **det ø*** (→ øen), K2C), men i nominaler med adjektiver anvendes efterhængt artikel i stedet for den fritstående (denne afvigelsestype ses dog kun hos gruppe 2, fx *så lille **drengen*** (→ **den lille dreng**) *klarede op til loftet*, K2C). Overforbrug af *den* og *denne* i stedet for efterhængt artikel kan skyldes transfer fra modersmålet (de litauiske demonstrative pronomener *tas* 'den', *šis/šitas* 'denne') eller engelsk (artiklen *the*) (jf. Axelsson 1994: 44).

Begge informantgrupper har en tendens til at udelade ubestemt artikel (fx *så der var et barn som tog sig selv på **ulvekostume*** (→ et ulvekostume), K4A) og at bruge en redundant bestemthedsmarkør i et bestemt nominal (fx *bedstefars **loftet*** (→ loft), K3C). Brug af bestemthedsmarkør i stedet for ubestemthedsmarkør er hyppigere i gruppe 2's data (fx *han jagte hans hund og med **gaffelen*** (→ en gaffel), K1C), men ses også hos gruppe 1 (fx *i **loftet** den lille dreng finder tunge **tung metaldøren*** (→ en tung metaldør) og *åbner dem*, K4A).

Til sidst vil vi redegøre for udviklingen i brug af bestemthedsystemet der sker når man går fra lavere til højere niveau. Når man sammenligner gruppe 2's og gruppe 1's præstationer, kan man konstatere at der ikke kun sker et markant fald i andelen af afvigende nominaler, men også en kvalitativ ændring i afvigelserne. Nedenfor gennemgår vi en række afvigelsestyper der næsten udelukkende forekommer hos informanterne på lavere niveau:

1. hybride former som indeholder både bestemt og ubestemt artikel, fx ***en** dagen* (→ **en** dag), M1B, ***en** blå brevet* (→ **et** blå brev), K1A. De forekommer kun hos 1. årsstuderende.
2. efterhængte artikler i bestemte nominaler med adjektivattributter, fx *lille **drengen*** (→ **den** lille dreng), K2C
3. nøgen form ental i stedet for flertal, fx *mens Max var i sit værelse der **begynde at vokse træ*** (→ træ**er**) der, K3B
4. genitivkonstruktioner uden genitivs -s, fx ***bedstefar hus*** (→ bedstefar**s** hus), K1C
5. brug af efterhængt artikel i ubestemte nominaler med ordet *mange/meget* eller adjektiviske attributter, fx *mange **dyret*** (→ mange dyr), K1A, *rigtig **ulvene*** (→ rigtige ulve), K1C
6. nominaler med multiple afvigelser, fx *næste dag **drengen** går gik at **besøg hans bedstefaren huset*** (→ sit bedstefars hus), K1C, *bedstefar sat og **begyndte** fortælle af **rejse*** (→ sine rejser), M1B

5. Konklusion

På baggrund af vores analyse kan vi konkludere at de litauiske studerendes indlæring af bestemthedsmarkering (i den tidlige fase) er betinget af nominalers morfosyntaktiske kompleksitet: Nominaler uden grammatiske bestemthedsmarkører indlæres først (1. stadie), dernæst følger artikelformer uden adjektivmodifikation (2. stadie) og bestemte nominaler med

adjektiviske attributter samt bestemte nominaler med sammenfattere indlæres senest (3. stadie). Vores data bekræfter Nyquists (2013) resultater for finske elevers L2 svensk at bestemt form ental indlæres tidligere end ubestemt form ental. Dog volder ubestemt form ental ikke længere besvær for de litauiske informanter på højere niveau.

Sværhedshierarkiet for de studerende på højere niveau ser dog lidt anderledes ud. Her er bestemte nominaler med adjektivmodifikation steget markant i korrekthed (til 95 %) mens nøgne substantiver optræder med lavere korrekthed end de andre nominaltyper. Det tyder på at især nøgen form ental er svær at indlære fordi den ikke har entydig semantik og kan bruges i flere forskellige kontekster. Bestemte nominaler med sammenfattere forbliver den sværeste nominaltype, formentlig fordi de er strukturelt forskellige fra andre bestemte nominaler i dansk.

L1 transfer er en anden faktor der synes at gøre sig gældende i litauerne indlæring af bestemthed. Det manifesterer sig på flere måder: Informanterne laver flere fejl i brug end fejl i form, og den hyppigste afvigelsestype er bortfald af (u)bestemthedsmarkør og overforbrug af demonstrativer/bestemte artikler i stedet for efterhængte artikler. Med andre ord har litauerne (selv på højere niveau) af og til svært ved at afgøre om et nominal skal markeres som bestemt eller ubestemt, da de ikke har denne skelnen på deres modersmål. Det bekræfter at indlæring af bestemthed i L2 afhænger af om der findes en grammatikaliseret bestemthedskategori i ens L1.

Om forfatterne

Birutė Spraunienė, ph.d., Institut for sprog og kulturer i Østersøregionen, Center for Nordiske Studier ved Vilnius Universitet.

Emil Jonas Slott, lektor, Institut for sprog og kulturer i Østersøregionen, Center for Nordiske Studier ved Vilnius Universitet.

Litteratur

- Agebjörn, A. (2021): *Learning of Definiteness by Belarusian Students of Swedish as a Foreign Language*. Göteborg: Göteborgs Universitet.
- Axelsson, M. (1994): *Noun phrase development in Swedish as a second language: a study of adult learners acquiring definiteness and the semantics and morphology of adjectives*. Stockholm: Stockholm University.
- Ekerot, L.J. (2020): *Ordföljd, tempus, bestämdhet. Föreläsningar om svenska som andraspråk*. Malmö: Gleerups.
- Ellis, R. & G. Barkhuizen (2006): *Analysing learner language*. Oxford: Oxford

University Press.

- Eriksson, A. & E. Wijk-Andersson (1988). *Swedish Nouns and Articles in German and Polish Students' Swedish Writing*. FUMS Rapport 141. Uppsala: Uppsala universitet.
- Hansen, E. & L. Heltoft (2011): *Grammatik over Det Danske Sprog, I-III*. Det Danske Sprog- og Litteraturselskab & Syddansk Universitetsforlag.
- Kořaczek, N. (2020): Bestämndhet och indirekta anaforer i polska studenters inläraarsvenska. *European Journal of Scandinavian Studies* 50(1), 134-152. <https://doi.org/10.1515/ejss-2020-0007>
- Nyqvist, E.L. (2013): *Species och artikelbruk i finskspråkiga grundskoleelevers inläraarsvenska. En longitudinell undersökning i årskurserna 7-9*. Doktorsavhandling. Åbo universitet.
- Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability theory and generative entrenchment. *Bilingualism, Language, and Cognition* 1, 1-20. <https://doi.org/10.1017/S1366728998000017>
- Pienemann, M. & Håkansson G. (1999): A unified approach towards the development of Swedish as L2: a processability account. *Studies in Second Language Acquisition* 21(3), 383-420. <https://doi.org/10.1017/S0272263199003022>

Struktur og konstituens i forskellige grammatiske tilgange

Sten Vikner & Ken Ramshøj Christensen

Abstract

We will show how to use ‘our own’ open source programme *SVG Syntax Tree Generator* to generate structures like (1b) below (<https://tildeweb.au.dk/au132769/syntaxtree/>).

Then we will show a way of arguing in favour of (or against) the existence of units/constituents in syntactic analysis, using constituency tests as found across various approaches, e.g. substitution or movement, cf. e.g. Diderichsen (1946: 163), Hansen & Heltoft (2011: 95–101), Vikner & Jørgensen (2017: 149–153).

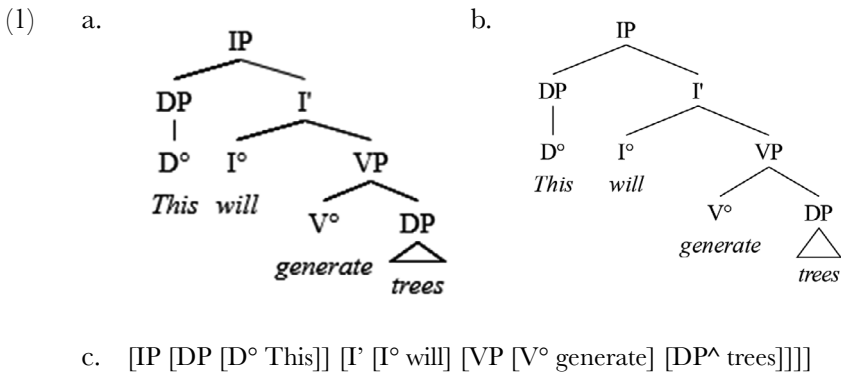
In the course of this discussion, we will address certain weaknesses in the analyses of Diderichsen (1946) and of Hansen & Heltoft (2011), and we will sketch how such problems are avoided in our own generative analysis. We will also argue that if a string of words can be moved or substituted, then it has to be taken to be a constituent in the syntactic analysis. In other words, we strongly advocate that positive evidence from constituency tests should be taken seriously.

Nøgleord

struktur, konstituent, konstituentest, verbalfrase (VP), verbalhypotagme, verbalhelhed, træstruktur, generativ tilgang, sætningsskema, pseudoklønning, substitution, flytning

1. *SVG Syntax Tree Generator*

I denne artikel¹ vil vi diskutere og sammenholde forskellige grammatiske tilgange til sætningers interne struktur. I den forbindelse vil vi påpege nogle aspekter af analyserne i Diderichsen (1946) og i Hansen & Heltoft (2011) som vi finder problematiske, og vi vil opridsse hvordan disse problemer kan undgås ved at antage en generativ analyse. Men først vil vi i dette afsnit helt kort introducere det program som vi har brugt til at tegne træstrukturerne i denne artikel.



*Syntax Tree Generator*² er et open source-program fra 2011, lavet af Miles Shang, og det kan generere syntaktiske træstrukturere i form af **png**-filer, der dog desværre let bliver pixeleret, se (1a).

Vi har modificeret dette program til et ”nyt” open source-program *SVG Syntax Tree Generator*³, der genererer syntaktiske træstrukturere i form af **svg**-filer. Med andre ord anvender programmet nu vektorgrafik, og derfor bliver resultatet ikke pixeleret, se (1b).

1 Tak til Patrick Clicqué-Göbel, Lars Heltoft, Eva Skafte Jensen, Henrik Jørgensen, Sune Sonderberg Mortensen, Lennart Westergaard, Carl Vikner og en anonym fagfællebedømmer, til publikum ved Grammatiknetværkets konference i oktober 2023 og ikke mindst til de studerende i vores kurser om syntaks. Forskningen bag denne artikel var en del af et forskningsprojekt ved Aarhus Universitet, finansieret af Forskningsrådet for Kultur og Kommunikation: *Ligheder og forskelle mellem sætninger og nominalgrupper - sammenlignende syntaks på tværs af teoretiske tilgange*.

2 Programmet *Syntax Tree Generator* findes på <http://mshang.ca/syntaxtree/>.

3 Programmet *SVG Syntax Tree Generator* findes på Sten Vikners website, <https://tildeweb.au.dk/au132769/syntaxtree/>, på Ken Ramshøj Christensens website, <https://tildeweb.au.dk/au572/syntaxtree/>, og også på <https://syntaxtree.abitcreative.co/>.

Begge programmer skal have input i form af en struktur med firkantede parenteser, som i (1c).

Vi anvender selv programmet i vores undervisning i engelsk grammatik, også fordi der er en pædagogisk pointe i at de studerende får fornemmelse for forbindelsen mellem på den ene side en notation med firkantede parenteser og på den anden side den tilsvarende træstruktur.

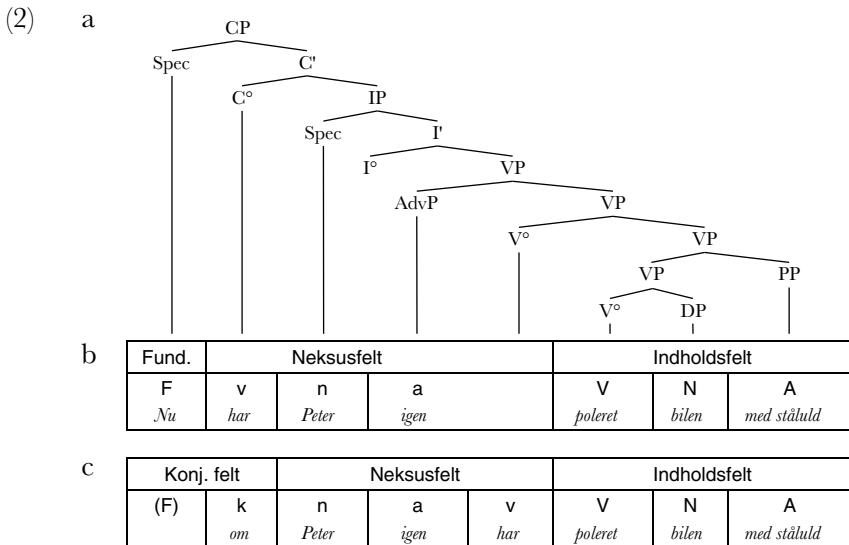
Sammenlignet med at tegne træer i *Word* er *SVG Syntax Tree Generator* en del hurtigere, og så er der indbygget symmetri. Til gengæld er der elementer man desværre ikke kan styre på samme måde som i *Word*, fx formatering af enkelte ord i træet: understregning, fed skrift, særlige farver, eller en særlig font, som i (3b) nedenfor.

I appendix til artiklen har vi angivet input til træerne i (16)-(22).

2. Struktur og konstituens

2.1 Træanalyser vs. feltanalyser

I denne artikel vil vi som nævnt diskutere og sammenholde forskellige grammatiske tilgange til sætningers interne struktur.

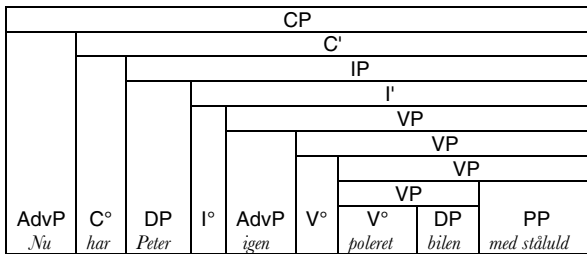


Diagrammerne i (2a-c) med både en generativ træanalyse og Diderichsens feltanalyse (1946: 186) har vi ofte brugt til at vise hvor mange paralleller der er mellem de to tilgange og dermed også hvordan indsigt fra den ene tilgang kan overføres til og anvendes i den anden (fx Bjerre & al. 2008:

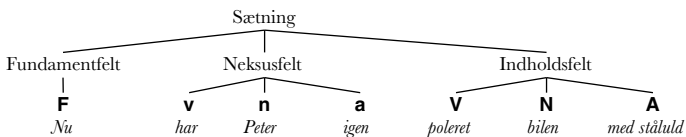
146; Nyvad 2014: 122, Vikner 1999: 86, Vikner 2015: 430, Vikner 2019: 178-179, Vikner & Jørgensen 2017: 139, se også Christensen 2005: 52).

I denne forbindelse anfører vi altid at der er mange vigtige forskelle mellem den generative analyse i (2a) og en Diderichsen-analyse som i (2b,c), jf. resten af denne artikel, men vi understreger at det kun er en notationsforskel, ikke en vigtig forskel, om en given analyse illustreres i en træstruktur eller i en feltstruktur (fx Vikner 2015: 431). Man kan omskrive den generative træstruktur i (2a) som feltstrukturen i (3a) uden noget tab af information. Det samme gælder omskrivningen af Diderichsen-feltstrukturerne i (2b,c) som træstrukturerne i (3b,c) (som også foreslået så tidligt som i Brandt 1974: 73). Med andre ord er (2a) og (3a) notationelle varianter af hinanden, ligesom (2b) og (3b) er notationelle varianter af hinanden, og (2c) og (3c) er notationelle varianter af hinanden.⁴

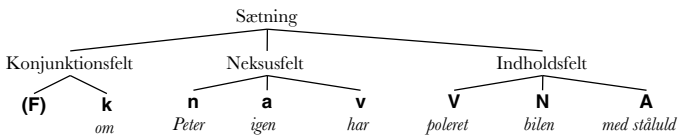
(3) a.



b.



c.



4 Selvom vi således mener at en træstruktur og en feltstruktur kan være notationelle varianter af hinanden, har der på intet tidspunkt været tale om at se Diderichsen-analysen som en notationel variant af vores generative analyse, eller vice versa – i modsætning til hvad Heltoft (2016: 72) lidt misvisende siger med henvisning til Vikner (2015: 430). De to tilgange har nemlig forskellige forudsigelser uafhængigt af valget af notation.

Forskellen i notation mellem en træstruktur og en feltstruktur er således ikke en vigtig forskel mellem den generative analyse i (2a) og en Diderichsen-analyse som i (2b,c) (eller en analyse efter Hansen & Heltoft (2011), som i (16c) nedenfor).

De vigtige forskelle mellem de to typer analyser skal derimod opregnes separat, og de består bl.a. af

- distinktionen mellem pladser og felter hos Diderichsen (1946), og hvad deraf følger,
- distinktionen mellem kerner og fraser i vores generative struktur, og hvad deraf følger,
- begrænsningen til binær forgrening i vores generative struktur vs. ingen tilsvarende begrænsning hos Diderichsen (1946) og Hansen & Heltoft (2011),
- og ikke mindst selve de konstituentter som sætningen deles op i i de forskellige analyser.

2.2 *Konstituentter i sætningsstrukturanalyserne*

Vi vil derfor nu fokusere på den detaljerede konstituentstruktur som er illustreret i (2) og (3), især hvad angår verbalhelheder/-hypotagmer. Vores diskussion er baseret på de elleve forsøg på pseudoklønninger i (5a-k), som igen har udgangspunkt i basissætningerne i (4a,b). Med udtrykket 'basissætning' mener vi at hvis pseudoklønningerne i (5a-k) ikke havde fundet sted, ville de transitive eksempler (5a,b,d,e,g,i,k) have været identiske med (4a) og de intransitive (5c,f,h,j) med (4b). Med andre ord kan de pseudokløvede sætninger ses som mere komplekse versioner af basissætningerne.

- (4) a. ... idet jeg altid gerne har villet kunne fejre min fødselsdag i fuld offentlighed.
 b. ... idet jeg altid gerne har villet kunne feste i fuld offentlighed.
- (5) a. Det jeg altid gerne har villet kunne fejre ___ i fuld offentlighed, er **min fødselsdag**.
 b. *Det jeg altid gerne har villet kunne fejre min ___ fuld offentlighed, er **fødselsdag i**.
 c. Det jeg altid gerne har villet kunne ___ i fuld offentlighed, er **at feste**.

- d. Det jeg altid gerne har villet kunne ____ i fuld offentlighed, er **at fejre min fødselsdag.**
- e. *Det jeg altid gerne har villet kunne ____ min fødselsdag i fuld offentlighed, er **at fejre.**
- f. Det jeg altid gerne har villet kunne ____, er **at feste i fuld offentlighed.**
- g. Det jeg altid gerne har villet kunne ____, er **at fejre min fødselsdag i fuld offentlighed.**
- h. Det jeg altid gerne har villet ____, er **at kunne feste i fuld offentlighed.**
- i. Det jeg altid gerne har villet ____, er **at kunne fejre min fødselsdag i fuld offentlighed.**
- j. *Det jeg altid gerne har villet ____ i fuld offentlighed, er **at kunne feste.**
- k. *Det jeg altid gerne har villet ____ i fuld offentlighed, er **at kunne fejre min fødselsdag.**

Anvendelsen af en anden konstituentstest, nemlig flytning til fundamentfeltet, ville have givet et parallelt resultat.

Argumentation bag sådanne konstituentstests er at hvis to eller flere ord enten sammen kan flyttes til et andet sted i sætningen eller sammen kan substitueres med ét element, så er det fordi disse to eller flere ord tilsammen udgør en enhed, altså en konstituent. Til gengæld kan man ikke konkludere noget om to eller flere ords konstituentstatus hvis disse ord hverken kan flyttes sammen til et andet sted i sætningen eller sammen kan substitueres med ét element. Dette kunne nemlig skyldes at disse ord ikke udgør en konstituent, men det kunne muligvis også have andre grunde.

Derudover vil vi også straks indrømme at man i vores sætningsstrukturanalyse kan blive nødt til at antage eksistensen af en konstituent som ikke kan flyttes eller substitueres, selvom antagelsen af en sådan konstituent naturligtvis så er mindre overbevisende end hvis den kunne flyttes eller substitueres. Et eksempel på dette er at vi mener at (4a) indeholder mindst tre forskellige VPer, nemlig ikke bare *fejre min fødselsdag i fuld offentlighed*, hvilket understøttes af (5g), og *kunne fejre min fødselsdag i fuld offentlighed*, hvilket understøttes af (5i), men også *villet kunne fejre min fødselsdag i fuld offentlighed*, selvom det er meget svært at finde understøttelse i form af konstituentstests for denne sidste VP.

Et andet eksempel på elementer som er konstituent i sætningen men

som kun sjældent vil kunne bestå en konstituenttest, er de enkelte ord. Vi må anerkende at sætninger i sidste ende er bygget op af enkelte ord, og dermed slipper vi ikke uden om at anerkende at disse enkeltord også er konstituentter.

Vi er bevidste om at vi nedenfor i diskussionerne af eksemplerne i (5) gør det som Heltoft (2016: 74) kritisk beskriver på følgende måde:

- (6) Som nævnt i indledningen opfatter den generative grammatik Diderichsens felter og pladser som ækvivalente med konstituentstrukturer og konstituentter.

(Heltoft 2016: 74)

Dette anser vi imidlertid ikke for at være et problem. Når sætningsskemaerne hos Diderichsen (1946) indeholder felter og pladser, se (2b,c), så mener vi at disse felter og pladser klart er bestanddele af og dermed konstituentter i Diderichsens sætningsanalyse. Se fx definitionen af *konstituent* i *Nudansk Ordbog* (Becker-Christensen & Petersen 2010) som *et element i en sammensat helhed = bestanddel, led⁵*. Med andre ord, når sætningsskemaet er bygget op af felter og pladser, så må vi kunne sige at felter og pladser udgør byggesten i sætningens struktur.

Hertil kommer at både Diderichsen (1946) og Hansen & Heltoft (2011) ofte nævner *verbalhelheder/verbalhypotagmer*, og flytninger og substitutioner af disse. (7)-(12) nedenfor dokumenterer at verbalhelheder (som vi ville betegne som VP'er) omtales både ofte og udførligt i Hansen & Heltoft (2011), og vi anerkender hermed at Hansen & Heltoft (2011) er opmærksomme på relevansen og værdien af de citerede data og de anvendte konstituentstests.

Ikke desto mindre leder man forgæves efter sådanne konstituentter, dvs. verbalhelheder, i de konkrete analyser der angives i Diderichsen (1946) eller i Hansen & Heltoft (2011), idet de hverken findes i sætningsskemaerne (som (2b,c) ovenfor eller (16b,c) etc. nedenfor) eller i de alternative 'indholdsledsnotationsstrukturer' (se fx Hansen & Heltoft 2011: 911 øverst).

5 Vi er enige i *Nudansk Ordbog*'s (Becker-Christensen & Petersen 2010) antagelse af at *konstituent*-begrebet er det samme inden for og uden for sprogvidenskaben. Ordbogen har nemlig kun én definition af begrebet *konstituent*, og der anføres to eksempler på denne ene definition, hvoraf det ene er fra sprogvidenskaben og det andet ikke er: I. *sætningens konstituentter*; II. *Hidtil har vi antaget at maksimering er den vigtigste konstituent i rationalitetsbegrebet*.

- (7) ... et **Hypotagme med verbal Kerne** (*mødte en Mand i Gaar*) har en helt anden Bygning end et Hypotagme med et Genstandsord som **Kerne** (*en gammel Mand*) ...

(Diderichsen 1946: 141)

- (8) Infinitivet med dets Bestemmelser udgør et Hypotagme, der samlet kan staa som Fundament:
Sende Børnene hjem turde han ikke ...

(Diderichsen 1946: 166)

- (9) Et **VERBALHYPOTAGME** er et hypotagme med et verbum som kerne.
(Hansen & Heltoft 2011: 277)

- (10) VP = verb phrase, engelsk for **verbalhelhed**.
(Hansen & Heltoft 2011: 374)

- (11) Det udtryk der svarer til tema og rema er opdelingen af sætningen i subjekt og **verbalhelhed**, og argumentet for denne opdeling er at verbalhelheden behandles som en syntaktisk-topologisk enhed. [...]:

spille spar fem ud kan du altså ikke
**spille kan du altså ikke spar fem ud*
hente os en kasse øl gad han da godt
**hente gad han da godt os en kasse øl*

(Hansen & Heltoft 2011: 1198)

- (12) Der sker altså en semantisk henvisning til verbalhelhedens indhold (til dens remadel [...]), og rettere: til det forhold som denne betegner. Den er obligatorisk når det gælder valensled:

*Sara maledede døren. *Det gjorde jeg vinduet.*
*Lise opførte sig skrækeligt. *Det gjorde Jens eksemplarisk.*
*Hun rejser til Odense. *Det gør jeg hjem.*

Derimod kan frie led udskiftes i anaforsætningen:

Ida slår græs i denne uge. Det gjorde jeg i sidste uge.
Han renser pladen med terpentint. Jeg gør det med cellulosefortynder.
Vi står som regel på ski i Norge, men vi har en gang gjort det i Østrig

(Hansen & Heltoft 2011: 1766)

Som allerede sagt finder vi det problematisk at de verbalhelheder som omtales og påvises i (7)-(12), ikke genfindes i de konkrete analyser i Diderichsen (1946) og Hansen & Heltoft (2011). Vi finder det derudover også uheldigt at de verbalhelheder som omtales og påvises i (7)-(12), er i konflikt med de konstituent/bestanddele som findes i de konkrete analyser (fx feltet **V**). Vi mener nemlig at sådanne konflikter absolut bør undgås, dvs. at hvis man kan flytte eller substituere en streng af ord, så bør denne streng af ord også udgøre en konstituent i sætningsstrukturanalysen.

Et konkret eksempel på en sådan konflikt er at den verbalhelhed som kan pseudoklives i (5d), *fejre min fødselsdag*, går på tværs af grænsen mellem feltet **V** og objektsfeltet i de konkrete analyser i Diderichsen (1946) og Hansen & Heltoft (2011) (dette ses også i diskussionen af (19) nedenfor). Konflikten er altså mellem på den ene side den verbalhelhed som består af *fejre min fødselsdag* (som understøttes af konstituentests) og på den anden side feltet **V** i sætningsskemaet, der består af *har villet kunne fejre*. Hvis svaret på denne sidste type indvending er at der ikke er nogen grund til at forvente overensstemmelse mellem konstituent som verbalhelheder på den ene side og så sætningsskemaerne på den anden side, så er spørgsmålet hvad der så kan anerkendes som evidens for eller imod pladser og felter i sætningsskemaerne i Diderichsen (1946) og Hansen & Heltoft (2011). Som eksempel på en sådan evidens i forbindelse med en verbalhelhed kan anføres at en streng af ord kun kan udgøre en verbalhelhed hvis den er kompatibel med de forskellige flytnings- og substitutionstests, jf. hvorledes Diderichsen's (1946: 166) eksempel i (8) ovenfor understøtter antagelsen af verbalhelheden *sende børnene hjem*. Hertil kommer at hvis sætningsstrukturanalyserne ikke direkte siger noget om konstituent som verbalhelhederne, er det ikke klart for os hvordan man så kan redegøre for data som i (4)/(5), hvilket vi vil uddybe i afsnit 3 nedenfor.

Et sidste og måske mindre punkt er at vi mener at kunne vise at der er brug for at kunne have en verbalhelhed/VP inden i en anden verbalhelhed/VP inden for en og samme sætning, hvilket vi kommer tilbage til i (18)-(22) nedenfor. Så vidt vi kan se, er dette hverken en mulighed hos Diderichsen (1946) eller hos Hansen & Heltoft (2011), idet der her højst kan være én verbalhelhed i hver sætning (uanset antallet af verber):

- (13) Dansk sætningsstruktur er således opbygget at der altid skal være et subjekt som en **verbalhelhed** udsiges om. Dette afspejler blandt

andet en semantisk opbygning af sætningen i et emne (subjektet) som der siges noget om i verbalhelheden, den grammatikaliserede kommentar. Subjektet som grammatikaliseret emne kaldes TEMA, verbalhelheden som grammatikaliseret kommentar kaldes REMA.

(Hansen & Heltoft 2011: 73)

Dette står således i modsætning ikke alene til vores generative analyse, men også til fx den norske og den svenske referencegrammatik, som begge godt kan have en verbalhelhed/VP inden i en anden verbalhelhed/VP inden for en og samme sætning:

- (14) Verbet utgjør kjernen i **verbfrasen**. Verbfrasen kan bestå av det tempusbøyde verbet alene (*gutten spiller, jenta spilte*), eller av det tempusbøyde verbet pluss utfyllinger av forskjellige slag (*gutten spiller fotball, jenta spilte vakkert*).

Det tempusbøyde verbet kan også ta en ny verbfrase som utfylling (*gutten har spilt fotball, jenta hadde spilt ferdig*) og vi sier da at de to verbformene (*har + spilt, hadde + spilt*) danner et sammensatt verbal.

(Faarlund, Lie & Vannebo 1997: 465)

- (15) **Verbfrasen** har ett verb som huvudord. Den kan utgöras av ett huvudord ensamt, men den kan också innehålla olika bestämmningar: objekt (av olika slag), egentliga subjekt, predikativ (av olika slag) och adverbial (av olika slag). [...]

Om huvudordet är en infinit verbform fungerar verbfrasen vanligen nominalt. Med en supinumform som huvudord fungerar den som objekt till *ha*. [...]

Jag hade *tecknat av två kvinnor som stod modell*. (S)

(Teleman, Hellberg & Andersson 1999: 253–254)

Vores evidens for den samme antagelse som i (14)/(15), altså at man kan have en VP inden i en anden VP, er at man på dansk kan sige fx både (5g), som understøtter antagelsen af VP'en *fejre min fødselsdag i fuld offentlighed*, og (5i), som understøtter antagelsen af VP'en *kunne fejre min fødselsdag i fuld offentlighed*.

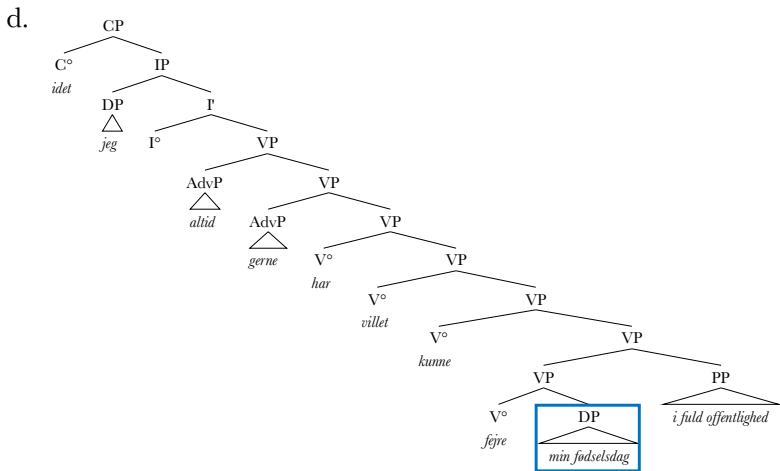
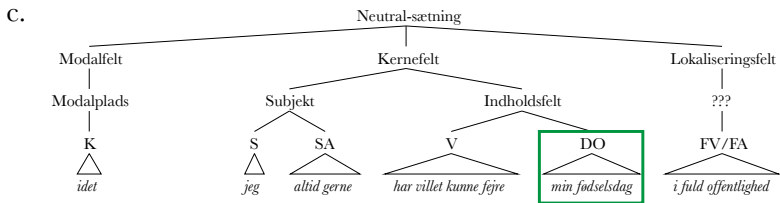
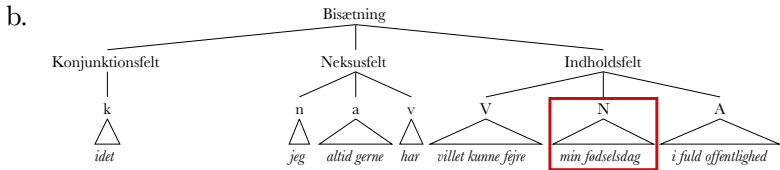
I det følgende afsnit, afsnit 3, vil vi prøve at analysere og illustrere hele denne problematik i detaljer ved hjælp af data i (5) illustreret i strukturanalyser (tegnet i *SVG Syntax Tree Generator*).

3. Detaljerede analyser af pseudoklønningerne i (5)

3.1 Den pseudokløvede konstituent indeholder ingen verber

I (16) nedenfor viser vi tre analyser af den samme pseudoklønning, (16a), af basissætningen i (4a).

(16) a. Det jeg altid gerne har villet kunne fejre ___ i fuld offentlighed, er **min fødselsdag**.



(16b) er en funktionel analyse efter Diderichsen (1946), (16c) er et forsøg på en funktionel analyse efter Hansen & Heltoft (2011), og (16d) er en ge-

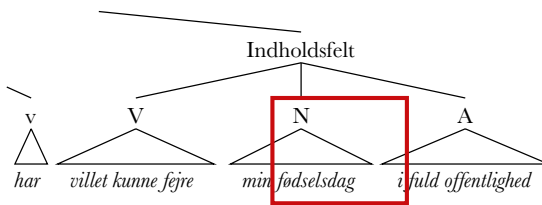
nerativ analyse efter Bjerre & al. (2008), Vikner (2015), Vikner & Jørgensen (2017), m.fl. (16b,c,d) er med andre ord tre forskellige analyser af den basissætning (4a) som vi netop prøver at udforske nærmere ved hjælp af pseudokløvningen i (5a) = (16a). De tre farvede kasser illustrerer hvilken del af basissætningen der har undergået pseudokløvning.

Min fødselsdag udgør **N** i (16b), **DO** i (16c) og objekts-DP'en i (16d). Således er der ingen af de tre analyser der ville forvente nogen problemer i (16a), som jo også er fuldt acceptabel.

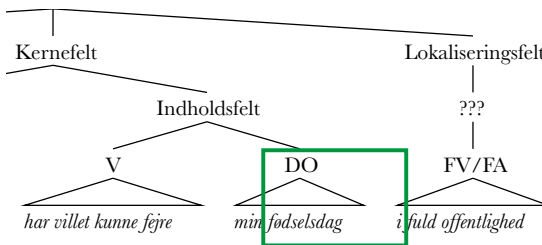
Lad os nu gå videre til noget som alle analyser ligeledes er enige om, nemlig at det ikke er muligt at pseudokløve kun en del af objektet sammen med kun en del af den følgende PP⁶.

- (17) a. * Det jeg altid gerne har villet kunne fejre min ___ fuld offentlighed, er **fødselsdag i**.

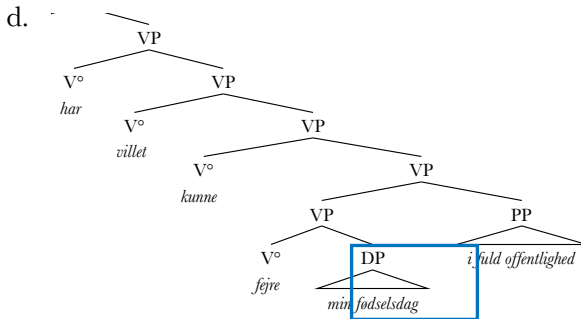
b.



c.



6 Bemærk at hvor (16) indeholder tre strukturanalyser for hele sætningen, så indeholder (17) og også (19), (20), (22) nedenfor af pladshensyn kun et udsnit af de samme strukturanalyser (lige som (21) nedenfor kun indeholder et udsnit af strukturanalyserne for den sætning som analyseres i sin helhed i (18)).



Fødselsdag i udgør ikke en konstituent men derimod dele af to forskellige konstituenters i alle tre analyser (17b,c,d), og derfor forventer alle tre analyser at der vil være problemer i (17a) = (5b), som jo også er helt uacceptabel. Bemærk derudover at en eventuel konstituent [*fødselsdag i*] også ville være i konflikt med analysen i (16) ovenfor, idet *fødselsdag* ikke samtidig kan være en del af konstituenten [*min fødselsdag*] og en del af konstituenten [*fødselsdag i*].

I næste afsnit går vi videre til at se på pseudoklønninger der også omfatter hovedverbet.

3.2 Den pseudokløvede konstituent indeholder et verbum: hovedverbet

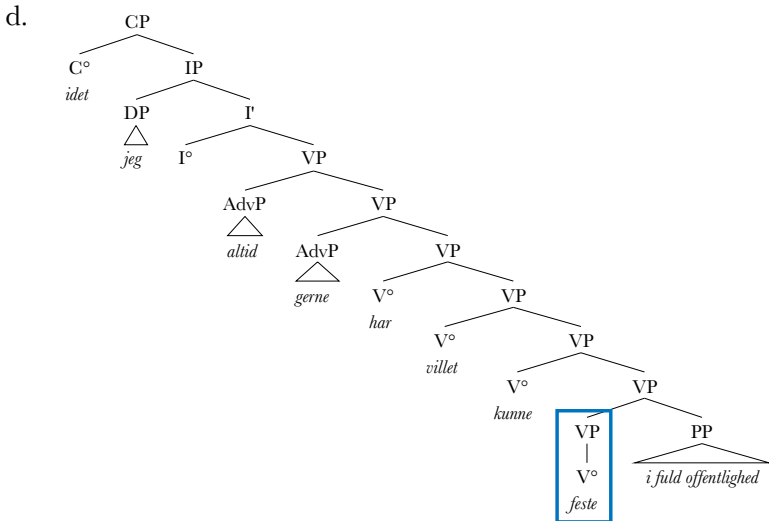
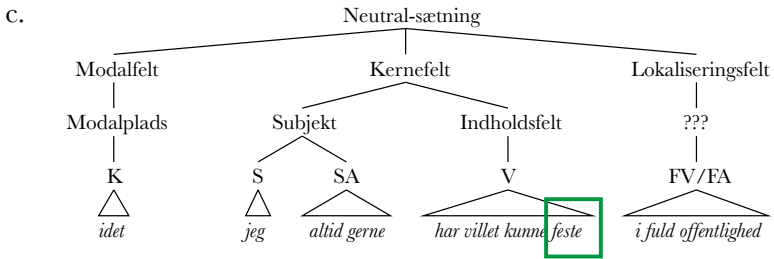
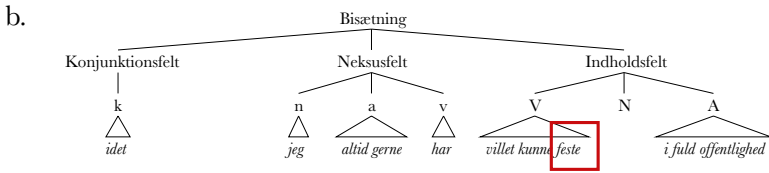
I (18) nedenfor har vi igen analyseret den samme basissætning, (4a), med de tre forskellige analysetilgange, ligesom i (16) ovenfor, denne gang dog med den ændring at det der pseudokløves, indeholder et verbum.

I (18b,c) udgør *at feste* ikke nogen selvstændig konstituent: I (18b) udgør *at feste* kun en del af **V** (som også omfatter *villet kunne*). I (18c) udgør *at feste* også kun en del af **V** (som også omfatter *har villet kunne*).

I (18d) udgør *at feste* derimod både en VP og et V° (hvilket betyder at der i (18d) er en VP som ikke omfatter *i fuld offentlighed*). VP er vigtig her fordi pseudoklønning (ligesom de fleste andre konstituentstests) kun er mulig for fraser/XP'er (såsom VP), og ikke for kerner/X° (såsom V°).

Således er det kun (18d) der ikke forventer nogen problemer i den acceptable (18a) = (5c).

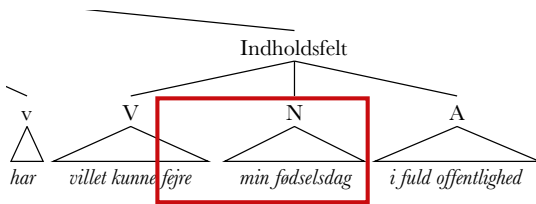
(18) a. Det jeg altid gerne har villet kunne ___ i fuld offentlighed, er **at feste**.



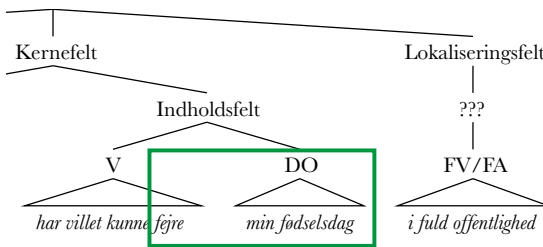
Lad os nu se hvad der sker hvis pseudokløvningen anvendes på et transitivt hovedverbum og dets objekt. Som det fremgår, er det kun i den generative analyse, (19d), at verbet danner en konstituent sammen med dets objekt.

(19) a. Det jeg altid gerne har villet kunne ___ i fuld offentlighed, er **at fejre min fødselsdag**.

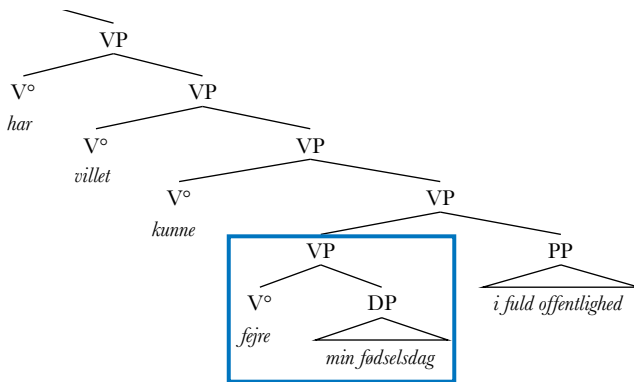
b.



c.



d.



I (19b,c) udgør *at fejre min fødselsdag* ikke nogen selvstændig konstituent: I (19b) udgør *at fejre min fødselsdag* en del af **V** (som også omfatter *villet kunne*) **samt** **N**. I (19c) udgør *at fejre min fødselsdag* ligeledes en del af **V** (som også omfatter *har villet kunne*) **samt** **DO**.

I (19d) udgør *at fejre min fødselsdag* derimod en VP (hvilket betyder at der i (19d) altså er en VP som ikke omfatter *i fuld offentlighed*). VP er vigtig her fordi pseudoklønning kun er mulig for fraser/XP'er, og ikke for kerner/X°.

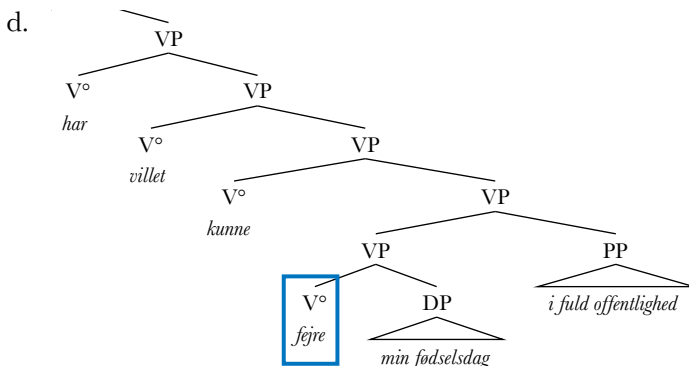
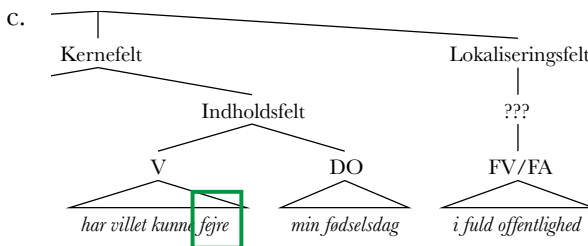
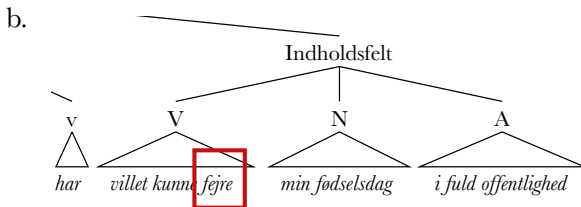
Således er det kun (19d) der ikke forventer nogen problemer i den acceptable (19a) = (5d).

Vi vil nu gå videre til en pseudoklønning der indeholder et transitivt hovedverbum men ikke objektet, som i den uacceptable sætning i (20a) = (5e).

I (20b,c) udgør *at fejre* ikke nogen selvstændig konstituent: I (20b) udgør *at fejre* en del af **V** (som også omfatter *villet kunne*). I (20c) udgør *at fejre* ligeledes en del af **V** (som også omfatter *har villet kunne*).

I (20d) udgør *at fejre* kun et **V°** og ikke nogen VP, og det er derfor sætningen er umulig idet pseudoklønning som sagt kun er mulig for fraser/XP'er, og ikke for kerner/X°.

- (20) a. * Det jeg altid gerne har villet kunne ___ min fødselsdag i fuld offentlighed, er **at fejre**.

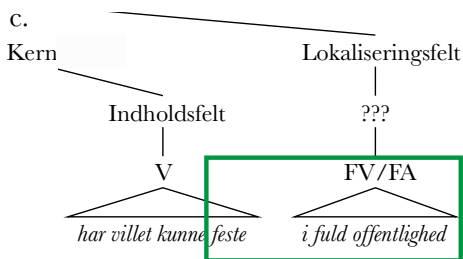
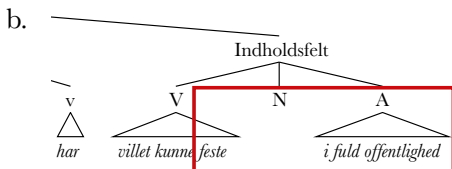


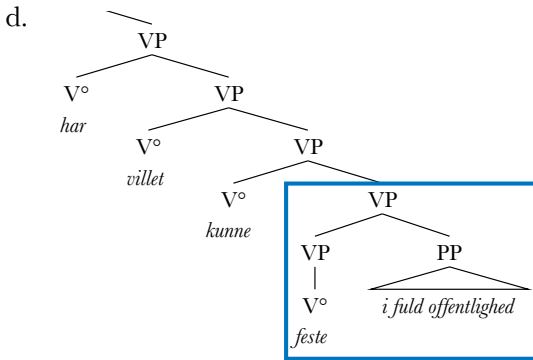
Således forventer alle tre analyser problemer i den unacceptable (20a), selv om det kun er (20d) der giver en analyse af den unacceptable (20a) som er forskellig fra den tilsvarende analyse af den acceptable (18a) ovenfor. I (18d) udgør *feste* en VP i sig selv, mens dette ikke er tilfældet for *fejre* i (20d).

Men hvad nu hvis pseudokløvningen indeholder et intransitivt hovedverbum og et adverbial, som i (21a)? I (21b,c) udgør *at feste i fuld offentlighed* ikke nogen selvstændig konstituent: I (21b) udgør *at feste i fuld offentlighed* en del af **V** (som også omfatter *villet kunne*) **samt A**. I (21c) udgør *at feste i fuld offentlighed* ligeledes en del af **V** (som også omfatter *har villet kunne*) **samt Lokaliseringsfeltet**.

I (21d) udgør *at feste i fuld offentlighed* derimod en konstituent, nemlig en VP (hvilket betyder at der i (21d) altså både er mindst en VP som omfatter *i fuld offentlighed* og en VP som ikke gør det). VP er vigtig her fordi pseudokløvning kun er mulig for fraser/XP'er, og ikke for kerner/X°. Således er det kun (21d) der ikke forventer nogen problemer i den acceptable (21a) = (5f).

- (21) a. Det jeg altid gerne har villet kunne ____, er **at feste i fuld offentlighed**.

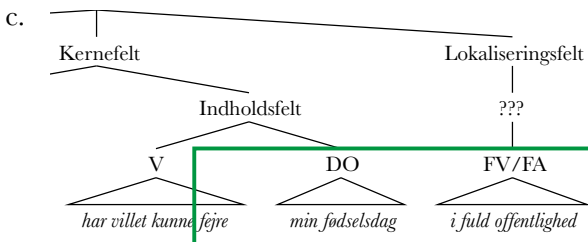
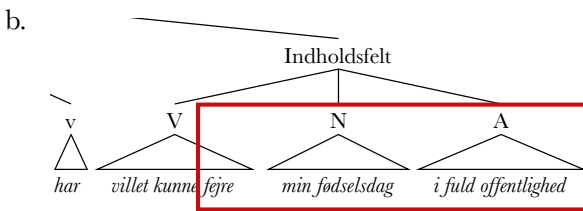


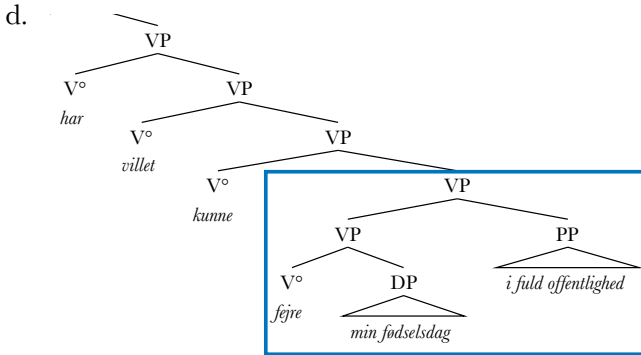


Det sidste eksempel vi vil se på, indeholder en pseudoklønning af et transitivt hovedverbum og dets objekt, ligesom i (19a), men nu også et adverbial.

I (22b,c) udgør *at fejre min fødselsdag i fuld offentlighed* ikke nogen selvstændig konstituent: I (22b) udgør *at fejre min fødselsdag i fuld offentlighed* en del af **V** (som også omfatter *villet kunne*) **plus N samt A**. Bemærk at *fejre min fødselsdag i fuld offentlighed* ikke udgør et **Indholdsfelt** i (22b), det ville nemlig være *villet kunne fejre min fødselsdag i fuld offentlighed*. I (22c) udgør *at fejre min fødselsdag i fuld offentlighed* ligeledes en del af **V** (som også omfatter *har villet kunne*) **plus DO samt Lokaliseringsfeltet**.

- (22) a. Det jeg altid gerne har villet kunne ____, er **at fejre min fødselsdag i fuld offentlighed**.





I (22d) udgør *at fejre min fødselsdag i fuld offentlighed* derimod en VP (hvilket betyder at der i (22d) altså både er mindst en VP som omfatter *i fuld offentlighed* og en VP som ikke gør det). Således er det kun (22d) der ikke forventer nogen problemer i den acceptable pseudoklønning i (22a) = (5g).

Nu har vi diskuteret eksemplerne i (5a-g) i detaljer⁷, og på basis af denne diskussion går vi nu til til vores konklusion.

4. Konklusion

Denne artikel begyndte med at vise hvordan man kan bruge open source-programmet *SVG Syntax Tree Generator* til at generere træstrukturer som i (1b) og i (16)-(22). Vi viste derefter hvordan sådanne træstrukturer kan hjælpe med anskueliggørelsen når man vil argumentere for eksistensen af en konstituent ved hjælp af pseudoklønning, som kun er en af mange konstituentests der anerkendes på tværs af de forskellige tilgange, jf. Diderichsen (1746: 163), Hansen & Heltoft (2011: 95–101), Vikner & Jørgensen (2017: 149–153) og mange andre.

Vi har prøvet at vise at det ikke blot er muligt men gavnligt at tage evidens fra konstituentests alvorligt, evidens af den type som fremgik både af (11)-(12) citeret fra Hansen & Heltoft (2011) og af vores egne data i (5) og (18)-(22). Selvom evidens af denne type ikke genspejles af analyserne i Diderichsen (1946), som fx (16b), eller af analyserne i Hansen & Heltoft (2011), som fx (16c), så findes der syntaktiske analyser som også inddrager denne type evidens, og her vil vi selvfølgelig fremhæve den generative analyse som vi selv anvender, som fx (16d).

⁷ Hvis man gerne vil have den fulde historie, dvs. inklusive analysen af eks. (5h-k), hvor der er to verber i den streng af ord der bliver pseudokløvet, henviser vi til diskussionen i et appendiks i det separate dokument <https://tildeweb.au.dk/au132769/papers/vikn24b.pdf>.

Endnu et eksempel på hvordan vores analyse direkte spiller sammen med denne type evidens, kan ses ved at betragte passagen i (12) nærmere (der som sagt er citeret fra Hansen & Heltoft 2011: 1766), gentaget i uddrag her som (23):

- (23) Der sker altså en semantisk henvisning til verbalhelhedens indhold (til dens remadel [...]), og rettere: til det forhold som denne betegner. Den er obligatorisk når det gælder valensled:

*Sara maledede døren. *Det gjorde jeg vinduet.*

[...] Derimod kan frie led udskiftes i anaforsætningen:

Ida slår græs i denne uge. Det gjorde jeg i sidste uge.

I vores generative analyse er et objekt som *døren* i *male døren* en obligatorisk del af den VP der er bygget op omkring hovedverbet *male* i *Sara maledede døren* (jf. analysen af *fejre min fødselsdag* i (19d) ovenfor), hvorimod der findes en VP bygget op omkring hovedverbet *slå* i *Ida slår græs i denne uge* som **ikke** omfatter *i denne uge* (jf. analysen af *fejre min fødselsdag i fuld offentlighed* i (19d) og (22d) ovenfor). Derfor er det umuligt for *gøre det* (som selv er en VP) at referere til *male* uden at inkludere *vinduet*, hvorimod det er helt acceptabelt for *gøre det* at referere til *slå græs* uden at inkludere *i denne uge*.

I denne artikel har vi i diskussionen af forskellige tilgange til sætningers interne struktur gjort opmærksom på aspekter af strukturanalyserne i Diderichsen (1946) og i Hansen & Heltoft (2011) som vi finder problematiske, nemlig at i hvert fald nogle af de enheder som deres sætningsskemaer er bygget op af (fx feltet **V**), er i konflikt med de enheder (fx VP'er/verbalhelheder) som kan påvises ved hjælp af konstituentests, og som er nødvendige for at redegøre for fx mulighederne for pseudoklønning, som argumenteret i afsnit 3. Som også vist i afsnit 3, kan disse problemer undgås i den generative analyse⁸, og mange (men dog ikke alle) af de konstituenters som de generative strukturanalyser består af, kan påvises ved hjælp af konstituentests.

I hele den forudgående diskussion har vi således gjort os til fortalere for følgende:

8 Hertil kommer at denne generative analyse kan redegøre for tvetydige sætninger som *tjenestepigen blev beskyldt for tyveri af sit herskab* (Vikner 2019), og den kan overføres til andre sprog, fx engelsk (Vikner 2016).

- (24) a. Hvis man rent faktisk **kan** flytte eller substituere en streng af et eller flere ord, så **bør** denne streng af ord udgøre en konstituent i sætningsstrukturanalysen. Det er det vi mener når vi taler om at tage positiv evidens fra konstituentests alvorligt.
- b. Derudover kan man i sætningsstrukturanalysen blive nødt til at antage eksistensen af en konstituent som ikke kan flyttes eller substitueres, selvom antagelsen af en sådan konstituent naturligvis så er mindre overbevisende end hvis man havde kunnet flytte eller substituere den.
- c. Desuden må det selvfølgelig kræves at eksistensen af en sådan konstituent ikke er i konflikt med eksistensen af de konstituentter som rent faktisk **kan** flyttes eller substitueres, jf. diskussionen af (17) ovenfor.

Vi ser det hele som en del af en fortsat debat om hvad analyse af sætningsstruktur er eller bør være, herunder

- (25) a. Har vi det optimale antal af hierarkiske niveauer?
 b. Giver analysen os de konstituentter vi har brug for?
 c. Har vi brug for de konstituentter analysen giver os?
 d. Hvordan kan vi undersøge det? Hvad tæller som evidens?

Appendiks: Input til træerne i (16)-(22)

(16b), (17b), (19b), (20b), (22b): [Bisætning [Konjunktionsfelt [^k idet]]
 [Neksusfelt [^n jeg][^a altid gerne][^v har]]][Indholdsfelt [^V villet kunne fejre]
 [^N min fødselsdag] [^A i fuld offentlighed]]]

(16c), (17c), (19c), (20c), (22c): [Neutral-sætning [Modalfelt [Modalplads
 [^K idet]]]][Kernefelt [Subjekt [^S jeg][^SA altid gerne]]][Indholdsfelt
 [^V har villet kunne fejre] [^DO min fødselsdag]]][Lokaliseringsfelt [??
 [^FV/FA i fuld offentlighed]]]]]

(16d), (17d), (19d), (20d), (22d): [CP [C° [idet]] [IP [^DP jeg] [I' [I° []
 [VP [^AdvP altid] [VP [^AdvP gerne] [VP [V° har] [VP [V° villet] [VP
 [V° kunne] [VP [VP [V° fejre] [^DP min fødselsdag]] [^PP i fuld offent-
 lighed]]]]]]]]]]

(18b), (21b): [Bisætning [Konjunktionsfelt [^k idet]][Neksusfelt [^n jeg] [^a altid gerne][^v har]][Indholdsfelt [^V villet kunne feste] [N []] [^A i fuld offentlighed]]]

(18c), (21c): [Neutral-sætning [Modalfelt [Modalplads [^K idet]]][Kernefelt [Subjekt [^S jeg][^SA altid gerne]][Indholdsfelt [^V har villet kunne feste]]][Lokaliseringsfelt [??? [^FV/FA i fuld offentlighed]]]]]

(18d), (21d): [CP [C° [idet]] [IP [^DP jeg] [I° [I° []] [VP [^AdvP altid] [VP [^AdvP gerne] [VP [V° har] [VP [V° villet] [VP [V° kunne] [VP [V° feste]] [^PP i fuld offentlighed]]]]]]]]]]

Om forfatterne

Sten Vikner, professor i engelsk lingvistik, dr.phil.habil., Inst. f. Kommunikation og Kultur, Aarhus Universitet.

Ken Ramshøj Christensen, lektor i engelsk lingvistik, ph.d., Inst. f. Kommunikation og Kultur, Aarhus Universitet.

Litteratur

Becker-Christensen, Christian & Jan Heegård Petersen, (red.) (2010): *Politikens nudansk ordbog med etymologi*, 4. udg. København: Politikens Forlag, <https://www.ordbogen.com/da/dictionaries/>.

Bjerre, Tavs, Eva Engels, Henrik Jørgensen & Sten Vikner (2008): Points of convergence between functional and formal approaches to syntactic analysis. *Working papers in Scandinavian syntax* 82, 131–166, <https://tildeweb.au.dk/au132769/papers/bjer08a.pdf>.

Brandt, Per Aage (1974): *Tegn, sætning, subjekt*. Grenå: GMT.

Christensen, Ken Ramshøj (2005): Interfaces: Negation - Syntax - Brain. Ph.d.-afhandling, Aarhus Universitet, <https://tildeweb.au.dk/au572/Papers/krc-phd.pdf>.

Diderichsen, Paul (1946): *Elementær dansk Grammatik*. København: Gyldendal.

Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo (1997): *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.

Hansen, Erik & Lars Heltoft (2011): *Grammatik over det danske sprog*. København: Det Danske Sprog- og Litteraturselskab.

- Heltoft, Lars (2016): Topologisk og kombinatorisk syntaks som alvorligt ment skelnen. *Ny Forskning i Grammatik* 23, 71–91. <https://doi.org/10.7146/nfg.v24i23.24645>.
- Nyvad, Anne Mette (2014): Neurolingvistiske perspektiver på sammenslyngede sætninger. I: *Problemer og perspektiver i dansk syntaks – med Kristian Mikkelsen som anledning*, Ole Togeby, Sten Vikner & Henrik Jørgensen (red.). Odense: Syddansk Universitetsforlag, 116–136. <http://oletogeb.dk/filer/Mikkelsen-Togeby-Vikner-Joergensen-2014.pdf>.
- Teleman, Ulf, Staffan Hellberg & Erik Andersson (1999): *Svenska akademis grammatik - 3. Fraser*, Stockholm: Norstedts.
- Vikner, Sten (1999): Ledstillingen i Dansk og Government & Binding. I: *Sætningsskemaet i generativ lingvistik*. Per Anker Jensen & Peter Skadhauge (red.). Kolding: Syddansk Universitet, 83–110, <https://tildeweb.au.dk/au132769/papers/vikn99a.pdf>.
- Vikner, Sten (2015): Dansk sætningsstruktur – et generativt vs. et Diderichsen-perspektiv. *MUDS - Moderne om udforskningen af dansk sprog* 15, 429–450, <https://tildeweb.au.dk/au132769/papers/vikn15a.pdf>.
- Vikner, Sten (2016): English VPs and why they contain more than just verbs. I: *Let us Have Articles betwixt us – Papers in Historical and Comparative Linguistics in Honour of Johanna L. Wood*. Sten Vikner, Henrik Jørgensen & Elly van Gelderen (red.), Aarhus: Aarhus Universitet, 439–64, <https://doi.org/10.7146/aul.119.107>.
- Vikner, Sten (2019): Sætningsstruktur og strukturelle flertydigheder. *Ny Forskning i Grammatik* 26, 176–190, <https://doi.org/10.7146/nfg.v0i26.116001>.
- Vikner, Sten & Henrik Jørgensen (2017): En formel vs. en funktionel tilgang til dansk sætningsstruktur, *NyS – Nydanske sprogstudier* 52–53, 135–168, <https://doi.org/10.7146/nys.v1i52-53.24954>.

Det middeldanske baggrundsfelt: Subjekts- og letledspladser

Lennart Westergaard

Abstract

According to Diderichsen (1941), Middle Danish word order corresponds closely to Modern Danish word order, although he acknowledges certain deviations. However, Heltoft (2003) contends that Middle Danish word order had a different structure than Modern Danish word order. He argues that Middle Danish word order iconically expressed a focus background distinction. Constituents which appear between the finite verb and the negation are said to be coded as backgrounded or non-focused constituents. In this paper, I propose a compromise. Based on a corpus study, I argue that within Heltoft's background field, Middle Danish had a designated subject position. This position is located immediately after the finite verb and succeeded by light or background pronouns.

Nøgleord

subjekt, topologi, middeldansk, lette pronomener

1. Introduktion¹

I sin disputats *Sætningsbygningen i Skaanske Lov* argumenterer Diderichsen (1941: 34-37) for, at man i middeldansk grundlæggende har samme ledstilling som i moderne dansk. Sætningen kan inddeles i tre felter, for-, central- og slutfelt, med respektive funktionsbetegnelser, fundament, aktual- og indholdsfelt. Inden for aktual- og indholdsfeltet har sætningsledene en fast rækkefølge, således at verbaler (v) kommer før nominaler (n), som igen står før adverbialer (a) i de to felter (i i deklarativ ledstilling, jf. Hansen & Heltoft 2011: 1569):

¹ Jeg skylder Torben Juul Jensen og en anonym bedømmer tak for værdifulde kommentarer.

- | | | | | | |
|-----|-----------|---|-------------|---|--------------|
| (1) | forfelt | – | centralfelt | – | slutfelt |
| | fundament | – | aktualfelt | – | indholdsfelt |
| | F | – | v n a | – | V N A |

Ifølge Diderichsen er der blot flere undtagelser til strukturen i middeldansk, end der er i moderne dansk. Det drejer sig om undtagelser mht. vægtprincippet (1. og 2.) og særlige forbindelser svarende til forskellige former for inkorporering (3., 4. og 5.):

1. Tryksvage pronomener kan placeres før subjekter ”af tungere Karakter”:
 (2) uiti þæt faþur barns. (som citeret i Diderichsen 1943: 36)
 ’bevise det (obj) faderens barn (subj)’
 ’da skal barnets fader bevise det’²
2. Tunge subjekter kan slutstilles:
 (3) bõtæ ok þre mark hin, ær þem a (som citeret i Diderichsen 1943: 174)
 ’bøde også tre mark (obj) den, som dem ejer (subj)’
 ’da skal også ham, der ejer dem, bøde tre mark’
3. I de tilfælde, hvor forbindelsen mellem verbalet og et sætningsled svarer til en enhedstryksforbindelse i moderne dansk, kan sætningsleddet stå sammen med verbalet og dermed bl.a. foran subjektet:
 (4) liggi (...) wiþær hans friþ (som citeret i Diderichsen 1943: 37)
 ’ligge ved hans fred (subj)’
 ’da skal han tabe sin fred’
4. I de tilfælde hvor et personligt objekt kan betragtes som et indholdsmæssigt eller logisk subjekt, kan det stå før det egentlige subjekt:

2 Oversættelserne af eksemplerne fra de gamle danske love er enten inspireret af eller kopieret i sin helhed fra Kroman & Juul (1945) oversættelse. Jeg har tilpasset dem til moderne retskrivning.

- (5) Gar børnum ysilðæ ofna hændær. (som citeret i Diderichsen 1943: 37)
 'Går børn (obj) usselt (subj) på hænder'
 'Kommer børnene ud for trang'

5. I passivkonstruktioner, hvor et upersonligt subjekt udtrykker ét begreb sammen med verbet, kan subjektet stå sidst i sætningen:

- (6) at hanum war stæfnd þing (som citeret i Diderichsen 1943: 37; 175)
 'at ham (obj) var stævnet ting (subj)'
 'at der var stævnet ham ting'

På trods af disse undtagelser regner Diderichsen med, at sætningsledde-
 ne allerede i middeldansk havde en fast rækkefølge svarende til moderne
 dansk.

Dette har Heltoft opponeret imod (1995: 139-161; baseret på Peder-
 sen 1993: 200-204; sml. Heltoft 2003: 43; 2019: 161-164). Heltoft (1995:
 143-145) påpeger, at et eksempel som (7) ikke kan forklares med nogen af
 Diderichsens undtagelser. Subjektet (*arfui hans* 'hans arvinger') står her i
 slutfeltet (efter det infinitte verbal *dylia* 'nægte'). Da pronomenet står i for-
 feltet, kan subjektet ikke være fortrængt af pronomenet (vs. 1). Subjektet
 er ikke tungt (vs. 2). Der er ikke tale om en form for inkorporering (vs. 3).
 Der er ikke et objekt, der fungerer som indholdssubjekt (vs. 4). Endelig er
 der ikke nogen passiv-konstruktion, som kunne forklare subjektets place-
 ring (vs. 5):

- (7) Æn vm thæt vil dylia arfui hans. at hanum døthum.
 'men hvis det vil nægte arving hans (subj) ved ham død'
 'men om en af hans arvinger vil bestride det efter hans død'
 (Skånske Kirkelov ca. 1350, som citeret i Heltoft 1995: 143)

Han argumenterer i stedet for, at den middeldanske ledstilling er væsens-
 forskellig fra den moderne danske: Centralfeltspacering er med til at an-
 give en fokus-baggrundsdistinktion, ikke sætningsledsfunktioner. Hvad
 der placeres før negationens plads, er baggrundsinformation, hvad der
 placeres efter denne, er i fokusdomænet. Led, der er i fokus, forstås som
 led, som afsenderen forventer, modtageren kunne være uenig i ('what he
 – the speaker – wishes the hearer to contest, should he disagree or hesitate

in accepting the utterance” Heltoft 2003: 52)³. Baggrund defineres som kodet ikke-fokus (ibid. 54):

	k	forfelt	V_{fin}	centralfelt			V_{inf}	slutfelt
				baggrund	neg.	fokus		
(8)		Thænnæ sten 'denne sten	ma kan		æi ej	eld ild	skathe skade'	
		'denne sten kan (selv) ild ikke skade'						
(9)		han 'han	giorthe gjorde	bondanum bonden	æy ej	mera schatha mere skade	j thy af hoggi i det afhug	
		'han gjorde ikke bonden mere skade i det afhug'						

Figur 1: Middeldansk sætningskema efter og med eksempler fra Heltoft (2019a: 162; 2003: 43)

Som indikeret i oversættelsen antages det, at subjektet *eld* ('ild') er i fokus i (8). I (9) er objektet *bondanum* ('bonden') placeret i baggrundsfeltet, og ifølge Heltofts analyse er det dermed markeret som ikke-fokus.

I modsætning til, hvad Diderichsen i sin analyse kommer frem til, har Middeldansk ifølge denne analyse ikke nogen dedikeret subjektspads (Heltoft 2019: 134; 2001: 193; 2003: 43). Dette mener jeg ikke er helt korrekt, og jeg vil foreslå et kompromis mellem Diderichsens og Heltofts analyse: Selvom middeldansk ledstilling er væsensforskellig fra moderne dansk ledstilling ved primært at udtrykke informationsstrukturelle betydninger, kan man i middeldansk regne med en subjektspads. Kun subjektet kan stå umiddelbart efter det finitte verbal og afgrænses her til højre af en letledsplads, hvor ikke-fokuserede pronomener står.

Jeg vil dermed argumentere for følgende generalisering:

(10) finit verbal > subjekt > letledsplads

Hypotesen om en subjektspads implicerer dermed også en hypotese om en middeldansk letledsplads. En sådan findes ifølge Heltoft (2005: 152) dog heller ikke endnu i middeldansk (men jf. Pedersen 1993: 202-4 og referencerne dér for samme plads i gammelnorsk og fornsvensk). Hel-

3 Sigtet med artiklen er ikke at diskutere Heltofts informationsstrukturelle analyse, og for en mere udførlig diskussion af fokusbegrebet henvises derfor til Heltoft (2003: 51-54) og referencerne der.

toft antager, at pronomener i et eksempel som (11) blot er placeret i baggrundsfeltet og ikke på en mulig middeldansk letledsplads:

- (11) æn wilt thu **mic thæt** æi sighæ (...)
 'men vil du **mig det** ej sige'
 'men vil du ikke sige **mig det** (...)'
 (Christinalengenden 1300, som citeret i Heltoft 2005: 152)

Artiklens argumentation for en middeldansk subjektsplads, er derimod også en argumentation for en middeldansk letledsplads. Den middeldanske letledsplads er dog strukturelt forskelligt fra dens moderne danske efterfølger, da de ikke-fokuserede pronomener også kan placeres her, når sætningen indeholder et infinit verbal (jf. (11) som påpeget i Heltoft 2005: 152). Dette er som bekendt ikke en mulighed for ikke-fokuserede pronomener i moderne dansk:

- (12) *men vil du **mig det** ikke sige, så (...)

Da vi ikke med sikkerhed kan antage, at pronomener, der placeres på letledspladsen faktisk er lette i middeldansk, vil jeg omtale dem som baggrundspromener, dvs. pronomener, der i den konkrete ytring ikke er fokuserede, og som kan placeres i letledspladsen i moderne dansk.

Resten af artiklen er struktureret som følger: I afsnit 2 vil jeg meget kort diskutere to metodisk-teoretiske aspekter ved sætningsskemaanalysen: 1. hvordan man kan rekonstruere et sætningsskema uden at basere analysen på maksimaludfyldte eksempler, og 2. hvad der skal til for, at en plads tæller som en plads. I afsnit 3 præsenterer jeg min korpusanalyse, der afprøver generaliseringen i (10). I afsnit 4 kommer jeg kort ind på nogle diakrone konsekvenser af analysen for subjektets udviklingshistorie. I afsnit 5 sammenfatter jeg de væsentligste pointer.

2. Topologisk analyse

Hvor i sætningen et led placeres, kan være afgørende for leddets og sætningens tolkning. For eksempel: Om en NP placeres før en evt. negation, er afgørende for, om vi tolker den som subjekt eller objekt i moderne dansk:

- (13) Bo spiste ikke **fisken**.
- (14) Bo spiste **fisken** ikke.

Sætningsskemaet er en strukturel beskrivelse af det udtrykssystem, der er med til at kode sådant indhold. Det gengiver, hvor forskellige led kan (og dermed også ikke kan) eller skal placeres. Som diskuteret i afsnit 1 er udtrykssystemet ikke begrænset til kodning af syntaktiske funktioner, men kan også kode informationsstruktur m.m. (Diderichsen 1966[1943]: 54-55).

Idet sætningsskemaet er en strukturbeskrivelse, skal det være maksimalt generelt, dvs. mere komplekse sætninger skal kunne beskrives på samme måde som enklere sætninger og omvendt. En sådan generalitet opnår man typisk ved at basere analysen på maksimaludfyldning, dvs. bygge analysen på så komplekse sætninger som muligt og så abstrahere leddenes indbyrdes rækkefølge herfra (Diderichsen 1966[1943]: 54-55). Dette indebærer dog et problem for historiske analyser. Her kan det være svært at finde (mange) gode eksempler med maksimaludfyldning.

Man kan i stedet basere analysen på en antagelse om transitive relationer: Hvis rækkefølge $A > B$ og rækkefølgen $B > C$ kan etableres, og vi ikke finder eksempler på rækkefølgen $C > A$ (som ikke kan forklares på anden vis, fx placering i topikaliserings- eller fokuseringspladser), kan vi deducere skemaet: $A > B > C$.

Et andet spørgsmål er, hvornår man kan regne med, at en plads er en plads. Som minimum må man her kræve, at en topologisk veldefinerbar plads kun tillader én type led. I det tilfælde vil pladsen have et symbolsk potentiale, dvs. angive, at et led, der placeres her, er af den type, pladsen tillader. Pladsen vil da udtrykke, at leddet fx har en specifik syntaktisk eller informationsstrukturel funktion. Om samme type led også kan optræde på andre pladser eller felter, er sekundært. I moderne dansk kan de fleste led jo også stå i forfeltet og ikke bare på deres pladser senere i sætningen, hvor deres sætningsledsfunktion kodes.

3. Korpusundersøgelsen

I det følgende præsenteres først korpusundersøgelsens materiale. Derefter diskuterer jeg min operationalisering af hypotesen om generalisering (10). Til sidst præsenteres resultaterne.

3.1. *Materialie*

Materialet for korpusundersøgelsen dækker perioden fra ca. 1300 til ca. 1500 og indeholder 278.572 løbende ord. Siden jeg er interesseret i kategoriale forskelle (*langue*: (u)mulighed), snarere end ususmønstre (*parole*: tendenser), har jeg vægtet mængde over balance. Mit korpus består af alle de tekster fra den pågældende periode, der er tilgængelige på tekstnet.dk eller på DUDS (duds.nordisk.ku.dk/tekstresurser/dsst/teksterne/). Rationalet er, at jo flere tekster, jeg inkluderer, desto større er sandsynligheden for, at jeg finder modeksempler til min hypotese. Materialet er derfor et bekvemmelighedssample.

Følgende tekster indgår i materialet. Digitale udgaver af teksterne kan tilgås på enten tekstnet.dk eller på DUDS (duds.nordisk.ku.dk/tekstresurser/dsst/teksterne/). Her kan der derudover findes yderligere tekstkritisk information:

- Eriks lov AM 455,12 (1300) (Tekstnet)
- Skriftmaalsbøn K48 (1300) (Tekstnet)
- Skrifter af Harpestreng nks 66,8, k48, nks70r og k4 (1300-1480) (Tekstnet)
- Flensborg stadsret (1300) (Tekstnet)
- Kongetal AM 28,8 (1300) (Tekstnet)
- Landegrænse mellem Danmark og Sverige AM 28,8 (1300) (Tekstnet)
- Legendefragment Christinalegende Camb. Add. 3827,1 (Tekstnet) (1300) (Tekstnet)
- Legendefragmenter Camb. Add. 3827,1 (1300) (Tekstnet)
- Legendefragmenter k48 (1300) (Tekstnet)
- Lucidarius AM 76, 8^o (1350; håndskrift fra ca. 1460) (Tekstnet)
- Runekrønike AM 28, 8, fragment (1300) (Tekstnet)
- Sjællandske Kirkelov AM 455,12 (1300) (Tekstnet)
- Skånske lov AM 28, 8 (1300) (Tekstnet)
- Valdemars lov AM 455,12 (1300) (Tekstnet)
- Jyske Lov NKS 295,8 (1325) (Tekstnet)
- Skaanske Kirkelov AM 28,8 (1350-1400) (Tekstnet)
- Julemærker NKS 66,8 (1375-1400) (Tekstnet)
- Rydårbogen E don. Var. 3,8 (1400) (Tekstnet)
- Gesta Danorum b77 (1400-1450) (Tekstnet)

Siæla Trøst (1425) (DUDS)
 Peder Laales ordsprog NKS 813x,4 (1450) (Tekstnet)
 Lægebog fra top til taa D600 (1450-1500) (Tekstnet)
 Lægebog GKS 3487,8 (1475-1500) (Tekstnet)
 Uddrag af Karl Magnus Krønike (1480) (DUDS)

Mange af de citerede tekster mangler sidetal i den digitale udgave. Eksemplerne, der citeres i det følgende, kan dog fremfindes ved at bruge søgefunktionen i den digitale udgave.

Teksterne har jeg indlæst i korpusværktøjet *AntConc* (Anthony 2022). Ingen yderligere tekstbehandling har været nødvendig.

3.2. Operationalisering af hypotesen, korpus-søgning og sortering

Pladser i sætningsskemaet er, som navnet antyder, skematiske, dvs. de har ikke selv materiel form. Deres betydning beror i stedet på leddenes indbyrdes rækkefølge. Det gør det umuligt at søge på en plads i en søgemaskine, der udelukkende tillader at søge på grafemer. Man kan kun søge på materiale, der kan stå på pladsen. Jeg har derfor valgt pronomenet *han* som stedfortræder for letledspladsen. Ved at søge på *han* og dets kasusvarianter kan det dermed afklares, hvad der kan stå på spørgsmålstegnets plads i (15):

(15) finit verbal > ? > *han/ham*
 (letledsplads)

De følgende tre mønstre er logisk mulige:

1. ? = udelukkende subjekter
2. ? = subjekter og adverbialer
3. ? = subjekter, adverbialer og objekter

Mønster 1. og 2. ville sandsynliggøre⁴ hypotesen om en subjektspads, da disse to mønstre ville indikere at den eneste NP-type, der kan stå før baggrundspronomenene, var subjektet. Mønster 1. ville derudover betyde,

⁴ Da hypotesen er en hypotese om umuligheden af en struktur, kan den i princippet udelukkende bekræftes ved hjælp af negativ evidens. Da det er umuligt at fremskaffe sådan evidens på baggrund historisk materiale, kan hypotesen kun sandsynliggøres.

at der kun er én plads før baggrundspronomenene. Dermed ville fundet også implicere en generalisering for adverbialernes ledstilling, nemlig at disse lige som objekter ikke kan placeres før baggrundspronomenene. Mønster 3. ville afkræfte hypotesen.

For at foregribe konklusionen ser det ud til, at mønster 1. gælder. En sådan tolkning af materialet kræver dog en yderligere kvalitativ analyse, og jeg vil derfor præsentere og diskutere resultater, der taler for mønstre 1. og 2., individuelt i afsnit 3.3 og 3.4.

For at finde alle forekomster af rammen (15) har jeg fremsøgt alle forekomster af subjekts- og objektsformerne af pronomenet *han* (dvs. *han* og *hanum*⁵) gennem søgestrengen ”han^{*}”⁶. Søgningen gav 9.544 fund. Falske positiver som *hand* (’hånd’), *hans*, det feminine akkusativpronomen *hana* (’hende’) m.m. har jeg frasorteret.

Da det udelukkende er sætninger, hvor pronomenet er placeret i baggrundsfeltet, der er relevante for undersøgelsen, har jeg også frasorteret eksempler, hvor pronomenet entydigt er placeret i forfeltet, dvs. før et finit verbal som *ær* (’er’) i (16), eller entydigt i slutfeltet, dvs. efter et infinit verbal som *teftet* (’smagt’) i (17):

- (16) **han** ær ey rættir cristin man
 ’han er ej ret kristen mand’
 ’han er ikke en god kristen mand’
 (SjT, 1425)

- (17) saa ath aldri fangæ thee full teftet **hanum**
 ‘så at aldrig får de fult smagt **ham**’
 ’så at de aldrig får smagt **det** til fulde’
 (Luc, 1350)

Jeg har dog beholdt sådanne eksempler, hvor det er tvetydigt, om pronomenet er placeret i centralfeltet eller i slutfeltet som i (18), hvor pronomenet står efter det finitte verbal *hauæ* (’have’ i konjunktiv), men der ikke er noget til at afgøre, om pronomenet står i centralfeltet eller slutfeltet:

5 Subjekt- og objektformen af *han* har følgende ortografiske varianter i materialet: *han*, *hann*, *hanom*, *hanum* og *hannum*, og jeg antager, at dette er nogenlunde dækkende mht. den ortografiske variation.

6 Asterisken er en jokerposition for ét eller flere grafemer inden for samme grafemstreng.

- (18) tha hauæ **han** sin eghæn thræl
 'da have **han** sin egen træl'
 'da skal **han** have sin egen træl'

Jeg har vurderet, at dette ikke udgør et problem, da en sådan grovmasket søgning favoriserer modhypotesen, altså mønster 3.

I mange tilfælde kan det være svært at afgøre, om et pronomen er baggrund eller fokus, og dermed, om det er placeret i baggrunds- eller i fokusfeltet. Jeg har derfor ikke skelnet mellem baggrunds- og fokuspronomen og behandlet alle pronomen som potentielt værende baggrundspronomen (men se den kvalitative del af analysen i afsnit 3.4). Dette kunne være et problem, da en objekt-NP i princippet kan optræde før et fokuseret subjektspronomen, uden at det ville modsige generalisering (10), da det er udelukket, at pronomenet er placeret i baggrundsfeltet (altså letledspladsen i (10)), hvis det er fokuseret (baggrundsfeltet udtrykker jo netop 'baggrund' og er derfor uforenelig med fokuserede led). Unøjagtigheden kommer dog igen modhypotesen til gode.

Derimod har jeg frasorteret eksempler, hvor pronomenet ikke er det første pronomen i centralfeltet. Generaliseringen i (10) og den topologiske ramme i (15) er en generalisering, der udelukkende vedrører positionen mellem finit verbal og det første evt. pronomen. Pronomen, der optræder efter andre pronomen og andet materiale, siger derfor ikke noget om, hvad der kan stå i den relevante position i (15). I (19) står *wæl* ('godt') før *hanum* ('ham'), men efter *han*. *Wæl* er dermed ikke relevant for rammen i (15) og dermed heller ikke for generalisering (10):

- (19) tha ma han wæl **hanum** logh fæstæ at thet ær hans hemæ fød.
 'da kan han godt **ham** lov fæste at det er hans hjemmefødt'
 'da kan han godt præstere **ham** edsbevis for, at det er hjemmefødt
 hos ham.'
 (EL, 1300)

Endvidere har jeg frasorteret eksempler, hvor det finitte verbal mangler, fx pga. ellipse:

- (20) oc halft takæ hans børnæ. oc halft **han** mæth theræ stiyp mothær.
 'og halvdelen tager hans børn og halvdelen **han** med deres sted-
 mor'

'Og den ene halvdel tager hans børn, og den anden tager **han** med deres stedmor'
(FS, 1300)

Kun de pronomener er relevante, hvor pronomenet udgør et sætningsled. I (21) er *hanum* ikke et argument til *pløghær* ('pløge'), men indgår i en prædikativ forbindelse med *u witindæ* ('uvidende'), som forklarer, hvorfor pronomenet står sammen med participiet (og altså efter to NP'er):

- (21) Æn pløghær nokær man annærs manz hæfdh **hanum u witindæ**.
tha hauæ han sit ærfthæt for giorth.
'men pløjer nogen mand anden mands besiddelse **ham uvidende**, da have han sit arbejde forspildt'
'Men pløjer en mand en andens mands besiddelse **uden sit vidende**, da har han forspildt sit arbejde.'
(EL, 1300)

Eksempler, hvor pronomenet indgår i en sådan non-finit konstruktionstype, har jeg også frasorteret.

Der er synkretisme mellem nominativ og akkusativ i substantiver og pronomener. I (22) er fx hverken *kunugs umbuzman* eller *han* morfologisk entydigt markeret som subjekt eller (akkusativ-)objekt:

- (22) tha ma **kunugs umbuzman han** ey meræ kræftæ æn hanz houæthlot
'da må **kongens ombudsmand ham** ej mere kræve end hans hovedlod'
'da må **kongens ombudsmand** ikke kræve mere af **ham** end hans hovedlod'
(VL, 1300)

I praksis har det dog ikke været svært at afgøre sætningsledsfunktion på baggrund af indholdet i de allerfleste eksempler. Fx kan vi i et eksempel som (22) på baggrund af den sociokulturelle kontekst gå ud fra, at det er kongens ombudsmand, der kan kræve penge af en bonde, og ikke omvendt. I enkelte eksempler som fx (23) har det ikke været muligt at afgøre den syntaktiske funktion:

- (23) Æn særær **han han** swa at han liggær i siangæ æftær ællær berær,
 tha (...)
 'Men særer **han/ham** således at han ligger i seng efter eller slår, da
 ...'
 'Men særer eller slår **han ham** således, at han ligger i sengen der-
 efter, da ...'
 (JL, 1325)

Også sådanne uklare eksempler har jeg frasorteret.

Efter denne sortering havde jeg 3.148 forekomster tilbage. Disse dan-
 ner grundlaget for analysen i 3.4. Af de 3.148 forekomster er det i 966
 eksempler, at der også optræder en subjekt- eller objekt-NP efter det finitte
 verbal. Disse danner grundlaget for analysen i 3.3.

3.3. Resultater 1: NP'er og pronomener

Jeg har fundet 77 sætninger, hvor en subjekt-NP optræder efter et finit
 verbal og før et pronomen som i (22). Til sammenligning optræder sub-
 jekt-NP'en i 108 sætninger efter pronomenet som i (24). Er NP'en et ob-
 jekt, optræder det efter pronomenet i 780 eksempler som i (25) og kun i
 1 eksempel kunne man overveje, om objekts-NP'en optræder før et pro-
 nomen som i (26). Dette ene eksempel ville i så fald være et modbevis til
 hypotesen om generalisering (10):

- (24) bristær **hanum næfnd** þa bæri han iarn
 'brister **ham (obj) nævn (subj)** da bære han jern'
 'Brister **nævnet** for **ham**, da skal han bære jern, '
 (SL, 1300)
- (25) Læggær man thænnæ steen. a. døth lyk. tha tapær **han sin kraft**.
 'Lægger man denne sten på død krop, da taber **den (subj) sin
 kraft (obj)**.'
 'Lægger man denne sten på et lig, da taber **den sin kraft**.'
 (HS, 1300)
- (26) [kontekst: Men er der ingen ægtebørn, da tager slegfredbarnet, der
 blev tinglyst, arven efter faderen.]
 æn fangær **nokæt han af fathærns arf**. tha æruær thatæt æftær
 syskæn iauænt with athælkunæ børn.'

‘men får **noget han af faderens arv**, da arver det efter søskene lige med adel-kone (dvs. en ægteviet hustru) børn.’

‘Men får **det nogen arv efter faderen**, da arver det efter søskende lige med ægtebørn’

(JL, 1325)

Noget i (26) er en del af det diskontinuerte objekt *nokæt ... af fathærns arf* (‘noget af faderens arv’). Da det står før subjektspronomenet *han*, kunne det dermed ses som modeksempel til hypotesen. *Noget* er dog ikke nogen prototypisk NP og derudover står det dissociert fra sit attribut (*af fathærns arf*). Endvidere er en (kontrast-)fokus-læsning af *han* oplagt: Paragraffen kontrasterer slegfredbørns rettigheder med ægtebørnenes. Det er derfor ikke et overbevisende modeksempel, og da det er det eneste mulige modeksempel, har jeg valgt at se bort fra det.

Tabel 1 sammenfatter resultaterne:

	NP>pron	pron>NP
SubjNP	77	108
ObjNP	0 (1)	780 ⁷

Tabel 1: Fordeling af position af NP-type relativt til pronomener

Jeg anser det dermed som sandsynliggjort, at middeldansk havde en subjektspads, i og med subjekt-NP'er er den eneste NP-type, der kan stå mellem finit verbal og baggrundspronomen.

3.4. Resultater 2: Ikke-nominelt materiale før pronomener

I denne del præsenteres resultaterne for den ikke-nominelle del af analysen. I det hele taget er det meget sjældent (kun i 25 ud af 3.148 sætninger, hvor pronomener kan tolkes som værende placeret på letledspladsen), at der optræder andet end subjekter mellem det finite verbal og det første pronomener. Ved en gennemlæsning af de 3.148 sætning har jeg kun kunnet finde de følgende adverbialer:

7 Forskellen er signifikant (χ^2 -test: $p < 0.001$). P-værdien indikerer dog ikke andet, end at forskellen mellem subjekters og objekters position med al sandsynlighed ikke skyldes tilfældighed. Den statistiske test kan ikke endeligt afgøre, at det ikke i princippet er muligt, at objekter kan placeres for baggrundspronomen, men at det er yderst sjældent.

<i>ej</i> :	13x
<i>oc</i> ('og, også'):	7x
<i>atær</i> ('atter', 'igen'):	1x
<i>siden</i> :	1x
<i>e</i> ('altid'):	1x
<i>a hand</i> ('(sværge) på hånden'):	1x
<i>mæth næfnd</i> ('med nævning'):	1x

Adverbierne *ej*, *oc* ('og, også') og muligvis *atær* ('atter', 'igen') er fokusoperatore, dvs. gennem deres semantik er de med til at udpege fokus. Når pronomenet er placeret efter disse adverbier, kan vi derfor med en vis sikkerhed antage, at det er i fokus og dermed ikke placeret i baggrundsfeltet.

Pronomenerne i de resterende 4 eksempler kan betragtes som fokuserede eller tunge pronominalfraser. I (27) er pronomenet *hanum* udvidet med en relativsætning (*som steen hawer*) og dermed entydigt tungt og altså ikke placeret i letledspladsen i baggrundsfeltet:

- (27) ok gywe siden **hanum** som steen hawer then puluer
 'og give siden **ham** som sten har den pulver'
 'og giv så **ham**, der har galdesten, det pulver.'
 (LB 1450-1500, s. 35)

I (28) er pronomenet *hanum* nok fokuseret. Denne tolkning af eksemplet understøttes af Kromann & Iuuls oversættelse, hvor fokuset ekspliciteres med fokusoperatoren *både og*:

- (28) Hor sac ma ængiæn man giuæ manz kunæ num bonden siælf. Æn
 wil han sæctæ hennæ.
 tha weriæ sic mæth næfnd. **hanum** i gen oc biscop.
 'da værge sig med Nævn **ham** overfor og biskop'
 'Horsag må ingen mand rejse mod en mands kone uden bonden
 selv. Men vil han rejse sag mod hende, da skal hun værge sig med
 nævn, både mod **ham** og mod biskoppen.'
 (SjKL 1300)

Tolker man *i gen* ('igen, overfor') som postposition, kunne *hanum igen* også dreje sig om en PP. I så fald ville eksemplet ikke være et modeksekempel af syntaktiske grunde.

Også i (29) er det oplagt at tolke *hanum* som fokusbærer, da *hanum* kontrasterer med *den*, der i førsteomgang blev anklaget for mordet:

- (29) warthær man dræpin oc ængin man gar with. tha ma hin ær søkia wil nj mannum ofna hændær søkia mæth iærne. oc sithæn late han .j. ban sætia wil han sithæn nocro manne. a hand søkia tha swæri .a. hand **hanum** mæth tolf lagfastum mannum.
 'da sværge han på hånd **ham** med tolv lovfaste mænd'
 'Bliver en mand dræbt, og ingen tilstår, da skal sagsøgeren rejse jernbyrdssag mod ni mænd, og derefter skal han lade (drabsmanden sætte) i ban. Vil han derefter rejse sag mod nogen, da skal han sværge imod **ham** med tolv lovfaste mænd'
 (SkKL 1350)

(30) er det sidste eksempel, hvor jeg finder et adverbial, der ikke er en fokusoperator, før det første baggrundspronomen. Paragraffen har to aktører, en træll og trællens ejer. En fokustolkning er for så vidt en mulighed, da handlinger og pligtigheder af de forskellige aktører sammenlignes:

- (30) Hwat sum bondæn biuthær sin thræl at gøær ællær rathær, anti bæria ællær hoggæ ær annæt nokæt at gøær, hærwirkæ ællær bindæ nokær man ællær thylict, tha bötæ e **han** æmfullêlich sum han hafthæ thet syælf giorth
 'da bøde altid **han** lige.så.fuldt som han havde det selv gjort'
 'Det, som en bonde byder eller tilråder sin træll at gøre, hvad enten det er at slå eller hugge eller noget andet, at øve hærværk eller binde en mand eller deslige, det skal **han** altid bøde ligeså fuldt for, som hvis han selv havde gjort det.'
 (VL 1300)

Om end jeg mener, fokuslæsningen er en plausibel tolkning af eksemplet, medgiver jeg gerne, at en sådan fokuslæsning ikke er helt så entydigt som ved de øvrige eksempler.

Det tyder dermed på, at det ikke bare forholder sig på den måde, at subjekter er den eneste NP-type, der kan placeres mellem finit verbal og baggrundspronomen, men også at der ikke er andet materiale, der kan stå her.

Også i beslægtede sprog som fornsvensk (Berger 2022: 277-280) og mo-

derne tysk (Helbig & Buscha 2001: 485-6) lader det til, at subjekt-NP'er er den eneste NP-type, der kan optræde før baggrundspronomen.

Det skal understreges, at denne analyse uden videre er kompatibel med Heltofts analyse af det middeldanske centralfelt som et udtrykssystem, som primært koder informationsstruktur.

4. Diakrone konsekvenser for subjektsplassens udviklingshistorie

Analysen har konsekvenser for analysen af, hvordan subjektsplassen er opstået. Ifølge Heltoft (1995; 158-161; 2001) opstår subjektsplassen i forbindelse med opkomsten af subjekt-verbum katatakse, dvs. det forhold, at et finit verbal kræver et subjekt, og omtolkningen af subjektet som aktualsubjekt, dvs. som led i kodning af illokutionær ramme: Som aktualled er subjektets tilstedeværelse på subjektsplassen med til at udtrykke interrogativ illokution i modsætning til fx deklarativer i såkaldt dagbogsstil (jf. Hansen & Heltoft 2011: 1690-1):

- (31) – Hvad er klokken?
 – Aner det ikke.
 (Hansen & Heltoft 2011: 1691)

Svaret kan kun læses som deklarativ. Interrogativtolkning kræver et subjekt på subjektsplassen:

- (32) – Aner **du** det ikke?

Heltofts forklaring er en meget overbevisende forklaring på, hvorfor subjektets topologiske potentiale *reduces*, dvs. hvorfor subjektet skal stå sammen med det finitte verbal og ikke længere kan stå på en senere plads i sætningen (uden vikar på subjektsplassen eller forfæltet i hvert fald).

Problemet med denne forklaring er dog, at *opkomsten* af subjektsplassen prædaterer opkomsten af Heltofts betingende faktorer, som først opstår i løbet af middeldansk (Heltoft 2001: 199). Om end subjektsplassen ikke har præcis samme semantik, som den har i dag (den koder fx ikke aktualsubjekter), kan man ifølge min analyse regne med en subjektsplass igennem hele den middeldanske periode. Det er derfor ikke en ny subjektsplass, der opstår, men subjektet mister pladser, hvor det kunne pla-

ceres tidligere som fx den ikoniske fokusplads illustreret i (3). Med andre ord er det, vi mangler, en forklaring på, hvordan subjektsplassen opstår, altså hvad der fører til, at kun subjektet kan placeres mellem finit verbal og letledsplassen. Dette sker ifølge min analyse for middeldansk.

5. Afslutning

I denne artikel argumenterer jeg for, at man inden for baggrundfeltet i middeldansk kan regne med en subjektsplass. Her kan kun subjekter stå, og konstruktionen koder dermed NP'er, der placeres her, som subjekter. Analysen indikerer derudover, at man også kan regne med en letledsplass i middeldansk. Til sidst har jeg argumenteret for, at subjektsplassens opkomst stadig mangler en forklaring.

Om forfatteren

Lennart Westergaard, cand.mag. i lingvistik, ph.d.-studerende i nordiske sprog, Språk- och litteraturcentrum, Lunds Universitet.

Citerede primærkilder

- EL=Eriks Sjællandske Lov (AM 455,12°) (1300), udg. Boeck, S.S. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- FS=Flensborg Stadsret (Stb.2) (1300), udg. Marita Akhøj Nielsen. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- HS=Henrik Harpestreng: Harpestrengs stenbog (Sthlm. K4) (1480). Udg. af Marita Akhøj Nielsen. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- JL=Jyske Lov (NKS 295,8°) (1325). Udg. Jonathan Adams og Marita Akhøj Nielsen. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk.
- LB=Lægebog fra top til taa (D600) (1450-1500) (Tekstnet). Udg. af Marita Akhøj Nielsen. Det Danske Sprog- og Litteraturselskab 2015. tekstnet.dk
- Luc = Lucidarius (AM 76, 8°) (1350). Udg. Af Carl Joakim Brandt. København: Det Berlingske Bogtrykkeri. Digitaliseret af Det Danske Sprog- og Litteraturselskab 2024. tekstnet.dk.
- SjT=Sjælens Trøst (Ups. C. 529) (1425). Udg. af Niels Nielsen (1937). København: Universitets-Jubilæets Danske Samfund v., J. H. Schultz Forlag.

- SjKL=Sjællands Kirkelov (AM 455, 12°) (1300). Udg. af Boeck, S.S. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- SkKL=Skånske Kirkelov (NkS 66,8°) (1350) Udg. af Marita Akhøj Nielsen. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- SL=Skånske Lov (AM 28, 8°). Udg. af Martin Sejer Danielsen. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk
- VL=Valdemars lov (AM 455,12°) (1300) Udg. af Boeck, S.S. Det Danske Sprog- og Litteraturselskab 2021. tekstnet.dk

Litteratur

- Anthony, Laurence (2022): *AntConc* (Version 4.1.2) [Computer Software]. Tokyo, Japan: Waseda University. Tilgængelig på <https://www.laurenceanthony.net/software>
- Berger, Michael (2022): *Objekt i svenskans historia*. Lund: Lunds Universitet.
- Diderichsen, Paul (1941): *Sætningsbygningen i Skaanske Lov. Fremstillet som Grundlag for en rationel dansk Syntaks*. København: Munksgaard.
- Diderichsen, Paul (1966[1943]): Logische und topische Gliederung des germanischen Satzes. I: *Paul Diderichsen. Helhed og Struktur. Udvalgte sprogvitenskabelige Afhandlinger*. Anders Bjerrum et al. (udg.). København: C.E.G. Gads forlag, 52-63.
- Hansen, Erik & Lars Heltoft (2011): *Grammatik over det Danske Sprog*. København: Det Danske Sprog- og Litteraturselskab. Syddansk Universitetsforlag.
- Helbig, Gerhard & Joachim Buscha (2001): *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Berlin: Langenscheidt.
- Heltoft, Lars (1995): Grammatikaliseringprocesser i dansk syntakshistorie. I: *Sproghistorie i 90'erne*. Bente Holmberg (red.). København: Selskab for Nordisk Filologi, 125–168.
- Heltoft, Lars (2001): Recasting Danish subjects: Case system, word order and subject development. I: *Grammatical relations in change*. Jan Terje Faarlund (red.). Amsterdam, Philadelphia: John Benjamins, 171-204.
- Heltoft, Lars (2003): Iconic and categorical focus systems in Scandinavian. I: *Structures of Focus and Grammatical Relations*. Jorunn Hetland & Valeria Molnár (red.). Berlin/Boston: De Gruyter, 41-86.
- Heltoft, Lars (2005): Ledsætning og letled i dansk. OV-rækkefølgens rester. I: *Grammatikalisering og struktur*. Lars Heltoft, Lene Schøsler & Jens Nørgård-Sørensen (red.). København: Museum Tusulanum, 145-166.

- Heltoft, Lars (2019): Syntaks. I: *Dansk Sproghistorie: Bøjning og bygning*. Ebba Hjorth, Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen, Laurids Kristian & Fahl (red.). Aarhus: Aarhus Universitetsforlag, 129-225.
- Kroman, Erik & Stig Iuul (1945): *Danmarks gamle love paa nutidsdansk*. København: G. E. C. Gads Forlag.
- Pedersen, Karen Margrethe (1993): Letledsreglen og lighedsreglen – Novation, ekspansion og resistens. I: *Jyske Studier*. Karen Margrethe Pedersen & Inge Lise Pedersen (red.). København: C.A. Reitzel, 199–218.