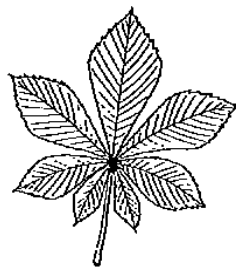
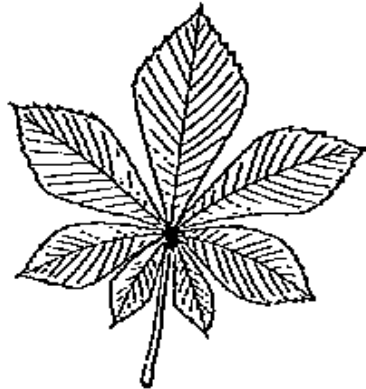


19. Møde
om Udforskningen af Dansk Sprog

NORDISK
Institut for Kommunikation og Kultur
Aarhus Universitet

13.-14. oktober 2022





19. Møde om Udforskningen af Dansk Sprog

Aarhus Universitet
13.-14. oktober 2022

Redigeret af Kirstine Boas, Inger Schoonderbeek Hansen, Tina Thode Hougaard og Ea
Lindhardt Overgaard
Aarhus 2023

NORDISK

Institut for Kommunikation og Kultur

Aarhus Universitet

www.muds.dk

Layout: Redaktionen

Tryk: Fællestrykkeriet, AUtryk

Aarhus Universitet

2023

ISBN 978-87-91134-01-3

FORORD

NORDISK, Institut for Kommunikation og Kultur, Aarhus Universitet, tog i 1986 initiativ til Møderne om Udforskningen af Dansk Sprog. Der har været afholdt møder hvert andet år siden da.

Det 19. møde fandt sted 13.-14. oktober 2022, og nogle år med coronarestriktioner, og et online MUDES i 2020, var gensynsglæden stor da vi endelig fysisk kunne mødes igen og afholde konference på vanlig vis i Århus. I alt var 52 forskere fra ind- og udland tilmeldt konferencen, og sammen med studerende og andre gæster deltog de aktivt i de 39 foredrag der blev afholdt.

Vores motto, "Ingen lingvistik uden humor", blev også holdt i hævd ved det 19. møde. Denne rapport indeholder 26 fagfællebedømte artikler med udgangspunkt i foredragene ved årets møde. En stor tak skal lyde til både konferencedeltagerne, oplægsholderne, bidragyderne og fagfællebedømmerne for de mange gode oplæg og artikler. Rapporten indeholder desuden 12 resuméer af oplæg hvor oplægsholder(e)n(e) f.eks. har fået udgivet bidraget andetsteds.

Omslagets farve på nærværende rapport afviger stærkt fra alle de forgange rapporters mere eller mindre farvestrålende omslag. Det er nemlig sidste gang rapporten trykkes fysisk. Det er nye tider, nye måder at publicere på. Det er jo vejen frem, men også lidt sørgeligt. Derfor er omslaget *sårt!*

Århus, august 2023

Kirstine Boas, Inger Schoonderbeek Hansen, Tina Thode Hougaard og Ea Lindhardt Overgaard

INDHOLD

Program		11
Deltagerne		14
Astrid Ag	Resumé: ”Og så fordi jeg ikke bruger de helt rigtige begreber” – om sprogbrug og uddannelsesfrafald	21
Liv Moeslund Ahlgren	Hvem handler hadbeskederne om? Magtmanipulerende sprog i danske hadbeskeder	23
Kirstine Boas	Tone-stød og tonal accent i nyere fynsk dialekt	37
Ocke-Schwen Bohn	Resumé: Internordisk sprogforståelse: Hvor svært kan det være?	51
Simon Borchmann	Resumé: Er handlingsanvisende diskursinitiale verballøse talehandlinger elliptiske? En afprøvning af hypotesen om minimalellipse	53
Kristoffer Friis Bøegh, Inger Schoonderbeek Hansen & Mette-Marie Møller Svendsen	Tegn-et-kort-data fra Aarhus og København – en perceptuel-dialektologisk undersøgelse	57
Christel Delia Christensen	Resumé: Subjektellipser på gravstene og i patientjournaler	75

Ken Ramshøj Christensen & Anne Mette Nyvad	Økuller? Det er ikke altid lige (u)acceptabelt at forlade en indlejring.	77
Birgit Eggert	Identitet i en stavemåde? Om stavevariation i nyfødtes fornavne	93
Yonatan Goldshtein	Resumé: Grundtræk af en prosodisk teori om stødet	107
Asgerd Gudiksen	Anna Pedersen og Udvalg for Folkemål	109
Anna Sofie Hartling, Michael Nguyen & Gustav Styrbjørn Johannessen	<i>At</i> eller ikke <i>at</i> i ledsætninger indledt af konjunktion/præposition	121
Peter Juel Henrichsen	Diktatoriske befølelser. Om ord og uord i Det Centrale Ordregister	133
Maja Agermose Hinrichsen	En sprogholdningsundersøgelse blandt fans af Rikke Thomsen	149
Annika Høeg	Hvad kan nonsensord fortælle os om stødets fonologi? Præsentation af et ph.d.-projekt	165
Torben Juel Jensen, Marie Maegaard & Nicolai Pharao	Hvordan vurderes jyske participiumsformer?	177
Viggo Bank Jensen	<i>Eftertid</i> . En grammatikterms historie og historiens relevans for nyere dansk verbalteori.	197

Julia Jørgensen & Maria Jørgensen	Interaktion på noder	213
Maria Jørgensen	Resumé: Om svar på specificerende HV-spørgsmål i dansk	227
Nicholas Mikkelsen	<i>Altså Bertel Haarder han ku ikk li noget med universiteter.</i> Deltagelse i historiefortælling i en dansk hverdagssamtale	229
Ehm Miltersen	Resumé: "Vi er nogle der bruger det!" - Et historisk perspektiv på danske kønsinklusive pronomener fra 1942 til 2021	241
Maja Mittag & Line Burholt Kristensen	Resumé: "Har du set Nor, efter <i>de</i> har fået et nyt job?" – Et acceptabilitetsstudie af singulært <i>de</i>	245
Malene Monka, Marie Maegaard, Nikolaj Pharao	Resumé: Sprogholdninger i center og periferi i Nordjylland og Sønderjylland	247
Anne Mette Nyvad og Ken Ramshøj Christensen	"Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!" Når performansen halter.	249
Ea Lindhardt Overgaard	1500-talssalmer og danske sprogmodeller	265
Sidsel Holm Rasmussen & Camilla Søballe Horslund	Resumé: Den danske alveopalatale sibilant [ɕ]: Fonologi, ortografi og psykologisk realitet af et muligt /ɕ/ fonem	279
Jørgen Schack & Eva Skafte Jensen	Uden og med <i>-t</i> . Adverbialer dannet af adjektiver på <i>-ig</i> og <i>-lig</i>	281

Irene Simonsen	Rækkefølgens betydning i bindestregs-nationalitetsadjektiver som <i>tysk-dansk</i> og <i>dansk-tysk</i>	293
Mikołaj Sobkowiak	Resumé: Nominalfrasens kompleksitet i dansk som L1 og L2 – en undersøgelse af polske læreres og danske gymnasieelevers skriftlige produktion	307
Andrzej Szubert	Konstruktioner af typen <i>en glædens dag</i> – en korpusanalyse	309
Søren Sandager Sørensen	Resumé: Klausale eksklamativer med hvor, hvilken og sikke – og deres ordstilling	319
Morten Tannert & Ulf Dalvad Berthelsen	Forfælsvariation i elevtekster fra 5. og 9. klassetrin	321
Jacob Thøgersen	Danskerne og det engelske sprog i 1990erne og 2020erne	343
Ole Togeby	<i>Lad os bare håbe!</i> Adverbialer, sætningsskema og tryk	361
Christophe Zerakitsky Vies	Resumé: <i>Altså</i> udtaler og funktioner i DanPASS-korpusset	369
Sten Vikner	Hvor kommer subjektet fra? -VP- intern oprindelse som grund til skopus-tvetydighed	371
Lennart Westergaard	<i>Jo</i> som dialogisk partikel i tidlig ældre nydansk	389
Thomas Widmann	Det Centrale Ordregister (COR): Hvad, hvor, hvorfor og hvordan?	403

PROGRAM

TORS DAG		
10.00-10.05	Åbning	
	<u>Richard Mortensenstuen</u>	<u>Mogens Zielerstuen</u>
	Ordstyrer: <i>Inger Schoonderbeek Hansen</i>	
10.05-10.35	Plenar v. Malene Monka, Marie Maegaard & Nicolai Pharao <i>Sprogholdninger i center og periferi i Nordjylland og Sønderjylland</i>	
	Ordstyrer: <i>Inger Schoonderbeek Hansen</i>	Ordstyrer: <i>Tina Thode Hougaard</i>
10.35-11.05	Kirstine Boas <i>(Hvorfor) taler de fynsk på Fyn?</i>	Ulf Dalvad Berthelsen <i>En kvantitativ analyse af forfælsvariation i elevtekster fra 5. og 9. klassetrin</i>
11.05-11.30	Kaffepause	
	Ordstyrer: <i>Malene Monka</i>	Ordstyrer: <i>Ulf Dalvad Berthelsen</i>
11.30-12.00	Kristoffer Friis Bøegh, Inger Schoonderbeek Hansen & Mette-Marie Møller Svendsen <i>Tegn-et-kort-data fra Århus og København: en perceptuel-dialektologisk undersøgelse</i>	Morten Tannert <i>En kvalitativ analyse af forfælsvariation i elevtekster fra 5. og 9. klassetrin</i>
12.00-12.30	Maja Agermose Hinrichsen <i>En sprogholdningsundersøgelse blandt fans af Rikke Thomsen</i>	Thomas Widmann <i>Det Centrale Ordregister – hvad er det, og hvad kan jeg bruge det til?</i>
12.30-13.00	Annika Høeg <i>Hvad kan nonsensord fortælle os om stødets fonologi? – præsentation af et ph.d.-projekt</i>	Peter Juel Henrichsen <i>COR-linkeren – trøst til den trætte sprogteknolog</i>
13.00-14.00	Frokost	
	Ordstyrer: <i>Kristoffer Friis Bøegh</i>	Ordstyrer: <i>Eva Skafte Jensen</i>
14.00-14.30	Yonatan Goldshtein <i>Grundtræk af en prosodisk teori og stødet</i>	Ea Lindhardt Toft <i>Hvordan måler man en genre?</i>
14.30-15.00	Sidsel Rasmussen & Camilla Søballe Horslund <i>Den danske alveopalatale sibilant [ɛ]: Fonologi, ortografi og psykologisk realitet af et muligt / ɛ/ fonem</i>	Lennart Westergaard <i>Jo findes jo allerede i gammeldansk – om den ældste dialogiske partikel</i>
15.00-15.30	Jacob Thøgersen <i>Danskerne og det engelske sprog – 1990'erne og 2020'erne</i>	Andrejz Szubert <i>Konstruktioner af typen "en glædens dag" – en korpusanalyse</i>
15.30-16.00	Kaffe	
	Ordstyrer: <i>Jacob Thøgersen</i>	Ordstyrer: <i>Kathrine Thisted Petersen</i>
16.00-16.30	Eva Skafte Jensen & Jørgen Schack <i>Uden og med -t</i>	Christophe Zerakitsky Vies <i>Altså udtaler og funktioner i DanPASS-korpusset</i>
16.30-17.00	Torben Juel Jensen, Marie Maegaard & Nicolai Pharao	Mikolaj Sobkowiak

	<i>Hvordan vurderes jyske participiumsformer?</i>	<i>Nominalfrasens kompleksitet i dansk som L1 og L2 – en undersøgelse af polske lørneres og danske gymnasieelevers skriftlige produktion</i>
17.00-17.30	Asgerd Gudiksen <i>Anna Pedersen og Udvalg for Folkemål</i>	Viggo Bank Jensen <i>Om betegnelsen "Eftertid" i dansk grammatikterminologi fra Rask til Mikkelsen</i>
17.30-18.00	Birgit Eggert <i>Identitet i en stavemåde. Om stovevariation i nyfødtes fornavne</i>	-

FREDAG		
	<u>Richard Mortensenstuen</u>	<u>Mogens Zielerstuen</u>
	Ordstyrer: <i>Ole Togeby</i>	Ordstyrer: <i>Maria Jørgensen</i>
09.00-9.30	Michael Nguyen & Anna Sofie Hartling <i>At eller ikke at efter konjunktion/præposition. En undersøgelse af et nutidigt avis-korpus</i>	Maja Mittag og Line Burholt Kristensen <i>"Har du set Nor_i, efter de_i har fået et nyt job?" Et acceptabilitetsstudie af singulært de.</i>
9.30-10.00	Christel Delia Christensen <i>Subjektsejllipse på gravstenene og i patientjournaler</i>	Astrid Ag <i>"Og så fordi jeg ikke bruger de helt rigtige begreber" – om sprogbrug og uddannelsesfald</i>
10.00-10.30	Anne Mette Nyvad & Ken Ramshøj Christensen <i>"Vi har fundet nogle sætninger på nettet, og dem er folk vild med!" Når performansen halter.</i>	Søren Sandager Sørensen <i>Klausale eksklamativer med "hvor", "hvilken" og "sikke" – og deres ordstilling</i>
10.30-11.00	Kaffe	
	Ordstyrer: <i>Michael Nguyen</i>	Ordstyrer: <i>Nicholas Mikkelsen</i>
11.00-11.30	Sten Vikner <i>Hvor kommer subjektet fra? Semantiske fordele ved en VP-intern oprindelse</i>	Maria Jørgensen <i>Om svar på specificerende HV-spørgsmål i dansk</i>
11.30-12.00	Anne Mette Nyvad & Ken Ramshøj Christensen <i>Ø-kuller? Hvor (u)acceptabelt er det at forlade indlejringen?</i>	Maria Jørgensen & Julia Jørgensen <i>Interaktion på noder – et eksperiment</i>
12.00-12.30	Simon Borchmann <i>Er handlingsanvisende diskursinitiale verbaltale talehandlinger elliptiske? En afprøvning af hypotesen om minimalejllipse</i>	Ocke-Schwen Bohn <i>Internordisk sprogforståelse: Hvor svært kan det være?</i>
12.30-13.30	Frokost	
	Ordstyrer: <i>Sten Vikner</i>	Ordstyrer: <i>Yonatan Goldshtein</i>
13.30-14.00	Ole Togeby <i>Adverbialer, tryk og a2-pladsen</i>	Liv Moeslund Ahlgren <i>Hvem handler hadbeskederne om?</i>
14.00-14.30	Irene Simonsen <i>Rækkefølgens betydning i bindestregs- og nationalitetsadjektiver som tysk-dansk og dansk-tysk</i>	Nicholas Mikkelsen <i>Deltagelse i historiefortælling i danske hverdagssamtaler</i>
	Ordstyrer: <i>Tina Thode Hougaard</i>	

14.30- 15.00	Plenar v. Ehm Miltersen <i>"Vi er nogle der bruger det!" – Et historisk perspektiv på danske kønsinklusive pronomener fra 1942 til 2021</i>
15.00- 15.15	Afslutning

DELTAGERNE

19. Møde om Udforskningen af Dansk Sprog

Astrid Ag
Københavns Universitet
astridag@hum.ku.dk

”Og så fordi jeg ikke bruger de helt
rigtige begreber”
– om sprogbrug og uddannelsesfrafald

Ulf Dalvad Berthelsen
Aarhus Universitet
udb@cc.au.dk

Forfelsesvariation i elevtekster fra 5. og
9. klassetrin (sammen med Morten
Tannert)

Kirstine Boas
Aarhus Universitet
boas@cc.au.dk

Tone-stød og tonal accent i nyere
fynsk dialekt

Ocke-Schwen Bohn
Aarhus Universitet
ocke-bohn@cc.au.dk

Internordisk sprogforståelse: Hvor
svært kan det være?

Simon Borchmann
Roskilde Universitet
sub@ruc.dk

Er handlingsanvisende diskursinitiale
verballøse talehandlinger elliptiske?
En afprøvning af hypotesen om
minimalellipse

Kristoffer Friis Bøegh
Aarhus Universitet
kfb@cc.au.dk

Tegn-et-kort-data fra Aarhus og
København – en perceptuel-
dialektologisk undersøgelse (sammen
med Inger Schoonderbeek Hansen &
Mette-Marie Møller Svendsen)

Ken Ramshøj Christensen
Aarhus Universitet
krc@cc.au.dk

Økuller? Det er ikke altid lige
(u)acceptabelt at forlade en indlejring
(sammen med Anne Mette Nyvad)

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”
Når performansen halter. (sammen med Anne Mette Nyvad)

Christel Delia Christensen
Aarhus Universitet
201408193@post.au.dk

Subjektellipser på gravstene og i patientjournaler

Philip Diderichsen
Københavns Universitet
cpd@hum.ku.dk

Birgit Hjort Eggert
Københavns Universitet
beggert@hum.ku.dk

Identitet i en stavemåde?
Om stavevariation i nyfødtes fornavne

Lea Friis Klamer
Københavns Universitet
wmf998@hum.ku.dk

Yonatan Goldshtein
Aarhus Universitet
yonatan@cc.au.dk

Asgerd Gudiksen
Københavns Universitet
gudik@hum.ku.dk

Anna Pedersen og Udvalg for Folkemål

Inger Schoonderbeek Hansen
Aarhus Universitet
jysis@cc.au.dk

Tegn-et-kort-data fra Aarhus og København – en perceptuel-dialektologisk undersøgelse (sammen med Kristoffer Friis Bøegh & Mette-Marie Møller Svendsen)

Anna Sofie Hartling
Dansk Sprognævn
ash@dsn.dk

At eller ikke *at* i ledsætninger indledt af konjunktion/præposition (sammen med Michael Nguyen & Gustav Styrbjørn Johannessen)

Lars Heltoft
Københavns Universitet
heltoft@hum.ku.dk

Peter Juel Henriksen
Dansk Sprognævn
pjh@dsn.dk

Diktatoriske befølelser. Om ord og uord i Det Centrale Ordregister

Maja Agermose Hinrichsen
Aarhus Universitet
majahinrichsen@gmail.com

En sprogholdningsundersøgelse blandt fans af Rikke Thomsen

Sidsel Holm Rasmussen
Aarhus Universitet
sira@cc.au.dk

Den danske alveopalatale sibilant [ɶ]:
Fonologi, ortografi og psykologisk realitet af et muligt /ɶ/ fonem (sammen med Camilla Søballe Horslund)

Camilla Søballe Horslund
Aarhus Universitet
horslund@cc.au.dk

Den danske alveopalatale sibilant [ɶ]:
Fonologi, ortografi og psykologisk realitet af et muligt /ɶ/ fonem (sammen med Sidsel Holm Rasmussen)

Tina Thode Hougaard
Aarhus Universitet
north@cc.au.dk

Asta Hvid Andersen
Københavns Universitet
hws287@hum.ku.dk

Annika Høeg
Stockholms Universitet
annika.hoeg@su.se

Hvad kan nonsensord fortælle os om
stødets fonologi? Præsentation af et
ph.d.-projekt

Eva Skafte Jensen
Dansk Sprognævn
esj@dsn.dk

Uden og med *-t*. Adverbialer dannet
af adjektiver på *-ig* og *-lig* (sammen
med Jørgen Schack)

Viggo Bank Jensen
Københavns Universitet
vbj@hum.ku.dk

Eftertid. En grammatikterms historie
og historiens relevans for nyere dansk
verbalteori.

Torben Juel Jensen
Københavns Universitet
tjuelj@hum.ku.dk

Hvordan vurderes jyske
participiumsformer? (sammen med
Marie Maegaard & Nicolai Pharao)

Maria Jørgensen
Aarhus Universitet
marijo@cc.au.dk

Om svar på specificerende HV-
spørgsmål i dansk

Interaktion på noder (sammen med
Julia Jørgensen)

Heidi McGhee
Aarhus Universitet
h.mcgee@cc.au.dk

Nicholas Mikkelsen
Aarhus Universitet
nicholas.mikkelsen@cc.au.dk

*Altså Bertel Haarder han ku ikk li
noget med universiteter*. Deltagelse i
historiefortælling i en dansk
hverdagssamtale

Ehm Miltersen
Aarhus Universitet
201405358@post.au.dk

“Vi er nogle der bruger det!” - Et
historisk perspektiv på danske
kønsinklusive pronomener fra 1942 til
2021

Maja Mittag
Københavns Universitet
majamittag@hum.ku.dk

Har du set Nor, efter *de* har fået et nyt job?” – Et acceptabilitetsstudie af singulært *de* (sammen med Line Burholt Kristiansen)

Liv Moeslund Ahlgren
Aarhus Universitet
kontakt@livmoeslund.dk

Hvem handler hadbeskederne om?
Magtmanipulerende sprog i danske hadbeskeder

Malene Monka
Københavns Universitet
monka@hum.ku.dk

Sprogholdninger i center og periferi i Nordjylland og Sønderjylland (sammen med Marie Maegaard & Nicolai Pharao)

Michael Nguyen
Dansk Sprognævn
mn@dsn.dk

At eller ikke *at* i ledsætninger indledt af konjunktion/præposition (sammen med Anna Sofie Hartling & Gustav Styrbjørn Johannessen)

Anne Mette Nyvad
Aarhus Universitet
amn@cc.au.dk

Økuller? Det er ikke altid lige (u)acceptabelt at forlade en indlejring (sammen med Ken Ramshøj Christensen)

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”
Når performansen halter. (sammen med Ken Ramshøj Christensen)

Ea Lindhardt Overgaard
Aarhus Universitet
elt@cc.au.dk

1500-talssalmer og danske sprogmodeller

Marianne Rathje
Dansk Sprognævn
mr@dsn.dk

Ole Ravnholt
Orsprog.dk
or@orsprog.dk

Jørgen Schack
Dansk Sprognævn
schack@dsn.dk

Irene Simonsen
Syddansk Universitet
simonsen@sdu.dk

Mikolaj Sobkowiak
Adam Mickiewicz Universitet i Poznan
miksobko@amu.edu.pl

Mette-Marie Møller Svendsen
Aarhus Universitet
mms@cc.au.dk

Andrzej Szubert
Adam Mickiewicz Universitet i Poznan
szubert@amu.edu.pl

Søren Sandager Sørensen
Aarhus Universitet
mails@sorensandagersorensen.dk

Uden og med *-t*. Adverbialer dannet af adjektiver på *-ig* og *-lig* (sammen med Eva Skafte Jensen)

Rækkefølgens betydning i bindestregs-nationalitetsadjektiver som *tysk-dansk* og *dansk-tysk*

Nominalfrasens kompleksitet i dansk som L1 og L2
– en undersøgelse af polske læreres og danske gymnasieelevers skriftlige produktion

Tegn-et-kort-data fra Aarhus og København – en perceptuel-dialektologisk undersøgelse (sammen med Kristoffer Friis Bøegh & Inger Schoonderbeek Hansen)

Konstruktioner af typen *en glædens dag* – en korpusanalyse

Klausale eksklamativer med *hvor*, *hvilken* og *sikke* – og deres ordstilling

Morten Tannert
VIA University
MTAN@via.dk

Forfælsvariation i elevtekster fra 5. og
9. klassetrin (sammen med Ulf Dalvad
Berthelsen)

Kathrine Thisted Petersen
Aarhus Universitet
norktp@cc.au.dk

Jacob Thøgersen
Københavns Universitet
jthoegersen@hum.ku.dk

Danskerne og det engelske sprog i
1990erne og 2020erne

Ole Tøgeby
Aarhus Universitet
norot@cc.au.dk

Lad os bare håbe!
Adverbialer, sætningskema og tryk

Sten Vikner
Aarhus Universitet
sten.vikner@cc.au.dk

Hvor kommer subjektet fra? -VP-
intern oprindelse som grund til
skopus-tvetydighed

Lennart Westergaard
Lunds Universitet
lennart.westergaard.1513@nordlund.lu.se

Jo som dialogisk partikel i tidlig
ældre nydansk

Thomas Widmann
Dansk Sprognævn
tw@dsn.dk

Det Centrale Ordregister (COR):
Hvad, hvor, hvorfor og hvordan?

Christophe Zerkitsky Vies
Københavns Universitet
christophe.vies@gmail.com

Altså udtaler og funktioner i
DanPASS-korpusset

Resumé:

”Og så fordi jeg ikke bruger de helt rigtige begreber”

– om sprogbrug og uddannelsesfrafald

Astrid Ag

Københavns Universitet

Når man som studerende vælger at afslutte en uddannelse før tid, skyldes det ofte flere forskellige faktorer. Uddannelsesfrafald er nemlig en kompleks problematik og bør derfor belyses fra forskellige vinkler. I projektet *Speaking Up* ønsker vi at undersøge om sprog spiller en rolle og i givet fald hvilken. Vi vil gerne forstå hvordan sprog kan fungere som en sorteringsmekanisme der får nogle studerende til at blive på en uddannelse og andre til at opgive den. I dette oplæg kigger jeg mere specifikt på hvilke sproglige stile, herunder brug af fagudtryk, studerende associerer med uddannelsesmæssig succes, og hvordan uddannelsesmæssig succes forhandles sprogligt blandt dem. Jeg fokuserer på interviewberetninger fra studerende om at manglende kendskab til fagbegreber har ført til at medstuderende ikke lytter til dem, og om beretninger fra studerende om (mere eller mindre bevidst) at bruge bestemte ord som andre ikke forstår. Jeg analyserer også optagelser af gruppearbejder hvor sådanne dynamikker er i spil blandt de studerende, ligesom jeg inddrager refleksioner fra deres undervisere. Fordi data indsamles ad to omgange med et års mellemrum, kan jeg relatere disse forhold til uddannelsesrelaterede konsekvenser og dermed belyse eventuelle sammenhænge mellem sprogbrug og uddannelsesfrafald.

Data består af etnografiske observationer, interviews, selvoptagelser og spørgeskemaer indsamlet blandt pædagog- og lærerstuderende på Professionshøjskolen Absalon i Vordingborg. Dette delprojekt i Vordingborg er en del af *Speaking Up*-projektet som med forskellige metodiske tilgange undersøger sprog som faktor for social mobilitet og konsekvenserne af at tale noget andet end rigsdansk, når man starter på en uddannelse. Projektet følger lærer- og pædagogstuderende og deres lærere og studievejledere i Hjørring, Jelling og Vordingborg.

Hvem handler hadbeskederne om?

Magtmanipulerende sprog i danske hadbeskeder

Liv Moeslund Ahlgren

Aarhus Universitet

1. Introduktion

Hvis man optræder i medierne, eller på anden vis er en offentligt kendt person, kan der være stor risiko for, at ens indbakker fyldes med ubehagelige eller nedgørende beskeder fra fremmede. I denne artikel vil jeg argumentere for, at disse *hadbeskeder* er komplekse sproglige fænomener: På overfladen indeholder de primært en grov kritik af modtageren og fremstiller denne på fornedrende vis; men samtidig er der forskellige sproglige elementer, der præsenterer afsenderen som overlegen og magtfuld. Jeg undersøger således, hvordan hadbeskeder samtidig handler om både modtageren og afsenderen. Mere konkret undersøger jeg de sproglige strategier, der bruges til at udøve magt, altså hvordan der gennem nedværdigelse af modtageren bliver manipuleret med samtaledeltagernes magtpositioner. Studiet bygger på en tekstnær analyse af 45 autentiske hadbeskeder sendt til danske kvinder. Resultaterne vil forhåbentligt ikke blot være relevant for dem, der er interesserede i genren, men også kunne udvide vores forståelse af magtmanipulerende sprog generelt.

Den eksisterende sprogvidenskabeligt funderede litteratur om hadbeskeder er sparsom (se dog Piñeiro-Otero & Martínez-Rolán, 2021; og Cloud, 2009), og nærværende studie er derfor i stedet funderet i teorier og metoder fra forskning i trusler og uhøflighed. Med udgangspunkt i Culpepers (2011) uhøflighedsteori undersøger Muschalik (2018), hvordan trusler kan bruges til at påpege et bestemt magtforhold mellem modtager og afsender, og hvordan dette magtforhold gennem dem trusselhandlingen kan ændres til afsenderens fordel. Dette gælder både trusler, som udsiges for at få modtageren til at udføre en handling, og trusler som udsiges som en slags straf eller gengæld for modtagerens handlinger (Muschalik, 2018: 131). Selvom man traditionelt har anset trusler for at kræve en bagvedliggende autoritet eller magtposition for at kunne være effektive (Harris, 1984; Limberg, 2009:1379), kan de også fungere, selvom afsenderen ikke besidder en høj magtposition – faktisk kan truslen ses som en måde at opnå magt (Bojsen-

Møller et al, 2020:22-23; Muschalik 2018:20). Trusler kan altså både fremhæve eksisterende magtstrukturer eller skabe dem gennem trusselhandlingen.

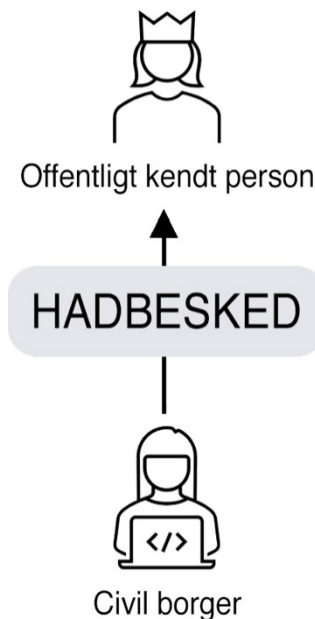
Lokal magtmanipulation er ikke kun en væsentlig funktion ved trusler men også ved andre former for uhøflig, krænkende eller aggressiv adfærd (Culpeper, 2011; Beebe, 1995). I dette studie argumenterer jeg således for, at denne type lokal magtmanipulation også foregår i hadbeskederne, der altså kan bruges som et værktøj til at vende et magtforhold om. Det er dog vigtigt at pointere, at hadbeskeder også altid må ses i et større perspektiv, hvor blandt andet sexistiske samfundsstrukturer også spiller en vigtig rolle. Studier af voldeligt eller krænkende sprog fokuserer ofte på at afdække sproglige strategier for hadtale (Assimakopoulos et al, 2017; Baumgarten et al., 2019; Geyer, 2018), eller hvordan f.eks. kvindefjendsk online kommunikation ofte opstår i samfund med stigende anti-feminisme og politisk højredrejning (Ging & Siaperá 2019). Piñeiro-Otero og Martínez-Rolán (2021) kortlægger således, hvordan mange beskeder til offentligt kendte kvinder på Twitter indeholder nedværdigende eller krænkende sprog, og konkluderer, at selvom hadbeskederne ikke altid indeholder decideret misogyn hadtale, er de stadig et udtryk for et stigende kvindehad, som særligt trives på platforme som Twitter. Den lokale magtmanipulation, der bliver undersøgt i denne artikel, er således kun en lille del af de komplekse fænomener, som en analyse af hadbeskeder kan belyse.

2. Om hadbeskeder

Før vi bevæger os videre, er det på sin plads at få klart opridset, hvad vi forstår ved 'hadbesked'. Desværre findes der ingen klar definition på dette, og de fleste undersøgelser af hadbeskeder bruger termen uden begrebsafklaring eller refererer til fænomener som *sproglige angreb* (Analyse & Tal, 2021), *digital chikane* (Dansk Journalistforbund, 2020) eller *uønskede henvendelser og hadefulde kommentarer* (Steffensen & Zuleta, 2022).

I denne artikel defineres hadbeskeder både ud fra deres indhold og kontekst. Hadbeskeder er beskeder, som på forskellig vis opfattes som ubehagelige af modtageren, og de indeholder oftest en form for kritik af modtageren, trusler eller uønskede komplimenter. Henvendelser med saglig kritik af modtagerens holdninger regner jeg ikke for at være hadbeskeder. Hadbeskeder er heller ikke nødvendigvis det samme som f.eks. en vred sms fra en ekskæreste, for de har et helt særligt modtager-afsender-forhold. Kontekstmæssigt er hadbeskeder nemlig beskeder, der er sendt fra en civil borger til en offentligt kendt person. Modtageren og afsenderen er derfor typisk fremmede for hinanden, og modtageren har typisk kulturel, politisk eller økonomisk magt i samfundet, som afsenderen ikke har (se Schneider, 2010:249-50 og Schneider & Barron,

2008:18 om *mikrosociale faktorer*). Figur 1 illustrerer dette magtforhold, som vi vil vende tilbage til i diskussionen i afsnit 6.



Figur 1: Modtager- og afsenderforhold. Kronen indikerer magt.

Hadbeskeder kan desuden sendes både gennem offentlige og private kanaler, dvs. både som f.eks. offentligt tilgængelige Facebook-kommentarer eller i private indbakker.

Hadbeskeder modtages både af politikere, debattører og journalister (Bruun-Hansen, 2022; Hagen, 2015; Kvist & Øtting, 2015; Steffensen & Zuleta, 2022); deltagere i TV-show (Astrup, 2021; Rabøl, 2021); og jeg har været i personlig kontakt med sportsudøvere, influencere, aktivister, forfattere og forskere, der også modtager hadbeskeder. Der er desuden nogle demografiske forskelle på hadbeskederes omfang og karakter. For eksempel modtager kvinder oftere beskeder, der handler om deres køn eller person, hvorimod mænds beskeder handler om holdninger og udtalelser (Bruun-Hansen, 2022; Hagen, 2015), og blandt lokalpolitikere udsættes kvinder og etniske minoritetspersoner for mest chikane (Steffensen & Zuleta, 2022). Fordi hadbeskeder samtidig ofte fører til selvcensur eller en fuldstændig udmeldelse af den offentlige debat (Herskind, 2022; Kvist & Øtting, 2015; Kaasgaard, 2020; Lauritsen et al., 2016; Steffensen & Zuleta, 2022), udgør denne demografiske skævvridning et demokratisk problem.

3. Modtager- og afsenderfremstilling i hadbeskeder

Som nævnt findes der stort set ingen sprogvidenskabelig forskning i hadbeskeder – undtagelsen er Clouds (2009) studie, der kortlægger identitetsskabelse og *framing* i 290 hadbeskeder. Her

skaber afsenderne deres egen identitet som kontrast til deres repræsentation af modtageren i beskederne. Modtageren bliver ofte fremstillet som elitær intellektuel, landsforræder og kønsforræder, hvilket således er med til at fremstille afsenderne som jordbundne patrioter og rigtige mænd (Cloud, 2009). Clouds analyse bygger på teori om, at sociale grupper kan skabe intern solidaritet ved at konstruere en fælles fjende og opstille sig som modsætning hertil (Burke, 1969; Butler, 1997; McCaffrey & Keys, 2000). Et farligt 'dem' giver således et styrket 'vi'. Udover at skabe bedre sammenhængskraft i grupper kan negativ repræsentation af andre også være med til at skabe positive selvbilleder, hvor et billede af modtageren som f.eks. nationalforræder samtidigt er et billede af afsenderen som patriot. Hadbeskederes karakteristiske af modtageren fungerer altså som en karakteristik af afsenderen, og Cloud (2009:458, 473) ser derfor hadbeskeder som et værktøj til at stille afsenderen i bedre lys.

Dette studie er også funderet i idéen om modtagerfremstilling-som-afsenderfremstilling, men jeg undersøger dog en anden form for afsenderfremstillinger end Cloud (2009). I stedet for at analysere, hvordan *du er klam* bliver til *jeg er ikke klam*, undersøger jeg i stedet, hvordan nedværdigelse af modtageren, uanset den kritiserede adfærd eller egenskab, er med til at fremstille afsenderen som overlegen og magtfuld. Jeg bygger på teori om præsuppositioner, altså noget, man antager er sandt eller velkendt forud for en ytrings udsigelse (Yule, 1996:25). Ved at spørge om nogen har haft en god ferie, antager man, at vedkommende rent faktisk har haft ferie – denne ferie er en forudsætning for at ytringen giver mening. Ligesom at ytringens indhold kan præsupponere bestemte forhold, kan deres illokutioner (altså deres kommunikative funktioner) det også. Når man for eksempel giver en ordre, præsupponeres det, at man har autoritet til at give en ordre, som modtageren vil følge, og denne autoritet bliver altså en forudsætning for, at ordren giver mening som sproghandling (McDonald, 2022:215). Ved at udføre bestemte sproghandlinger forudsætter man altså bestemte sociale positioner eller relationer – uanset om disse rent faktisk findes. Som jeg vil uddybe i denne artikel, indeholder hadbeskederne forskellige elementer, der på sådan vis positionerer afsenderen som særligt overlegen eller magtfuld.

4. Datagrundlag

Datagrundlaget for dette studie er 45 autentiske danske hadbeskeder. De er modtaget af offentligt kendte kvinder, herunder både politikere fra højre- og venstrefløjen, forfattere, TV- og radioværter, debattører, influencere, aktivister og forskere. Alle beskederne er sendt gennem private kanaler, dvs. på sms, e-mail, Facebook Messenger eller Instagrams Direct Message. Alle kan kategoriseres som kritiske eller negative hadbeskeder, hvilket betyder, at de rummer en eller anden form for kritik af modtageren (modsat hadbeskeder, der indeholder uønskede

komplimenter (Bruun-Hansen, 2022). De er indsamlet fra avisartikler, dokumentarer eller radioprogrammer, hvor de har været gengivet, eller ved at jeg personligt har henvendt mig til modtagerne og bedt om lov til at tilgå nogle af de hadbeskeder, de har modtaget privat. Fælles for alle beskederne er, at modtageren selv har vurderet, at den kunne kategoriseres som en hadbesked. Dette betyder, at beskeder, hvor kritikken af modtageren måske ikke er så tydelig, stadig vil kunne indgå i denne analyse, fordi de – uanset hvad vi som forskere mener – har haft en effekt på modtageren.

Alle beskeder, der er gengivet i denne artikel, er fuldt anonymiserede, men referencer til køn, race, religion, kropstype, udseende og lignende er dog bevaret. Alle originale grammatiske og ortografiske afvigelser er ligeledes bevaret så vidt muligt. Derudover er en advarsel på sin plads – beskederne gengivet i artiklen her indeholder voldsomt sprogbrug og kan virke ubehagelige.

5. Analyse

5.1 Modtagerrepræsentation

Før vi giver os i kast med at analysere, hvordan der i hadbeskederne sprogligt bliver manipuleret med afsenderens og modtagerens magtpositioner, vil jeg først tage jer med tilbage til mine tidligste observationer af hadbeskederne sproglige karakteristika. Beskederne varierer enormt meget i længde, stiludtryk, grovhed, kreativitet og indhold af skældsord. Eksempel (1-2) har således samme modtager, men er meget forskellige i deres sproglige udtryk:

- (1) Hvis du ikke kan lide lugten i bageriet kan du for helvede blive væk. Dit fede korpus kan ikke engang tåle myggestik og hvis Coruna ikke tager dig nu, så får du helt sikkert blodpropper og så må vi se om du kan fremstamme alt det lort du hele tiden render rundt i pressen med. Bliv væk fra lægen, klamme kælling.
- (2) Syns du det er dejligt at være tyk?

Det, beskederne har tilfælles, er dog, at de handler om modtageren, som bliver kritiseret og nedgjort. Dette fokus på modtageren er også noget, der kan ses sprogligt, blandt andet ved den høje forekomst af andenpersonspronomen. I tabel 1 ses det, hvordan *du* er det mest frekvente ord, og *du er* er den mest frekvente toordskombination, Til sammenligning er *jeg* det 13. mest frekvente ord. Derudover indeholder alle beskeder undtagen én et eller flere andenpersonspronomen.

Unigrams	Bigrams
du = 78	du er = 14
er = 51	det er = 9
en = 38	[navn] [navn] = 7
og = 38	du får = 5
at = 34	er en = 4
i = 28	til at = 4
dig = 27	er du = 4
for = 25	du ligner = 4
ikke = 24	ligner en = 4
det = 23	du ikke = 4
til = 20	for at = 4
så = 19	der er = 4
jeg = 18	din fede = 4
din = 17	for helvede = 4
på = 15	hvorfor er = 3
[navn] = 14	dit eget = 3
har = 14	en abe = 3
som = 13	prøv at = 3
af = 12	fy for = 3
et = 11	for satan = 3

Tabel 1: Etords- og toordskombinationer

Med 45 beskeder er der naturligvis ikke nok data til at kunne begive sig ud i korpuslingvistiske analyser – men distributionen af pronomener peger i hvert fald på, at modtageren (blandt andet repræsenteret gennem andenpersonspronomen) er helt central for hadbeskeder (Adolphs, 2008:10).

Rent indholdsmæssigt handler hadbeskederne altså tydeligt om modtageren. Men samtidig rummer de også forskellige sproglige elementer, der kan være med til at fremstille afsenderen på en særlig måde – selvom de ikke som sådan handler om afsenderen. I de følgende afsnit vil jeg vise, hvordan evalueringer, påbud, moraliseringer og reprimander i hadbeskederne bruges til dette.

5.2 Evalueringer

Eksempel (3) er en kort besked, der består af en kopulakonstruktion (*du er X*), hvor prædikatet indeholder en række adjektiver og adverbialer og skældsordet *smatso*. Derefter kommer en vokativ *dit X*-konstruktion, hvor det possessive pronomen ikke bærer nogen possessiv betydning (Julien, 2016), men derimod markerer en ekspressiv og nedgørende betydning (d'Avis & Meibauer, 2013:189; Havryliv, 2009) – her *lesbiske møgdyr*.

- (3) Du er bare en total usmagelig led smatso dit lesbiske møgdyr.

Der er flere elementer, der gør beskeden kraftfuld: den ekspressive *din X*-konstruktion; brugen af skældsord; og brugen af adjektiver og adverbialer, der styrker skældsordene. Derudover bevirker *smatso*, der typisk bruges om kvinder, og adjektivet *lesbiske* at modtageren bliver nedgjort på baggrund af sin gruppeidentitet (her: køn og seksualitet) (Meibauer, 2016:157).

Ved at beskrive andre negativt kan man samtidigt positionere sig selv som en modsætning hertil (McCaffrey & Keys, 2000; Cloud, 2009): Dette er særligt tydeligt i hadtale, hvor man kan konstituere sig selv som modpol til en *out-group* blot ved at kritisere denne gruppe (Butler, 1997; Burke, 1969). Hvis man siger, at en bestemt fremmed gruppe er på en bestemt måde, siger man samtidig, at ens egen gruppe *ikke* er på den måde. Noget lignende foregår med hadbeskeder som den i eksempel (3), hvor modtager nedgøres baggrund af en række forskellige forhold. Fordi beskrivelsen af modtageren er så negativ, bliver afsenderen konstrueret som en modsætning hertil, helt uden at skrive et eneste ord om sig selv.

Udover at negative evalueringer kan sætte afsenderen i bedre lys, kan de også positionere afsenderen som særligt magtfuld. Dette er særligt tydeligt med evalueringer som i eksempel (4), der med det emotive *dog* og eksklamativ-sætningen med *hvor* bærer præg af at være en reel bedømmelse snarere end en nøgtern beskrivelse. En sådan evaluering forudsætter altså overlegenhed hos afsenderen (McDonald, 2022, s. 215), som fremstilles som om denne har autoritet nok til at kunne gøre modtageren til objekt for kritisk bedømmelse. Det er ikke til at sige, om modtageren accepterer denne autoritet, men afsenderen bliver i hvert fald i teksten fremstillet som en, der er i position til at kunne evaluere og bedømme modtageren.

- (4) Klamme, dumme, selvfede, falsk slut! / Eeeeeejjj, hvor du dog ødelægger alt med dine ligegyldige ytringer! [...]

5.3 Påbud

I den næste hadbesked, jeg vil fremhæve (eksempel 5), finder vi ligeledes en vokativ *din X*-konstruktion med et kønnet skældsord (*so*) og modificerende adjektiver (*fede klamme*), og vi har derfor igen at gøre med en både nedgørende og ekspressiv konstruktion.

- (5) din fede klamme so..gå på slanke kur og hold din fede kæft i.

Beskeden indeholder derudover to imperativkonstruktioner (*gå på slanke kur* og *hold din fede kæft i*). Disse imperativkonstruktioner gør flere forskellige ting. For det første præsupponerer de, at den beskrevne adfærd rent faktisk er en realitet – for overhovedet at forstå ytringen må man acceptere, at modtageren er tyk og udtaler sig uretmæssigt. Derudover påbyder de modtageren at ændre sin adfærd, hvilket understreger, at adfærden er negativ. På den måde bliver dette påbud om adfærdsændring en indiskutabel kritik af modtageren.

I eksempel (6) bliver der ligeledes præsupponeret bestemte negative karakteristika ved modtageren, fordi hun påbydes at ændre på en konkret del af sit udseende. Her er påbuddet mere specifikt end i forrige eksempel, idet der bliver fremstillet en helt konkret handling, modtageren burde foretage sig. Det efterfølgende *på for tak*¹ får derudover påbuddet til at fremstå som en reel befaling.

- (6) Har du tid til at chatte? /// FÅR NU FJERNET DEN MEGA STORE [KENDETEGN VED ANSIGTET], på for tak.

Brugen af imperativ – og især imperativer der tager form af ordrer eller påbud – kan være med til at fremstille afsenderen som havende en særlig høj magtposition. Bestemte illokutionære handlinger kan præsupponere bestemte magtpositioner eller sociale relationer (McDonald, 2022:215), og ved at beordre modtageren til at ændre sin adfærd fremstilles afsenderen netop som én, der har den fornødne magt til at kunne dette. Igen er det ikke sikkert, at modtageren accepterer dette, men afsenderen bliver i hvert fald i teksten etableret som magtfuld.

5.4 Moralisering

Den næste hadbesked (eksempel 7) indeholder både evaluering og påbud, der fungerer på samme måde som i de forrige to eksempler. Modtageren bliver kritiseret direkte med *du er total*

¹ Der menes sandsynligvis *på forhånd tak*

overgearet jo; og indirekte fordi hun påbydes at [*skrue*] *lidt ned for charmen*. Som med de forrige beskeder (eksemplerne 3-6) er dette med til at stille afsenderen i et bedre lys og fremstille denne som havende en bestemt autoritet eller magt.

(7) Skru lidt ned for charmen dame. / Du er total overgearet jo

Ved at bruge partiklen *jo* bliver det det gjort klart, at observationen om modtager burde være åbenlys, og sammenlagt med f.eks. vokativen *dame* får beskeden karakter af at være belærende og nedladende. På den måde rummer beskeden et element af moralisering, altså at afsenderen belærer modtageren om, hvad der er ”rigtigt og forkert, godt eller dårligt” (Den Danske Ordbog). Helt konkret bliver der i beskeden udpeget et problem, som modtageren ikke selv har set (at hun er overgearet), og samtidig bliver der fremsat en løsning (at skrue ned for charmen), dvs. afsenderen er bedrevidende. Denne kombination af kritik og bedrevidenhed er med til at skabe et særligt magtforhold mellem samtaleparterne. Kritikken i sig selv sætter modtageren i en underlegen position, og det moraliserende element fremstiller afsenderen som en, der er i stand til at belære og som derfor er overlegen eller har en vis autoritet.

Et andet eksempel på moraliserende sprogbrug er eksempel (8), hvor modtageren kritiseres for at have brugt begrebet *worst case scenario*:

(8) Hvorfor snakker du ikke dansk så alle kan forstå hvad du siger / Tror du alle ved hvad worst case scenario betyder

Her præsupponerer spørgsmålet *Hvorfor snakker du ikke dansk så alle kan forstå hvad du siger*, at modtageren netop ikke har snakket dansk, og at dette er skyld i, at hun ikke kan forstås. Kritikpunktet præsenteres ikke som ny information, men modtagerens adfærd og de negative konsekvenser heraf er en forudsætning for overhovedet at forstå spørgsmålet. Ved at bruge *hvorfor* som spørgeord spørges der ind til årsagen eller motivet for modtagerens adfærd, hvilket forudsætter, at modtageren bevidst har valgt en forkert adfærd. På den måde bliver modtager kritiseret for at være bevidst om sin fejlagtige adfærd uden at ændre på den til andres fordel, og spørgsmålet bliver derfor nærmest en kritik af modtagerens moral. Dette giver også beskeden karakter af at være moraliserende, og afsenderen fremstilles som en der ved, hvad der er rigtigt og forkert, og som har ret til at belære modtager om det.

5.5 Opsange

Et sidste element i hadbeskederne, der fremstiller afsenderen som autoritær eller magtfuld, er opsange, hvor modtageren får skældud for sine handlinger. Eksempel (9) er en besked sendt i forbindelse med en medieoptræden, hvor modtageren har brugt ordet *proxy*.

- (9) DER ER FANME IKKE NOGET DER HEDDER 'PROXY' PÅ DANSK DIN LORTEKÆLLING!DET HEDDER S T E D F O R T R Æ D D E R!

Denne besked rummer de fleste af de elementer, der er gennemgået tidligere: Modtagerens adfærd bliver evalueret og kritiseret; der er et indirekte påbud om adfærdsændring; og modtageren bliver belært om et problem og får foreskrevet en løsning. Alt dette er med til at fremstille afsenderen som overlegen. Det, der giver beskeden præg af at være en opsang er, at den rent sprogligt er enormt ekspressiv, idet den er skrevet i versaler, der bruges bandeord som *FANME*, og der er en vokativ *din X*-konstruktion med skældsordet *LORTEKÆLLING*. Modtageren bliver ikke blot kritiseret for sin adfærd, hun bliver skældt ud. Aggressiv sproglig adfærd tolereres normalt kun, hvis der på forhånd findes et magthierarki mellem samtaledeltagerne (Culpeper, 2011:200). Ligesom med påbuddene, der præsupponerer en vis autoritet hos afsenderen, er opsange derfor også med til at fremstille afsenderen som havende en vis autoritet eller overlegenhed.

Udover at efterligne en bestemt magtstruktur kan opsange også fungere som en slags straf til modtageren. Fordi kritikken af modtageren leveres i en så aggressiv eller affektiv tone, bliver det kommunikeret, at modtageren selv er skyld i denne aggression (Culpeper, 2011:221). Modtagerens adfærd har således foranlediget en vrede eller frustration hos afsenderen, og opsangen bliver en måde at gøre gengæld eller straffe modtageren: Modtageren bliver sat på plads.

6. Diskussion: Hadbeskeder som magtmanipulation

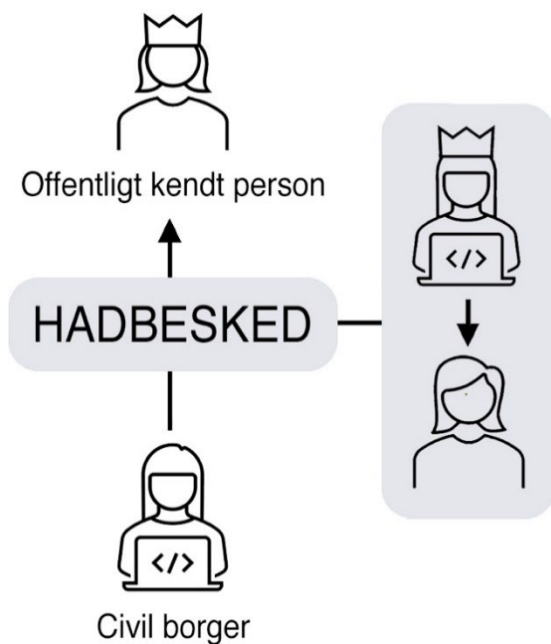
Min indledende observation i undersøgelsen af hadbeskederne var, at deres primære indhold var nedgørelse og kritik. Analysen af evalueringer, påbud, moraliseringer og reprimander afslører dog, at afsenderens selv fremstilling og -positionering også er central for hadbeskeder. De forskellige måder at kritisere og nedgøre modtageren bliver således en måde at hæve sig selv over modtageren på.

For at forstå baggrunden for denne proces må vi vende os mod hadbeskedernes kontekst. Indtil videre har vi behandlet deres sprog og indhold, men det er også nødvendigt at inddrage afsenderen og modtagerens sociale relation i vores analyse. Som beskrevet i afsnit 2, er had-

beskederes modtagere altid offentligt kendte personer som politikere, journalister, debattører, influencere, aktivister, TV- og radioværter, deltagere i TV-show, forfattere, forskere og så videre. Alle disse har på sin vis en politisk, social, kulturel eller økonomisk magt, som afsenderen – der oftest er en civil borger – ikke har. I Faircloughs (1989, s. 43) termer kan man sige, at modtageren har *magt bag diskursen* (*power behind discourse*), altså en mere eller mindre institutionel magtposition, der findes forud for interaktionen.

Modstykket til magt bag diskursen er *magt i diskursen* (*power in discourse*) (Fairclough, 1989:43), hvor man gennem sproget udøver magt og dermed opnår nogle (omend kortvarige) magtforhold samtaledeltagerne imellem. Magt i diskursen kræver derfor heller ikke en førsproglig magtposition. Det, vi ser i hadbeskederne, er således en lokal, sproglig magtudøvelse, hvor afsenderen ved at evaluere, påbyde, moralisere og skælde ud nedværdiger modtageren og indtager en højere magtposition selv.

Forud for hadbeskedens tilblivelse findes der således et hierarkisk magtforhold, hvor modtageren befinder sig i toppen. I selve hadbeskeden foregår der dog en lokal, sproglig magtudøvelse, og magtforholdet vendes om, så afsenderen kommer øverst, som illustreret i figur 2.



Figur 2: Modtager- og afsenderforhold. (Kronen indikerer magt).

Her kunne man fristes til at udlede, at hadbeskederes formål er at kritisere og nedværdige modtageren på en måde, der giver afsenderen en magt, som de mangler uden for interaktionen, altså at hadbeskeder er en reaktion på et ulige magtforhold og et forsøg på at omvende dette. Deres ofte ekspressive udtryk, som vi især så i eksemplet med reprimanden (eksempel 4) vidner om stor frustration eller afmagt, og hadbeskederne kan være en måde hvorpå denne afmagt vendes

til magt. Her skal vi dog passe på, at vi ikke bevæger os ud i at konkludere noget om afsenderens intentioner eller hensigter – dels fordi disse kategorier oftest er analytiske efterrationaliseringer (Culpeper, 2011, s. 49; Kissling, 1991:445), dels fordi dette studie blot er en sproglig analyse af 45 hadbeskeder, og ikke eksempelvis en stor sociologisk eller psykologisk undersøgelse af afsenderens bevæggrunde og følelsesliv. Hvad vi med større sikkerhed kan konkludere, er dog, at der foregår en lokal, sproglig magtudøvelse, hvor kritikken og nedværdigelsen af modtageren bliver grundlaget for, at afsenderen fremstilles som magtfuld. Med andre ord: Hadbeskederne handler lige så meget om afsenderen som om modtageren.

7. Afsluttende bemærkninger

Lad os afslutningsvist vende tilbage til det spørgsmål, der bliver stillet i artiklens titel: Hvem handler hadbeskederne om? Det lette svar er, at hadbeskeder handler om modtageren. Alle de 45 beskeder, jeg har behandlet i denne undersøgelse, indeholder en eller anden form for kritik af modtageren, og derudover er der en enormt høj frekvens af andenpersonspronomen, hvilket indikerer, at netop modtageren (*du*'et) er vigtig for genren. En dybere analyse af hadbeskederens sproglige karakteristika afslører dog en del elementer, der også fremstiller afsenderen på forskellige måder – særligt elementer, der får afsenderen til at virke overlegen eller magtfuld. På den måde kan man sige, at hadbeskeder også handler om afsenderen.

I forlængelse af denne konklusion kan man desuden spørge: *Hvordan* handler hadbeskederne om afsenderen? I artiklen har jeg beskrevet, hvordan hadbeskederens evalueringer, påbud, moraliseringer og reprimander alle er sproghandlinger, der præsupponerer bestemte sociale positioner. Særligt præsupponerer de, at afsenderen har en vis autoritet eller magt, og ved at bruge disse sproghandlinger bliver afsenderen derfor fremstillet som magtfuld. Fordi afsenderen ikke nødvendigvis er magtfuld forud for interaktionen, bliver hadbeskederne en måde at manipulere samtaledeletagernes magtpositioner.

En undersøgelse af hadbeskederens sproglige karakteristika kan på denne måde bidrage til en specifik genreviden og samtidig udvide vores forståelse for, hvordan man generelt kan bruge sproget til at udøve magt.

Litteratur

Adolphs, Svenja (2008) *Corpus and Context: Investigating pragmatic functions in spoken discourse*. John Benjamins Publishing Company.

Analyse & Tal (2021) *Angreb i den offentlige debat på Facebook*. Analyse & Tal. Sidst tilgået den 03.01.2023 på <https://strapi.ogtal.dk/uploads/966f1ebcfa9942d3aef338e9920611f4.pdf>

Hvem handler hadbeskederne om? Magtmanipulerende sprog i danske hadbeskeder

- Assimakopoulos, Stavros, Fabienne H. Baider, og Sharon Millar (2017) *Online Hate Speech in the European Union*. Cham: Springer International Publishing.
- Astrup, Elisabeth (2021) »Det er slemt det her«: Hadbeskeder til X-factordeletagere udløser hård kritik af tonen på nettet. *Politiken*. Sidst tilgået den 03.01:2023 på <https://politiken.dk/kultur/medier/art8135921/Hadbeskeder-til-X-factordeletagere-udløser-hård-kritik-af-tonen-på-nettet>.
- Baumgarten, Nicole, Eckhard Bick, Klaus Geyer, Ditte Aakær Iversen, Andrea Kleene, Anna Vibeke Lindø, Jana Neitsch, Oliver Niebuhr, og Esben Nedenskov Petersen (2019) Towards Balance and Boundaries in Public Discourse: Expressing and Perceiving Online Hate Speech (XPEROHS). I *International Journal of Language and Communication* 50:87–108.
- Beebe, Leslie M (1995) Polite fictions: Instrumental rudeness as pragmatic competence. I J. E. Alatis, C. A. Straehle, B. Gallenberger, og M. Ronkin (red.), *Linguistics and the Education of Language Teachers: Ethnolinguistic, Psycholinguistic and Sociolinguistic Aspects. Georgetown University Round Table on Languages and Linguistics*, 154–68. Georgetown: Georgetown University Press.
- Bojsen-Møller, Marie, Sune Auken, Amy J. Devitt, og Tanya Karoli Christensen (2020) Illicit Genres: The Case of Threatening Communications. I *Sakprosa* 12(1):1-53.
- Bruun-Hansen, Kerstin (2022) Voldtægtstrusler, stalkere og had: P3-værter åbner deres indbakke. *Journalisten*. Sidst tilgået den 03.01:2023 på <https://journalisten.dk/voldtaegtstrusler-stalkere-og-had-p3-vaerter-opfordrer-til-at-sloejfe-kommentarspor/>
- Burke, Kenneth (1969) *Rhetoric of motives*. University of California Press.
- Butler, Judith (1997) *Excitable Speech: A politics of the performative*. Routledge.
- Cloud, Dana L. (2009) Foiling the Intellectuals: Gender, Identity Framing, and the Rhetoric of the Kill in Conservative Hate Mail. I *Communication, Culture & Critique*, 2(4):457–479.
- Culpeper, Jonathan (2011) *Impoliteness: Using Language to Cause Offence*. Cambridge University Press.
- d’Avis, Franz og Jörg Meibauer (2013) Du Idiot! Din idiot! Pseudo-vocative constructions and insults in German (and Swedish). I B. Sonnenhauser og P. Noel Aziz Hanna (red.): *Vocative! Addressing between System and Performance*. Berlin/Boston: De Gruyter Mouton.
- Dansk Journalistforbund (2020) *Digital chikane*. Sidst tilgået den 03.01:2023 på <https://journalistforbundet.dk/sites/default/files/inline-files/Digital%20Chikane.pdf>.
- Den Danske Ordbog* = <https://ordnet.dk/ddo>. Sidst tilgået den 03.01:2023.
- DR2 (2014) *Ti stille, kvinde*. Danmarks Radio. Sidst tilgået den 03.01:2023 på https://www.dr.dk/drtv/serie/ti-stille-kvinde_221276.
- DR2 (2021) *Vi kommer og dræber dig*. Danmarks Radio. Sidst tilgået den 03.01:2023 på https://www.dr.dk/drtv/serie/vi-kommer-og-draeber-dig_256850.
- Fairclough, Norman (1989) *Language and power*. London: Longman.
- Geyer, Klaus (2018) Hadetalens ’grammatik’. I *Ny forskning i grammatik*, 25:16-32.
- Ging, Debbie og Siapera, Eugenia (2019) *Gender hate online: understanding the new anti-feminism*. Londres: Palgrave Macmillan.
- Hagen, Aina Landsverk (2015) *Meningers mot—Netthat og ytringsfrihet i Norge*. Oslo: Cappelen Damm.
- Harris, Sandra (1984) The form and function of threats in court. *Language & Communication* 4(4):247–71.

- Havryliv, Oksana (2009) *Verbale Aggression. Formen und Funktionen am Beispiel des Wienerischen*. Frankfurt/Main: Peter Lang.
- Herskind, Ida (2022) Anne Kirstine Cramon trækker sig på grund af den hårde tone. *Politiken*. Sidst tilgået den 03.01:2023 på <https://politiken.dk/kultur/medier/art8619127/Anne-Kirstine-Cramon-trækker-sig-på-grund-af-den-hårde-tone>.
- Julien, Marit (2016) Possessive predicational vocatives in Scandinavian. I *The Journal of Comparative Germanic Linguistics*, 19(2):75–108.
- Kissling, Elizabeth Arveda (1991) Street Harassment: The Language of Sexual Terrorism. I *Discourse & Society*, 2(4):451–460.
- Kvist, Gitte og Eric Øtting. (2015). Hård tone på Facebook får politikere til at opgive debatten. *DR Østjylland*. Sidst tilgået den 03.01:2023 <https://www.dr.dk/nyheder/regionale/oestjylland/haard-tone-paa-facebook-faar-politikere-til-opgive-debatten>.
- Kaasgaard, Stinne (2020) Jeg var ved at græde ved tanken om, at jeg skulle være i Danmark bare to sekunder mere. Sidst tilgået den 03.01:2023 *Eurowoman*. <https://www.alt.dk/artikler/emma-holten>.
- Lauritsen, Michael, Freja Toftegaard, og Anne Phillipsen (2016) Du fortjener at dø: Mindst 36 folketingspolitikere har fået trusler på livet. *DR Syd- og Sønderjylland*. Sidst tilgået den 03.01:2023 på <https://www.dr.dk/nyheder/regionale/syd/du-fortjener-doe-mindst-36-folketingspolitikere-har-faaet-trusler-paa-livet>.
- Limberg, Holger (2009) Impoliteness and threat responses. *Journal of Pragmatics* 41:1376–94.
- McCaffrey, Dawn, og Jennifer Keys (2000) Competitive framing processes in the abortion debate: Polarization-vilification, frame saving, and frame debunking. I *Sociological Quarterly*, 41:41–62.
- McDonald, Lucy (2022) Cat-Calls, Compliments and Coercion. I *Pacific Philosophical Quarterly*, 103(1):208-30.
- Meibauer, Jörg (2016) Slurring as Insulting. I Rita Finkbeiner, Jörg Meibauer, og Heike Wiese (red.), *Linguistik Aktuell/Linguistics Today*, 143–166. Amsterdam: John Benjamins Publishing Company.
- Muschalik, Julia (2018) *Threatening in English: A Mixed Method Approach*. Amsterdam: John Benjamins Publishing Company.
- Piñero-Otero, Teresa og Martínez-Rolán, Xabier (2021) Say it to my face: Analysing hate speech against women on Twitter. *Profesional de la información*, 30:5, 1-16.
- Rabøl, Laura Byager (2021) Du er grim, Skrid hjem, Du synger falsk, MGP var i lørdags. *TV2*. Sidst tilgået den 03.01:2023 på <https://underholdning.tv2.dk/2021-03-12-du-er-grim-skrid-hjem-du-synger-falsk-mgp-var-i-loerdags>.
- Schneider, Klaus P. (2010) Variational Pragmatics. I Miriam Fried, Jan-Ola Östman, og Jeff Verschueren (red.), *Variation and Change: Pragmatic Perspectives*, 239–67. Amsterdam: John Benjamins Publishing Company.
- Schneider, Klaus P. og Anne Barron (2008) Where pragmatics and dialectology meet: Introducing variational pragmatics. I Klaus P Schneider og Anne Barron (red.), *Variational pragmatics: A focus on regional varieties in pluricentric language*. Amsterdam: John Benjamins Publishing Company:1-32.
- Steffensen, Tinne og Lumi Zuleta (2022) *Chikane og trusler i lokalpolitik*. København: Institut for Menneskerettigheder.
- Yule, George (1996) *Pragmatics*. Oxford: Oxford University Press.

Tone-stød og tonal accent i nyere fynsk dialekt

Kirstine Boas

Aarhus Universitet

Indledning

I denne artikel præsenterer jeg en undersøgelse af hyppigheden af tone-stød og tonale accenter blandt en gruppe ældre informanter i to landsbyer på Østfyn. Empirien indsamlede Silke Julie Flodin og jeg i 2020 i forbindelse med vores bachelorprojekt (Boas og Flodin 2021). I denne artikel bringer jeg en ny og opdateret bearbejdning af dette. Stød og accent i fynsk har ikke været undersøgt længe (Andersen 1958), og fraværet af dette glimrer blandt andet i det nye store opslagsværk, *Dansk Sproghistorie*, hvor blot jysk, sjællandsk og vestjysk stød behandles i kapitlet om udtale i yngre dansk dialekt. Tonal accent i dansk er heller ikke omtalt her (Brink 2018:239 ff.). Undersøgelsen bidrager med en beskrivelse af tone-stød og tonal accent, og dens forskningsspørgsmål lyder som følger: Hvordan kan vi fonetisk beskrive de fænomener Poul Andersen for 65 år siden kaldte henholdsvis tone-stød og tonal accent? Hvordan påvirker disse tonale fænomener lyden af fynsk? Og i hvilken grad er de ældre fynboer bevidste om brugen af disse dialektale træk?

Tone-stød og tonal accent på Østfyn

I det gamle Vindinge Herred på Østfyn, hvor både Langå og Ørbæk ligger i dag, fandt Poul Andersen både tone-stød og tonale accenter på ord hvor der i rigsmålet ville være stød (Andersen 1958:34). Det beskrev han i *Fonemsystemet i Østfynsk* hvor han bl.a. undersøger stød og tonegang i en- og tostavelsesord på Fyn. Andersens beskrivelser af de tonale accenter er mere grundige og eksemplificerede sammenlignet med hans beskrivelser af tone-stød, men der er dog målbare forskelle mellem de to fænomener.

På Østfyn fandt han at stød som fænomen er almindeligt forekommende, men at det som noget særligt for dette område var karakteriseret ved fonetisk at være så svagt, at udefrakommende ofte ville erklære at det slet ikke var realiseret. Tone-stødet kan dermed have en ofte meget svag lukkefase på vokalen såfremt denne overhovedet kan identificeres. Lukkefasen, den knirkende lyd, ledsages af en hastigt stigende tonebevægelse som falder igen. En kraftigt

stigende og derpå faldende toneforløb er dermed også karakteristisk for tone-stødet (Andersen 1958:34 f.). Jeg vil bruge IPA-tegnet for stød, ʔ, til at angive tone-stødene.

Poul Andersen beskriver den tonale accent på Fyn som en stigende eller højt intoneret tonal accent. Han noterer denne med et ´. Jeg vil også transskribere den stigende accent med et ´ foran den pågældende stavelse¹. Den tonale accent har ikke stød, og derfor hverken lukkefase eller knirk, og tonen falder ikke igen, sådan som Andersen beskrev det ved tone-stødet (Andersen 1958:34 ff.). Den tonale accent forekommer både på Øst- og Sydfyn hvorimod tone-stødet kun er blevet identificeret i det gamle Vindinge Herred på Østfyn. Den tonale accent forekommer fortrinsvis i enstavelsesord med høj vokal (eller vokalen o), fx *bi*, *spis* og *sol*. Traditionelt kunne dette forekomme på hele Fyn, men for den østfynske egn gjaldt det dog også at lavere vokaler havde tonale accenter (Andersen 1958:36), fx *jævn*, *løgn* og *dreng*. Ordene der indgår i denne undersøgelse, har derfor både høje, halvhøje, halvlave og lave vokaler. Det kan således efterprøves om der stadig er tonal accent på alle vokalhøjder i dette område. Det er desuden karakteristisk for den østfynske tonale accent at vokalen og de efterfølgende konsonanter lyder korte, og det korte udtryk på vokalen styrkes yderligere hvis der ikke er en efterfølgende konsonant (Andersen 1958:36).

Der indgår kun enstavelsesord i undersøgelsen fordi Poul Andersen fandt tone-stød og tonale accenter på disse (Andersen 1958). Ord med en markant stigende tone, en eventuel lukkefase og et derefter faldende toneforløb vil blive rubriceret som tonestød. Ord med en stejlt stigende tone (der ikke falder igen), eller en højt intoneret tone, vil blive rubriceret som tonale accenter. Ord med ikke-stød er grundet pladsmangel i denne omgang ikke talt med selvom der er en betydelig mængde af disse i vores data.

Informanter og undersøgelsens design

Informanterne i undersøgelsen er fire pensionister mellem 70 og 80 år, henholdsvis et ægtepar fra Langå, som ligger syd for stødgrænsen, og et ægtepar fra Ørbæk, nord for stødgrænsen. De to pensionister fra Ørbæk har hele livet igennem boet inden for en 10 kilometers omkreds, og de bor desuden på en gård der havde været mandens families eje i mange generationer. Kvinden i ægteparret fra Langå har ligeledes altid boet på Østfyn, og manden har, med undtagelse af et års værnepligt i København, også altid boet på Østfyn. Informanterne var dermed gode til netop at teste det tætteste vi kommer på ældre, originalt østfynsk i dag. Vi lavede semistrukturerede interviews der forgik i informanternes eget hjem. De bestod i at de fire informanter skulle

¹ For overblikkets skyld noterer jeg derfor ikke enhedstryk.

identificere objekter på to billeder og læse en medbragt tekst op (Figur 1). De skulle læse den op to gange: De blev eksplicit bedt om at læse den op til deres partner så naturligt som muligt (med forventningen om at de her ville være mere dialektale). Efter den første læsning blev de bedt om at gentage oplæsningen, men denne gang skulle de få det til at lyde så meget som muligt som sproget i TV-avisen (med forventningen om et intenderet standarddansk). Til dette svarede tre ud af fire informanter ”altså som rigsdansk?” hvilket indirekte indikerede en bevidsthed om det modsatte, dvs. dialekt. Det vil jeg komme tilbage til i afsnittet *Fynsk med vilje*. Ordene på både billederne og i teksten er ord hentet fra Poul Andersens *Fonemsystemet i Østfynsk*, herunder især fra hans dialektkort. Det er dermed ord vi forventede der kunne forekomme tone-stød eller tonal accent i (Andersen 1958). Med undtagelse af *korn* og *land* er alle de fremhævede ord i Figur 1 karakteriseret ved en stigende tonal accent. Da det hos Andersen er uklart om nogle af disse også kan have tone-stød, og om *korn* og *land* har stød som i rigsmålet, eller tone-stød, sigtes der i denne artikel mod at undersøge om stødet realiseres som i rigsmålet, med en stigende-faldende tone (tone-stød) eller med stigende eller højt intoneret tone (tonal accent). Identificeringen af objekter på billederne slog desværre fejl da informanterne generelt viste sig at have større ekspertviden end interviewereren, og derfor ikke så en ’ko’ eller et ’skib’, men en ’jerseykvie fra Thisted’ og ’en tremasters som dem der ligger til ved Marstal’. Oplæsningerne var til gengæld succesfulde, og da alt blev optaget, kunne al lyd på testordene analyseres og behandles i talebehandlingsprogrammet *Praat* (Boersma & Weenink 2023) efterfølgende. Spektrogrammerne lavet i Praat understøttede desuden hvad der kunne høres på lydklippene.

På en **vej** mod Kerteminde gik en **ung dreng**. Med sin **kniv**, som i øvrigt var spids, skar han en **jævn** bid af et æble, og sagde til sin sultne kæreste: **spis** det! Du mangler **sul** på kroppen. De var gode til at passe på hinanden.

Hun så en **bi** sidde på en **gul** blomst, og gik lidt nærmere. Bag blomsten sad en lille **fugl**, der spiste bær. Da hun rejste sig, kunne hun se både **korn**, **skov** og **sol** over **Fyn** og et flag i rødt og hvidt, der vejrede. Hun sagde til sig selv: hvor er det et dejligt **land**. Kæresteparret begyndte at gå hjemad, for de havde en dej til at stå at hæve. Da de nåede deres hus, lå der en pakke foran døren. Hun spurgte: Hvem skal bære den ind? Dig eller mig? Den tjans tog han, og sagde at det ikke ville være hårdt for ham. I virkeligheden var pakken tung, så det var en **løgn**. I pakken lå, så vidt han vidste, hans fødselsdagsgave til hende: en lille **hund**. Hun blev lykkelig for gaven, og de fejrede den med at spise nybagte boller.

Figur 1: Oplæsningsteksten. De fremhævede ord har Poul Andersen noteret tone-stød eller tonal accent ved.

Findes der stadig tone-stød og tonal accent?

Jeg vil gennemgå nogle af de markerede ord fra Figur 1 for at få et indtryk af om der stadig bliver brugt tonale accenter og tone-stød på Østfyn, og hvilke typer enstavelsesord disse synes at falde på, i hvilken grad informanterne kodeskifter, og hvilken effekt de tonale stød har på lyden af (øst)fynsk. Jeg har struktureret gennemgangen ved at starte med de høje vokaler og ende ved de lave. Jeg vil undervejs bringe eksempler på tonegange, stigninger i Hz, fonetiske transskriptioner i tilpasset IPA, samt hvilke fonemer de tonale accenter eller tone-stødene falder på. Fordelingen af standarddansk stød (blot nævnt som *stød*), ikke-stød, tone-stød og tonale accenter kan ses i Tabel 1 og 2. I disse indgår ordet *skov* ikke hos manden fra Ørbæk da han desværre konsekvent udtalte det som *skovl*.

	Mand fra Langå	Kvinde fra Langå	Mand fra Ørbæk	Kvinde fra Ørbæk
Vej	Ikke-stød	Tonal accent	Stød	Tonal accent
Ung	Tonal accent	Ikke-stød	Ikke-stød	Tonal accent
Dreng	Tone-stød	Ikke-stød	Stød	Tone-stød
Kniv	Tone-stød	Stød	Tone-stød	Tonal accent
Jævn	Tone-stød	Tone-stød	Tone-stød	Tonal accent
Spis	Tonal accent	Stød	Tone-stød	Tonal accent
Sul	Tone-stød	Tone-stød	Tonal accent	Tone-stød
Bi	Tonal accent	Ikke-stød	Stød	Tonal accent
Gul	Tonal accent	Ikke-stød	Ikke-stød	Tone-stød
Fugl	Tone-stød	Tone-stød	Tonal accent	Tone-stød
Korn	Tone-stød	Tone-stød	Tonal accent	Tone-stød
Skov	Ikke-stød	Stød	----	Stød
Sol	Tone-stød	Tone-stød	Tone-stød	Tone-stød
Fyn	Ikke-stød	Tone-stød	Tonal accent	Tonal accent
Land	Stød	Stød	Stød	Ikke-stød
Løgn	Stød	Tone-stød	Stød	Tonal accent
Hund	Stød	Stød	Tone-stød	Tone-stød

Tabel 1: Realisationer af tone-stød og tonale accenter i dialektoplæsningerne.

	Mand fra Langå	Kvinde fra Langå	Mand fra Ørbæk	Kvinde fra Ørbæk
Vej	Tonal accent	Stød	Tonal accent	Ikke-stød
Ung	Ikke-stød	Tonal accent	Tone-stød	Tone-stød
Dreng	Stød	Stød	Stød	Ikke-stød
Kniv	Ikke-stød	Stød	Stød	Tonal accent
Jævn	Stød	Stød	Stød	Tonal accent
Spis	Tonal accent	Stød	Stød	Tone-stød
Sul	Stød	Tone-stød	Stød	Tone-stød
Bi	Ikke-stød	Stød	Stød	Ikke-stød
Gul	Stød	Ikke-stød	Ikke-stød	Ikke-stød
Fugl	Stød	Stød	Stød	Ikke-stød
Korn	Stød	Tone-stød	Stød	Tonal accent
Skov	Stød	Tone-stød	---	Stød
Sol	Stød	Stød	Stød	Ikke-stød
Fyn	Stød	Tone-stød	Tonal accent	Tonal accent
Land	Stød	Stød	Stød	Stød
Løgn	Stød	Stød	Stød	Stød
Hund	Tonal accent	Tone-stød	Stød	Ikke-stød

Tabel 2: Realisationer af tone-stød og tonale accenter i de intenderet standarddanske oplæsninger.

Ordet *bi* er ifølge Andersen et af de ord tonal accent fortrinsvis forekommer på. Det har en høj vokal. Kvinden fra Ørbæk har tonal accent på ordet som kommer til udtryk ved en stejlt stigende tone med lyden [ˈbi]. Vokalen lyder kort, hvilket også stemmer overens med Andersens observationer af ord med tonale accenter (Andersen 1958:36). Manden fra Langå har en højt intoneret udtale, og dermed også en tonal accent. Begge udtaler er fra informanternes dialektoplæsninger.

Ordet *spis* indgår ikke som en del af Andersens dialektkort, men han skriver alligevel at det er et af de imperativer med en høj vokal, der også ofte er tonal accent på (1958:36 f.). Kvinden fra Ørbæk og manden fra Langå har tonale accenter i dialektoplæsningen, som begge er højt intonerede udtaler med korte afviklinger af vokalen. Disse kan fonetisk beskrives som [ˈsbis] med kort vokal. Kvinden fra Ørbæk har i sin intenderede standarddanske oplæsning et tone-stød med et stigende-faldende toneforløb. Der er intet identificerbart knirk, men vokalen lyder betydeligt længere end hendes tonale accent [sbi:ʔs].

Ordet *kniv* har en høj og kort vokal. Kvinden fra Ørbæk har en tonal accent her grundet en markant tonal stigning på [i] i sin dialektoplæsning. I sin intenderede standarddanske oplæs-

ning stiger tonen i *kniv* til 280 Hz i en sætning hvor den gennemsnitlige grundtonefrekvens er 250 Hz. Toneforløbet falder igen. Dette giver et højt intoneret stigende-faldende tone-stød: [kni[?]w]. Vokallængden er ikke entydigt lang eller kort her. Mændene fra både Langå og Ørbæk har ligeledes tone-stød i deres dialektoplæsninger, som begge er karakteriseret ved et stigende-faldende toneforløb uden et knirk.

Ved ordet *sul* havde tre informanter tone-stød i deres dialektoplæsninger, og to informanter havde det i de standarddanske oplæsninger. Alle tone-stødene har et stigende-faldende toneforløb der toppede midt på vokalen, og her kan der ej heller identificeres et knirk. Jeg opfatter vokalen som lydende lang, især hos kvinden fra Ørbæk [su:[?]l]. Manden fra Ørbæk havde i sin dialektoplæsning en tonal accent som havde et stejlt stigende toneforløb på en kortlydende vokal.

Hos manden fra Ørbæk er ordet *fugl* højt intoneret da dets gennemsnitlige grundtonefrekvens er på 212 Hz (målt på den stemte sekvens). *Fugl* udtales uden et identificerbart knirk hos denne informant. Hans tone stiger på [u], og ordet når kort før midten af den kortlydende vokal sin maksimale tonehøjde på 223 Hz. Han udtaler dermed ordet: [ˈful]. Til sammenligning har han i sin standarddanske oplæsning en gennemsnitlig grundtonefrekvens på 167 Hz i den stemte del af ordet, og en maksimal frekvens på 176 Hz. Her har han et standarddansk stød. Der er to informanter, kvinden fra Ørbæk og manden fra Langå, som bruger den gamle fynske udtale i deres dialektoplæsninger, og denne pointe vil jeg komme tilbage til i afsnittet nedenfor om kodeskift. For nu skal det dog bemærkes at der er en markant stigende tone som falder igen hos disse to informanter, og derfor rubriceres disse også som tone-stød. Kvinden fra Langå i sin dialektoplæsning og manden fra Langå i sin intenderede standarddanske oplæsning har begge markant stigende toner som falder igen, og disse tæller derfor også som tone-stød.

Ordet *hund* bliver udtalt med tone-stød i de to dialektoplæsninger fra Ørbæk. De har en markant tonestigning på vokallyden, og desuden en markant nasalering af vokalen [ũ]. Manden fra Ørbæk laver en nasalvokal, så det lyder således: [hũ[?]]. Ordet *Fyn* bliver konsekvent udtalt med tonal accent af parret fra Ørbæk, og med et tone-stød af kvinden fra Langå. De tonale accenter har begge markante tonestigninger på [y], og vokalen afvikles hurtigt. Ved tone-stødet er toneforløbet stigende-faldende, og vokalen har et længere udtryk: [fy:[?]n].

Ordet *sol* har et halvhøjt vokalfonem. Det udtales med et tone-stød i alle dialektoplæsninger og med enten stød eller ikke-stød i alle standarddanske oplæsninger. Vokalerne i tone-stødene er lange, og de transskriberes derfor: [so:[?]l]. I dialektoplæsningen af *korn* har tre informanter tone-stød med et stigende-faldende toneforløb: [ko:[?]ɳ]. Ordet *jævn* har en halvhøj vokal [ɛ] og blev udtalt med tone-stød med stigende-faldende tone der toppe sidst på vokalen hos tre

informanter i deres dialektoplæsninger. Kvinden fra Ørbæk havde en tonal accent på ordet i sin dialektoplæsning, og det var karakteriseret ved et stigende toneforløb.

Ordet *ung* har en halvlav vokal [ɔ] og bliver af manden fra Langå og kvinden fra Ørbæk udtalt med en højt intoneret tonal accent i deres dialektoplæsninger. I sin standarddanske oplæsning har kvinden fra Langå også en tonal accent, men den er karakteriseret ved en kraftigt stigende tone. Parret fra Ørbæk har i deres standarddanske oplæsninger begge tone-stød med et stigende-faldende toneforløb. *Dreng* og *land* har korte og lave vokaler. *Dreng* har tone-stød hos manden fra Langå og kvinden fra Ørbæk, og de har et stigende-faldende toneforløb. *Land* har hverken tonal accent eller tone-stød i denne dataindsamling.

I det ovenstående har jeg gennemgået de fleste ord fra oplæsningsteksten (alle ord kan ses i Tabel 1 og 2) med forventningen om at der var mulighed for enten tone-stød eller tonal accent. Ordene blev læst op otte gange hver, og det giver i alt 134 forekomster hvis de to *skovl*-læsninger modregnes. I 27 af dem var der en tonal accent, og i 35 ord var der en stigende-faldende tone, hvorfor disse er blevet rubriceret som et tone-stød. Der var mod forventning ikke noget identificerbart knirk blandt tone-stødene, og ved nærmere eftersyn efter uregelmæssigheder i spektrogrammerne i Praat bekræftedes dette også. 27 tonale accenter ud af 134 mulige giver en andel på 20 %, og for tone-stødene var der 35 ud af 134 mulige en andel på 26 %. Disse tal viser at der stadig både er tonale accenter og tone-stød på Østfyn blandt de ældre borgere. Over halvdelen af ordene har høje vokaler, og den anden halvdel er fordelt på halvhøje og lave vokaler. De høje vokaler trækker generelt flest tone-stød og tonale accenter, men blandt de lave vokaler er der også eksempler på ord som fx 'vej' og 'skov' der også trækker mange af de dialektale træk. Generelt er der en tendens til at tone-stødet ligger på vokalerne i fynsk, også i de ord hvor stødet falder på konsonanter i standarddansk; dette stemmer overens med Andersens observationer (Andersen 1958:34 ff.). Blandt de tonale accenter lyder vokalerne til at få et komprimeret og intenst udtryk, hvilket ikke var overraskende. Vokalerne lyder til gengæld længere blandt tone-stødene, blandt hvilke jeg ikke kunne identificere nogle knirk (Andersen 1958:36). Deles observationerne op i henholdsvis de dialektale og de intenderede standarddanske oplæsninger, får vi informationer om informanternes evne til at kodeskifte, og dermed deres grad af bevidsthed om egen brug af dialekt, hvilket artiklen senere vil diskutere.

Fynsk med vilje

Da vi både indsamlede en oplæsning på et så naturligt sprog som interviewsituationen tillod og en oplæsning som viste hvor standarddansk informanterne kunne tale, er det muligt også at

vurdere de ældre informanternes evne til at kodeskifte (Woolard 2005). Blandt de tonale accenter er 10 ud af 67 mulige realiseret i den intenderede standarddanske oplæsning, hvilket udgør en andel på 15 %. I dialektoplæsningen er 17 ud af 67 mulige ord med tonale accenter, hvilket udgør en andel på 25 %. Blandt tone-stødene bliver ni af disse realiseret i den intenderede standarddanske oplæsning, hvilket svarer til 13 %. Betydeligt højere er andelen blandt dialektoplæsningerne hvor 26 ud af 67 mulige tone-stød realiseres; det svarer til en andel på 39 %. Jeg havde forventet at der var flere tone-stød i de intenderede standarddanske oplæsninger sammenlignet med dialektoplæsningerne fordi tone-stødene teoretisk har antydningen af den lukkefase der forekommer i det standarddanske stød. Denne viste sig dog ikke at indtræffe blandt disse informanternes tone-stød.

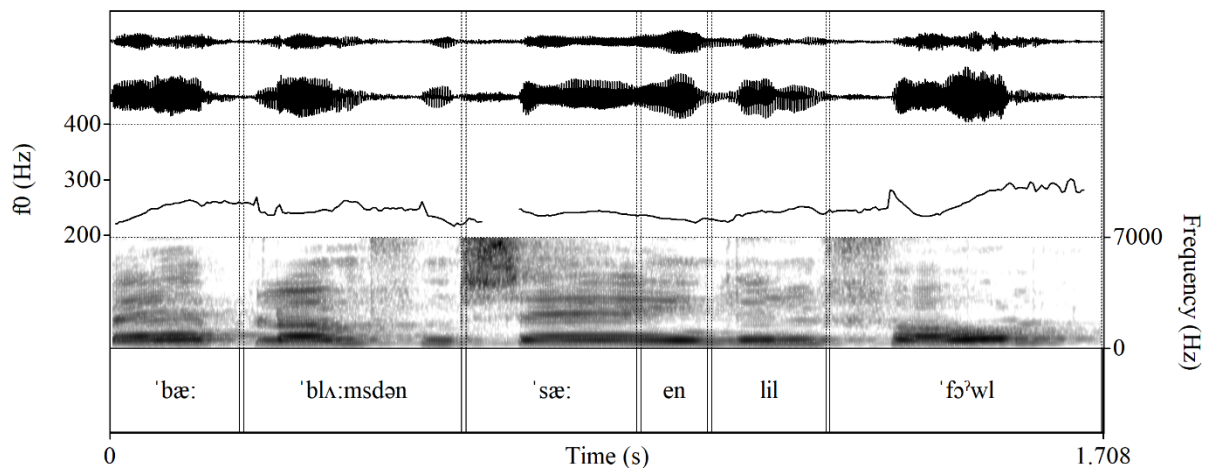
Tallene viser at de ældre østfynboere til en vis grad evner at kodeskifte og dermed er bevidste om deres egen dialekt; i dialektoplæsningerne har de flere forekomster af dialektale træk sammenlignet med de intenderede standarddanske. Deres intenderede standarddanske oplæsninger har dog stadig en betydelig mængde dialektale træk, og det peger på at meget af det tonale til en vis grad er resistent. Med det sagt er der dog ikke tale om et diskret skift mellem koder, men i høj grad en gradsforskel. Det skal bemærkes at den reelle andel af lokale og regionale dialektale træk formentlig er højere fordi ikke-stød-varianten ikke indgår her selvom den traditionelt forekommer på både Syd- og Østfyn (Andersen 1958). En anden indikation på evnen til bevidst at kodeskifte ses hos både kvinden fra Ørbæk og manden fra Langå som i deres dialektoplæsninger begge viser kendskab til den gammelfynske udtale [fɔ^ʔwl], som de udtaler med en lang diftong der svarer til den ældre udtale af ordet i området jf. Ømålsordbogen (2000). Den gammelfynske udtale bruger de ikke i de intenderede standarddanske oplæsninger. Det ser ud til at de ældre er meget bevidste om dialektale træk, og brugen af disse når de forsøger at tale 'rigsdansk'. Der kunne med fordel følges op med undersøgelser af om de her intenderede standarddanske oplæsninger høres som standarddanske eller fynske informanter. Diskussionen af det receptive aspekt ville være et væsentligt bidrag til spørgsmålet om kodeskiftet blandt fynboerne.

Den (øst)fynske klang

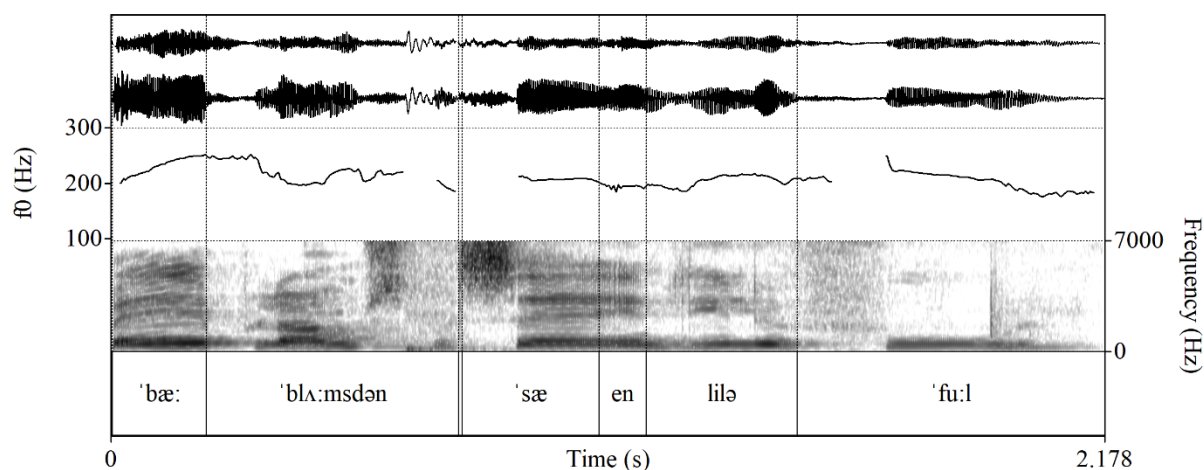
I Anders Wedel Berthelsens populærvidenskabelige bog, *Fynsk så det synger*, siger Asgerd Gudiksen om den syngende kvalitet i fynsk at denne “først og fremmest hænger sammen med den måde, hvorpå det fynske stød udtales” (Berthelsen 2017:41 f.). Stødene der i denne artikel er omtalt, er tonale uden knirk hvilket bidrager til et melodisk og mindre hakkende udtryk. Inge Lise Pedersen skriver om de stigende toner på fynsk:

Man skal ikke lade sig forlede af de karakteristiske stigende toner til at betragte fynsk som et tonesprog som svensk-norsk og østslesvigsk. Stigende tone på Fyn må snarest betragtes som en rent tonal realisation af stødaccenten. De stigende eller højtliggende toner (ofte med interval svarende til en lille tert) præger imidlertid sætningsintonationen som helhed, og bidrager til at gøre den mere livlig og syngende end mange andre steder (Pedersen 2002:46).

Poul Andersen skriver også at de tonale accenter er ”medvirkende til at give østfynsk Sætningsmelodi dens karakteristiske stigende, eller som det almindeligvis udtrykkes ”syngende” og ”spørgende, forbavsede” Præg” (Andersen 1958:37). Lad os se nærmere på de højtliggende toner og sætningsintonationen, og om disse hænger sammen med det syngende udtryk.



Figur 2: Kvinde fra Ørbæks dialektoplæsning: ”Bag blomsten sad en lille fugl”.



Figur 3: Kvinde fra Ørbæks intenderede standarddanske oplæsning: ”Bag blomsten sad en lille fugl”.

Grundtoneskalaen er i Figur 2 og 3 afbilledet logaritmisk for mere akkurat at gengive forholdet mellem frekvens (lyden målt i Hz) og pitch (hvor højt eller lavt vi opfatter lyd). Den logaritmiske grundtoneskala giver dermed gode forudsætninger for at analysere de akustiske aspekter af lydproduktion. Kvinden fra Ørbæk udtaler *fugl* med et tone-stød. Ordet har en gennemsnitlig grundtonefrekvens på 267 Hz og en maksimal frekvens på 296 Hz i sin dialektoplæsning, hvilket svarer til en stigning på 10,9 % (se Figur 2 og Tabel 3). Det er den højest intonerede sætning i dette data, og det er også en af de sætninger der har det mest melodiske udtryk. I sin intenderede standarddanske oplæsning udtaler hun *fugl* stødløst med en gennemsnitlig grundtonefrekvens på 201 Hz og en maksimal frekvens på 244 Hz. Stigningen er i denne oplæsning på 21,4 % (Figur 3 og Tabel 3). I Figur 2 ses det at de højtliggende toner synes at trække intonationen op, og den falder derfor ikke så stødt gennem ordet som den gør i den intenderede standarddanske oplæsning (hvor ’fugl’ udtales stødløst) (Figur 3). Manden fra Langå viser et lignende mønster i sætningen *Da hun rejste sig, kunne hun se både korn, skov, sol over Fyn* (Tabel 4). Sætningen i hans dialektoplæsning har en gennemsnitlig grundtonefrekvens på 109 Hz mens den ligger på 91 Hz i den intenderede standarddanske oplæsning. Sætningen i dialektoplæsningen oplæses dermed med en frekvens der er 19,5 % højere sammenlignet med frekvensniveauet når han forsøger tale ’rigsdansk’. Ordet *korn* udtaler han med et tonalt stød, og ordet har en gennemsnitlig grundtonefrekvens på 116 Hz i dialektoplæsningen. Han når en maksimal frekvens på 127 Hz, hvilket udgør en tonal stigning på 10,9 %.

	Helsætning – gennemsnitlig grundtonefrekvens	Fugl – gennemsnitlig grundtonefrekvens	Fugl – maksimal frekvens	Forskel i procent	Angivelse af stødforhold i fugl
Dialektoplæsning	248 Hz	267 Hz	296 Hz	10,9 %	Tonestød
Intenderet standarddansk oplæsning	205 Hz	201 Hz	244 Hz	21,4 %	Stødløs
Forskel i procent	21 %	33 %	21,3 %		

Tabel 3: Kvinde fra Ørbæk: ”Bag blomsten sad en lille fugl”.

	Helsætning - gennemsnitlig grundtonefrekvens	Korn - gennemsnitlig grundtonefrekvens	Korn – maksimal frekvens	Forskel i procent	Angivelse af stødforhold i korn
Dialektoplæsning	108 Hz	116 Hz	128 Hz	10,9 %	Tonestød
Intenderet standarddansk oplæsning	90 Hz	92 Hz	119 Hz	29,4 %	Stød
Forskel i procent	19,5 %	26 %	8 %		

Tabel 4: Mand fra Langå: ”Da hun rejste sig, kunne hun se både korn, skov og sol over Fyn”.

Der tegner sig dermed en tendens til at der i sætninger med tonale stød er et højt gennemsnitligt toneleje, og at markant stigende toner dermed er med til at trække tonelejet op. Grundtonefrekvensen falder ikke så kraftigt ved tone-stødene som i udtalerne med stød eller ikke-stød. Det vil altså sige at jo flere tone-stød og tonale accenter der er, des højere vil den gennemsnitlige grundtonefrekvens blive for hele sætningen.

Opsamling

I denne artikel har jeg fremlagt en undersøgelse af ældre østfynboers brug af traditionelle lokale og regionale fynske dialekttræk, nærmere bestemt det tone-stød og de tonale accenter Poul Andersen beskrev tilbage i 1958. Ud af 134 mulige realisationer var der på tværs af alle oplæsninger 27 eksempler på tonal accent og 35 eksempler på tone-stød. Samlet set udgør det 46 % af de mulige realisationer af disse lokale fynske træk, altså knap halvdelen, hvilket er en betydelig del. De tonale accenter udgjorde en femtedel af dataene, og tone-stødene over en fjerdedel. Tone-stødene og de tonale accenter synes i det store hele at falde på vokallydene. Tone-stødene havde ikke identificerbare knirk, men til gengæld var den stigende-faldende tonegang karakteristisk for dem alle. Blandt tone-stødene og de tonale accenter var ordene med høje vokaler hyppigst repræsenteret blandt de ord der havde mange af de her undersøgte dialektale træk, men ord med både halvhøje, halvlave og lave vokaler havde også i nogen grad tone-stød og tonale accenter. Generelt havde de tonale accenter kortlydende vokaler, og tone-stødene langtlidende vokaler. Der var betydeligt flere tone-stød og tonale accenter i informanternes dialektoplæsninger sammenlignet med deres intenderede standarddanske oplæsninger. Det betyder at informanterne til en vis grad evner at kodeskifte, og at de er bevidste om at kunne fralægge sig de træk de ikke forbinder med 'rigsdansk'. Det var overraskende at der var flere tone-stød end tonale accenter i dialektoplæsningerne fordi dette træk teoretisk synes at ligge tættere op ad det standarddanske knirk. Jeg har i undersøgelsen også vist at mængden af tonale accenter og tone-stødene synes at hænge sammen med højden på den gennemsnitlige grundtonefrekvens, forstået på den måde at jo flere tonale accenter og tone-stød der er i en sætning, des højere bliver den for helsætningen gennemsnitlige grundtonefrekvens. Dette har jeg vist gennem to eksempler på sammenligninger mellem sætninger læst op på henholdsvis dialekt og intenderet standarddansk. Dette er dog en mindre kvalitativ undersøgelse, og spørgsmålet om hvorvidt de tonale fænomener fører til en højere gennemsnitlig grundtone og opfattes som mere syngende og melodisk, er noget jeg vil belyse yderligere i mit ph.d.-projekt.

Litteratur

- Andersen, Poul (1958) *Fonemsystemet i Østfynsk*. København: Schultz.
- Berthelsen, Anders Wedel (2017) *Fynsk så det synger – nu med brunsviger*. Odense: Forlaget Skulkenborg.
- Boas, Kirstine og Silke Julie Flodin (2021) Synger de stadig? Akustiske beskrivelser af det østfynske tonestød. Upubliceret bachelorprojekt på uddannelsen for Nordisk Sprog og Litteratur på Aarhus Universitet.
- Boersma, Paul & Weenink, David (2023) *Praat: doing phonetics by computer* [Computer program]. Version 6.3.09, downloaded 2. marts 2023 fra <http://www.praat.org/>.
- Brink, Lars (2018) Stød, tryk og tonegang. I Ebba Hjorth, Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen og Laurids Kristian Fahl (red.). *Dansk Sproghistorie. Ord for ord for ord*. Aarhus: Aarhus Universitetsforlag: 229-259.
- Woolard, Kathryn A (2005) Codeswitching. I Alessandro Duranti (red.). *A companion to linguistic anthropology*. Blackwell Publishing Ltd.
- Pedersen, Inge Lise (2019) Kap. 1: Hovedtrækkene i de danske dialekter frem til 1950. I Georg S. Adamsen (red.). *Dialekter i rigt mål*. Modersmål-Selskabet. 25-69.
- Pedersen, Inge Lise (2002) Hovedtrækkene i de danske dialekter. I *Dialekter – sidste udkald?* Modersmål-Selskabet & CA Reitzels Forlag. Odense. 39-81.
- Ømålsordbogen. En sproglig-saglig ordbog over dialekterne på Sjælland, Lolland-Falster, Fyn og omliggende øer* (1-) 1992-. København: Center for Dialektforskning & Universitets-Jubilæet Danske Samfund.

Resumé:

Internordisk sprogforståelse: Hvor svært kan det være?

Ocke-Schwen Bohn

Aarhus Universitet

Er dansk, norsk og svensk så nært beklægtede med hinanden, at sprogbrugerne nemt kan forstå hinanden på tværs af sproggrænserne? For at måle sværhedsgraden af internordisk sprogforståelse lavede vi flere undersøgelser, hvor vi sammenlignede dansktalendes evner til at forstå engelske, norske og svenske ytringer.

I alt fem deltagergrupper, tre med dansk som modersmål og to med dansk som andetsprog, lyttede til danske ytringer (alle grupper) og til enten norske, svenske, eller engelske ytringer. Deltagernes opgave var at svare så hurtigt som muligt, om hver enkelt ytring var forkert (fx ”Alle hunde har femten ben”) eller korrekt (fx ”Varme og kulde er modsætninger”).

Vi tolkede længden af responstiden og antal af forkerte svar som forståelsesomkostninger. For deltagere med dansk som modersmål viste det sig, at engelsk var næsten lige så nemt at forstå som dansk, mens det tog signifikant længere tid og var mere fejlbehæftet at forstå norsk og især svensk. Som forventet tog det deltagere med dansk som andetsprog i gennemsnit lidt længere tid at svare på danske ytringer, mens antallet af korrekte svar var den samme som hos deltagere med dansk som modersmål. Deltagere med dansk som andetsprog viste stor variation i responstid og i antal korrekte svar på norske og svenske ytringer. For flere andetsprogsdeltagere var norsk lige så nemt/svært at forstå som dansk, mens forståelsesomkostninger ved svensk var meget højere end ved dansk.

Vores undersøgelser tyder på, at det er en sandhed med stærke modifikationer at påstå, at dansktalende nemt kan forstå norsk og svensk. Især svensk volder problemer for dem, som har dansk som modersmål, og endnu mere for dem, som har dansk som andetsprog. Vi er p.t. i gang med at undersøge, hvilke ytringsegenskaber (leksikalske eller fonologiske) bidrager til de forskellige gruppers forståelsesomkostninger.

Resumé:

Er handlingsanvisende diskursinitiale verballøse talehandlinger elliptiske? En afprøvning af hypotesen om minimalellipse

Simon Borchmann

Roskilde Universitet

I dette oplæg vil jeg på grundlag af et lille korpus af handlingsanvisende verballøse talehandlinger afprøve en ellipsebaseret analyse af diskursinitiale verballøse ytringer. Beskrivelser af ellipse har typisk koncentreret sig om komplekse udtryk der både omfatter enheder der opfylder det fuldstændighedskriterium beskrivelsen opererer med (Eikmeyer 1985), og enheder der ikke gør, fx:

- 1) Vigtigst er nok at vi løber HØJRE om Damhussøen for at spare 1,5 km og HØJRE om Mosen langs Pilesvinget. (<https://copenhagenmarathon.dk/faellestraening-30-januar/>)

Det sidste led i sideordningen er uden verbum (Wilder 2018). Men det optræder i en ko-tekst der omfatter et forudgående sætningsformet udtryk. Dermed er der et grundlag for en rekonstruktion af den fuldstændige variant man i ellipsebaserede beskrivelser antager er forbundet med det elliptiske udtryk (Klein 1985). Således kan man alene ved at støtte sig til ko-teksten nå frem til følgende fuldstændige variant: at vi løber højre om Mosen langs Pilesvinget.

Der finder imidlertid også verballøse ytringer som optræder uden ko-tekst, fx:

- 2) og så til venstre nede for enden (Passager til taxachauffør)

De kaldes diskursinitiale (Yanofsky 1978, Barton 1990), og sådanne ytringer er der væsentligt færre der har behandlet i beskrivelser af ellipse. Nogle udgrænser dem endog, netop med den begrundelse at det er ko-teksten der forsyner os med et grundlag for rekonstruktion – ikke bare af verbet, men også af markeringer af tempus og aspekt (Bondarenko 2019). Når der ikke er en ko-tekst, er der altså ikke noget grundlag for rekonstruktion, og derfor kan sådanne ytringer

ikke anses for elliptiske. Se også Brink & Schack (2014) der dog begrundet udgrænsningen på anden vis. Kleins (1985) og Merchants (2004, 2006) beskrivelser af ellipse omfatter ikke desto mindre diskursinitiale verballøse ytringer. Kleins analyse af links um! er dog ikke begrundet. Til forskel herfra fremhæver Merchant at diskursinitiale verballøse ytringer "shows the same kinds of grammatical dependency effects that we find in fully sentential structures" (Merchant 2006:76):

(Enan) kafe (parakalo)!
a coffee.ACC please

Her påpeges det at akkusativens er den kasusmarkering vi vil forvente af det relevante verbum:

Ferte mou (enan) kafe (parakalo)!
bring.IMP me a coffee.ACC please

Merchant (2004, 2006) giver eksempler på sådanne kasusoverensstemmelser på en række andre sprog (se også Hall 2018), og i forlængelse heraf foreslår han en ellipsebaseret analyse af diskursinitiale verballøse ytringer. Analysen kaldes minimalellipse og forudsætter at konteksten idet den gør en handling salient, tillader en udeladelse af do it. Det illustreres med følgende eksempel: "[Seeing someone trying to pound in a nail with a screwdriver] no no - with a hammer!" (Merchant 2004:726). Analysen suppleres af en scriptbaseret variant der tillader udeladelse af mere specifikke verber, fx bring når man bestiller noget på en restaurant.

I oplægget præsenterer jeg en afprøvning af varianter af den minimalellipsebaserede analyse på dansk. Analysen anvendes blandt andet på eksempel (2) og følgende forekomster:

- 3) bare hånden derover ikke (Svæveflyveinstruktør til svæveflyvelev)
- 4) ind bagved ham nu (Svæveflyveinstruktør til svæveflyvelev)
- 5) lige en lille smule mere fart på (Svæveflyveinstruktør til svæveflyvelev)
- 6) så bare ganske stille og roligt (Svæveflyveinstruktør til svæveflyvelev)
- 7) bagvægt (erfaren mountainbikerytter til uerfaren mountainbikerytter)
- 8) udfladning (Svæveflyveinstruktør til svæveflyvelev)
- 9) alt det der sådan levrede eller hvad man skal kalde det geleagtige noget væk med det (jæger til sin hustru)

Konklusionen er at analysen ikke kan generaliseres til handlingsanvisende verballøse talehandlinger, men at den som en heuristik (Craenenbroeck & Temmerman 2018) kan være ganske informativ med hensyn til deres grammatiske og semantiske egenskaber som selvstændige talehandlinger.

Litteratur

- Barton, Ellen (1990) *Nonsentential Constituents: A theory of grammatical structure and pragmatic interpretation*. John Benjamins Publishing Company.
- Bondarenko, Antonina (2019) A corpus-based contrastive study of verbless sentences: quantitative and qualitative perspectives. *Studia Neophilologica*, 91:2, 175-198. DOI: 10.1080/00393274.2019.1616221
- Brink, Lars & Jørgen Schack (2014) Tale er sølv. Om begrebet syntaktisk ellipse. *Danske Studier*, 45-75.
- van Craenenbroeck, Jeroen & Tanja Temmerman (2018) Ellipsis in natural language: Theoretical and empirical perspectives. Tanja Temmerman & Jeroen van Craenenbroeck (eds.) *The Oxford Handbook of Ellipsis*. Oxford University Press.
- Eikmeyer, Hans-Jürgen (1985) Ellipsen und Analysestrategien in inkrementellen Sprachverarbeitungsmodellen. I Reinhard Meyer-Hermann & Hanne Rieser (eds.). *Ellipsen und fragmentarische Ausdrücke*, Band 2, 1–25. Tübingen: Niemeyer.
- Hall, Allison (2018) Fragments. Tanja Temmerman & Jeroen van Craenenbroeck (eds.) *The Oxford Handbook of Ellipsis*. Oxford University Press.
- Klein, Wolfgang (1985) Ellipse, Fokusgliederung und thematischer Stand. I Reinhard Meyer-Hermann & Hanne Rieser (eds.). *Ellipsen und fragmentarische Ausdrücke*, Band 1, 1–24. Tübingen: Niemeyer.
- Merchant, Jason (2006) "Small Structures": A Sententialist Perspective. I Ljiljana Progovac, Kate Paesani, Eugenia Casielles, & Ellen Barton (eds). *The Syntax of Nonsententials: Multidisciplinary perspectives*, 73-91. Amsterdam, NE: John Benjamins.
- Merchant, Jason (2004) Fragments and Ellipsis. *Linguistics and Philosophy*, 27:6, 661-738.
- Napoli, Donna (1982) Initial Material Deletion in English. *Glossa* 16, 85-111.
- Yanofsky, Nancy (1978) NP Utterances. *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, 491-502. Chicago: Chicago Linguistic Society.
- Wilder, Christopher (2018) Conjunction Reduction and Right Node Raising. Tanja Temmerman & Jeroen van Craenenbroeck (eds.) *The Oxford Handbook of Ellipsis*. Oxford University Press.

Tegn-et-kort-data fra Aarhus og København – en perceptuel-dialektologisk undersøgelse

Kristoffer Friis Bøegh, Inger Schoonderbeek Hansen & Mette-Marie Møller Svendsen
Aarhus Universitet

1. Indledning¹

I denne artikel præsenterer vi en første pilotundersøgelse i dansk sociodialektologisk (jf. Kristiansen 2018) sammenhæng hvori vi anvender en perceptuel-dialektologisk metode til at afdække sprogbrugeres opfattelser af og holdninger til sproglig variation: *the draw-a-map task* (jf. fx Preston 2018:185 f.), eller på dansk: 'tegn-et-kort-metoden'. Vi benytter tegn-et-kort-metoden til at anlægge et nyt, sammenlignende sociolingvistisk-geografisk perspektiv på den sproglige virkelighed som den opleves – og potentielt opleves forskelligt – af sprogbrugere i Danmark. Konkret har vi undersøgt en respondentgruppe i hhv. Aarhus og København og deres respektive opfattelser af det nutidige danske sproglandskab igennem et eksperiment i to dele: først fra et *bottom up*-perspektiv (sprogbrugernes) og dernæst fra et *top down*-perspektiv (sprogforskerens). Undersøgelsen har følgende forskningsspørgsmål: Hvilke sproglige forskelle og ligheder findes der i Danmark ifølge sprogbrugerne? Hvilke betegnelser anvender de om den opfattede sproglige variation? Og kommer opfattelserne af den sproglige virkelighed forskelligt til udtryk i tegn-et-kort-dataene fra Aarhus og fra København?

I afsnit 2 redegør vi for baggrunden for undersøgelsen. I afsnit 3 beskriver vi undersøgelsens metoder og data. I afsnit 4 præsenterer og diskuterer vi undersøgelsens resultater. I afsnit 5 bringer vi en opsamling og perspektiverer til fremtidig forskning.

2. Perceptuel dialektologi i dansk sociodialektologi

Danmark har en lang forskningstradition inden for sproglig variation og sprogforandring (jf. fx Pedersen 2003; Bøegh et al. 2023). Eksperimentelle undersøgelser af sprogbrugeres holdninger

¹ Stor tak til Malene Monka og Kristine Køhler Mortensen som stod for indsamlingen af de københavnske data til undersøgelsen, og til deres hold i E22 på Dansk, Københavns Universitet, som leverede besvarelsene. Tak også til vores hold i F22 i hhv. *Sprogbrugsanalyse* og *Sproglig variation* på Nordisk, Aarhus Universitet, som ligeledes leverede data; til vores praktikanter Ilaria Iannaccone og Erica Morone som assisterede med at digitalisere de håndtegnede kort; samt til publikum til MUDS 19 og fagfællen for gode kommentarer.

til udvalgte sproglige varieteter har vist at der findes forskel på sprogbrugeres bevidste og underbevidste sprogholdninger (bl.a. Kristiansen 1991, 2009; Quist 2005; Maegaard 2007; jf. desuden Kristiansen 2020). Forholdet mellem disse to sprogholdningstyper har Tore Kristiansen (2009:170) kaldt for paradoksalt, for det viser sig at selv om sprogbrugernes bevidste eller eksplicit udtrykte holdninger til lokale varieteter (dvs. dialekter) generelt er positive, så peger deres underbevidste eller implicit udtrykte holdninger væk fra det lokale, dvs. i retning af København, der længe har været opfattet som det altdominerende sproglige normcenter (dvs. rigsdansk). Det senmoderne Danmark fremstår således som et af de mest homogene sprogsamfund i Europa, og det er, lyder argumentet, de underbevidste sprogholdninger, og den standardsprogsideologi de afspejler, der matcher den observerede sprogforandring, herunder den omfattende standardiseringsproces det danske sprogsamfund har gennemgået siden begyndelsen af det 20. århundrede (jf. Kristiansen 2020).

Resultaterne af flere nyere studier tyder imidlertid på at sprogbrugere opfatter den sproglige virkelighed mere forskelligt end det hidtil har været antaget i dansk sociodialektologi, alt afhængigt af hvor i Danmark sprogbrugerne kommer fra. Dette udfordrer ideen om standard-sprogets hegemoni. Det gælder især sprogholdningsundersøgelser udført i Jylland: i Nordjylland og i Sønderjylland, som del af projektet *Dialekt i periferien* (Maegaard & Quist 2020, se Maegaard et al. 2020 for en samlet fremstilling), samt i Midtvestjylland (Hansen & Goldshtein 2021; Hansen 2022). Disse undersøgelser tegner et mindre entydigt sprogideologisk billede (jf. desuden fx Ladegaard 1999; Svenstrup & Thøgersen 2009) og peger dermed i retning af en mere kompleks sociolingvistisk virkelighed. En del af forklaringen skal øjensynligt findes i denne type eksperimenters metodologi, herunder de sproglige stimuli. Fx kan undersøgelsesresultaterne af sprogideologiske spørgsmål være forskellige alt efter i hvor høj grad de sproglige stimuli er tilpasset lokale forhold, dvs. stemmer overens med respondenternes sproglige virkelighed; af den grund kommer bl.a. resultatet med ét, københavnsk-orienteret normcenter ikke så entydigt frem i de seneste som i tidligere undersøgelser (jf. Hansen 2022).

Den perceptuelle dialektologi fokuserer på almindelige sprogbrugeres opfattelser af og holdninger til sproglig variation ud fra et sociolingvistisk-geografisk perspektiv (se fx Preston 2018 for en oversigt). Hvor man i dansk sociodialektologi især har benyttet en kombination af forskellige eksperimentelle undersøgelsesformer og etnografiske feltarbejder, har tegn-et-kort-metoden stået stærkt internationalt inden for den perceptuelle dialektologi (jf. fx Preston 1999, 2018; Cramer 2021; Montgomery 2022). Metoden har potentiale til at lade sprogbrugere komme endnu mere til orde, og den kan dermed bidrage til yderligere at nuancere billedet af sproglig variation og sprogforandring i Danmark.

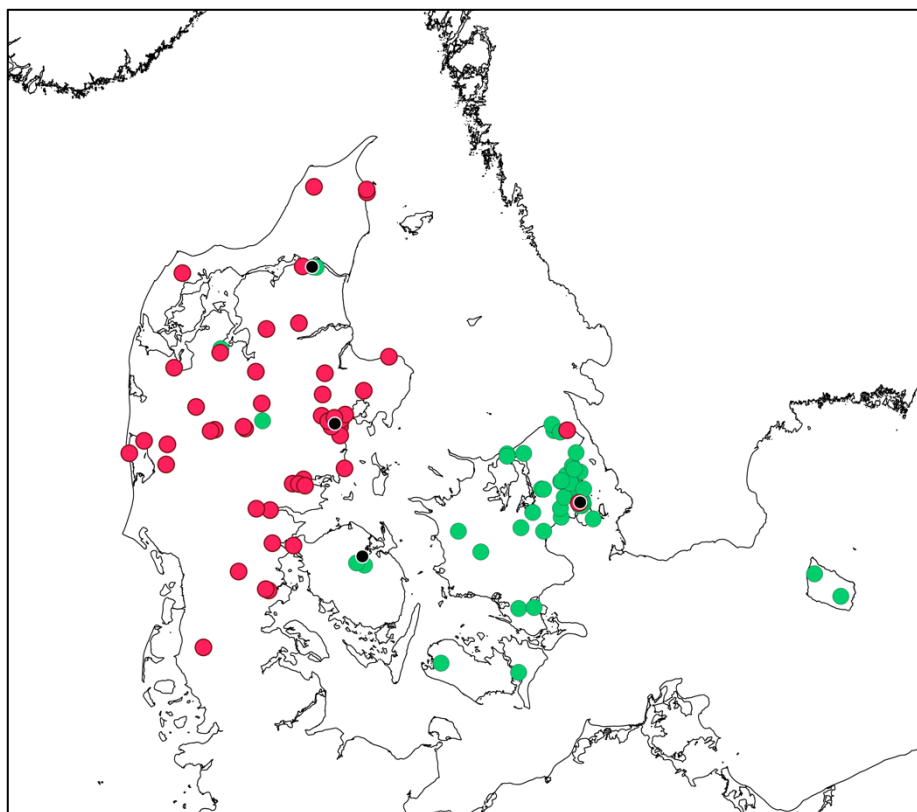
3. Metoder og data

Tegn-et-kort-metoden blev introduceret i 1980'erne af Dennis Preston (fx 1982, 1986, 1989). Metoden går ud på at en gruppe respondenter får udleveret et kort hver over et givent geografisk område, i dette tilfælde et danmarkskort, hvorpå de indtegner deres opfattelse af hvor der findes sproglige forskelle og ligheder, og hvordan disse forskellige sprogområder/-varieteter kan betegnes og karakteriseres. Altså kan metoden bruges til at undersøge hvilke geografiske og sociale varieteter af dansk talesprog der er perceptuelt fremtrædende for sprogbrugere, hvordan de afgrænser disse forskellige sprogområder, og hvad de forstår som kendetegn (socialt, lingvistisk) ved et givent område eller en given varietet. I tillæg til at indsamle individuelle kort anvender vi i denne undersøgelse en videreudvikling af tegn-et-kort-metoden i digitalt regi (jf. Montgomery & Stoeckle 2013; Montgomery 2022). Den går ud på at benytte GIS-software (dvs. et geografisk informationssystem) til at aggregere data fra individuelle besvarelser og yderligere udforske mere generelle mønstre i materialet og visualisere resultaterne. I det følgende præsenterer vi de to respondentgrupper, vores eksperiment og databehandlingen.

3.1. Respondenter

Vores respondenter var 47 BA-studerende fra Nordisk sprog og litteratur på Aarhus Universitet (AU) og 51 BA-studerende fra Dansk på Københavns Universitet (KU). I gennemsnit var respondenterne i starten af 20'erne (AU: 23,4 år; KU: 22,2 år), og kønsfordelingen var meget ens for de to respondentgrupper (AU: 79 % kvinder, 17 % mænd, 4 % uden angivet køn; KU: 78 % kvinder, 18 % mænd, 4 % uden angivet køn). Figur 1 viser hvor respondenterne angav de kommer fra. Enkelte respondenter ($n = 3$) stammer fra Jylland selv om de indgår i KU-datasættet, og vice versa ($n = 2$), men det overordnede mønster er at AU-respondenterne kommer fra Jylland og KU-respondenterne fra det østlige Danmark.² Vi antager derfor at resultaterne overordnet set afspejler hhv. et vestdansk og østdansk perspektiv på sproglig variation. Da der er tale om et pilotforsøg, valgte vi nogle for os oplagte respondentgrupper, nemlig vores egne og kollegers danskstuderende. Grupperne må dermed forudsættes at være delvis forudindtagne i forhold til emnet; omvendt danner de et godt udgangspunkt for en planlagt opskalering med fx yngre udskolings- eller gymnasieelever.

² En enkelt respondent der angav at være fra Færøerne, har også bidraget til KU-datasættet, men vedkommende er ikke repræsenteret på kortet i figur 1.



Figur 1: Oversigtskort over respondenternes selvrapporterede hjemegne. Røde cirkler = AU-respondenter; grønne cirkler = KU-respondenter.

3.2. Eksperimentets to dele

Respondenterne fik hver udleveret to danmarkskort. Standardmetoden inden for perceptuel dialektologi er at benytte et kort med minimal markering, almindeligvis kun landegrænser og større byer (jf. Bounds & Sutherland 2018). Vores kort til eksperimentet viste den dansk-tyske grænse og København, Aarhus, Aalborg og Odense, men var ellers blankt (se figur 2).

Formålet med eksperimentets del 1 var at undersøge sprogbrugernes egne afgrænsninger af og bevidste holdninger til sproglig variation. Respondenterne blev instrueret i at indtegne på det ene kort hvor de mente der findes sproglige forskelle og ligheder, give de forskellige sprogområder (eller: varieteter) en betegnelse, give deres mening til kende om sprogområderne og gerne give eksempler på kendetegn for de forskellige områder. Dertil blev de bedt om at notere baggrundsinformation om deres alder, køn, hjemegn og sproglige baggrund. Respondenterne fik 15 min. til at besvare del 1.

Formålet med eksperimentets del 2 var at undersøge hvad sprogbrugerne forstår ved en række gængse betegnelser for sprogvarieteter (jf. fx Svenstrup & Thøgersen 2009). Instruktio- nen til del 2 blev først vist efter endt besvarelse af del 1. Respondenterne blev instrueret i at indtegne områder for følgende varietetsbetegnelser: *københavnsk*, *jysk*, *dialekt* og *rigsdansk*.

denterne havde indtegnet på deres kort, i form af vektor-polygoner, dvs. digitale repræsentationer af de håndtegnede inddelinger. Alle polygoner blev tagget med deltagernes sociodemografiske data og deres valgte betegnelser for sprogområder/-varieteter. Det samlede datasæt består af 1.289 polygonobjekter. I de fleste tilfælde svarer ét polygonobjekt til én håndtegnet inddeling, men der er også tilfælde hvor ét polygonobjekt svarer til flere håndtegnede inddelinger (typisk i eksperimentets del 2, hvor fx *dialekt* oftest blev indtegnet flere forskellige steder af samme respondent).

Vi brugte QGIS til at producere kortvisualiseringer ud fra disse aggregerede data. Vi samlede de forskellige polygonobjekter i polygongrupper ud fra de betegnelser respondenterne selv havde indtegnet på deres kort – dog med sted- og varietetbaserede betegnelser slået sammen hvor dette var muligt (så fx *Fyn* og *fynsk* behandles samlet, dvs. under én betegnelse: *fynsk*). Vi farvelagde polygongrupperne og gav dem 3 % opacitet per polygonobjekt. Altså er hvert enkelt polygonobjekt stort set gennemsigtigt, men når polygonobjekter overlappes, kan en høj ”enighedsgrad” (jf. *agreement level* eller *recognition rate*, se fx Montgomery 2012; Preston 2018:186) blandt respondenterne aflæses i form af en tydelig farvekoncentration i visualiseringen. En lavere enighedsgrad afspejles i form af en mindre tydelig farvekoncentration (jf. fx Montgomery & Stoeckle 2013:81).

4. Resultater

4.1. To perspektiver på den sproglige virkelighed i Danmark

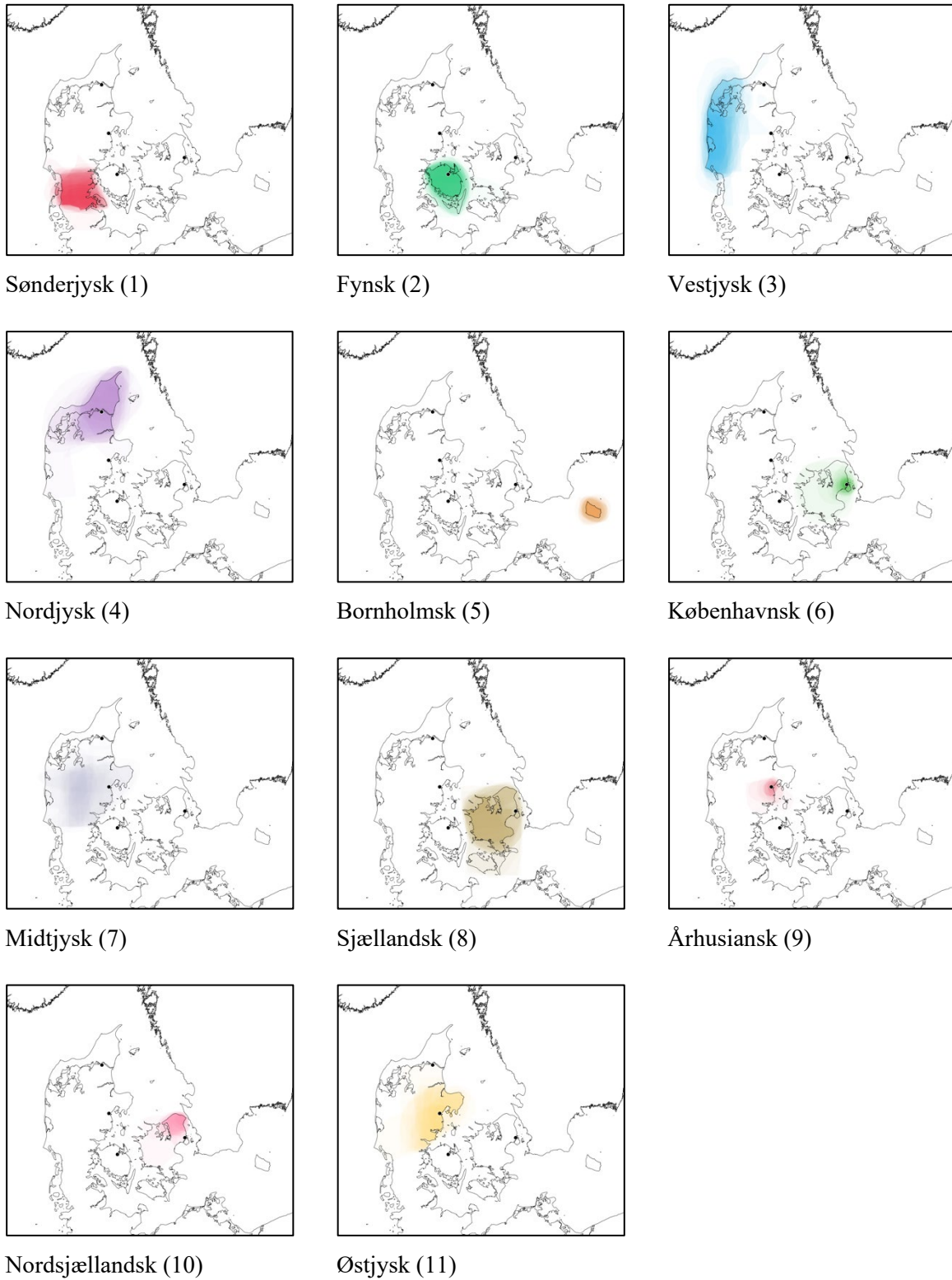
I dette afsnit præsenterer og diskuterer vi resultaterne af eksperimentets del 1, altså undersøgelsens *bottom up*-perspektiv på sproglig variation i Danmark. Grundet pladsbegrænsninger fokuserer vi primært på og præsenterer udelukkende kortvisualiseringer af de opfattede sprogområder/-varieteter (herefter forenklet til *varieteter*) med høj enighedsgrad – dvs. perceptuelt fremtrædende varieteter, hvilket vi her definerer som varieteter med 10 eller flere individuelle belæg per respondentgruppe (jf. i øvrigt diskussionen i Cramer 2021:3). En oversigt over varieteterne med høj enighedsgrad vises i tabel 1.

AU-respondenter (n = 47)				KU-respondenter (n = 51)			
Nr.	Varietet	n	%	Nr.	Varietet	n	%
1	Sønderjysk	45	95,7	1	Sønderjysk	46	90,2
2	Fynsk	43	91,5	2	Bornholmsk	40	78,4
3	Vestjysk	40	85,1	3	Fynsk	40	78,4
4	Nordjysk	39	83,0	4	Københavnsk	39	76,5
5	Bornholmsk	36	76,6	5	Nordjysk	32	62,8
6	Københavnsk	35	74,5	6	Nordsjællandsk	20	39,2
7	Midtjysk	27	57,5	7	Århusiansk	19	37,3
8	Sjællandsk	24	51,1	8	Jysk	18	35,3
9	Århusiansk	23	48,9	9	Vestjysk	18	35,3
10	Nordsjællandsk	22	46,8	10	Rigsdansk	16	31,4
11	Østjysk	21	44,7	11	Sydsjællandsk	13	25,5
				12	Sjællandsk	11	21,6

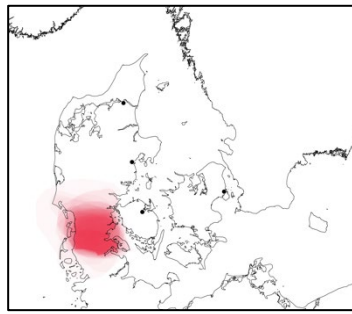
Tabel 1: Varieteter med høj enighedsgrad hos hhv. AU- og KU-respondenterne.

Resultaterne viser at de to respondentgrupper skelner mellem et sammenligneligt antal varieteter (total: n = 906; AU: n = 476, 52,5 %; KU: n = 430, 47,5 %). Varieteterne med høj enighedsgrad som ses i tabel 1 er ligeledes sammenlignelige (AU: n = 355, 74,6 %; KU: n = 312, 72,6 %). I vores KU-data er der dog kun fem varieteter med mere end 50 % enighed hvor antallet for AU-respondenterne er otte. I vores AU-data er faldet desuden mere skarpt fra perceptuelt fremtrædende varieteter til øvrige varieteter: fra 44,7 % for *østjysk* til 19,1 % (n = 9) og nedefter for de næste varieteter (*sydsjællandsk*, *Aalborg*, *nordvestjysk*). Til sammenligning er næste varietet efter *sjællandsk* (21,6 %) i KU-dataene *Vestegnen* (n = 9, 17,6 %) og *Amager* og *Lolland-Falster* (n = 8, 15,7 %).

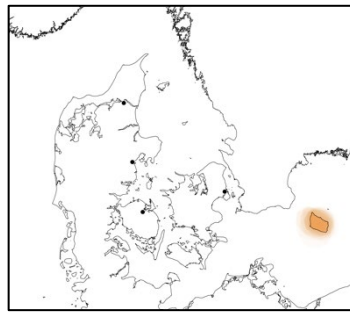
Figur 3 og 4 viser kort med de forskellige perceptuelt fremtrædende varieteter visualiseret separat for hhv. AU- og KU-respondenterne; dernæst viser figur 5 og 6 de samme varieteter samlet på ét kort for hver respondentgruppe. For sammenlignelighedens skyld matcher de valgte farver for de varieteter der optræder hos begge respondentgrupper (så fx *bornholmsk* konsekvent er farvet orange, osv.). Omvendt har varieteterne *midtjysk* (kun med høj enighedsgrad i AU-resultaterne) og *jysk* (KU) samme farve (grå) på de forskellige kort; det samme er tilfældet for *østjysk* (AU) og *rigsdansk* (KU) (gul). En sidste farve (støvet lyseblå) benyttes kun for *sydsjællandsk* (KU).



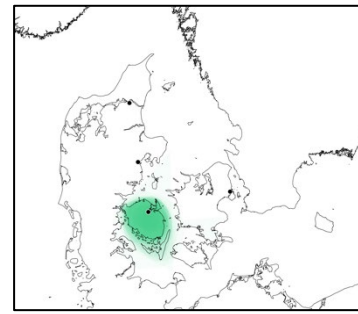
Figur 3: AU-respondenternes perceptuelt fremtrædende varieteter. Varieteterne er rangordnet fra øverst til venstre (sønderjysk) til nederst til højre (østjysk) jf. tabel 1.



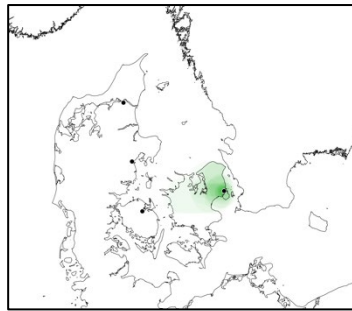
Sønderjysk (1)



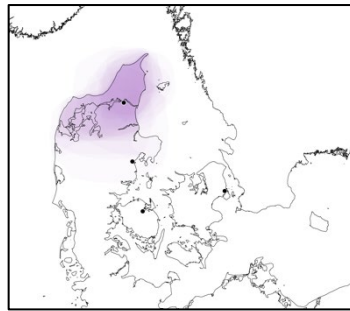
Bornholmsk (2)



Fynsk (3)



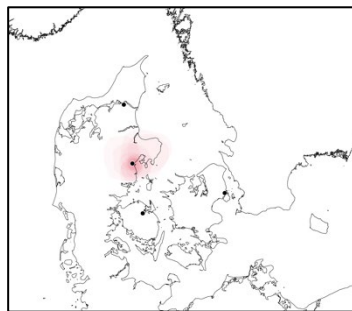
Københavnsk (4)



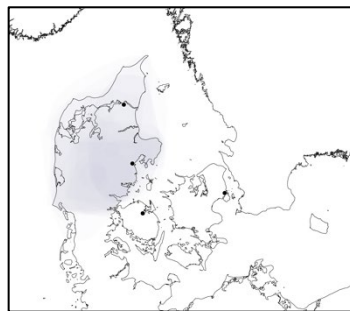
Nordjysk (5)



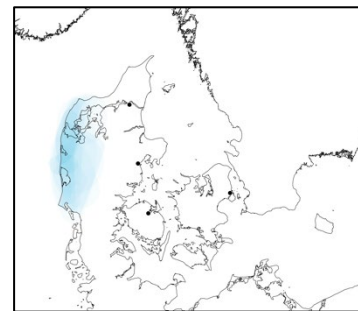
Nordsjællandsk (6)



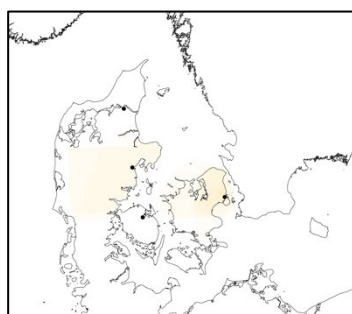
Århusiansk (7)



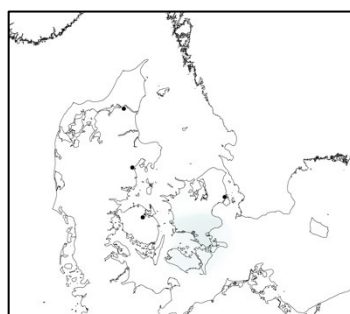
Jysk (8)



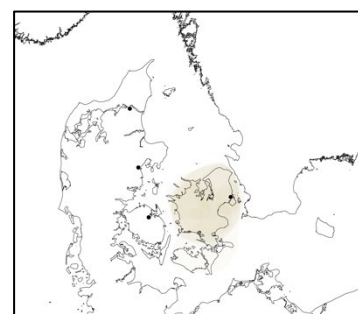
Vestjysk (9)



Rigsdansk (10)

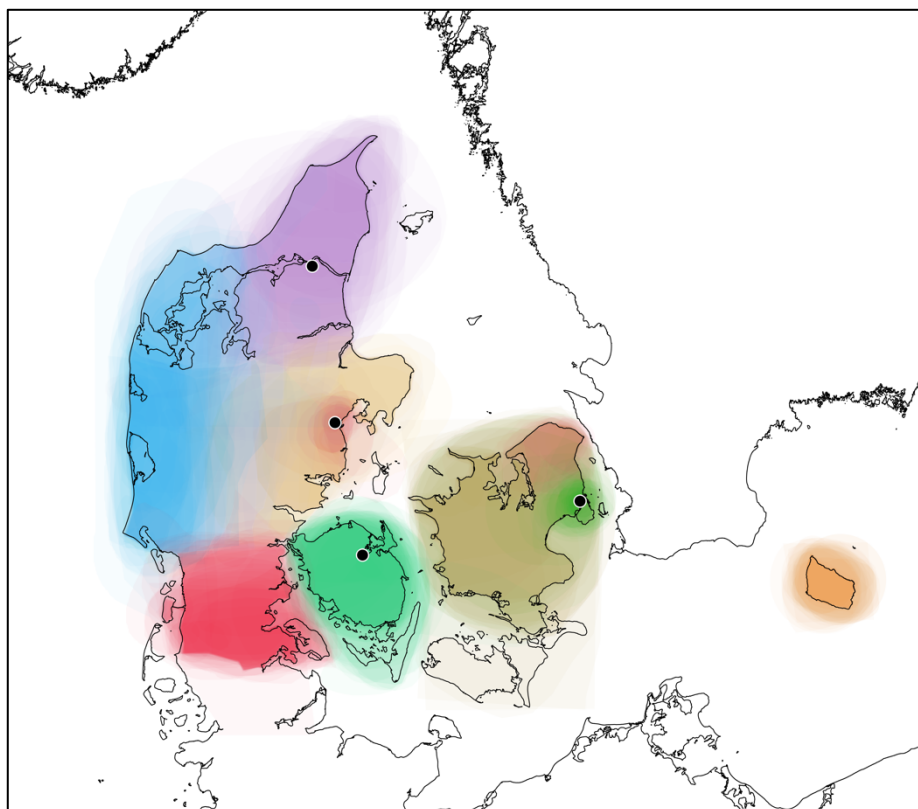


Sydsjællandsk (11)

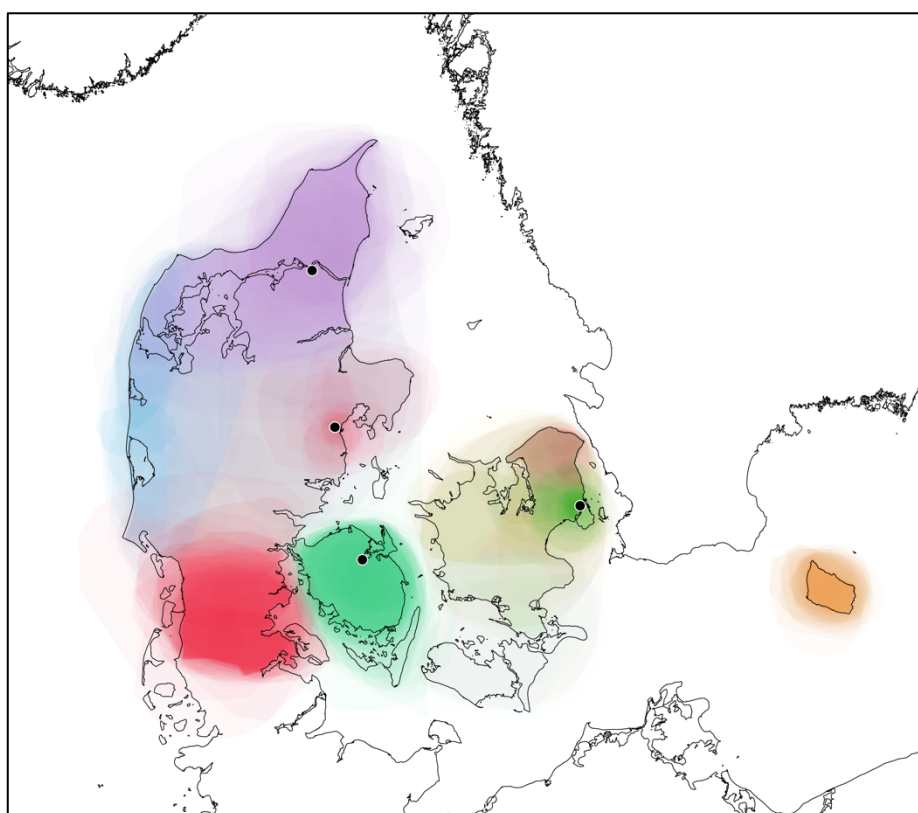


Sjællandsk (12)

Figur 4: KU-respondenternes perceptuelt fremtrædende varieteter. Varieteterne er rangordnet fra øverst til venstre (sønderjysk) til nederst til højre (sjællandsk) jf. tabel 1.



Figur 5: AU-respondenternes perceptuelt fremtrædende varieteter visualiseret samlet.



Figur 6: KU-respondenternes perceptuelt fremtrædende varieteter visualiseret samlet.

Ud fra tabellen og de forskellige kortvisualiseringer kan vi konstatere at respondenterne ikke opfatter Danmark helt ensartet sprogligt set. Der er i høj grad tale om stedbaserede, regionale inddelinger som passer godt på brede regionale skel, men inddelingerne svarer ikke til de traditionelle dialektinddelinger for fx Jylland som man vil kunne se i Jysk Ordbog (jf. desuden fx Bøegh 2022 for et supplerende eksempel). Resultaterne understøtter dermed en grundlæggende pointe fra bl.a. Preston (fx 1982, 1986, 1989): Der er ikke nødvendigvis direkte sammenhæng mellem de varieteter sprogforskere og sprogbrugere identificerer og betragter som mest fremtrædende (jf. også Cramer 2021:3 f.). 9 ud af i alt 14 varieteter med høj enighedsgrad (64,3 %) går igen hos AU- og KU-respondenterne. Overordnet set er der dermed tale om sammenlignelige billeder, hvis vi ser bort fra forskelle i rangordninger af individuelle varieteter og varieteter der ikke figurerer hos begge respondentgrupper.

For begge respondentgrupper er *sønderjysk* den mest fremtrædende varietet. Respondenternes perception af denne varietet modsvarer flere videnskabelige undersøgelser af den status sønderjysk har i det danske sprogsamfund som stadig aktiv, om end til en vis grad også nivellerende dialekt (jf. Monka 2013; Maegaard & Quist 2020). Montgomery (2012) forklarer nogle lignende resultater i Storbritannien (jf. desuden resultater for USA i fx Preston 1986) med henvisning til at de varieteter og sproglige områder der er de mest veletablerede i den kollektive nationale bevidsthed (det kan være pga. reelle sproglige forskelle, men også fx stigmatisering), opleves som kulturelt fremtrædende af sprogbrugerne, som derfor konsekvent indtegner dem. Også *fynsk*, *bornholmsk*, *nordjysk* og *københavnsk* kan siges at være veletablerede hos begge respondentgrupper. Ud over København (*københavnsk*) er Aarhus (*århusersk*) den eneste anden by der opnår en høj enighedsgrad.

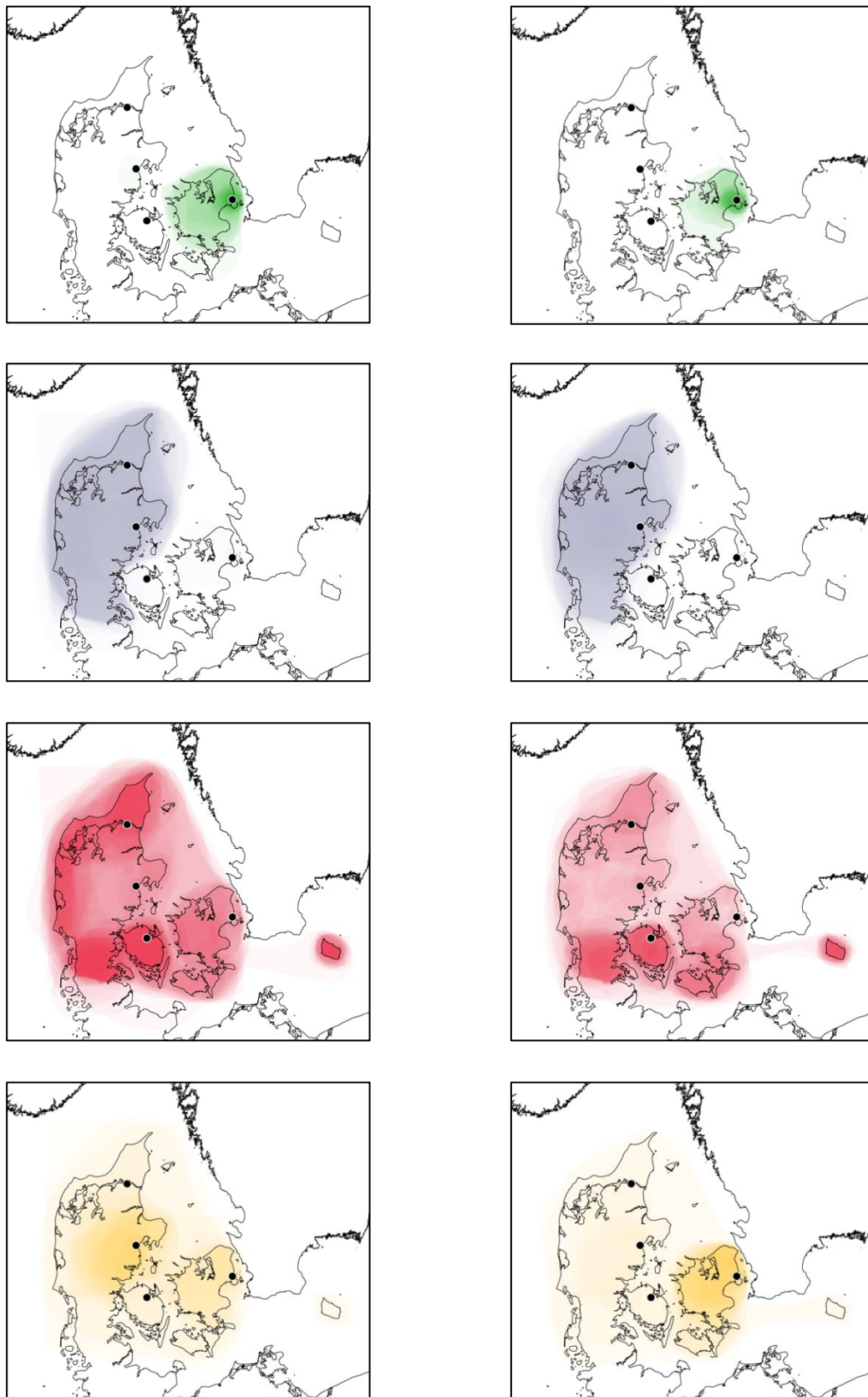
Resultaterne viser i tråd med tidligere forskning at sprogbrugere generelt er mere opmærksomme på sproglig variation i områder de selv har tæt kendskab til (fx Preston 1986; Montgomery 2012, 2022:171 f.). Dette afspejles mest tydeligt i forhold til varieteter med lavere enighedsgrad, men dog med multiple belæg (n = 2-9), hvor der ses en større grad af variation. Dette gælder også i forhold til respondenternes valgte betegnelser. KU-respondenterne har fx i mange tilfælde flere inddelinger i hovedstadsområdet; i tillæg til *Vestegnen* og *Amager* har de fx *nørrebro*, *vesterbro*, *frederiksberg* og *urban*. Tilsvarende identificerer AU-respondenterne hele 11 jyske byer (fx *Randers*, *Frederikshavn*, *Skive*) hvorimod eneste jyske by med mere end et enkelt belæg i vores KU-data er *Aalborg* (ud over Aarhus, som allerede nævnt). Desuden fungerer nogle af de mindre fremtrædende varieteter præciserende i forhold til de mere fremtrædende varieteter; fx skelner enkelte AU-respondenter (n = 2) mellem *vest-* og *østsønderjysk*,

hvilket forklarer hvorfor enighedsgraden for *sønderjysk* ikke er 100 % for AU-respondenterne som samlet gruppe.

Der er også variation i forhold til det overordnede billede, dvs. på baggrund af varieteterne med høj enighedsgrad, mellem AU- og KU-respondentgrupperne; fx er *vestjysk* hhv. 3. (AU) og 9. (KU) mest fremtrædende varietet, og omvendt er *nordsjællandsk* hhv. 6. (KU) og 11. (AU) mest fremtrædende varietet. Visuelt fremstår Jylland mere klart inddelt i forskellige områder på baggrund af AU-dataene, mens indtrykket på baggrund af KU-dataene er at Jylland fremstår mere ”udflydende”, og vice versa. *Vestjysk* er et eksempel på en varietet som er perceptuelt fremtrædende hos begge respondentgrupper, men hvor området både fortolkes forskelligt i forhold til udbredelse og varierer tydeligt i forhold til farvekoncentration. AU-respondenterne skelner desuden mellem *midtjysk* og *østjysk*, hvor KU-respondenterne skelner mellem *sydsjællandsk* og desuden de overordnede kategorier *jysk* og *rigsdansk* (se mere i afsnit 4.2). Også disse resultater kan ses som udtryk for et lokalt orienteret fokus hos respondenterne; fx kan *jysk*, som kun er fremtrædende hos KU-respondenterne, tolkes som en for generisk kategori til at vække genklang hos AU-respondenterne. Tilsvarende figurerer *sjællandsk* hos begge respondentgrupper, men er mere fremtrædende i AU-resultaterne.

4.2. Sprogforskernes betegnelser

I dette afsnit ser vi nærmere på undersøgelsens *top down*-perspektiv, dvs. respondenternes placering af sprogforskernes betegnelser, dvs. *københavnsk*, *jysk*, *dialekt* og *rigsdansk*, i eksperimentets del 2. Spørgsmålet om hvad sprogbrugerne forstår ved sprogforskernes betegnelser, er interessant da det har vist sig at være langt fra entydigt, især ikke for betegnelsen *rigsdansk* (se fx Svenstrup & Thøgersen 2009; Maegaard & Quist 2020; Hansen 2022). Vores undersøgelsesmetode er derfor særdeles relevant idet vi får sprogbrugernes direkte input til hvor på danmarkskortet de mener disse varieteter findes. Hvor vi i afsnit 4.1 diskutererede resultater på baggrund af varierende enighedsgrad, så udfyldte alle respondenter (med enkelte undtagelser) opgaven i eksperimentets del 2 og placerede dermed alle fire varieteter. Resultaterne vises i figur 7. Bemærk at i de tilfælde hvor en varietet også optrådte i afsnit 4.1, svarer farvevalget her til det i de tidligere figurer; altså er *københavnsk* farvet grøn, *jysk* grå og *rigsdansk* gul; *dialekt* er farvet rød (samme farve som *sønderjysk* i afsnit 4.1).



Figur 7: Respondenternes placering af sprogforskningens betegnelser københavnsk (grøn), jysk (grå), dialekt (rød) og rigsdansk (gul). AU-resultaterne vises til venstre, KU-resultaterne til højre.

Som det fremgår af kortvisualiseringerne, er AU- og KU-respondentgrupperne på den ene side generelt enige om de geografisk lokaliserbare betegnelser *københavnsk* og *jysk*, med nogle relativt små forskelle. AU-respondenterne markerer oftere hele Sjælland som *københavnsk*, mens flere af KU-respondenterne er mere præcise, til trods for at flere af dem også angiver Nordsjælland og dele af Sydsjælland. Begge respondentgrupper angiver at *jysk* tales i hele Jylland. For KU-respondenternes vedkommende afviger det fra deres svar på opgave 1 hvor *jysk* er forskelligt fra *sønderjysk*, hvis afgrænsning endda ligger længere mod nord (jf. afsnit 4.1). På den anden side er begge grupper kun delvis enige om de umiddelbart ulokaliserbare betegnelser *dialekt* og *rigsdansk*. Ifølge respondenterne findes *dialekt* i hele Danmark, og de udpeger også særskilt Sønderjylland, Bornholm og Fyn som de steder hvor man taler (mest) *dialekt*. I tillæg ser vi tilsyneladende en nærhedseffekt: AU-respondenterne peger på *dialekt* i Jyllands vestlige og nordlige egne, og i en grad vi ikke genfinder hos KU-respondenterne der imidlertid, men i lidt mindre grad, udpeger det sydlige Sjælland. I tidligere studier anvender informanter ofte betegnelsen *dialekt* ('at tale bredt') i opposition til *rigsdansk* ('at tale fint') som hinandens modsætninger (fx Monka 2013; Hansen 2022), og set i dét lys kan respondenternes placeringer af *dialekt* tolkes som en markering af modsætningsforholdet mellem *dialekt* (substandard) og deres eget sprog (jf. også fx Preston 1999:xxxv; Montgomery 2012), og dermed en distinktion mellem 'center' og 'periferi'.

Et bemærkelsesværdigt resultat ses i forbindelse med respondenternes placering af *rigsdansk*. Som det fremgår af figur 7, kan vi observere to forskellige tendenser. Begge respondentgrupper angiver at *rigsdansk* tales i hele landet, og det stemmer i og for sig overens med Brink & Lunds (1975, 2018:197) definition hvor *rigsdansk* betragtes som ideologisk ulokaliserbart (jf. også figur 4, kort 10 i afsnit 4.1). Farvekoncentrationen er imidlertid forskellig på de to kort, dvs. at AU- og KU-respondentgrupperne forstår noget forskelligt ved *rigsdansk*. AU-respondenterne fremhæver primært et østjysk område hvor de mener der tales *rigsdansk*, og dernæst nærmest hele Sjælland. Imens fremhæver KU-respondenterne hovedparten af Sjælland, inkl. København, som det sted hvor de mener folk taler *rigsdansk*. Det betyder at begge respondentgrupper er enige om at *rigsdansk* ikke er geografisk neutralt, men altså lokaliserbart. Dette stemmer umiddelbart overens med resultaterne af de undersøgelser der sætter spørgsmålstegn ved ideen om København som eneste, entydige sproglige normcenter i Danmark (jf. også fx Svenstrup & Thøgersen 2009; Hansen & Goldshtein 2021; Hansen 2022), og det til trods for at vores eksperiment fremdrager respondenters bevidste opfattelser af sproglige varieteter i Danmark og deres indbyrdes hierarki. Hvor *rigsdansk* placeres, tyder vores resultater på, afhænger altså af hvor respondenterne kommer fra.

5. Opsamling og fremtidig forskning

I denne artikel har vi præsenteret et perceptuel-dialektologisk pilotforsøg hvori vi har afprøvet tegn-et-kort-metoden til at belyse sprogbrugeres opfattelser af og holdninger til sproglig variation i Danmark. Resultaterne af undersøgelsen tegner et billede af et sociolingvistisk sproglandskab som overordnet set er mere komplekst end antaget i tidligere forskning. Respondenterne identificerede sproglig variation i form af nogle overordnede, geografisk lokaliserbare varieteter, og anvendte tilsvarende betegnelser om disse opfattede sproglige forskelle og ligheder der til en vis grad stemmer overens med sprogforskernes betegnelser. Fx var AU-respondenternes inddeling af sproglig variation i Jylland mere finmasket end KU-respondenternes, og det omvendte gjaldt for KU-respondenternes inddeling af Sjælland. Respondenterne var dermed tilbøjelige til at beskrive det område de selv kommer fra, mere nuanceret og præcist. Undersøgelsen vender ikke fuldstændig op og ned på det sprogideologiske billede fra tidligere sociodialektologisk forskning, men den føjer et nyt perspektiv til den igangværende udforskning af sociolingvistisk variation og forandring i Danmark, bl.a. ved at kaste nyt lys på respondenters opfattelse af nogle af de mest gængse betegnelser vi sprogforskere anvender, herunder den ulokaliserbare betegnelse *rigsdansk*.

Grundet pladsbegrænsninger har vi kun præsenteret en delmængde af resultaterne fra vores tegn-et-kort-eksperiment. Vi har bl.a. undladt at diskutere de holdningsdata respondenterne indtegnede på deres kort. Respondenterne delte relativt stereotype forestillinger om landsdelene når de blev spurgt til det; fx er Nordsjælland for en respondent ”lyden af penge”, og Sønderjylland lyder ”bondsk”. Dette materiale fortjener en nøjere granskning. Resultaterne af undersøgelsen er baseret på besvarelser fra 98 respondenter, og i fremtidig forskning bør projektet derfor skaleres op, og perspektiver fra flere og andre typer respondentgrupper bør inddrages (jf. også fx Montgomery 2022:165). Vi ønsker bl.a. at undersøge folkeskoleelever, der kan antages at være mere umiddelbare i deres tilgang til at besvare tegn-et-kort-opgaven end BA-studerende, og det vil ligeledes være interessant at inddrage andre grupper af voksne sprogbrugere og her undersøge om der er forskellige opfattelser ud fra om respondenterne kommer fra landet (fx Vestjylland, Nordjylland, Sønderjylland) eller fra byen (fx Aalborg, Odense), om der er forskel på hhv. ældre og yngre respondenters besvarelser etc. Nærværende pilotundersøgelse er derfor et første skridt i retning af at indføre tegn-et-kort-metoden som et nyt tiltag i sociodialektologien i Danmark.

Litteratur

- Bounds, Paulina & Charles J. Sutherland (2018) Perceptual basemaps reloaded: The role basemaps play in eliciting perceptions. I *Journal of Linguistic Geography* 6:145-166.
- Brink, Lars & Jørn Lund (1975) *Dansk Rigsmål. Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København*. Bd. 2. København: Gyldendal.
- Brink, Lars & Jørn Lund (2018) Yngre nydansk. I Ebba Hjorth et al. (red.). *Dansk Sproghistorie*. Bd. 2. København & Aarhus: Det Danske Sprog- og Litteraturselskab & Aarhus Universitetsforlag:197-228.
- Bøegh, Kristoffer Friis (2022) Om dialektcykelturen i 1940 og udforskningen af artikelgrænsen mellem øst- og vestjysk. I *Ord & Sag* 42:31-44.
- Bøegh, Kristoffer Friis, Peter Bakker, Inger Schoonderbeek Hansen & Carsten Levisen (2023) The Beginning of Quantitative Sociolinguistics in the Nineteenth Century: The Dane Anker Jensen (1878–1937) and his pioneering study “The Linguistic Situation in the Parish of Aaby, Aarhus County” (1898). I *Historiographia Linguistica* 49(2/3). DOI: 10.1075/hl.00114.boe
- Cramer, Jennifer (2021) Mental maps and perceptual dialectology. I *Language and Linguistics Compass* 15(2):1-15.
- Hansen, Inger Schoonderbeek & Yonatan Goldshtein (2021) Popular geopolitics and the conceptualization of linguistic norm centres. I *Journal of Postcolonial Linguistics* 5:204-225.
- Hansen, Inger Schoonderbeek (2022) Findes der en særlig jysk sprogidentitet? I *NyS* 61(3):73-125.
- Jysk Ordbog (2000-) Sidst tilgået den 11.01.2023 på <http://jyskordbog.dk/jyskordbog/ordbog.html>
- Kristiansen, Tore (1991) Sprogholdninger hos folkeskolelærere, unge mennesker og personalechefer på Næstvedegnen. I *Danske Folkemål* 33:51-62.
- Kristiansen, Tore (2009) The macro-level social meanings of late-modern Danish accents. I *Acta Linguistica Hafniensia: International Journal of Linguistics* 41:167-192.
- Kristiansen, Tore (2018) Sociodialectology. I Charles Boberg, John Nerbonne & Dominic Watt (red.). *The Handbook of Dialectology*. Hoboken: Wiley-Blackwell:106-122.
- Kristiansen, Tore (2020) Sprogets væsen og det gode sprogsamfund – som jeg har set det. I *Danske Talesprog* 20:7-32.
- Ladegaard, Hans Jørgen (1999) Dansk rigsmål i et socialpsykologisk perspektiv: En diskussion og empirisk analyse af rigsmålsbegrebet. I *Danske folkemål* 3:3-43.
- Maegaard, Marie (2007) Udtalevariation og -forandring i københavnsk: En etnografisk undersøgelse af sprogbrug, sociale kategorier og social praksis blandt unge på en københavnsk folkeskole. I *Danske Talesprog* 8. København: Hans Reitzels Forlag.
- Maegaard, Marie, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (2020) *Standardization as Sociolinguistic Change. A Transversal Study of Three Traditional Dialect Areas*. New York: Routledge.
- Maegaard, Marie & Pia Quist (2020) Perception, Recognition, and Indexicality. Experimental Investigations of Variation in Northern Jutland, Southern Jutland, and in Bornholm. I Marie Maegaard, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (red.). *Standardization as Sociolinguistic Change. A Transversal Study of Three Traditional Dialect Areas*. New York: Routledge:145-168.

- Monka, Malene (2013) Sted og sprogforandring: en undersøgelse af sprogforandring i virkelig tid hos mobile og bofaste informanter fra Odder, Vinderup og Tinglev. I *Danske Talesprog* 13. København: Museum Tusulanum.
- Montgomery, Chris (2012) The effect of proximity in perceptual dialectology. I *Journal of Sociolinguistics* 16:638-668.
- Montgomery, Chris & Philipp Stoeckle (2013) Geographic information systems and perceptual dialectology: A method for processing draw-a-map data. I *Journal of Linguistic Geography* 1:52-85.
- Montgomery, Chris (2022) Perceptual Dialectology. I Ruth Kircher & Lena Zipp (red.). *Research Methods in Language Attitudes*. Cambridge: Cambridge University Press:160-182.
- Pedersen, Inge-Lise (2003) Traditional dialects of Danish and the de-dialectalization 1900–2000. I *International Journal of the Sociology of Language* 159:9-28.
- Preston, Dennis Richard (1982) Perceptual dialectology: Mental maps of United States dialects from a Hawaiian perspective. I *University of Hawaii Working Papers in Linguistics* 14 (2):5-49.
- Preston, Dennis Richard (1986) Five visions of America. I *Language in Society* 15:221-240.
- Preston, Dennis Richard (1989) *Perceptual dialectology: Non-linguists' view of aerial linguistics*. Dordrecht: Foris.
- Preston, Dennis Richard (1999) Introduction. I Dennis Richard Preston (red.). *Handbook of Perceptual Dialectology*. Amsterdam: John Benjamins Publishing Company:xxiii-xl.
- Preston, Dennis Richard (2018) Perceptual dialectology. I Charles Boberg, John Nerbonne & Dominic Watt (red.). *The Handbook of Dialectology*. Oxford: Wiley & Sons:177-203.
- QGIS Development Team (2023) *QGIS Geographic Information System*. Sidst tilgået den 11.01.2023 på <http://qgis.osgeo.org>
- Quist, Pia (2005) *Stilistiske Praksisser i Storbyens Heterogene Skole. En etnografisk og sociolingvistisk undersøgelse af sproglig variation*. København: Museum Tusulanum.
- Svenstrup, Christoph Hare & Jacob Thøgersen (2009) On the discursive construction of ”rigsdansk”, ”jysk”, ”københavnsk” and other varieties among adolescents in western Jutland. I Marie Maegaard et al. (red.). *Language attitudes, standardization and language change: Perspectives on themes raised by Tore Kristiansen on the occasion of his 60th birthday*. Oslo: Novus forlag: 181-210.

Resumé:

Subjektellipser på gravstene og i patientjournaler

Christel Delia Christensen

Aarhus Universitet

Subjekt er obligatorisk i dansk sætningsstruktur, men der findes alligevel eksempler på, hvor subjektet er udeladt. Jeg har undersøgt fænomenet subjektellipser på gravstene og patientjournaler, da jeg observerede, at det forekommer ganske hyppigt. Det er data, som jeg selv har samlet via netværk og ture på århusianske kirkegårde, og dette data har jeg behandlet, hvori jeg har bestemt, om det drejer sig om situationelle, kontekstuelle eller strukturelle ellipser i givne situationer. Jeg har defineret subjektellipser ud fra Lars Brink og Jørgen Schacks definitioner. Væsentligt er det, at der skal være en fornemmelse af afkortethed, og man bør kunne rekonstruere det eliderede.

Litteratur

Brink, Lars & Schack, Jørgen (2014) Tale er sølv: Om begrebet syntaktisk ellipse. I *Danske Studier*, 45-75.

Jensen, Eva Skaftø (2014) Subjektellipse eller informationsstrukturel organisering i sms. I Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.) 15. møde om Udforskningen af Dansk Sprog:171-189.

Rathje, Marianne (2013) Dropper lige subjektet. Subjektellipser i kommentarfeltet på ekstrabladet.dk. I Simon Borchmann et al. (red.) *Gode ord er bedre end guld. Festskrift til Henrik Jørgensen*. Århus: Aarhus Universitet.

Togeby, Ole (2007) *Fungerer denne sætning? – Funktionel dansk sproglære* (1. e-bogsudgave). Aarhus: Academia.

Økuller? Det er ikke altid lige (u)acceptabelt at forlade en indlejring

Ken Ramshøj Christensen & Anne Mette Nyvad

Aarhus Universitet

1. Indledning: Knuder i generativ grammatik¹

I engelsk og dansk, og et væld af andre sprog, er det muligt at ekstrahere eller flytte et led ud af en indlejret sætning, så det optræder i begyndelsen af hovedsætningen, potentielt på tværs af en række sætningsgrænser, fx i spørgsmål, sml. a og b i (1) og (2), hvor det fremflyttede led og dets 'basisposition' er understreget. Sådanne sætninger kaldes på dansk "sætningsknuder" (Hansen 1984: 74, Hansen & Heltoft 2011: 333, 1812).

- (1) a. She said [that she knew [that he really liked chili]].
b. What did she say [that she knew [that he really liked ___]]?
- (2) a. Hun sagde [at hun vidste [at han virkelig godt kunne lide chili]].
b. Hvad sagde hun [at hun vidste [at han virkelig godt kunne lide ___]]?

Men der er også sætningstyper der tilsyneladende ikke (ret gerne) tillader ekstraktion, fx:

- (3) **Relativsætning**
- a. I saw the man [who baked that cake].
b. *Which cake did you see the man [who baked ___]?
(Carnie 2006: 48, eks. (6a-b))

¹ Tak til Eva Skaftø Jensen for konstruktiv kritik på en tidligere udgave af denne artikel, og tak til tilhørerne på MUDS 19 (14. oktober 2022) for spørgsmål og kommentarer. Forskningen blev delvist støttet af Det Frie Forskningsråd (Bevillings-ID: DFF-9062-00047B).

(4) **Hv-spørgsmål**

- a. Bill asked [what₁ Mary loves __₁]?
- b. *Who₂ did Bill ask [what₁ __₂ loves __₁]?
(Carnie 2006: 48, eks. (7b,d))

(5) **Adverbiel ledsætning**

- a. Did you leave [because Mary talked about that topic]?
- b. *Which topic did you leave [because Mary talked about __]?
(Szabolcsi & Lohndal 2017: 5, eks. (6b))

Udover sætninger såsom (3)-(5) er der også andre strukturer der ikke tillader ekstraktion. Mest velkendt er nok koordination (sideordning), der ikke tillader ekstraktion af det ene af de to koordinerede led:

(6) **Koordination**

- a. I liked [Mary and John].
- b. *Who did you like [Mary and __]?
- c. *Who did you like [__ and John]?
(Carnie 2006: 48, eks. (6a-c))

Siden Ross (1967) har man i generativ grammatik traditionelt omtalt sætninger og strukturer, der helt eller delvist blokerer ekstraktion, som fx (3)-(6), som *øer* (engelsk *islands*) (se fx Carnie 2006). Standardantagelsen er at disse øer er ugrammatiske (eller meget lidt acceptable) på tværs af alle sprog pga. visse universelle begrænsninger. Det er dog gentagne gange fremkommet modeksempler fra de fastlandsskandinaviske sprog (Bondevik m.fl. 2020, Christensen m.fl. 2013a, Christensen & Nyvad 2014, Engdahl & Ejerhed 1982, Erteschik-Shir 1973, Kush m.fl. 2019, Lindahl 2017, Müller 2019). Nedenfor gennemgår vi resultaterne fra vores undersøgelser af danske *hv*-spørgsmål, adverbielle ledsætninger og relativsætninger. Vi sætter spørgsmålstegn ved at disse konstruktioner er syntaktiske 'øer' på tværs af alle sprog, og diskuterer hvordan vi kan redegøre for dem uden at smide al syntaktisk teori overbord. Men først:

2. Modeksempler i traditionel dansk grammatik

I 1894 og 1911 beskriver Mikkelsen sætningsknuder som "sammenslyngede" sætninger (Mikkelsen 1894: 441, Mikkelsen 1911: 669). Det er interessant at der blandt Mikkelsens eksempler

er adskillige eksempler på flytninger fra øer: I (7) er der relativering ud af en relativsætning, og i (8)-(10) er der flytning fra et indlejret spørgsmål:

(7) **Topikalisering ud af en relativsætning**

- a. Han er en Person, som du næppe kan nævne nogen
[der kan maale sig med ___].
(Mikkelsen 1894: 422)

(8) **Topikalisering ud af et *hv*-spørgsmål**

- a. Ham₂ ved jeg nok [hvem₁ ___₂ er ___₁].
b. Dem₂ har jeg aldrig hørt [hvor₁ ___₂ er ___₁].
(Mikkelsen 1894: 322, se også Mikkelsen 1911: 671)

(9) **Relativering ud af et *hv*-spørgsmål**

- a. Det er en Dreng, som₂ jeg ikke kan finde paa
[hvad₁ jeg skal stille op ___₁ med ___₂].
b. Hvad er det, som₂ du ikke ved [hvor₁ ___₂ er ___₁]?
(Mikkelsen 1894: 422)
c. Hvilken Opgave er det, som₂ du ikke ved
[hvorledes₁ du skal løse ___₂ ___₁]?
(Mikkelsen 1911: 671)

(10) ***Hv*-flytning ud af et *hv*-spørgsmål**

- a. Hvad₂ ved du ikke [hvor₁ ___₂ er ___₁]?
(Mikkelsen 1894: 422)
b. Hvilken opgave₂ ved du ikke [hvorledes₁ du skal løse ___₁ ___₂]?
(Mikkelsen 1911: 671)

Ifølge Mikkelsen (1894: 322, 1911: 671) er relativering at foretrække fremfor *hv*-flytning. Han angiver eksempel (9)b-c og skriver ”ikke så gerne” om de tilsvarende (10)a-b. Det er værd at bemærke at Mikkelsen eksplicit angiver hvis noget er umuligt, så ”ikke så gerne” betyder sandsynligvis ikke ”umuligt” men nok bare ”mindre godt”.

Erik Hansen (1984: 75) angiver også en række eksempler der viser at ekstraktion fra visse øer er muligt i dansk:

(11) **Topikalisering ud af en relativsætning**

- a. Her har jeg en god ven [der bor ___].
- b. Det kender jeg da en fyr [der gør ___ hver dag].

(12) **Topikalisering ud af et *hv*-spørgsmål**

- a. Den₂ har jeg længe ville spørge dig om [hvor₁ ___₂ hører hjemme ___₁].
- b. De 50 kr.₂ er det ligemeget [hvornår₁ jeg får ___₂ igen ___₁].
- c. Konfliktforskning₂ er der en del tvivl om [hvad₁ ___₂ egentlig er ___₁].

Interessant nok skriver Hansen (1984: 75) også at det ikke er muligt at ”danne knude med hvilken som helst sætning” og angiver eksempel (13) nedenfor, der har flytning fra en adverbial ledsætning, som ugrammatisk:

- (13) *Den blev jeg ærligt talt sur [da han forlangte penge for ___].

Hvad der bestemmer om ekstraktion er mulig eller ej, er ikke helt klart. ”Om det er muligt afhænger af ledsætningens art og den måde den er indordnet i oversætningen på” (Hansen 1984: 75). Ifølge Hansen & Heltoft (2011: 1418) kan der ”ikke formuleres faste regler for fra hvilke ledsætninger materiale kan stilles på førstepladsen gennem sætningsknude.” De giver også en række eksempler på sætningsknuder der bl.a. involverer ø-konstruktioner med relativsætninger og kløvninger (2011: kap. II §76.2 og kap. XX §17). Men kontra Hansen (1984: 75) giver Hansen & Heltoft (2011: 1418) også følgende eksempler med flytning fra en adverbial ledsætning, som de skriver forekommer sjældent og ”og helt overvejende kun i talesprog”:

(14) **Topikalisering ud af en adverbial ledsætning**

- a. Den vase får du ballade [hvis du taber ___].
- b. Det skal du sige til [så snart du hører noget om ___].
- c. Den ligning er han meget stolt [fordi han kan løse ___].
- d. De sko kan jeg godt vente [mens du reparerer ___].

Ifølge Aage Hansen (1967: 110) er sætningsknuder ”outrerede” udtryk for ”ubehjælpssomhed, skødesløshed eller sproglig frækhed”. Mikkelsen (1894) angiver derimod eksemplerne uden at sige noget om at de skulle være stilistisk afvigende, og ifølge Hansen & Heltoft (Hansen &

Heltoft 2011: 1814) er der imidlertid intet uforklarligt eller uanalyserbart ved den danske sætningsknode, og den strider hverken mod fornuft eller almen grammatik”. Ifølge Nyvad m.fl. (2017) er sætningsknuder noget man mest finder i talesprog og uformelle, talesprogsagtige tekster. Kort sagt ser det i hvert fald ud til at dansk ikke følger universelle forbud mod ekstraktion fra øer der enten er relativsætninger, *hv*-spørgsmål eller en adverbial ledsætning. Men kunne eksemplerne ovenfor bare være isolerede tilfælde der ikke rigtigt afspejler hvordan dansk egentlig ser ud nu? For at kunne svare på det spørgsmål er vi nødt til at ty til eksperimentelle, kvantitative metoder.

3. Eksperimentelle undersøgelser af danske øer

3.1 Relativsætninger

At ekstraktion fra en relativsætning i engelsk er ugrammatisk (eller i bedste fald, meget uacceptabel, jf. (3)), er velkendt i den teoretiske litteratur. Det er også blevet bekræftet eksperimentelt (Christensen & Nyvad 2022). Det samme gælder for *hv*-flytning fra et indlejret *hv*-spørgsmål i engelsk (Christensen & Nyvad 2019). Så her synes der at være overensstemmelse mellem den teoretiske og empiriske litteratur. Selvom der, som nævnt, gentagne gange er fremført modeksempler til antagelsen om universelle ø-konstruktioner med data fra de fastlandsskandinaviske sprog, er de ofte blevet behandlet som ’falske’ modeksempler. Ideen er at der fx slet ikke er tale om ’rigtige’ relativsætninger, men i stedet en slags komplementsætning, nemlig en *small clause* der svarer til et frit prædikat (Kush & Lindahl 2011, Kush m.fl. 2013). Christensen & Nyvad (2014) undersøgte denne antagelse og fandt ikke belæg for at sige at relativsætninger kategorisk blokerer for ekstraktion, eller at ekstraktion afhænger af om matrixverbet kan tage en *small clause* som objekt. På dansk er *som* tvetydigt eftersom det kan introducere enten en subjekts- eller objektsrelativsætning eller et frit prædikat, mens *der* kun indleder subjektsrelativer:

- (15) a. Jeg foretrak/kyssede alle dem [**som/der** kender dig]. (subj.-rel.)
b. Jeg foretrak/kyssede alle dem [**som/*der** du kender]. (obj.-rel.)
c. Jeg foretrak/??kyssede dem [**som/*der** samtalepartnere]. (frit præd.)

I undersøgelsen blev 112 deltagere præsenteret for en række sætninger som (16)-(17) som de skulle vurdere på en skala fra 1 (helt uacceptabel) til 7 (helt acceptabel). Den ene halvdel af sætningerne havde et matrixverbum der i et forstudie scorede højt med et frit prædikat som

komplement, fx *anmelde*, den anden halvdel havde et matrixverbum der scorede lavt, fx *genere*. Hver sætning optrådte med og uden fremflytning og med *som* eller *der*:

- (16) Marie har måske anmeldt/generet en butiksejer
[som/der ansatte den slags personer].
- (17) Den slags personer har Marie måske anmeldt/generet en butiksejer
[som/der ansatte ___].

Resultaterne viste stor variation i acceptabiliteten. Fremflytning reducerer kraftigt acceptabiliteten, men effekten afhænger også (men i væsentligt mindre grad) af matrixverbets leksikalske frekvens og af hvor mange gange forsøgsdeltagerne så det i forsøget (en repetitionseffekt).² Selvom disse effekter er små, er de statistisk signifikante og kan tages som tegn på at ekstraktion fra en relativsætning faktisk er grammatisk mulig, ikke bare en 'fejl'. Hvis en sætningskonstruktion er ugrammatisk (dvs. umulig), burde manipulationer på det leksikalske niveau ikke kunne forbedre acceptabiliteten (Hofmeister & Sag 2010). Den lave acceptabilitet skyldes sætningens høje strukturelle kompleksitet og påvirkes af ordenes hyppighed (se også Christensen & Nyvad 2023b, 2023a).

Der er ofte stor forskel på hvor højt forskellige sætningstyper scorer i acceptabilitetsbedømmelser i kontrollerede forsøg, og hvor ofte de optræder i naturlig sprogbrug. Selvom objektsekstraktion fra en relativsætning i ovenstående forsøg (Christensen & Nyvad 2014) i gennemsnit reducerede acceptabiliteten fra 5,9 til 2,6 på skalaen fra 1-7, er det fx meget let at finde eksempler på internettet. Der er knap en halv million eksempler bare med *men det er der ingen/mange der/som* på Google (søgning med anførselstegn og site.dk). Et af problemerne med kontrollerede forsøg er netop at der skal kontrolleres for en lang række faktorer på tværs af en lang række sætninger for at holde 'alt andet lige', hvilket gør det sværere at gøre eksemplerne helt naturlige trods ihærdige forsøg.

² Ordfrekvens spiller ofte en rolle. Den altafgørende faktor der afgør om danskere laver særskrivningsfejl af sammensatte ord (fx *næse operation* i stedet for *næseoperation*) er frekvens: lavere frekvens giver flere fejl (Andersen & Diderichsen 2011). Der er også en sammenhæng mellem leksikalsk frekvens og sætningers acceptabilitet (Christensen & Nyvad 2023a).

3.2 Hv-spørgsmål

I en række forsøg, både med acceptabilitetsbedømmelse på en skala fra 1 til 5, med binær grammatikalitetsbedømmelse og med hjerneskaning med fMRI, undersøgte Christensen m.fl. (2013a, 2013b) *hv*-flytning ud af indlejrede *hv*-spørgsmål. Ligesom ved relativsætninger reduceres acceptabiliteten kraftigt ved lang flytning (dvs. ved knudedannelse) fra en indlejret deklarativsætning, der jo helt ukontroversielt anses for at være grammatisk, fx (19): Det er muligt at spørge enten til det indlejrede objekt, *hvad* i (19)a, eller til et indlejret adverbial, *hvor* i (19)b, men det er signifikant værre at flytte *hvor* end *hvad*. (Bemærk at *hvor* ikke kan læses som et hovedsætningsadverbial: **Hvor ved hun mon det?* **Hun ved det i stuen.*)

- (18) Ved hun godt at man kan leje noget dér?
- (19) a. Hvad ved hun godt [at man kan leje __ dér]?
b. Hvor ved hun godt [at man kan leje noget __]?
- (20) a. Hvad₂ ved hun godt [hvor₁ man kan leje __₂ __₁]?
b. Hvor₂ ved hun godt [hvad₁ man kan leje __₁ __₂]?
- (21) *Ved hun godt [hvor₂ hvad₁ man kan leje __₁ __₂]?

Prisen for *hv*-flytningen i (19) var en gennemsnitlig reduktion på 28,2 % i acceptabiliteten (fra 4,78 på 5-punktsskalaen i (18) til 3,43). Ulig flytningen i (19) er der ikke forskel på hvor acceptabelt det er at flytte *hvad* eller *hvor* ud af et indlejret *hv*-spørgsmål som i (20). Prisen for at flytte ud af *hv*-øen er en yderligere reduktion på 28,6 % i forhold til (19) til 2,45. Selvom det er ret lavt, er det signifikant bedre end den ugrammatiske kontrolsætningstype i (21), der scorede i bund med 1,48 på skalaen fra 1 til 5. Det er interessant at bemærke at der var små men statistisk signifikante repetitionseffekter på acceptabiliteten både ved knuderne i (19) og ved ø-ekstraktionerne i (20) (de blev vurderet en lille smule højere over tid), mens det ikke var tilfældet for (21). Det tyder også på at flytning fra en *hv*-ø er grammatisk mulig, men svær pga. kompleksiteten, hvilket kan afhjælpes en smule ved gentaget brug. Igen passer det fint sammen med Mikkelsens (1894) observation, jf. diskussion af (10) ovenfor, og igen viser data at dansk har en mere fleksibel ordstilling end ofte antaget, navnlig i den generative litteratur. Teoretisk set, burde *hv*-flytning fra en *hv*-ø ikke være mulig. Det vender vi tilbage til i afsnit 4 nedenfor.

3.3 Adverbielle ledsætninger

Poulsen (2008) var den første til at undersøge ekstraktion fra adverbielle ledsætninger eksperimentelt i dansk. Hans undersøgelse fokuserede på *hvis*-sætninger og viste meget lave acceptabilitetsbedømmelser ved ekstraktion, uafhængigt af semantisk kohæsion mellem matrix- og underordnet sætning (kontra Jensen 2002, se også Erteschik-Shir 1973, 2006, Van Valin & LaPolla 1997). Der var altså ikke signifikant forskel på *Den whisky bliver hun fuld hvis hun bærer inden kampen* og *Den kam bliver hun fuld hvis hun taber inden kampen*. Poulsen (2008) konkluderede derfor (forsigtigt) at ekstraktion fra adverbielle ledsætninger nok bør betragtes som simple talefejl.

Nyvad m.fl. (2022) undersøgte relativering ud af tre forskellige slags adverbielle ledsætninger i engelsk. Vi brugte relativering da det er velkendt at topikalisering er væsentligt mere markeret på engelsk end på dansk, og vi ville gerne kunne sammenligne de to sprog på lige vilkår. 235 deltagere med engelsk som modersmål skulle vurdere en række sætninger med og uden fremflytning fra indlejrede sætninger (se de danske eksempler i (22)-(24) nedenfor) på en skala fra 1 til 7: komplementsætninger med *that* 'at' og tre slags adverbielle ledsætninger: betingelsessætninger med *if* 'hvis', tidsadverbialer med *when* 'da' og betingelsessætninger med *because* 'fordi'. (Forsøget havde også en række kontrolsætninger, fx klart ugrammatiske sætninger med flytning fra et koordineret led, såsom *This is the dog that I was ashamed because I actually lost both the cat and ___ in one day.*)

Igen udløste ekstraktion i sig selv en signifikant reduktion af acceptabiliteten, men overraskende nok blev flytning fra *that*- og *if*-sætninger vurderet lige højt (4,97 på 7-punktsskalaen). Flytning fra en *when*-sætning var væsentligt mindre acceptabelt (vurderet til 3,92), og det var endnu værre med flytning fra *because*-sætninger (vurderet til 3,39), men dog signifikant bedre end flytning af et koordineret led (vurderet til ca. 1,8) (sml. også (5)b og (6)b ovenfor). Der var også en signifikant repetitionseffekt for ekstraktion fra *when*-sætninger. Kort sagt viser disse resultater at engelske adverbielle ledsætninger, navnlig *if*-sætninger, opfører sig væsentligt mindre ø-agtigt end forventet.

Nyvad m.fl. (2023) oversatte det engelske forsøg til dansk og gennemførte præcis det samme forsøg på dansk med 335 deltagere. Også her skulle forsøgsdeltagerne vurdere en række sætninger som (23)-(24) på en skala fra 1 (helt uacceptabel) til 7 (helt acceptabel). Hver sætning blev præsenteret med en kontekst (22) for at gøre det så naturligt og talesprogsagtigt som mu-

ligt. Sætninger blev vist både uden relativering, som i (23), og med relativering (eller kløvning), som i (24).³

- (22) Kontekst: I det sidste træningsprogram jeg udarbejdede for Emma, ville jeg gøre det så godt som umuligt for hende og inkluderede derfor endnu et sæt virkelig brutale pull-ups.
- (23) a. Det er åbenlyst, at jeg blev overrasket over,
[at hun faktisk gennemførte dét program].
- b. Det er åbenlyst, at jeg ville blive overrasket,
[hvis hun faktisk gennemførte dét program]
- c. Det er åbenlyst, at jeg blev overrasket,
[da/fordi hun faktisk gennemførte dét program].
- (24) a. Det er dét program som jeg blev overrasket over,
[at hun faktisk gennemførte ____].
- b. Det er dét program som jeg ville blive overrasket,
[hvis hun faktisk gennemførte ____].
- c. Det er dét program som jeg blev overrasket,
[da/fordi hun faktisk gennemførte ____].

Baseret på de tidligere forsøg med fremflytning i dansk og på de mange tilfælde med ekstraktioner fra øer i traditionel dansk grammatik forventede vi at få sammenlignelige resultater i de danske og engelske forsøg. Som forventet, er det relative acceptabilitetshierarki næsten det samme: Relativering fra *hvis*-sætninger er signifikant bedre end fra *da*-sætninger og *fordi*-sætninger, som dog ikke er signifikant forskellige. Interessant nok er der ikke nogen effekt af ekstraktion fra en *at*-sætning. Endnu mere interessant og ret overraskende er at ekstraktion fra *hvis*-sætninger kun vurderes til 3,45, og ekstraktion fra *da*- og *fordi*-sætninger vurderes endnu lavere: ca. 2,8 på skalaen fra 1 til 7. Det er noget lavere end de engelske niveauer, og danske adverbial-øer forekommer altså til at være mere ø-agtige. Forsøget blev kørt tre gange: relative-

³ For at undersøge om konteksten evt. kunne være unaturlig og dermed have en negativ effekt på bedømmelserne, kørte vi det samme forsøg igen bare uden kontekst, og det gav stort set samme resultat. Der var 11 kontekster/sæt med "åbenlyst" og 13 med "klart". Det gjorde vi bevidst for variationens skyld. Vi medgiver at det kunne være relevant for resultatet, men vi kørte forsøget igen (for tredje gang) hvor der ikke var relativering, så sætningerne begyndte med "Så jeg blev overrasket..." i stedet for "Det er åbenlyst...". Resultatet var igen stort set det samme.

ring med kontekst (som vist ovenfor), relativering uden kontekst og topikalisering med kontekst (*Det program blev jeg overrasket...*), og med stort set det samme resultat.

Naturlige eksempler med flytninger fra *hvis*- og i mindre grad *da*-sætninger er lette nok at finde, mens det ikke er tilfældet for *fordi*-sætninger (både i dansk og tilsvarende engelsk) (Müller & Eggers 2022), så de lave acceptabilitetsniveauer er lidt overraskende, navnlig for *hvis* og *da*. I Nyvad et al. (2023) foreslås det at det måske skyldes (i) den velkendte negative effekt af at der er tale om et kontrolleret forsøg (som beskrevet ovenfor), og (ii) at der er mange flere der taler engelsk som andetsprog end dansk. Måske er engelsktalende mere vant til variation i sprogsystemet end dansktalende er, og derfor sprogligt mere 'tolerante'. Det skal dog siges at forsøgene ikke viser at folk med engelsk som modersmål tillader hvad som helst: Forsøgene med flytning fra relativsætninger og *hv*-spørgsmål viser netop at det ikke er tilfældet (Christensen & Nyvad 2019, 2022). Relativsætninger og *hv*-spørgsmål er øer der er lettere at undslippe i dansk end i engelsk, mens det modsatte ser ud til at være tilfældet for adverbelle ledsætninger.

3.4 Koordination

Umiddelbart kunne det måske godt se ud som om de fleste flytninger fra øer er nogenlunde acceptable i dansk (eller i hvert fald ofte bedre end i engelsk), og at alt er relativt. Men det er nu ikke rigtigt tilfældet. Vi har gentagne gange nævnt, men ikke rigtigt diskuteret, flytning fra koordination. I (25) ses danske eksempler der svarer til de engelske i (6) ovenfor:

- (25) a. Jeg kan godt huske [Susi og Leo].
b. *Hvem kan du godt huske [Susi og ___]?
c. *Hvem kan du godt huske [___ og Leo]?

At flytning af det ene eller det andet sideordnede led er ugrammatisk, og ikke bare lettere uacceptabelt af forskellige årsager, er blevet bekræftet eksperimentelt.⁴ Sætninger som (26)b og (27)b er blevet vurderet til hhv. 1,32 og 1,44 i gennemsnit på en skala fra 1 til 7 i forsøg med hhv. 212 og 335 deltagere (se Nyvad m.fl. 2022 for engelsk).

⁴ Det er til muligt at flytte *dele* af *begge* sideordnede led samtidigt (såkaldt *across-the-board movement*):

- (i) Den slags mad kan jeg [[godt lide ___] men [ikke rigtigt tåle ___]].
(ii) Jeg gad godt vide hvilken slags film [hun hader ___] men [han elsker ___].

Økuller? Det er ikke altid lige (u)acceptabelt at forlade en indlejring

- (26) a. Men jeg kender da allerede hunden og barnet.
b. *Men barnet jeg kender hunden og ___ da allerede.
(Christensen & Nyvad 2023a)
- (27) a. Det er åbenlyst, at jeg blev rasende over at
han alligevel læste både [journalen og dagbogen] i mit fravær.
b. *Det er dagbogen, som jeg blev rasende {over at / hvis / da / fordi}
han alligevel læste både [journalen og ___] i mit fravær.
(Nyvad m.fl. 2023)

4. Regler og sammensværgelser

I den generative litteratur er der traditionelt blevet argumenteret for at der er bestemte universelle regler eller begrænsninger der blokerer for flytning fra de forskellige slags øer. Ross (1967) foreslog en række regler for øer, herunder *The Coordinate Structure Constraint* (CSC), der blokerer for flytning af et sideordnet led eller dele af et sådant led. Ross foreslog også *The Complex Noun Phrase Constraint* (CNPC), der blokerer for flytning fra relativsætninger og sætninger der er komplement til et substantiv (fx *påstanden (om) at Jorden er flad* og *den ide at han vil starte en robotfabrik*). Chomsky (1964) foreslog at flytning fra et indlejret *hv*-spørgsmål blokeres af *The Wh-Island Condition*.⁵ Huang (1982: 505) *Condition on Extraction Domain* siger (kort sagt) at det kun er muligt at ekstrahere fra en sætning der ikke er del af et subjekt, eller er (del af) et adverbialt led. Senere er disse begrænsninger eller regler (undtagen CSC) blevet forsøgt samlet under Chomskys (1973) *Subjacency Condition*, der i en nøddeskal siger at *hv*-flytning maksimalt kan krydse (dvs. flytte igennem) en enkelt sætningsgrænse eller en DP-grænse. I (3)b krydser fremflytningen af *which cake* først DP-grænsen ved *the* og derefter sætningsgrænsen ved subjektet *you*, og dermed er *Subjacency* overtrådt. Ved flytning fra et indlejret spørgsmål som fx i (4)b skal *who* krydse to sætningsgrænser, hvilket også overtræder *Subjacency*. I nyere teoretisk generativ grammatik forklares ø-fænomener som overtrædelser af det mere generelle princip *Locality* (Chomsky 1995, Rizzi 1990). I *hv*-øerne og relativsætnings-øerne ovenfor krydser det fremflyttede *hv*-element et andet fremflyttet *hv*-element. Den lange flytning til starten af matrixsætningen er derfor ikke lokal, hvilket gør sæt-

⁵ Ross (1967: 27–28) argumenterede derimod for at *The Wh-Island Condition* ikke kategorisk blokerer for ekstraktion, men at det snarere er et gradspørgsmål. Han giver bl.a. følgende eksempel: *Which books did he tell you why/?whether/?when he wanted to read*. Desuden spiller det en væsentlig rolle om den indlejrte sætning er finit eller ej. Infinitte sætninger er mere åbne for ekstraktion.

ningen ugrammatisk. Flytningen er kun mulig hvis der kan 'mellemlandes' i en tom position ved venstrekanten af den indlejrede sætning, som i (1)b.

Som det tydeligt fremgår af mængden af data fra dansk ovenfor, kan der ikke være tale om universelle og ubrydelige syntaktiske regler der 'forbyder' eller decideret blokerer ekstraktion. Der findes også mange forsøg på at forklare fænomenet vha. ikke-syntaktiske faktorer. Fx er det foreslået at lange ekstraktioner generelt, og flytninger fra i hvert fald indlejrede spørgsmål og relativsætninger i særdeleshed, er svære at processere pga. at arbejdshukommelsen overbebyrdes. Der er simpelthen for mange diskursreferenter (Gibson 1998, 2006) og/eller syntaktiske konstituenten (Hawkins 1994, 2004, Christensen m.fl. 2013a) mellem udgangspunktet for flytningen og endestationen i starten af hovedsætningen. For at forklare forskellen på fx dansk, der ser ud til at tillade flytning fra øer, og engelsk, der i langt mindre grad gør, er det foreslået at der i dansk er hvad man kunne kalde et 'ekstra forfelt' i indlejrede sætninger – en ekstra tom plads der kan bruge som 'flugtvej' fra øen (Nyvad m.fl. 2017, Vikner m.fl. 2017). Det er således muligt at 'flygte' fra øen, men det er dyrt i arbejdshukommelse pga. 'den lange vej' fra start til slut, og resultatet er at sætningen ikke vurderes til at være helt acceptabel.

For flytninger fra adverbelle ledsætninger og relativsætninger er det bl.a. foreslået at det styrende princip er informationsstruktur. Ifølge Abeillé m.fl. (2020) skyldes ø-effekter en diskurs-konflikt (*Focus-Background Conflict*), der opstår når et element der er en del af baggrundsinformation, bliver fokaliseret ved fremflytning. Denne ide bygger på en lang tradition hvor man forsøger at forklare ø-effekter med primært pragmatiske faktorer: Den indlejrede sætning skal være 'dominant' for at tillade ekstraktion (Erteschik-Shir 1973); sætninger der er 'baggrund', er øer (Goldberg 2005); det fremflyttede element skal være 'relevant' for kernebudskabet i ytringen (Chaves & Putnam 2020). Men der er ingen af dem der rigtigt forklarer mønsteret. Udfordringen er at nogle ekstraktioner er (næsten) helt acceptable, mens andre i varierende grad er uacceptable. (Det er også et problem for en ren syntaktisk forklaring.) Adverbelle ledsætninger er typisk (men ikke altid) baggrund (se fx Abeillé m.fl. 2020: 3, Chaves & Putnam 2020: 235), og *hv*-flytning er fokusmarkering, mens relativisering ikke er det, så det burde give et lige dårligt resultat uanset hvilken type konjunktion der indleder sætningen.

Selvom det sikkert ville være mere tilfredsstillende at kunne pege på en enkelt faktor som udslagsgivende, skal forklaringen sandsynligvis snarere findes i en interaktion mellem syntaktisk struktur (og arbejdshukommelse), (leksikalsk) semantik og pragmatik (Nyvad m.fl. 2022, 2023). Som beskrevet i afsnit 3.3 ovenfor, afhænger prisen (faldet i acceptabilitet) for at forlade en adverbial ledsætning af dens funktion: Flytning fra *hvis* (årsag) er bedre end fra *da* (tid) eller *fordi* (årsag), der dog stadig er væsentlig bedre end ugrammatiske kontrolsætninger. Desuden

er der forskel på engelsk *when* og *because*, men altså ikke på dansk *da* og *fordi*, og den sprogforskning er ikke umiddelbar let at forklare med informationsstruktur.

Adverbielle ledsætninger med *hvis*, *da* og *fordi* er alle potentielle øer fordi de jo netop er adverbielle, ikke en del af argumentstrukturen. Men de har forskellige særlige egenskaber. *Hvis* (ligesom engelsk *if*) danner sandsynligvis en *hv*-ø (Bhatt & Pancheva 2006, Haegeman 2010, Larson 1985), men *hvis* (ulig engelsk *if*) kan indgå i 'konjunktionsstabling' (rekursiv underordning, fx *hvis at det ikke virker..., det er bare fordi at..., det er mig som at der...*), og derfor er der teoretisk mulighed for en flugtvej (et smuthul). *Da* tillader ikke indlejret V2 (**da jeg havde ikke smadret vasen*), hvilket klart signalerer underordning (*da* kan ikke indlede en sætning med hovedsætningsledstilling og indleder obligatoriske en underordnet sætning). Men *da* tillader ikke stabling (jf. **da at jeg gik*), der ellers ser ud til at være kendetegnende for underordnede sætninger i dansk, og det kunne måske tyde på at der ikke er mulighed for en flugtvej, og at der dermed er tale om en ø der er svær at undslippe. Ligesom *hvis* kan *fordi* også stables (*fordi at du gik*), men på den anden side har *fordi* nogle træk tilfælles med hovedsætninger: Fx tillader de indlejret V2 og epistemiske modaludtryk (*fordi kagen har han sandsynligvis/helt sikkert/jo allerede spist*). Der kan således være noget i den syntaktiske struktur der gør ekstraktion mere besværligt og mindre acceptabelt.

Flytningen er altså i sig selv mulig, men krævende for arbejdshukommelsen og afhængig af funktion og informationsstruktur. Der er interaktion (en 'sammensværgelse') mellem disse faktorer, og derfor er det endnu ikke helt klart hvor meget de hver især bidrager med, men det er jo et empirisk spørgsmål.

5. Konklusion

Dansk er specielt spændende fordi syntaktiske øer gentagne gange har vist sig ikke at opføre sig som de burde hvis begrænsningerne for ø-ekstraktioner var absolutte og universelle. Vi har givetvis brug for både syntaksen, funktionen og performanssystemerne for at kunne begynde på at forklare hvorfor det ikke altid er lige (u)acceptabelt at forlade en indlejring. Grammatikken muliggør, arbejdshukommelsen begrænser og pragmatikken sandsynliggør.

Litteratur

- Abeillé, Anne, Barbara Hemforth, Elodie Winckel, og Edward Gibson (2020) Extraction from Subjects: Differences in Acceptability Depend on the Discourse Function of the Construction. *Cognition* 204:104293.
- Andersen, Margrethe Heidemann, og Philip Diderichsen (2011) Sjældne og sære sammensætninger. *NyS, Nydanske Sprogstudier*, nr. 41:40–65.
- Bhatt, Rajesh, og Roumyana Pancheva. 2006. Conditionals. I Martin Everaert og Henk van Riemsdijk (red.). *The Blackwell Companion to Syntax*. Malden, MA, USA: John Wiley & Sons, Ltd., 638–87.
- Bondevik, Ingrid, Dave Kush, og Terje Lohndal (2020) Variation in Adjunct Islands: The Case of Norwegian. *Nordic Journal of Linguistics* 44 (3):1–32.
- Carnie, Andrew (2006) Island Constraints. I Keith Brown (red.), *Encyclopedia of Language & Linguistics (Second Edition)*. Oxford: Elsevier, 48–51.
- Chaves, Rui P. og Michael T. Putnam (2020) *Unbounded Dependency Constructions: Theoretical and Experimental Perspectives*. Oxford: Oxford University Press.
- Chomsky, Noam (1964) *Current Issues in Linguistic Theory*: Hague: Mouton.
- Chomsky, Noam (1973) Conditions on Transformation. I Stephen R. Anderson og Paul Kiparsky (red.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, 232–86.
- Chomsky, Noam (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.
- Christensen, Ken Ramshøj, Johannes Kizach, og Anne Mette Nyvad (2013a) Escape from the Island: Grammaticality and (Reduced) Acceptability of *Wh*-Island Violations in Danish. *Journal of Psycholinguistic Research* 42 (1):51–70.
- Christensen, Ken Ramshøj, Johannes Kizach, og Anne Mette Nyvad (2013b) The Processing of Syntactic Islands – An fMRI Study. *Journal of Neurolinguistics* 26 (2):239–51.
- Christensen, Ken Ramshøj, og Anne Mette Nyvad (2014) On the Nature of Escapable Relative Islands. *Nordic Journal of Linguistics* 37 (1):29–45.
- Christensen, Ken Ramshøj, og Anne Mette Nyvad (2019) No escape from the island: On extraction from complement *wh*-clauses in English. I Ken Ramshøj Christensen, Henrik Jørgensen, og Johanna L. Wood (red.). *The Sign of the V – Papers in Honour of Sten Vikner*. Aarhus: Dept. of English, School of Communication & Culture, Aarhus University, 95–112.
- Christensen, Ken Ramshøj, og Anne Mette Nyvad (2022) The Island Is Still There: Experimental Evidence For The Inescapability Of Relative Clauses In English. *Studia Linguistica* 76 (3):1–25.
- Christensen, Ken Ramshøj, og Anne Mette Nyvad (2023a) Frequency, animacy, structural complexity, and acceptability. Ms. indsendt til bedømmelse.
- Christensen, Ken Ramshøj, og Anne Mette Nyvad (2023b) Sætningskompleksitet, frekvens og acceptabilitet. *Ny Forskning i Grammatik* 30: under udgivelse.
- Engdahl, Elisabet, og Eva Ejerhed (red.) (1982) *Readings on unbounded dependencies in Scandinavian languages*. Umeå: Universitetet i Umeå; Almqvist & Wiksell International.
- Erteschik-Shir, Nomi (1973) On the nature of island constraints. Ph.d.-afhandling. Cambridge, MA: MIT.
- Erteschik-Shir, Nomi (2006) Bridge Phenomena. I Martin B.H. Everaert og Henk C. van Riemsdijk (red.). *The Blackwell Companion to Syntax*. Malden, MA, USA: Blackwell Publishing, 284–94.
- Gibson, Edward (1998) Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68 (1):1–76.

- Gibson, Edward (2006) Sentence Comprehension, Linguistic Complexity In. I *Encyclopedia of Cognitive Science*. American Cancer Society.
- Goldberg, Adele E. (2005) *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Haegeman, Liliane (2010) The Internal Syntax of Adverbial Clauses. *Lingua* 120 (3):628–48.
- Hansen, Erik (1984) *Dæmonernes port: støttemateriale til undervisningen i nydansk grammatik*. 3. udg. København: Reitzel.
- Hansen, Erik, og Lars Heltoft (2011) *Grammatik over det Danske Sprog. bind I-III*. København: Det Danske Sprog- og Litteraturselskab.
- Hansen, Aage (1967) *Moderne Dansk. Bind I*. København: Grafisk Forlag.
- Hawkins, John A. (1994) *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. (2004) *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hofmeister, Philip, og Ivan A. Sag (2010) Cognitive constraints and island effects. *Language* 86 (2):366–415.
- Huang, Cheng-Teh James (1982) Logical Relations in Chinese and the Theory of Grammar. Ph.d.-afhandling, Cambridge, MA: MIT.
- Jensen, Anne (2002) Sætningsknuder i dansk. *NyS, Nydanske Sprogstudier* 29: 105–24.
- Kush, Dave, og Filippa Lindahl (2011) On the escapability of islands in Scandinavian. I *85th Annual Meeting of the Linguistic Society of America*. Pittsburgh, Pennsylvania.
- Kush, Dave, Terje Lohndal, og Jon Sprouse (2019) On the Island Sensitivity of Topicalization in Norwegian: An Experimental Investigation. *Language* 95 (3):393–420.
- Kush, Dave, Akira Omaki, og Norbert Hornstein (2013) Microvariation in islands? I Jon Sprouse og Norbert Hornstein (red.). *Experimental Syntax and Island Effects*. Cambridge: Cambridge University Press, 239–64.
- Larson, Richard K. (1985) Bare NP adverbs. *Linguistic Inquiry* 16 (4):595–621.
- Lindahl, Filippa (2017) Extraction from Relative Clauses in Swedish. Ph.d.-afhandling. Göteborg: Göteborg Universitet.
- Mikkelsen, Kristian (1894) *Dansk Sproglære, Med Sproghistoriske Tillæg: Haandbog For Lærere Og Viderekomne*. København: Lehman & Stages Forlag.
- Mikkelsen, Kristian (1911) *Dansk ordføjnings-lære med sproghistoriske tillæg. Håndbog for viderekomne og lærere*. København: Lehman & Stages Forlag.
- Müller, Christiane (2019) Permeable islands. A contrastive study of Swedish and English adjunct clause extractions. Ph.d.-afhandling. Lund: Lund Universitet.
- Müller, Christiane, og Clara Ulrich Eggers (2022) Island Extractions in the Wild: A Corpus Study of Adjunct and Relative Clause Islands in Danish and English. *Languages* 7 (2):125.
- Nyvad, Anne Mette, Ken Ramshøj Christensen, og Sten Vikner. 2017. CP-recursion in Danish: A cP/CP-analysis. *The Linguistic Review* 34 (3): 449–77. <https://doi.org/10.1515/tlr-2017-0008>.
- Nyvad, Anne Mette, Christiane Müller, og Ken Ramshøj Christensen (2022) Too True to Be Good? The Non-Uniformity of Extraction from Adjunct Clauses in English. *Languages* 7 (4):244.

- Nyvad, Anne Mette, Christiane Müller, og Ken Ramshøj Christensen (2023) Moving away from the island. Extraction from adjunct clauses in Danish. Under udarbejdelse.
- Poulsen, Mads (2008) Acceptability and Processing of Long-Distance Dependencies in Danish. *Nordic Journal of Linguistics* 31 (1):73–107.
- Rizzi, Luigi (1990) *Relativized minimality*. Cambridge, MA: MIT Press.
- Ross, John Robert (1967) Constraints on variables in syntax. Ph.d.-afhandling. Cambridge, MA: MIT.
- Szabolcsi, Anna, og Terje Lohndal (2017) Strong vs. Weak Islands. I Martin Everaert og Henk van Riemsdijk (red.). *The Companion to Syntax*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1–51.
- Van Valin, Robert D. Jr, og Randy J. LaPolla (1997) *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Vikner, Sten, Ken Ramshøj Christensen, og Anne Mette Nyvad (2017) V2 and cP/CP. I Laura Bailey og Michelle Sheehan (red.). *Order and structure in syntax I: Word order and syntactic structure*. Berlin: Language Science Press, 313–24.

Identitet i en stavemåde?

Om stavevariation i nyfødtes fornavne

Birgit Eggert

Københavns Universitet

1. Introduktion¹

Personnavne har altid optrådt i mange forskellige stavemåder. Eksempelvis optræder mandnavnet *Christen* i den danske folketælling fra 1801 i stavemåderne *Christen*, *Kristen*, *Chresten*, *Kresten* og *Cresten* (jf. Eggert 2009:104). Denne variation skyldes både usikkerhed fordi man ikke skrev særlig ofte, og manglen på en fælles stavenorm. I Danmark kom den første officielle retskrivningsordbog i 1872 (Grundtvig 1872), og i 1910 oprettede man Stednavneudvalget som blandt andet skulle fastsætte retskrivningsformer af danske stednavne (jf. Olesen 2022:11-14), men personnavne er aldrig blevet underkastet en retskrivning i Danmark. Selvom der til en vis grad er almindelige måder at stave de mest benyttede navne på, har variationen i stavemåder af personnavne helt frem til i dag været meget stor, og den adskiller sig ind imellem ganske meget fra den ortografiske norm som bruges til resten af sproget.

Som det fremgår af det følgende, ønsker flere og flere danskere i dag unikke navne. En måde at skabe unikhed på er at stave allerede velkendte navne på en lidt anderledes måde end vanligt, og antallet af stavemåder af allerede eksisterende navne, er stigende i Danmark. Dette stigende antal stavevarianter er ikke tidligere undersøgt for dansk navnemateriale, og det vil jeg gøre nogle indledende forsøg på i det følgende ved at sætte det i relation til det stigende antal unikke fornavne som er givet til nyfødte de senere år.

2. Teoretisk baggrund

2.1. Stavevariationens betydning

På et tidspunkt i løbet af 1900-tallet blev specifikke stavevarianter identitetsbærende på den måde at den enkelte navnebærer altid brugte den samme variantform af et navn. Hedder man for eksempel *Kristina*, kan man ikke længere finde på at skrive *Christina* om sig selv. Det er

¹ Dette er en forkortet, dansk udgave af artiklen ”Different spellings as uniqueness in today’s naming in Denmark” som er planlagt til udgivelse i *Beiträge zur Namenforschung* 2023.

uvist præcist hvornår denne udvikling er sket, og måske er den sket over en længere årrække (jf. Eggert 2017).

Stavevarianter som bevidste valg om variation har ikke haft noget stort fokus i den nordiske navneforskning og mig bekendt heller ikke i navneforskning uden for Norden; det er et relativt nyt emne i faglitteraturen. I en undersøgelse af forældres navnevalg til nyfødte børn i den svenske by Göteborg i 2007 kom Emilia Aldrin (2011a:41) frem til at 25 % af forældrene havde tænkt over stavemåden af barnets navn. I hendes videre forskning konkluderer Aldrin (2011b:89-90) på et materiale fra samme by i perioden 2007-2009 at "Genom att välja att stava namnet på ett lite oväntat sätt menar jag att dessa föräldrar positionerar sig som originella. ... Bland dem som svarade att de funderat på olika sätt att stava namnet motiverade 5 % (7 av 153) det slutgiltiga valet med att det var mer "ovanligt"."

For nogle år siden lavede jeg selv en lille undersøgelse af dansk, statistisk navnemateriale over navne givet til nyfødte i perioden 1985-2011 og kom frem til at man generelt foretrak stavemåder med *c* og *ch* fremfor stavemåder med *k* for lyden [k^h] (Eggert 2013), men undersøgelsen er ikke rettet mod varianter som er til for variationens skyld, men kun hvilke stavevarianter som er de foretrukne. I sin undersøgelse af forældres navnevalg til nyfødte i Danmark i 2012 spurgte Katrine Kehlet Bechsgaard (2015) ikke forældrene direkte om valget af stavemåde, men om forældrene synes det var vigtigt at navnet så godt ud på skrift. Det ser imidlertid ikke ud til i undersøgelsen at det havde nogen stor betydning for de adspurgte forældre. I en fremstilling om personnavne i Tyskland skrives desuden at "Die onymische² Abweichung von der orthographischen Norm ist fast Normalität (und generiert eine Vielzahl an Namenvarianten, mit denen sich die Namenträger stark identifizieren)" (Nübling et al. (2015:87). Det vil altså sige at navnenes stavemåder er meget væsentlige og et særkende for navnebærerne.

2.2. Udviklingen mod en høj grad af individualitet i navngivningen

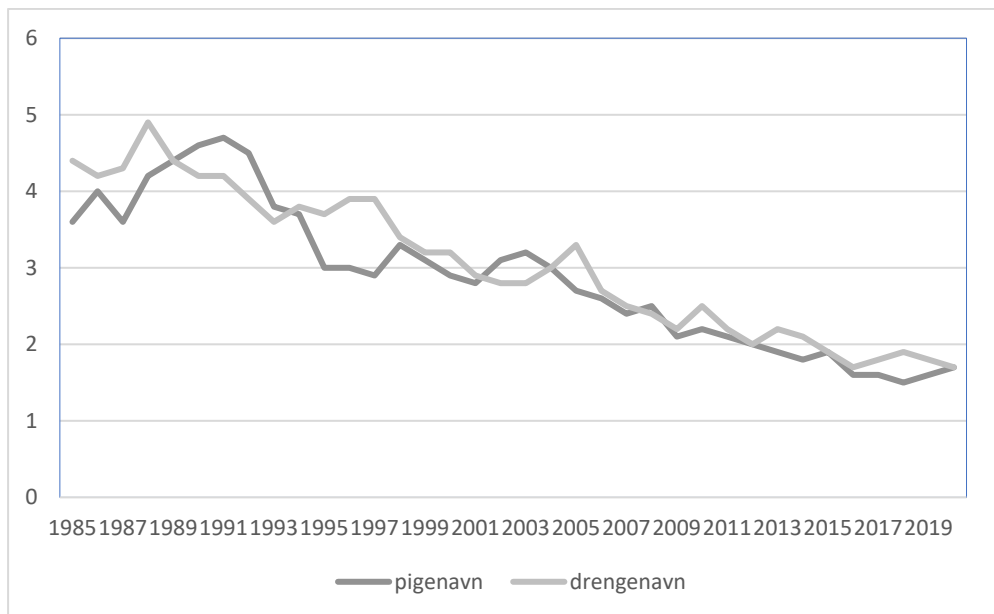
Forskellige navnestatistikker viser udvikling mod en mere varieret navngivning og mod flere unikke navne gennem de seneste ca. 200 år (Eggert 2010). I 2015 kom Danmarks Statistik med en opgørelse som viste at næsten alle nyfødte, 96 %, fik et unikt navn i Danmark når man så på deres fulde navn, dvs. alle fornavne, alle mellemnavne og efternavnet, og at dette tal var steget flere procent i løbet af bare 10 år (NYT fra DSt 2015).

I Danmark kan man ifølge Navneloven bære så mange fornavne man ønsker, så mange mellemnavne man ønsker, men kun et efternavn. Alle skal have mindst ét fornavn og et efter-

² "onymische" = onymiske, dvs. 'navnemæssige; som har med navne at gøre'.

navn senest seks måneder efter fødslen (Navneloven 2005). Fornavne er individualbenævnelser som består af det/de første navn(e) i det fulde navn. Mellemlavne er oftest familiebenævnelser som man også kan bære som efternavn, men kan også bestå af kønsmodsatte fornavne eller patronymnavne dannet til en forælders fornavn. Mellemlavne står efter fornavnene, men før efternavnet. Efternavnet er det sidste navn i en persons navnerække og er familiebenævnelser. Det er ofte, men ikke altid, nedarvet i familier gennem flere generationer.

Udviklingen mod en mere varieret navngivning har som nævnt været under udvikling længe. Ifølge en undersøgelse af de hyppigste navne i en landsdækkende folketælling fra 1801 bar ca. 30 % af kvinderne det hyppigste fornavn, *Anne*, og ca. 14 % af mændene bar det hyppigste fornavn, *Peder* (Eggert 2009). I løbet af 1900-tallet er der færre og færre som bærer de hyppigste fornavne (Eggert 2010), og udviklingen fortsatte ind i det 21. århundrede; i 2020 var det kun 1,7 % hos både pigerne og drengene som bar det hyppigste fornavn (se figur 1).



Figur 1: Nye navnebærere af det mest udbredte pige- og drengenavn som første fornavn i procent af nyfødte pr. år. Data fra DSt: Navne til nyfødte.

Danmarks Statistik har lavet toplister over de hyppigste fornavne til babyer hvert år siden 1985 (DSt: Navne til nyfødte). Disse tal gælder imidlertid kun det første fornavn for hver navnebærer; mange har flere fornavne, men det er altså kun det første fornavn som er talt med i det følgende. Desuden er forskellige stavemåder af samme navn talt sammen i Danmarks Statistiks lister uden det fortæller hvilke stavemåder det drejer sig om. Det gør det til et vanskeligt materiale at udnytte i forskning, men det kan efter min vurdering godt bruges her til illustration af udviklingen

i antallet af navnebærere over tid, da toplisterne er udfærdiget efter samme princip gennem alle årene.

2.3 Godkendte navne i Danmark

Frem til navneloven fra 2005 (Navneloven 2005) trådte i kraft, ændrede listerne over godkendte fornavne sig ikke meget. Til gengæld kunne man få navne og navneformer godkendt “efter omstændighederne” således at ansøgeren fik lov at bruge navnet, men uden at det pågældende navn blev optaget på listen over godkendte navne. Dette betyder at sjældne navne og navneformer ikke var særligt synlige for offentligheden.

Med navneloven fra 2005 bortfaldt muligheden for at få fornavne godkendt “efter omstændighederne”, og alle navne som blev godkendt efter ansøgning, kom ind på listen over godkendte navne som siden er blevet opdateret flere gange årligt. I 2005 automatgodkendte man desuden alle fornavne som blev båret af mindst 25 personer i Danmark. Det betød at der, da loven trådte i kraft den 1. april 2006, var 4.309 godkendte drengenavne og 6.362 godkendte pigenavne. Samtidig begyndte man at publicere listerne over godkendte fornavne på internettet, så det blev let tilgængeligt for alle at se alle de godkendte fornavne. Derved kunne man nemt gennem disse lister blive inspireret til navne og navneformer som man ikke kendte i forvejen, og sjældne navne og navneformer kunne derved lettere spredes end tilfældet var før 2006. Den 15. juli 2021 var antallet af navne på disse lister vokset kraftigt således at der var 20.134 godkendte drengenavne og 24.726 godkendte pigenavne.

I Danmark skal alle nyfødte som nævnt ovenfor være navngivet senest 6 måneder efter fødslen, og derfor er det rimeligt at antage at alle navne som blev brugt til børn født til og med 2020, findes på listen over godkendte navne efter opdateringen den 15. juli 2021 (Godkendte fornavne), og det er derfor navne fra denne opdatering som bruges i undersøgelsen nedenfor.

3. Stigningen i nye stavemåder af fornavne 1995-2020

Den følgende undersøgelse er lavet på baggrund af data fra Danmarks Statistik med årlige lister over samtlige fornavne givet som første fornavn til nyfødte med angivelse af antal navnebærere det pågældende år; disse data er ikke offentligt tilgængelige i listeform, men de kan fremsøges ved søgning i Danmarks Statistiks *Navnebarometer* (DSt: NavneBaro).

Der er foretaget en opdeling i mandlige og kvindelige navnebærere hvilket skyldes den måde data er indrettet på. I Danmark skelnes der stadig skarpt mellem pige- og drengenavne; ifølge Navneloven 2005 (§13 stk. 2) må en dreng ikke tillægges et pigenavn og en pige ikke et drengenavn, og dette afspejles i statistikkerne.

3.1 Flere unikke fornavne

Det er forholdsvis forskelligt hvor mange børn som fødes hvert år i Danmark, og som det ses i tabel 1, var fødselstallet faldende i den undersøgte periode. Samtidig er antallet af forskellige fornavne stærkt stigende i samme periode (hver forskellige stavemåde tælles her som et selvstændigt navn). Derfor har jeg udregnet et tal for antal forskellige fornavneformer i forhold til antallet af navnebærere for at vise udviklingen mod brugen af flere forskellige navne mere entydigt. Dette tal ses i tredje række i tabel 1, og som det ses, stiger tallet stødt gennem undersøgelsesperioden fra 0,062 i 1995 til 0,104 i 2020 for drengenavnene vedkommende, det er en stigning på 68 %. For pigenavnene vedkommende stiger tallet fra 0,084 i 1995 til 0,120 i 2020, og her er der tale om en stigning på 43 %.

Dernæst har jeg set på hvor mange i hver af de undersøgte årgange som har fået et navn som ingen andre børn fik det samme år, og også dette tal er stigende. Den procentmæssige stigning går for drengenes vedkommende fra 4,08 % af årgangen i 1995 til 6,22 % af årgangen i 2020, altså en stigning på omkring 50 % i løbet af de ca. 25 år undersøgelsen dækker. Stigningen for pigernes vedkommende er lidt mindre. Den går fra 5,36 % af årgangen i 1995 til 7,28 % af årgangen i 2020. Der er altså gennemgående flere piger som får et unikt navn end der er drenge selvom stigningen for pigerne i de undersøgte 25 år “kun” er på omkring 35 %.

	1995	2000	2005	2010	2015	2020
Antal navnebærere	M: 35.967 F: 33.730	M: 34.582 F: 32.978	M: 33.023 F: 31.742	M: 32.173 F: 30.660	M: 30.485 F: 28.897	M: 30.979 F: 29.580
Antal forskellige navne-former	M: 2241 F: 2844	M: 2653 F: 3212	M: 2510 F: 3192	M: 2710 F: 3307	M: 3032 F: 3491	M: 3211 F: 3561
Navneformer pr. navnebærer	M: 0,062 F: 0,084	M: 0,077 F: 0,097	M: 0,076 F: 0,101	M: 0,084 F: 0,108	M: 0,099 F: 0,120	M: 0,104 F: 0,120
Antal navneformer med kun én navnebærer	M: 1468 F: 1807	M: 1703 F: 2022	M: 1568 F: 1961	M: 1644 F: 2008	M: 1739 F: 2024	M: 1927 F: 2152
% navnebærere med et unikt fornavn det år	M: 4,08 % F: 5,36 %	M: 4,92 % F: 6,13 %	M: 4,75 % F: 6,18 %	M: 5,11 % F: 6,55 %	M: 5,70 % F: 7,00 %	M: 6,22 % F: 7,28 %

Tabel 1: Første fornavn til nyfødte 1995-2020. Hver forskellige stavemåde tælles som et selvstændigt navn. M = drenge/ drengenavne; F = piger/pigenavne. Data fra Danmarks Statistik.

3.2. Nye stavemåder af udbredte fornavne

For at vurdere om stigningen i brugen af unikke fornavne delvist kan skyldes en stigning i brugen af nye stavevarianter af kendte fornavne, har jeg taget udgangspunkt i toptilisten fra Danmarks Statistik over første fornavn til børn født i Danmark i 2020 (DSt: Navne til nyfødte). Som det er nævnt ovenfor, regner Danmarks Statistik forskellige stavemåder sammen når de publicerer disse lister. Det gør dem egnede til at se nærmere på hvilke stavemåder af de enkelte navne som er brugt og talt sammen på listen det pågældende år.

Som nævnt ovenfor oplyser Danmarks Statistik kun at forskellige stavemåder er slået sammen, men desværre ikke hvilke stavemåder det drejer sig om. Men gennem listerne over alle de forskellige stavemåder som er brugt hvert år, jf. tabel 1, er det muligt at komme med et kvalificeret gæt på hvilke stavemåder de har slået sammen.

I listerne nedenfor er der tilføjet i parentes hvilke stavemåder som blev benyttet i 2020, og som udtalemæssigt i moderne standarddansk (jf. Petersen & al. 2021:11) efter min vurdering kan afspejle det samme navn. I nogle tilfælde kan man regne sig frem til hvilke stavevarianter der er talt sammen, hvis de pågældende varianter til sammen giver det tal som angives i Danmarks Statistisk offentlige liste. Men i nogle tilfælde er det svært at regne ud hvilke stavemåder som er talt med, og hvilke der ikke er. I de tilfælde har jeg gættet mig frem. Jeg har sat streg under dem som jeg tror, er talt med fordi antallet af navnebærere sammenlagt giver det antal som er angivet i Danmarks Statistiks liste. I det følgende har jeg valgt at se nærmere på udviklingen af stavevarianter som afspejler navnene *Freja*, *Sofia*, *Oliver* og *Lucas* da de har flest stavevarianter i toplisterne fra 2020.

- | | |
|--|---|
| 1. <i>Alma</i> 514 (Alma 514) | 1. <i>Alfred</i> 520 (Alfred 520) |
| 2. <i>Agnes</i> 456 (Agnes 456) | 2. <i>Oscar</i> 516 (Oscar 282, Oskar 234, Óskar 1) |
| 3. <i>Ella</i> 448 (Ella 448) | 3. <i>Carl</i> 503 (Carl 258, Karl 245) |
| 4. <i>Freja</i> 439 (Fraya 2, Fregya 1, Freia 6, Freiah 1, Freja 349, Freya 84, Freyja 6) | 4. <i>Noah</i> 503 (Noa 22, Nóa 1, Noah 481, Nooa 2) |
| 5. <i>Clara</i> 432 (Clara 353, Klara 79) | 5. <i>William</i> 497 (Viljam 1, Villiam 8, William 489) |
| 6. <i>Emma</i> 432 (Emma 432, Emmah 1) | 6. <i>Oliver</i> 463 (Olivar 1, Oliver 463, ?Olivér 1, Óliver 1, Olliver 3) |
| 7. <i>Sofia</i> 423 (Sofia 268, Sofia 1, Sofiia 1, Sofija 6, Sofiya 3, Sophia 144, Sophiya 1; Zofia 9, Zophia 2) | 7. <i>Aksel</i> 454 (Aksel 323, Axel 131) |
| 8. <i>Karla</i> 398 (Carla 119, Karla 279) | 8. <i>Arthur</i> 453 (Arthur 445, Artur 8) |
| 9. <i>Anna</i> 375 (Anna 375) | 9. <i>Valdemar</i> 432 (Valdemar 421, Waldemar 11) |
| 10. <i>Ellie</i> 363 (?Eli 4, ?Elie 1, Elli 16, Ellie 319, Elly 23) | 10. <i>Lucas</i> 429 (Lucas 223, Lucaz 1, Luckas 2, Lukas 202, Lúkas 1, Lukasz 1) |

De stavevarianter af de fire valgte navne som er brugt til nyfødte i 2020, omfatter ikke alle de varianter som kan afspejle den samme udtale. På listerne over godkendte fornavne (Godkendte navne) som de var den 15. juli 2021, fandtes de navneformer som er oplyst nedenfor, og som jeg har vurderet, kan have den samme udtale i moderne standarddansk. Brugen af disse stavevarianter som første fornavn til nyfødte har jeg undersøgt for perioden 1996-2020 gennem data fra Danmarks Statistisk (DSt: NavneBaro).

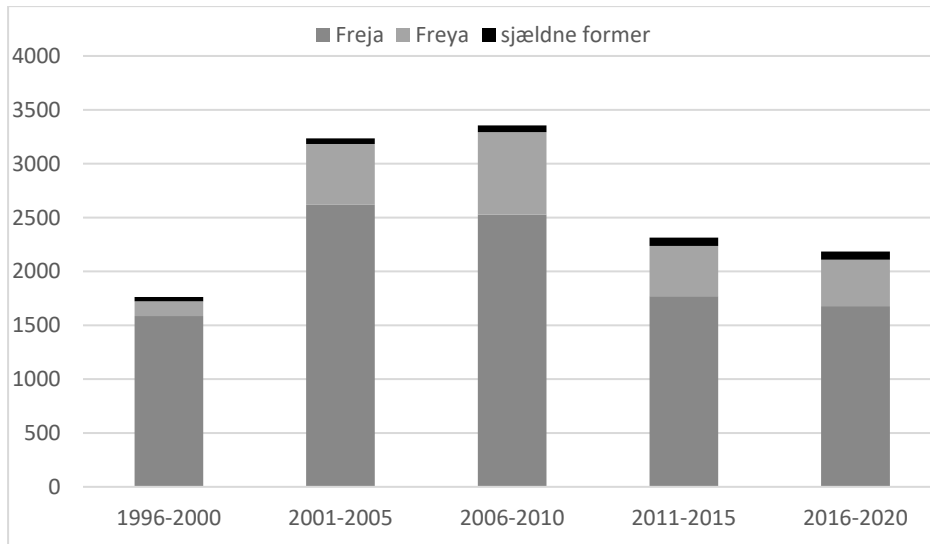
I listerne er der anført hvor mange gange den enkelte stavemåde er givet til nyfødte i perioden 1996-2020:

- *Freja*: Fraia (1), Fraja (11), Fraya (28), Fregya (1), Freiah (1), Freija (7), Freja (10.186), Freya (2.358), Freyah (2), Freyiia (0), Freia (156), Freyja (103).
- *Sofia*: Soffia (4), Soffia (3), Soffiah (0), Soffija (0), Sofhia (1), Sofia (5.079), Sofia (22), Sofiah (1), Sofiia (9), Sofija (62), Sofiya (26), Sofiyah (1), Sofya (7), Sophia (3.179), Sophija (0), Sophiya (3), Sophiyah (1), Zofia (95), Zofiia (0), Zophia (30), Zsofia (3), Zsófia (1)
- *Oliver*: Olivar (17), Oliver (15.984), Óliver (15), Olivér (6), Oliwer (28), Olliver (25), Olliwer (1)
- *Lucas*: Loucas (8), Loukas (6), Lucas (9.304), Lucasz (1), Lucaz (35), Luccas (11), Luchas (8), Luckas (60), Luckaz (3), Lukas (5.926), Lúkas (9), Lukasz (14), Lukaz (32), Lukhas (0), Luukas (1)

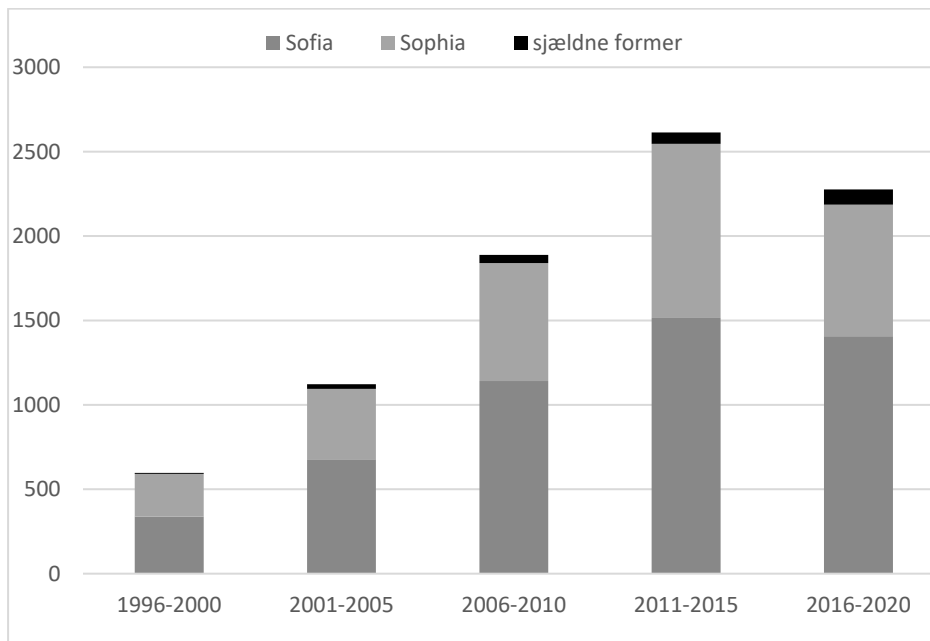
Som det ses, er en stor del af stavevarianterne meget sjældne. Enkelte varianter er slet ikke brugt til nyfødte 1996-2020 hvilket sandsynligvis skyldes at de er blevet godkendt efter ansøgning fra voksne, eller at de er brugt til nyfødte som et fornavn der ikke står først i navnerækken.

I diagrammerne i figur 2-5 ses fordelingen af den ene eller de to mest hyppige former af navnene i de mørkegrå og (evt.) mellemgrå søjler samt de sjældne stavevarianter (sort) som her er slået sammen. Tallene er opgjort for femårsperioder, og den sorte bjælke er smal ved alle navnene, men de ser alligevel lidt forskellige ud for pigenavnene i forhold til drengenavnene.

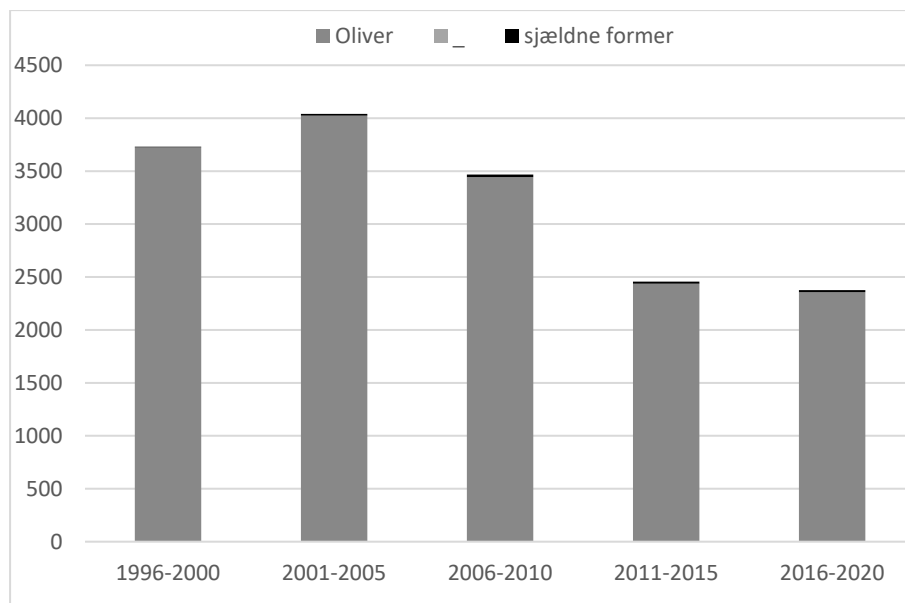
Hos pigenavnene, *Freja* og *Sofia*, bliver den sorte bjælke lidt tykkere for hvert tidsinterval, og selvom det samlede antal navnebærere falder for begge navne i slutningen af den undersøgte periode, er andelen af sjældne former alligevel stigende. Men det ser noget anderledes ud for de to undersøgte drengenavne, *Oliver* og *Lucas*. De har generelt en noget mindre andel af sjældne former, og ganske vist stiger andelen lidt i den undersøgte periode, men det er slet ikke lige så tydeligt som ved pigenavnene. På baggrund af denne lille undersøgelse er der altså noget der tyder på at opblomstringen af stavevarianter i højere grad er noget som gør sig gældende for pigenavnene end for drengenavnene.



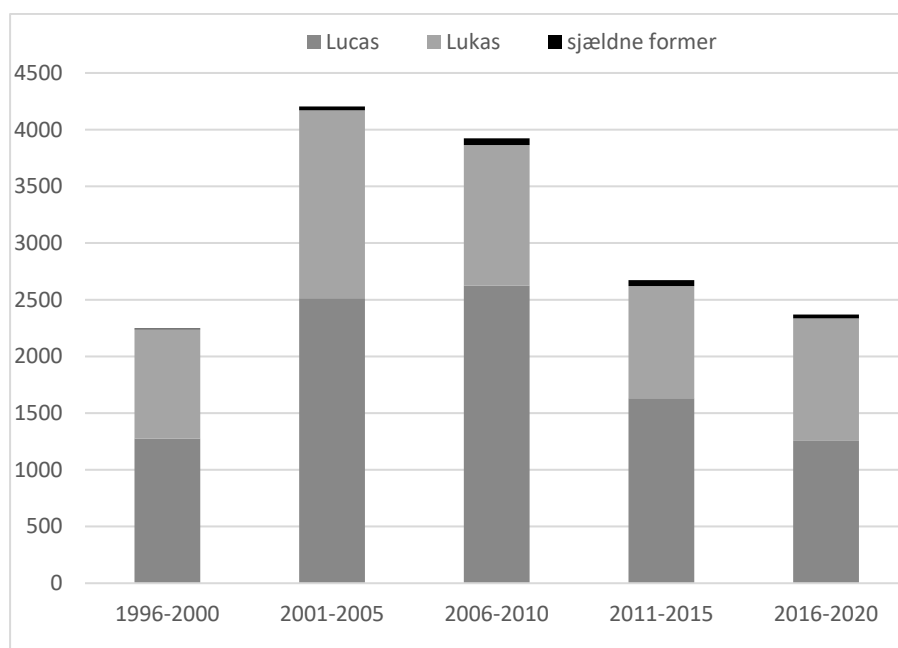
Figur 2: Fordelingen mellem formerne Freja (mørk grå), Freya (mellemgrå) og de sjældnere former (sort) i femårsperioder 1996-2020. De sjældne navneformers andel er: 1996-2000: 2,4 %; 2001-2005: 1,6 %; 2006-2010: 1,9 %; 2011-2015: 3,5 %; 2016-2020: 3,6 %. Data fra DSt: NavneBaro.



Figur 3: Fordelingen mellem formerne Sofia (mørk grå), Sophia (mellemgrå) og de sjældnere former (sort) i femårsperioder 1996-2020. De sjældne navneformers andel er: 1996-2000: 1,0 %; 2001-2005: 2,5 %; 2006-2010: 2,6 %; 2011-2015: 2,8 %; 2016-2020: 4,1 %. Data fra DSt: NavneBaro.



Figur 4: Fordelingen mellem formerne Oliver (mørk grå) og de sjældnere former (sort) i femårsperioder 1996-2020. De sjældne navneformers andel er: 1996-2000: 0,2 %; 2001-2005: 0,4 %; 2006-2010: 0,7 %; 2011-2015: 0,8 %; 2016-2020: 0,8 %. Data fra DSt: NavneBaro.



Figur 5: Fordelingen mellem formerne Lucas (mørk grå), Lukas (mellemgrå) og de sjældnere former (sort) i femårsperioder 1996-2020. De sjældne navneformers andel er: 1996-2000: 0,4 %; 2001-2005: 0,8 %; 2006-2010: 1,5 %; 2011-2015: 2,0 %; 2016-2020: 1,5 %. Data fra DSt: NavneBaro.

Emnet kalder på grundigere undersøgelser som tager hensyn til flere faktorer end der er gjort her: Hvor mange stavevarianter der er i brug, afhænger både af hvilke former navngiverne har ønsket, men også af hvilke varianter der er sprogligt mulige for at de afspejler den samme udtale

som den hyppigste variant. Det sidste vil naturligvis variere alt efter hvilke muligheder der er i dansk for hvor mange forskellige måder den enkelte lyd kan gengives stavemæssigt så sprogbrugerne kan genkende dem. Dette vil være en subjektiv bedømmelse, for mange af navnenes stavemåder afviger temmelig meget fra den retskrivning som benyttes til det øvrige sprog som det fremgår i Retskrivningsordbogen (RO). Hvis der er tale om navne som også er udbredte er i udlandet, vil mange af variantformerne skyldes stavemåder som kendes fra andre sprog, og som i større eller mindre grad kendes i Danmark selvom de afviger fra dansk retskrivning.

4. Konklusion

På baggrund af ovenstående undersøgelse kan det konkluderes at fornavnebrugen i Danmark er blevet mere forskelligartet i perioden 1995-2020. Det afspejles af tre forhold: for det første at der er færre nyfødte hvert år som får de hyppigste navne for årgangen; for det andet at der er sket en stigning i antallet af forskellige navne og navneformer (stavevarianter) som gives per barn i hver årgang; og for det tredje at der er flere og flere i hver årgang som får et unikt navn på den måde at de er ene om den pågældende navneform som første fornavn blandt nyfødte det år.

Undersøgelsen viser at antallet af stavevarianter af specifikke navne også ser ud til at være stigende. Der er kun undersøgt fire navne – to pigenavne og to drengenavne – og resultatet af denne indledende og meget lille undersøgelse af dette aspekt tyder på at der er forskel når det gælder navne til piger, og når det gælder navne til drenge. I hvert fald stiger andelen af sjældne stavevarianter gennem perioden 1996-2020 når det gælder stavevarianter til pigenavnene *Freja* og *Sofia*, mens billedet er mindre tydeligt for drengenavnene *Oliver* og *Lucas* hvor der bruges forholdsmæssigt færre af de sjældne stavevarianter, og hvor der ganske vist ses en stigning i perioden som ved pigenavnene, men også et lille fald til sidst, når det gælder *Lucas*. Undersøgelsen giver altså ikke et klart billede af udviklingen i brugen af flere stavevarianter. Dette bør undersøges nærmere med mere sofistikerede metoder end dem der er præsenteret her, hvis sådanne metoder kan opstilles.

Spørgsmålet er nu hvorfor et stigende antal personer ønsker unikke navne og sjældne stavemåder af fornavne. En del af forklaringen må være at samfundet bliver mere og mere internationaliseret. Folks berøringsflade med andre sprog og andre lande er større end nogen sinde før, og inspirationen kommer mange forskellige steder fra. En del af de sjældne navne og stavevarianter som er brugt i Danmark, er stavemåder som er almindelige i andre sprog end dansk, fx er *Lukasz* en almindelig polsk form af *Lucas* og *Zsofia* en almindelig ungarsk form af *Sofia*. Samtidig må valget af sjældne former betyde noget særligt for navngiverne og afspejle

et ønske om at signalere noget andet end det almindelige. Dermed kan man gennem navnenes stavemåder signalere identiteter som tilknytning til et bestemt land eller sprog, eller ønsker om anderledeshed og særlighed. Dette kan ses i lyset af den ovennævnte undersøgelse hvor Aldrin skriver at anderledes stavemåder kan afspejle et ønske om at positionere sig som original og være et bevidst valg for at navnet skal fremstå mere ualmindeligt. Dermed kan – i hvert fald en del af de anderledes stavemåder – indgå som en del af tendensen mod flere unikke fornavne.

Denne lille undersøgelse giver slet ikke et så dækkende billede som det kunne være ønsket, og der er opstået nogle spørgsmål som det kunne være interessant at finde svar på i fremtiden. Det kunne desuden være interessant at få afdækket om det, at listerne over godkendte navne er offentligt og let tilgængelige på internettet, gør at folk bliver inspirerede af dem til navne og navneformer som de ikke ellers ville have mødt. Samtidig ville det også være interessant hvis man kan finde ud af hvornår den præcise stavemåde af et navn blev tæt knyttet til den enkeltes identitet.

Litteratur

- Aldrin, Emilia (2011a) *Att välja namn. Namngivningsmönster i Göteborg 2007*. SoLiD nr 21 (= FUMS rapport nr 231). Uppsala.
- Aldrin, Emilia (2011b) *Namnval som social handling. Val av förnamn och samtal om förnamn bland föräldrar i Göteborg 2007–2009*. (Doctoral thesis, 2011).
- Bechsgaard, Katrine Kehlet (2015) *Hvorfor skal barnet hedde sådan? En undersøgelse af motiver bag valg af børnenavne og social variation i forhold til navnevalg*. Ph.d.-afhandling. København: Københavns Universitet.
- DSt: NavneBaro = Danmarks Statistik, Navnebarometer: online <https://www.dst.dk/da/Statistik/emner/borgere/navne/Baro>. (besøgt 19. maj 2023)
- DSt: Navne til nyfødte = Danmarks Statistik, Navne til nyfødte: online <https://www.dst.dk/da/Statistik/emner/borgere/navne/navne-til-nyfoedte>. (besøgt 19. maj 2023)
- Eggert, Birgit (2009) Almuens fornavne omkring år 1800. I: *Landbohistorisk Tidsskrift* 2009:1:89–125. (online: <https://tidsskrift.dk/landbohist/article/view/3817/3312>). (besøgt 19. maj 2023)
- Eggert, Birgit (2010) De hyppigste fornavne bliver sjældnere. I: *Mål og Mæle* 1/2010:26-30.
- Eggert, Birgit (2013) Carl eller Karl? Stavevarianter i dansk fornavnemode. I: Zakaris Svabo Hansen, Jógvan í Lon Jacobsen, Tina K. Jakobsen, Kristin Magnussen & Turið Sigurðardóttir (red.) *Frá Sturlungu til West Venture. Heiðursrit til Anfinn Johansen á 60 ára degnum*. (Annales Societatis Scientiarum Faeroensis. Supplementum 58). Tórshavn:55-68.
- Eggert, Birgit (2017) Hvad er det rigtige navn? Eksempler på personnavnevariationer gennem 300 år. I Tom Schmidt & Inge Særheim (red.) *Rapport fra den 16. nordiske navneforskerkongres på Kleppe, Norge, 8.-11. juni 2016*. (NORNA-rapporter 96). Uppsala:59-71.

- Godkendte fornavne = Familieretshusets lister over godkendte fornavne [opdateringen som den var den 15. juli 2021, er brugt i denne artikel]. Online: <https://familieretshuset.dk/navne/navne/godkendte-fornavne>
- Grundtvig, Svend (1872) *Dansk Haandordbog*. København.
- Navneloven 2005 = LOV nr. 524 af 24. juni 2005 *Navnelov*: <https://www.retsinformation.dk/eli/lta/2005/524>
(Navneloven tilpasses løbende, og gældende udgave pr. 6. januar 2022 som er bekendtgjort 14. september 2021 kan findes her: <https://www.retsinformation.dk/eli/lta/2021/1815>). (besøgt 19. maj 2023)
- Nübling, Damaris, Fabian Fahlbusch & Rita Heuser (2015) *Namen. Eine Einführung in die Onomastik*. Tübingen.
- NYT fra DSt 2015 = NYT fra Danmarks Statistik, 15. januar 2015, Nr. 21: Navne – hele befolkningen 1. januar 2015. Online publikation: <https://www.dst.dk/Site/Dst/Udgivelser/nyt/GetPdf.aspx?cid=18980>. (besøgt 19. maj 2023)
- Olesen, Rikke Steenholt (2022) Individ og kollektiv – At løfte Danmarks Stednavne. I Johnny G.G. Jakobsen (red., under medvirken af Birgit Eggert) *Danmarks Stednavne 100 år. Jubilæumsskrift i anledning af hundredåret for udgivelsen af bind 1: Samsøs Stednavne*. (Navnestudier 46). København:9-30. Online: <https://nors.ku.dk/publikationer/webpublikationer/navnestudier-46/>. (besøgt 19. maj 2023)
- Petersen & al. 2021 = Jan Heegård Petersen, Holger Juul, Nicolai Pharaos & Marie Maegaard (2021) *Udtalt. En introduktion til dansk fonetik*. København.
- RO = Retskrivningsordbogen, udgivet af Dansk Sprognævn: <https://dsn.dk/ordboeger/retskrivningsordbogen/>
(besøgt 19. maj 2023).

Resumé:

Grundtræk af en prosodisk teori om stødet

Yonatan Goldshtein

Aarhus Universitet

Det danske stød er en laryngal prosodi, der findes på bestemte stavelsers stemte rim (Fischer-Jørgensen 1989). Regelsættet for stødets distribution er berygtet for sin kompleksitet. Målet for en fonologisk teori om stødet må være at den: (1) indfanger stødets distribution i en strukturelt økonomisk model, (2) kan beskrive synkron og diakron variation, (3) kan bruges til sprogtypologisk sammenligning, og 4) har et nogenlunde gennemsigtigt teoretisk fundament. I dette foredrag skitserer jeg en ny fonologisk teori om det danske stød, der lever op til ovenstående krav. Denne teori diskuteres op imod Basbølls (2005) ikke-stød model.

Den præsenterede teori trækker på prosodisk fonologi (Itô & Mester 2012) og mere specifikt på de såkaldt 'metriske tilgange' (Köhnlein 2019) til accentuering. Hovedantagelsen er, at prosodi består af hierarkisk ordnede grupper af prosodiske konstituenten, og at accentuelle kontraster (fx stød vs. ikke-stød) kan forstås som et udtryk for underliggende forskelle i prosodisk struktur.

Konkret anser jeg stødet for en ordaccent, der markerer den mest prominente stavelse i det minimale prosodiske ord (min-ord). Monomorfemiske ord kan maksimalt have ét stød og består derfor maksimalt af et prosodisk min-ord. Sammensatte ord kan derimod bestå af flere prosodiske min-ord og kan derfor have mere end et stød. Uanset morfologisk kompleksitet har et ord kun en betonet stavelse. Betoning kan dermed ses som en ordaccent, der markerer den mest prominente stavelse i det maksimale prosodiske ord (max-ord). Lignende skelnen mellem min-ord og max-ord laves også i fx svensk (Myrberg & Riad 2015) og flere andre sprog (Itô & Mester 2021).

Hvor betoning kan findes i alle stavelsestyper, stiller stød en række krav til sin værstavelse. En stødstavelse skal have lang sonoritet og stødet foretrækker ordets sidste eller tredjesidste stavelse, mens det foretrækker ikke at findes i ordets næstsidste stavelse. Jeg argumenterer for, at også disse fonologiske restriktioner skal forstås ud fra det prosodiske hierarki.

Sonoritetskravet kan forstås i relation til begrebet mora (Basbøll 2005, Vazquéz-Larruscaín 2021), hvorimod kravet til stavelsens placering kan forstås i relation til den prosodiske fod (Itô & Mester 2021). Stødets forhold til foden viser nært slægtskab til den tonale accent i norsk og svensk (fx Bye under udgivelse), mens stødets forhold til mora afviger noget fra den tonale accents.

Litteratur

- Basbøll, Hans (2005) *The Phonology of Danish*. Oxford University Press.
- Bye, Patrik (under udgivelse). *The Evolution of Pitch Accent in North Germanic: Pragmatics, Phonology, Geography, and History*.
- Fischer-Jørgensen, Eli (1989) Phonetic Analysis of the Stød in Standard Danish. *Phonetica* 46(1-3):1-59.
- Itô, Junko & Armin Mester (2012) Recursive prosodic phrasing in Japanese. I Toni Borowski, Shigeto Kawahara, Shinya Takashito, Mariko Sugahara (red.) *Prosody Matters: Essays in Honor of Elisabeth Selkirk*. Equinox:280-303.
- Itô, Junko & Armin Mester (2021) Recursive Prosody and the Prosodic Form of Compounds. *Languages* 6(2):65-89.
- Köhnlein, Björn (2019) Why metrical approaches to tonal accent are worth pursuing. *Phonology* 36(3):531-535.
- Myrberg & Riad (2015) The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics* 38(2):115-147.
- Vazquéz-Larruscaín, Miguel (2021) Complementary length in Danish. Why not? *Nordlyd* 45(1):39-55.

Anna Pedersen og Udvalg for Folkemål

Asgerd Gudiksen

Københavns Universitet

1. Indledning

I et brev til Poul Andersen skriver Anna Pedersen den 6. marts 1927:

Mon De egentlig aner, hvad De har været med til, som medarbejder i Udvalg for Folkemaal? De har givet (el. vil da gøre det) et halvt hundrede mennesker her til lands et emne at skrive om som maaske er deres kæphest, papir til at skrive paa og penge oven i købet, hvad vil resultatet blive? Tænk paa alle tryksagerne som skulde gemmes i bibliotekerne. Ja, ja, nu er De advaret. Jeg spaar at Udvalg for Folkemaal kommer til at rive sine lader ned ... og bygge dem større

(Brev fra Anna Pedersen 6. marts 1927).

Poul Andersen var den daglige leder af Udvalg for Folkemaal (herefter UfF), som senere blev til Institut for Dansk Dialektforskning, og nu efter tre fusioner hedder Center for Dialektforskning. Anna Pedersen var meddeler og lokal medarbejder. I 1927 havde Poul Andersen og Anna Pedersen kendt hinanden omkring tre år, og på det tidspunkt var Anna Pedersen allerede kommet langt med at tilegne sig Danias lydskrift, så langt at hun senere samme år kunne skrive en artikel til tidsskriftet *Danske Folkemål* med titlen ”Om Gammelt og nyt i Strøby-dialekten” – med eksemplerne gengivet i Dania (jf. Pedersen 1927). Anna Pedersen lærte altså de professionelle optegneres arbejdsredskab, og også metodisk blev hendes arbejde for udvalget af professionelt tilsnit – og meget omfangsrigt (som citatet overfor forudsiger).

Med artiklen vil jeg nuancere det undertiden lidt stereotype billede af dialektmeddeleren som en enfoldig almuesmand der lader sig eksaminere uden at forstå ret meget af hvad det hele går ud på. Undervejs vil jeg forsøge at besvare spørgsmålene: Hvordan kom Anna Pedersen ind i arbejdet, hvad var hendes drivkraft og motivation, og var hendes interesse den samme som dialektologernes?

Der er blevet fremlagt en række eksempler på at meddelernes drivkraft og interesser kunne afvige fra dialekt- eller folkemindeforskernes, og at meddelerens stemme ikke altid kom til orde, fx er det blevet påvist at Karoline Graves beskrivelser af nordvestsjællandsk bondeliv blev redigeret hårdt af udgiveren Henrik Ussing (Ingemann 1980)¹. Med Anna Pedersen stillede sagen sig anderledes. Hun var interesseret i at lære dialektologisk metode, og det virker ikke som om rammerne begrænsede hende i væsentlig grad.

Kilderne er foruden Anne Pedersens optegnelser i Center for Dialektforsknings arkiv, de mange breve der blev udvekslet mellem Anna Pedersen og medarbejderne på kontoret igennem 45 år, og som nu indgår i Centrets brevsamling. Desuden har jeg også fundet relevante ting blandt hendes efterladte papirer, der også befinder sig i Centrets arkiv, bl.a. en selvbiografisk skitse.

2. Uff's historie kort²

Uff blev oprettet i 1909 med en ordbog over de danske øsmål som en central opgave. Baggrunden var at der allerede var udkommet ordbøger over jysk og bornholmsk, mens dialekterne på Sjælland, Møn, Lolland-Falster og Fyn var forsømt. Arbejdet kom dog først i fastere rammer i 1922 da Uff fik et lille kontor, og der bliver ansat to faste medarbejdere først deltids, og nogle år senere heltids, nemlig Poul Andersen og Ellen Raae, der senere i 1930'erne blev suppleret med fire fastansatte.

En påtrængende opgave for Poul Andersen og Ellen Raae var at udarbejde hjælpemidler til indsamlingen. I de første 10 år havde man især optegnet efter den såkaldte ordliste, en spørgeliste der indeholder 131 dialekttræk med tilhørende typeord hvis udtale eller bøjning skal notes af optegneren. Da man imidlertid skulle etablere grundlaget for en stor dialektordbog, måtte man tage andre metoder i brug. I 1925 fik man udarbejdet et billedhæfte, og i 1926 "Den Store Spørgeliste". Den Store Spørgeliste er egentlig et ret simpelt hjælpemiddel, hvis udformning i nogen grad blev bestemt af Udvalgets begrænsede resurser (Andersen 1932:172). Det er nemlig ikke en egentlig spørgeliste, men en omfattende ordliste opstillet efter begreber, fx "arbejde", "vejr og vind", "klæder" og "jorden". Man kunne næsten kalde den en lille begrebsordbog. Formålet med listen var at fremkalde sammenhængende beretninger om bondens livsverden sådan at ordene kom i deres naturlige sammenhæng, og den var tænkt som et hjælpemiddel både for optegneren i marken og for korresponderende meddelere. Det viste sig dog at være ret vanskeligt for meddelerne at arbejde med listen på egen hånd – Anna Pedersen var en af de få

¹ Se også Lilja 1996:188 ff.

² For længere versioner se Gudiksen & Hovmark 2009 (med henvisninger) og Andersen 1932 og 1939.

meddelere der kunne skrive de sammenhængende beskrivelser som man havde forestillet sig – og i løbet af 1930'erne udviklede man derfor egentlige etnologiske spørgelister til meddelejerne at arbejde med på egen hånd, mens ”Den Store Spørgeliste” kun brugtes ved direkte optegnelser i ”punktundersøgelserne”, der også fandt sted i 1930'erne og efterfulgtes af ”kompletteringen”, en omfattende supplerung af det allerede optegnede materiale. Efter anden verdenskrig aftog indsamlingen til fordel for forredaktion og fra omkring 1975 endelig redaktion. Fra 1957 og til 1988 udsendte Ømålsordbogens redaktion 50 spørgelister til en kreds af meddelere, især med spørgsmål vedrørende enkeltords forekomst og betydning. Anna Pedersen deltog i det meste, og UFF's historie er til en vis grad også Anna Pedersens historie. Anna Pedersen var dog næsten fyrrer før hun kom i kontakt med udvalget.

3. Anna Pedersens opvækst og uddannelse

Anna Pedersen blev født den 9. november 1885 i Strøby på Stevns, der ligger ca. 11 km. syd for købstaden Køge. I den korte, utrykte selvbiografiske skitse blandt hendes efterladte papirer fortæller hun at hendes far bl.a. var styrmand på en lille fragtbåd der sejlede landbrugsvarer til København. Det var ikke noget velhavende hjem, men man har indtryk af at familien ikke manglede noget i det daglige, og der var tale om at Anna skulle læse til lærerinde. Hendes mor var imidlertid svagelig og døde i 1902, og som den ældste af de fire søskende fik Anna Pedersen ansvar for hjemmet. Fem år senere døde også faderen. Hendes søskende var dog nu så gamle at de kunne klare sig selv, så i en alder af 22 var hun frit stillet og havde også en arv på 8000 kr. i ryggen. Nu havde hun mulighed for at få noget uddannelse ud over børneskolen og et kort højskoleophold. Hun bestemte sig til at tage til Ringe for at få den toårige friskolelærerindeuddannelse ved Ringe friskole. Det fremgår tydeligt af hendes erindringskitse at det ikke var det pædagogiske hun var interesseret i:

Særlige anlæg som lærerinde har jeg vist aldrig haft, det var både mine nærmeste og jeg selv på det rene med, og helt god samvittighed over for det, jeg indlod mig på, havde jeg ikke. Min tanke var blot at komme i lag med bøger og få lejlighed til at lære noget mere.

Selve uddannelsen – og den tilknytning til højskoleverdenen der fulgte med – blev hun glad for. Derimod kedede det hende at være lærerinde, da hun var blevet færdig:

Da jeg selv sad på skolebænk, var en skoledag aldrig lang, men uha hvor blev den lang, når man skulde være lærerinde.

Det blev da også kun en kortere årrække hun arbejdede som skolelærerinde, omkring 8 år, bl.a. på Fyn. I 1921 flyttede hun hjem igen til Strøby i et lille hus hun havde købt, og oprettede en privat pogeskole i hjemmet, dvs. hun gav læseundervisning til børn der endnu ikke var begyndt i skolen, og hun underviste også en overgang i syning. I 1933 afviklede hun Pogeskolen, og hvad hun siden lever af er lidt uklart, men der er i hvert fald ingen tvivl om at de honorarer hun får fra Uff for sit arbejde havde økonomisk betydning for hende.

4. Meddeler og optegner for Uff

Anna Pedersen kom i kontakt med Udvalget i slutningen af 1924, hvor hun reagerede på en opfordring udsendt til medlemmerne af foreningen Danmarks Folkeminder om at melde sig som meddeler, men hendes interesse for det gamle sjællandske begyndte tidligere. I 1913 har hun en lille artikel i Østsjællands Folkeblad, som er en reaktion på en tidligere artikel i avisen om sjællandsk skrevet af en ny lærer på egnen. Hun giver sin artikel en række eksempler på gamle stevnske og sjællandske ord, men derudover er der også nogle udtaleoplysninger der viser hendes sproglige iagttagelsesevne, bl.a. reflekterer hun over ”selvlyden e”:

E lyder i mange Ord, ikke som e i Rigsmålet, heller ikke som ie, snarere som ee, men den e-Lyd kan vist ikke udtrykkes skriftligt (Anna Pedersen Østsjællands Folkeblad 2/1 1913).

Hendes beskrivelse svarer ret præcist til hvordan diftongen der modsvarer langt *e* i standard-dansk, normalt er noteret i senere optegnelser, idet førstekomponenten enten er noteret med tegnet for et åbent *i* eller for et lukket *e*.

Ved sin første henvendelse medsender hun ”nogle udklip af Østsjællands Folkeblad og tillige egne optegnelser af ord fra Strøby sogn” (heriblandt indgår artiklen fra 1913). To måneder senere skriver hun igen (27/1 1925). Den ydre anledning er at hun ønsker at takke for et honorar hun uventet har modtaget fra Uff, men også en anden sag er vigtig for hende: hun vil gerne lære at bruge lydskrift: ”Mine optegnelser er jo ufuldkomne, fordi jeg ikke kan bruge lydskrift.”, skriver hun og fortsætter: ”og jeg kan ikke lære at bruge den uden mundtlig undervisning – maaske det saa endda vil knibe”.

Hun er også lidt usikker på hvad hun skal optegne, og brevet viser at hun endnu er langt fra den professionelle dialektologs synsvinkel, idet hun skriver at ”vi her paa Stevns ogsaa bruger en hel mængde ord, som er almindelige over hele Sjælland, men som man alligevel vrider

sig ved at tage med fordi de er optegnet af andre i andre egne”. Afslutningsvis glæder hun sig over at hun har ca. 50 udfyldte sedler, og blandt dem er der også ord som er ”alle optegneres fryd: de, som kun kendes og forstaas af os ’indfødte’ ”.

Men det trækker ud med undervisningen i lydskrift, formentlig fordi Poul Andersen har meget at se til. Året efter (24. februar 1926) klager hun over at hendes arbejde kun er ”halvgjort”, fordi der mangler lydskrevne former. Men samtidig er hun ikke uden selvfølelse og sans for værdien af sit arbejde:

Heller ikke synes jeg, at jeg kan opgive denne folkemaalsoptegnelse [trods den manglende færdighed i lydskrift, AG]. Den, som har hjemme i en landsby, og hvis slægt har levet der, kan hvad folkemaalsoptegnelse angaar, komme dybere til bunds end den mere kyndige, men fremmede (Anna Pedersen brev fra 24/2 1926).

I sommeren 1926 kom vejledningen endelig i gang. Poul Andersen indlogerede sig nogle dage på hotel Hafnia i Køge, så han kunne besøge Anna Pedersen dagligt, og arbejdet med lydskriften fortsætter dels pr. brev, dels – og især – mundtligt enten ved at Poul Andersen kommer til Strøby eller hun til København. Det meste af undervisningen i lydskrift kan vi altså ikke følge, men af Anna Pedersens spørgsmål i brevene får man dog et indtryk hvordan hun arbejder målbevidst og flittigt med at tilegne sig Dania. Hendes lydskriftsspørgsmål vedrører især vokal-kvalitet, stødforhold og syllabilitet og adskiller sig ikke fra hvad andre der er arbejder på at lære lydskrift, kan komme i tvivl om. Hun er fx usikker på om *lun* har ”rent u” i dialekten eller diftong, hvorfor der er længdetegn i *vɔwɪ* (komparativ af *våg* ’syg’), men ikke i *fowsɪ* (præsens af *fodse* ’sparke’), og hvor stødet skal placeres i ord som *graven* (bestemt form af *grav*) og *skabet* (bestemt form af *skab*), skal der noteres *gra:wɪn*, *ska:wð* eller *graw'n*, *skaw'ð*?

Senere i forløbet indeholder brevene ikke kun spørgsmål, men også passager der snarere er kommentarer. Den 28. oktober 1927 har hun fx en kommentar der drejer sig om et særligt sjællandsk stødfænomen, det der senere kom til at få termen ”sjællandsk kortvokalstød”:

Et helt andet spørgsmål er det med stødet i visse ord (Fætter, sætter, Latter, lægger o.s.v.) i dialekten her [...] Om dette tror jeg at kunne sige at dialekten (her) er forskellig fra rigsmaalet, idet der er stød i vocalen i disse ord. Vi gør et forsøg paa at gøre vocalen lang, det lykkes ikke rigtig, men et ”stød” faar vi da lavet (Anna Pedersen brev fra 28/10 1927).

Hvordan reaktionen har været på brevet er desværre uvist – der er ikke bevaret noget svar – men mon ikke oplysningerne har været læst med stor interesse? De få fonetiske beskrivelser af sjællandsk der forelå i 1927 (Kryger 1887, Nielsen 1902, Thorsen 1894), dækker nemlig tilfældigvis allesammen dialekter uden kortvokalstød, og fænomenet var altså ubeskrevet den gang.

I 1927 begyndte hun på arbejdet med Den Store Spørgeliste. Det var en krævende opgave, som hun først afslutter tre år senere, i 1930. Da har hun også skrevet henved 2000 sider i kvartarkformat og desuden 10.719 sedler (efter hendes egen opgivelse), hvor de centrale ord i skildringen af arbejdsprocesser osv. er udskrevet på sedler i lydskrevne citater der belyser ordets brug. I de følgende år udarbejder hun dele af den ”fraseologiske spørgeliste”, og ved ”punktundersøgelserne” i 1930’erne arbejder hun som optegner inden for det østsjællandske område på lige fod med de ansatte på Uff. På det tidspunkt har hun altså gennemgået en udvikling fra almindelig meddeler til betroet lokal medarbejder der deltager engageret i det intensive indsamlingsarbejde før anden verdenskrig. Efter anden verdenskrig aftager hendes optegnelsesarbejde i takt med at *Ømålsordbogen* går ind i redigeringsfasen, men hun har stadig kontakt med udvalget. Hun besvarer de ordspørgelister der udsendes fra 1957, indsender også af og til oplysninger uopfordret og bidrager også på anden måde til dialektforskningen, som vi skal se i de næste afsnit.

5. Anna Pedersen som meddeler ved strukturelle analyser

1939 er en milepæl i *Ømålsordbogen*s historie. Det år publicerer Poul Andersen nemlig for første gang en detaljeret plan for *Ømålsordbogen*, hvor han bl.a. fremlægger et forsøg på at reducere og normalisere tegninventaret på strukturel basis (Andersen 1939:31ff). Efter at have argumenteret for at en reduktion af tegninventaret er nødvendig skriver Poul Andersen:

Som Forberedelse til denne Reduktion af Tegnmængden og Fastlæggelse af Lydbetegnelsen i *Ømålsordbogen* har jeg gennem Aarene almindeligvis sammen med den Medarbejder, der har berejst vedkommende Dialektomraade, foretaget Efterhøring efter en ret omfangsrig Ordliste. Selve denne Efterhøring og den Diskussion af Enkeltheder, som den naturligt med det samme gav Anledning til, har hjulpet til at fæstne Betegnelsessættet [...] (Andersen 1939:31).

Forsøget omfatter de trykstærke vokaler i ”Stevnsdialekten”, og der er ikke tvivl om at den pågældende ”medarbejder” må være Anna Pedersen.

Forsøget er fremlagt i form af tre kolonner med overskrifterne ”Fonemet i Dansk” [dvs. rigsmål, AG], ”Stevnsdialektens fastlagte Betegnelse” og ”der dækker følgende Lydnuancer i det efterhørte Materiale”. Der er dog ikke tale om en egentlig strukturel opstilling af fonemsystemet i (øst)sjællandsk. ”Stevnsdialektens fastlagte Betegnelse” repræsenterer i mange tilfælde den typiske korrespondens og ikke det modsvarende fonem, fx er der udfør rigsmålsfonemet /e:/ anført diftongtegnet *j·ə*. Hvordan efterhøringen har gået for sig, fremgår ikke af brevene, heller ikke Anna Pedersens holdning til tegnreduktionen, men hun har forsynet lydskrevne sedler fra slutningen af 20’erne og begyndelsen af 30’erne med tilskrift i rødt hvor citatet er noteret med notationen fra Andersen 1939. Det har været et omfattende arbejde der tyder på at hun har være indforstået med ændringen i notationspraksis.

I 1960’erne er Anna Pedersens optegnelser igen udgangspunkt for strukturel analyse af østsjællandsk. Hendes sedler udgør nemlig grundstammen i det materiale som Inger Ejkskjær baserer *Fonemsystemet i Østsjællandsk* på (Ejkskjær 1970). Anna Pedersen var også selv inddraget i arbejdet som Inger Ejkskjær’s hovedmeddeler. Samarbejdet mellem de to var dog ikke helt harmonisk ved jeg fra samtaler med Inger Ejkskjær, og det fremgår også indirekte af præsentationen af meddelerne i bogen. Hendes andre, nyere meddelere, skriver Inger Ejkskjær, har ikke ”det fond af viden om hvad ’de gamle sagde’ som AP har erhvervet sig gennem sin mangeårige optegnerpraksis”, men:

På den anden side har disse nye meddelere fundet det langt morsommere at medvirke ved kommutationsproblemers løsning end AP, der aldrig rigtigt har kunnet godtage værdien af diskussioner af den art (Ejkskjær 1970:25).

Hvorfor Anna Pedersen var en noget uvillig meddeler for Inger Ejkskjær, er svært at afdække så mange år efter. En medvirkende årsag var måske at Anna Pedersen følte at hendes store arbejde blev underkendt, så meget desto mere som der også indgik arbejde med at forbedre og forenkle notationen. Eller måske synes hun bare der kom for lidt nyt ud af kommutationsprøverne? Muligvis havde Anna Pedersen heller ikke helt tiltro til at en jyde kunne beskrive den stevnske dialekt. Det var i hvert fald en fornemmelse Inger Ejkskjær havde (har hun fortalt mig), og den bekræftes til en vis grad af brevene (jf. citat fra brev 11/11 1934, i afsnit 7 nedenfor).

6. Anna Pedersens beskrivelser af den materielle folkekultur

Anna Pedersen har også bidraget med beskrivelser af den materielle folkekultur. Alene besvarelsen af Den Store Spørgeliste var på ca. 2000 sider (jf. afsnit 4, slutningen). Hun fik også

optaget artikler i Uff's tidsskrift *Danske Folkemål* om etnologiske emner, fx "Slagtning og slagteskikke. Optegnelser fra Stevns herred" (Pedersen 1928) eller "Farmors urtehave" (Pedersen 1929). Senere udgav hun tre bøger med undertitlen "Optegnelser fra Stevns og Bjæverskov³ Herreder": *Bondestue og Stegers* om kvindens arbejde (Pedersen 1944), *I Såtid og Høst* om mandens arbejde (Pedersen 1950) og *Højtid og Gæstebud* om skikkene ved gilder og højtider (Pedersen 1960). De to første bind er "tilrettelagt og udgivet i samarbejde med Poul Andersen". Poul Andersen pointerer i forordet til *Bondestue og Stegers* at "I Rækken af danske kulturhistoriske Beskrivelser vil denne lille Bog have sin Særkende derved, at Vægten er lagt på den Sproglige Side" (Pedersen 1944: 5), og tilsvarende hedder det i Anna Pedersens indledning: "Det skal ikke skjules, at det er de gamle ord, sproget i det hele, der ligger mig mest på hjerte". Interessen for det sproglige viser sig især ved at stevnsdialektens udtryk for genstande, arbejdsprocesser o.l. løbende er anført, men også ved at talemåder, ordsprog og varsler med tilknytning til de enkelte emner inddrages. Bøgerne er også karakteriseret ved at der ved siden af generelle beskrivelser også er citater og beretninger hentet fra optegnelserne (med angivelse af meddelelsen), og bøgerne giver herved også indtryk af variation i skik og ordforråd inden for området. I sine mindeord om Anna Pedersen har Poul Andersen givet en meget positiv karakteristik af de tre bøger:

Hos Anna Pedersen finder vi ingen lyrisk-romantisk Fremstilling af "de gode gamle dage" – der er ingen lykkeligt smilende og syngende Bondepiger med kridhvide høstærme på Marken [...] som vi kender det fra mange ældre Værker. Her finder vi en Skildring bygget på dokumentarisk Materiale [...] sikkert opfattet i Enkeltheden og altid kritisk vurderet – sorteret viden uden tågede Formodninger og apokryfe Forklaringer (Andersen 1974:2).

Poul Andersen var altså godt tilfreds med bøgerne, der også kom til at være nogle af de bøger som enhver ØMO-redaktør blev udstyret med som kontoreksempplar. Var Anna Pedersen tilfreds? Jeg tror det, dels på baggrund af bøgernes sprognære udformning, dels den måde hun omtaler sit arbejde for Uff på. Den seneste af bøgerne, *Højtid og Gæstebud*, der ikke angives at være udgivet i samarbejde med Poul Andersen, adskiller sig lidt fra de øvrige ved at indeholde en smule mere selvoplevet stof. Forskellen beror dog nok først og fremmest på emnet. Anna

³ I de to første bind dog skrevet Bjeverskov.

Pedersen var ikke opvokset på en gård (jf. afsnit 3) og havde derfor ikke samme grad første-håndsoplevelser med bondearbejdet som med fester og sammenkomster.

7. Motivation og drivkraft

Mange af brevene fra de intensive indsamlingsår er korte følgebrev eller kvitteringer for indkommet materiale eller for betaling. Undertiden kan der dog også være kommentarer til arbejdet eller personlige bemærkninger. Det gælder fx et brev Anna Pedersen sender 19. december 1928 (i en pakke med 63 kvartblade og 723 sedler). Brevet viser at hun nu ikke længere kun er ord-samlere der jagter sjældne ord, men optegneren der arbejder på at løse en opgave så godt som muligt, og at arbejdet interesserer hende meget:

Tak for de oplysninger jeg fik sidst. De har gavnet mig ikke saa lidt, selv om der endnu er ”langt frem”. Mange gange synes jeg jo, at arbejdet er for stort og optegneren for lille. Én ting er det jo at optegne gamle mærkelige ord alene, noget andet at faa sproget i fortid og nutid passet ind i rubrikker og give en forklaring, saa andre mennesker forstaar – især naar man kun er friskolelærerinde. Men ellers har jeg kun én gang før i mit liv været med til noget der var ligesaa interessant. Det var, da jeg som 24 aarig fik undervisning i botanik⁴ (Brev fra Anna Pedersen 19/12 1928).

Ordet interessant bruger hun også nogle år senere (1/4 1935), idet hun bemærker om en ny type spørgeliste at ”den er vel omtrent den interessanteste af alle de lister, jeg har haft”, og i et brev fra 1/12 1932 takker hun for ”mange timers fornøjeligt arbejde”. Et ufuldendt manuskript med titlen ”Om folkemåloptegnelse” vidner om at hun også har været optaget af metodiske spørgsmål. I et senere brev fra 11/11 1934 betoner hun et andet aspekt end det rent personlige udbytte. Hun er også drevet af en optagethed af stedet og af at bevare og formidle folkeminderne:

Da jeg begyndte at optegne folkeminder, var det vist ud fra en instinktmæssig følelse af at en Sjællænder dog muligvis kunde give et og andet til oplysning om sine landsmænd, som andre ikke kunde give (fx en Jyde om Sjællænderen). Det er hverken løgn eller latin, at den egentlige drivfjeder i mit optegningsarbejde, er lysten til at give et sandfærdigt billede af, hvad der er sjællandsk (specielt: stevnsk) (Brev fra Anna Pedersen 11/11 1934).

⁴ Af hendes erindringskitse fremgår det at hun var meget glad for botanikundervisningen på seminariet.

Citatet fra 1934 viser også at hun selv mener at have nogle særlige forudsætninger som lokal optegner (noget hun også er inde på i et brev fra 1926, se afsnit 4). Af samme brev og andre fremgår det dog også at hun kan føle sig underlegen og usikker på grund af sin kortvarige uddannelse. I brevet fra 1/12 1932 skriver hun fx også ”Jeg vilde dog blot ønske, at mit arbejde havde været en fuldendt optegnelse af en fuldt uddannet optegner – og ikke en lærlings famlende forsøg”, jf. også formuleringen ”kun er friskolelærerinde” i brevet fra 1928, citeret ovenfor.

8. Afslutning

I de første år efter Anna Pedersen er kommet i kontakt med Uff gennemgår hun en markant udvikling: hun lærer lydskrift og hendes indsamling ændrer karakter. I begyndelsen er hun optaget af at samle på sjældne ord, men ret hurtigt kommer hun til at deltage i det systematiske indsamlingsarbejde og bidrager med det som udvalget er interesseret i. Der fornemmes undertiden nogle brudflader, men man har ikke indtryk af at hun følte rammerne som en begrænsning; tværtimod ville hun selv gerne udvikle sig og lære dialektologisk metode. Gennem tilknytningen til Uff fik hun muligheder for skrivearbejde og intellektuel beskæftigelse som hun ellers ville have haft svært ved at opnå.

Anna Pedersens målbevidste arbejde med at lære Danias lydskrift og hendes rolle som optegner er usædvanlig. De fleste øvrige optegnelser i Dania er foretaget enten af medarbejdere ved Uff eller – i enkelte tilfælde – af filologisk uddannede personer med en mindre formaliseret tilknytning til Uff, der ligesom Anna Pedersen omtales som lokale medarbejdere.

De lokale medarbejdere var en lille gruppe med forskellige forudsætninger og interesser. Nogle var akademisk uddannede eller lærere, der beskrev deres barndomsdialekt, og andre havde ingen formel uddannelse. Fælles for dem er deres omfattende skriftlige bidrag, der især var baseret på ”Den Store Spørgeliste” og på de fraseologiske spørgelister (Andersen 1932:176 ff, Andersen 1939:59 ff). De fleste af bidragene har også en form for udtaleangivelse enten i Dania eller i mere eller mindre præcise uofficielle lydskrifter.

Da man omkring 1930 indledte punktundersøgelserne med det formål at få optegnelser i Dania jævnt fordelt over hele ømålsområdet, var den normale praksis den at de ansatte på Uff tog udgangspunkt i allerede foreliggende optegnelser og de brugte lokale medarbejdere som meddelere i de opfølgende fonetiske optegnelser og som hjælpere til at finde andre egnede meddelere på egnen. På den måde opnåede dialektologerne fra Uff ”en højere Grad af Forberedthed”, som Poul Andersen skriver i sin beretning fra 1932 (Andersen 1932:185), hvor han fortsætter: ”Vanskelighederne ved Arbejdets Tilrettelæggelse og Udførelse er til en vis Grad

allerede fremdraget af de lokale Medarbejdere under deres forberedende Arbejde”. I enkelte af punktområderne blev arbejdet tilrettelagt på en anden måde. På Vestfyn og på Hindsholm blev der ikke foretaget nogen egentlige punktundersøgelser fordi det allerede foreliggende materiale ansås for tilstrækkeligt, og på Stevns var det den lokale medarbejder Anna Pedersen, der selv stod for at finde andre meddelere og gennemføre de supplerende optegnelser fra andre sogne end Strøby. Gennem sit arbejde som optegner aflastede hun altså direkte de ansatte på Uff.

Ved siden af Anna Pedersen og de andre lokale medarbejdere hvis arbejde beskrives i Poul Andersens beretninger (jf. Andersen 1932, 1939), er der også de hundredvis af mere anonyme personer der har bidraget til Ømålsordbogens samlinger enten som korresponderende meddelere eller som meddelere ved direkte optegnelse. Med de sidstnævnte nærmer vi os den meddelertype der angives som typisk i kortere oversigter over dialektforskningens historie og metoder, en NORM (nonmobile old rural male) der ’eksamineres’ og ikke tager del i arbejdets tilrettelæggelse – bort set fra at mange af dem var F(emale). Det betyder dog ikke at meddelerens personlige egenskaber og interesser ikke havde betydning for optegnelsens kvalitet. I omtalen af de gode meddelere i Uff’s interne meddelerkartotek (jf. Gudiksen 2010) genfinder man egenskaber som også karakteriserede Anna Pedersen: forståelse for indsamlingen og sproglig sans.

Litteratur

- Andersen, Poul (1932) Beretning om Udvalg for Folkemaals virksomhed 1909-1932. I *Fortid og Nutid* 9: 161-210.
- Andersen, Poul (1939) Udvalg for Folkemaals Virksomhed 1933-38 (med Udkast til Plan for den kommende ”Ordbog over de danske Ømaal”). I *Danske Folkemaal* 12:17-64.
- Andersen, Poul (1974) Anna Pedersen – in memoriam. * 9. Novbr. 1885 – + 5. Marts 1972. I *Danske Folkemaal* 20:1-3.
- Ejskjær, Inger (1970) *Fonemsystemet i østsjælland på grundlag af dialekten i Strøby sogn*. København: Akademisk Forlag.
- Gudiksen, Asgerd (2010) Sedlen Stavbakterie findes dog vist ikke længere i Samlingerne. I *Nordiska studier i lexikografi 10. Rapport från Konferencen om lexikografi i Norden. Tammerfors 3-5 juni 2009*. Helsingfors: Nordisk Forening for Leksikografi: 130-139.
- Gudiksen, Asgerd & Henrik Hovmark (2009) Måske husker De noget alle andre har glemt? I Asgerd Gudiksen m.fl. (red.) *Dialektforskning i 100 år*. København: Afdeling for Dialektforskning og Museum Tusulanums forlag: 13-64.
- Ingemann, Henrik (1980) Karoline Graves. I *Folk og Kultur* 9: 53-74.
- Kryger, J.K. (1887) Bidrag til nordsjællandsk Lyd- og Bøjningslære. I *Blandinger til Oplysning om dansk Sprog i ældre og nyere Tid*. København: Universitets-Jubilæet Danske Samfund:347-397.

- Lilja, Agnete (1996) *Föreställningen om den ideala uppteckningen*. Uppsala: Dialekt- och folkminnesarkivet.
- Nielsen, S.P. (1902) Sproget i Tågeby. I *Dania* 9: 25–51.
- Pedersen, Anna (1927) Gammelt og Nyt i Strøby-dialekten. I *Danske Folkemaal* 1:138-143.
- Pedersen, Anna (1928) Slagtning og slagteskikke. Optegnelser fra Stevns herred. I *Danske Folkemaal* 2:103-107.
- Pedersen, Anna (1929) Farmors urtehave. I *Danske Folkemaal* 4:141-147.
- Pedersen, Anna (1944) *Fra Bondestue og stegers. Optegnelse fra Stevns og Bjeverskov Herreder* (Tilrettelagt og udgivet i samarbejde med Poul Andersen). København: Ejnar Munksgaard.
- Pedersen, Anna (1950) *I såtid og høst. Optegnelser fra Stevns og Bjeverskov herreder*. København: Ejnar Munksgaard.
- Pedersen, Anna (1960) *Højtid og gæstebud. Optegnelser fra Stevns og Bjeverskov herreder*. København: J. H. Schultz Forlag.
- Thorsen, P.K. (1894) *Sprogarten på Sejerø*. København: Universitets-Jubilæets Danske Samfund.
- Ømålsordbogen. En sproglig-saglig ordbog over dialekterne på Sjælland, Lolland-Falster, Fyn og omliggende øer* (1-) 1992-. København: Center for Dialektforskning & Universitets-Jubilæet Danske Samfund.

At eller ikke at i ledsætninger indledt af konjunktion/præposition

Anna Sofie Hartling, Michael Nguyen & Gustav Styrbjørn Johannessen
Dansk Sprognævn

1. Indledning

Må man skrive *fordi at*? Flertallet ville nok svare nej (formentlig bortset fra flertallet af nutidens sprogforskere). På internettet finder man både folk der beklager sig over konstruktionen, og folk der fraråder den.¹ Dette forbud er sejlivet arvegods fra præsriptive sprogfolk fra 1800-tallet og første halvdel af 1900-tallet, og det gælder ikke kun *fordi*, men også andre lignende ledsætningsinitiale elementer der kan stå på denne position foran *at*. I Håndbog i Nudansk anbefales det at man udelader *at* efter ”de fleste ledsætningsindledere”, og som eksempler nævnes *idet, hvis, når, selv om, som, som om, før, mens, hvem* og *hvor* (Jacobsen & Jørgensen 2013:84).

I to artikler fra hhv. 1975 og 1983 tager Erik Hansen dette præsriptive forbud under behandling (Hansen 1975, 1983). I artiklen fra 1975 fokuserer Hansen særligt på modsætningsparret *før* og *efter*: *Før at* er normstridigt, mens *efter at* følger normen (fx *Kør hjem før (at) det begynder at sne* og *Han gik hjem efter (at) lygterne var blevet tændt*). Der er her (mindst) to præsriptive regler i spil:

- (a) En underordnende konjunktion må ikke efterfølges af en anden underordnende konjunktion.
- (b) En præposition skal indlede en sætning med *at*.

Før at bryder med regel (a) fordi *før* betragtes som en konjunktion, mens *efter* uden efterfølgende *at* bryder med regel (b) fordi *efter* betragtes som en præposition.

Som Hansen (1975, 1983) gør rede for, er præsriptivisternes argumentation cirkel-sluttende når de forsøger at finde et grammatisk funderet forsvar for den sprogbrug som hører

¹ Se fx: <https://jyllands-posten.dk/debat/breve/ECE11796976/det-er-bare-fordi-at/>; <https://publicimage.dk/7-udbredte-grammatik-og-sprogfejl/>; <https://arosbusinessacademy.dk/25-formuleringer-du-skal-undga-sa-du-ikke-lyder-dum/>. Se også Håndbog i Nudansk (Jacobsen & Jørgensen 2013: 84).

under den socialt definerede rigsmålsnorm (altså bl.a. *fordi* og *før* uden *at* og *efter* med *at*): Man opstiller beskrivelsesmodeller der tager udgangspunkt i hvad der regnes for fint sprog, snarere end i sproget selv, og bagefter forklarer man noget i sproget selv med disse modeller.

I Nguyen & Hartling (2022) griber vi fat i problemet fra en ordklassevinkel. Det er nemlig grundlæggende uklart hvorfor *før* og *efter* ikke begge analyseres som enten præpositioner eller konjunktioner. Vi argumenterer for at ord som *før* og *efter* er præpositioner, uanset hvilken type led de indleder, herunder nominalsyntaxmer, præpositionsforbindelser og sætninger (både med og uden *at*).

Denne artikel handler om i hvor høj grad man i moderne dansk skrift- og talesprog følger de præskriptive regler for brugen af *at* i forbindelse med sådanne elementer (se (a) og (b) ovenfor). Mere specifikt handler artiklen om hvorvidt *fordi*, *før* og *efter* efterfølges af *at* i landsdækkende aviser i Dansk Sprognævns interne aviskorpus, Korp, og i talesprogskorpusset LAN-CHART. På baggrund af disse resultater vil vi diskutere om de præskriptive regler lever videre – måske i svækket form i forbindelse med enkeltordene *fordi*, *før* og *efter*.

Resten af artiklen er struktureret således: Vi gennemgår tidligere litteratur på området (afsnit 2), fremlægger en række hypoteser som skal afprøves (afsnit 3), gennemgår vores metode (afsnit 4) og de resultater vi når frem til (afsnit 5), og diskuterer hvordan hovedspørgsmålene skal besvares, og om hypoteserne holder vand (afsnit 6). I afsnit 7 sammenfatter vi resultaterne.

2. Tidligere litteratur på området

I Grammatik over det Danske Sprog (og andre steder, bl.a. Pedersen 2009) kaldes det *at* der følger efter en ”egentlig konjunktion” eller en subjektiv partikel, fx *bare*, *gid* og *mon*, for det indre *at* (Hansen & Heltoft 2019: 1657). Vi lader her de subjektive partikler ude af betragtning. Ifølge Hansen & Heltoft (2019:1657-1665) er det indre *at* styret af og har samme funktion i forhold til den konjunktion det står efter, som *at* i øvrigt har til matrixsætningens prædikat (dvs. når det står som eneste konjunktion). Det følger heraf at det indre *at*'s semantiske værdi tildeles af den konjunktion det efterfølger. I en sætning som *Dyrene skal ikke lide i kulden, fordi at man skal have lidt mere albuerum* er *at* altså styret af *fordi* hvis betydning det ”overtager” (eksemplet er fra Jyllands-Posten 27.01.2005, fundet i Korp). Desuden er det indre *at* ”bundet til den neutrale ledstilling og kan ikke stå i sætninger med deklarativ ledfølge” (Hansen & Heltoft 2019:1658). Vi har imidlertid en del eksempler i vores data på sætninger indledt af (især) *fordi*

at med deklarativ ledfølge, fx *fordi at han kom jo fra O_B_S hvor der kun var drenge* (LANCHART). Det ser vi ikke nærmere på i denne sammenhæng.²

Hansen & Heltoft (2019:1660) kommer i øvrigt også ind på det ulogiske ved at påstå at der skulle være forskel på *at*'et i *før at* og *efter at*; de argumenterer for at sætninger med *før at* og *efter at* er helt parallelle, og at begge tilfælde "uden vanskelighed" kan tolkes som "simple forekomster af netop det indre *at*" (se også Nguyen & Hartling 2022).

Hansen & Heltoft (2019:1660) angiver at det indre *at* har bevæget sig fra at være et sociolektalt og dialektalt træk (se også Pedersen 2009) "til nu at være ganske almindeligt også i talt rigsmål, især blandt yngre talere". Også Jacobsen & Jørgensen (2013:84) skriver at "[n]æsten alle bruger [...] *fordi at* i deres talesprog – også dem der påstår at de ikke gør det". I et spørgsmål-svar i *Nyt fra Sprognævnet* har vi i øvrigt rådgivet om at *efter* uden *at* i dag er helt gangbart, på lige fod med *efter at* (Hartling & Nguyen 2022).

Os bekendt er der ikke tidligere lavet kvantitative studier af 'konjunktion/præposition + *at*'. Boye et al.'s (2012) studium er imidlertid parallelt: De undersøger forekomsten af *at* i såkaldte komplementsætninger, dvs. "nominale ledsætninger, der kræves af et verbal eller et andet komplementtagende element[]" (Boye et al., 2012:42), fx *Han sagde (at) han ville gå hjem*. Som empirisk grundlag bruger de KorpusDK, som altovervejende er et skriftsprogskorpus, og talesprogskorpuset LANCHART. Vi vender kort tilbage til Boye et al.s undersøgelse i afsnit 7.

3. Hypoteser

Vi formoder at avistekster, og særligt tekster i landsdækkende aviser, som er dem vi undersøger, er repræsentanter for en relativt konservativ skriftsprogssnorm. De skrives primært af professionelle sprogbrugere – måske endda af de prototypiske "gode og sikre sprogbrugere" hvis "skriftlige sprogbrug" er den som dansk retskrivning skal hvile på (foruden på traditionen) (jf. *Bekendtgørelse om Dansk Sprognævns virksomhed og sammensætning* (<https://www.retsinformation.dk/eli/lta/2009/178>); se også Diderichsen & Schack 2015). Desuden undergår avisteksterne en redaktionel proces hvor formodentlig en redaktør og evt. en korrekturlæser sørger for at skriftsprogssnormen overholdes.³

² Der er skrevet en del om ledfølgen i ledsætninger, både i det talte og i det skrevne sprog. Om ledfølgen i ledsætninger med *fordi (at)*, se Gregersen (2019), Mikkelsen (2011) og Steensig (1998). Om placeringen af adverbialer (fx *ikke*) foran og efter det finite verbum, se Christensen et al. (2015) og Jensen (2011).

³ I 2021 afskedigede Berlingske ganske vist alle sine korrekturlæsere, og resurserne blev i stedet for allokeret til de indholdsredigerende, som er journalistuddannede. Ifølge chefredaktør Mette Østergaard var det ikke meningen med omstruktureringen at der skulle gås på kompromis med korrekturlæsningen (<https://journalisten.dk/berlingske-fyrer-alle-korrekturlaeserne/>).

Vi har altså en formodning om at forbuddet mod *fordi at*, *før at* og *efter* uden *at* stadig er gældende i de landsdækkende aviser. Med andre ord forventer vi følgende resultater af undersøgelsen af avisteksterne:

1. Proportionen af *fordi* er større end proportionen af *fordi at*.
2. Proportionen af *før* er større end proportionen af *før at*.
3. Proportionen af *efter* er mindre end proportionen af *efter at*.

Vi forventer endvidere en langt større andel af *fordi at*, *før at* og *efter* (uden *at*) i talesprogs-korpusset end i det aviskorpus vi vil bruge. Det skyldes bl.a. at man i talesproget formentlig ikke følger de præskriptive regler i samme grad som i aviser. Et sådant resultat vil altså måske kunne underbygge hypotesen om at det er de præskriptive regler der er på spil i skriftsproget: En spontan brug af *fordi at*, *før at* og *efter* (uden *at*) undertrykkes i skriftsproget, hvor man har tid til at komme i tanker om at det anses for dårligt sprog, og rette til. I netop den type skriftsprog vi undersøger, er der som nævnt typisk mere end én person om at overveje formen.

4. Metode

Vi undersøger vores hypoteser i to korpusser: skriftsprogskorpusset Korp og talesprogskorpusset LANCHART. Korp er som nævnt Dansk Sprognævnns interne korpus, og LANCHART er et ca. 13 mio. ord stort talesprogskorpus som huses af Sprogforandringscentret ved Københavns Universitet (DGCSS).

I den del af Korp vi undersøger, er der omtrent 261 mio. løbende ord. Dataene er fra landsdækkende aviser fra perioderne 2004-2005 og 2009-2016.⁴ Vi søgte efter følgende streng:

- *fordi/før/efter (+ at) + han/hun/man + (adverbium) + finit verbum*

Med denne søgerække ville vi begrænse antallet af irrelevante forekomster. Vi valgte at søge på pronominer (i nominativ) på subjekspladsen i stedet for substantiver og prænominale elementer (som fx adjektiver og artikler) for at undgå irrelevante eksempler hvor et nominalsyn-tagme fungerer som styrelse, fx *før/efter folketingsvalget* eller *før/efter det vigtige folketingsvalg*.

⁴ At vores skriftsprogsdata er fra netop disse perioder, skyldes at vi søgte med ordklassetagging, og Korpdataene er kun ordklassetagget i disse perioder.

Den del af LANCHART-korpuset vi undersøger, indeholder ca. 10,3 mio. løbende ord. Det svarer til det samlede LANCHART-korpus, fraregnet data fra udvandrere og deres efterkommere i USA, Canada og Argentina. Langt de fleste interviews i delkorpuset er udført i 2000'erne og 2010'erne, men der er også data fra 1970'erne, 1980'erne og 1990'erne. I dette korpus søgte vi efter følgende:

- *fordi/før/efter + 0-2 tokens (+ at) (+0-2 tokens) + han/hun/man*

Med de 0-2 tokens i søgestrengen har vi taget højde for at det er talesprog med alt hvad det indebærer af tøven, gentagelser osv. Foruden regulære ord dækker tokens netop også over andre typer ytringer, fx udråb, minimalrespons (fx et bekræftende *mm*), selvafbrydelser, tøvelyde, latter m.m.

Alle Korp- og LANCHART-eksempler blev manuelt overført til Excel. Her fjernede vi efterfølgende irrelevante eksempler: dels ufuldstændige eksempler hvor der ikke tydeligt var en komplet ledsætning, herunder anakolutier; dels – i talesprogsdataene – eksempler hvor der ikke følger et finit verbum (foruden eventuelle tøvelyde, gentagelser o.l.) efter *han/hun/man*, og sætninger hvor *før* og *efter* slet ikke indleder en ledsætning (fx *det vil hun ikke høre efter hun vil have de øl og de piller som hun vil have*). Blandt skriftsprogseksemplerne er der dubletter fordi aviserne indimellem udgiver ens eller næsten ens tekster. Alle dubletter blev talt med fordi det givne eksempel kan siges at være godtaget til udgivelse flere gange – af flere redaktører. Under alle omstændigheder var antallet af dubletter så lille at det ikke ville gøre nogen forskel for den statistiske analyse om de blev talt med eller ej.

5. Resultater

En sammentælling af forekomster på baggrund af den ovenfor beskrevne søgning ses i tabel 1 og 2. Vi ser først på resultaterne for de skriftsproglige data i tabel 1:

Sætningsinitiale elementer	Antal forekomster	Antal forekomster i %
<i>fordi</i>	28.102	99,96 %
<i>fordi at</i>	11	0,04 %
<i>før</i>	10.657	99,92 %
<i>før at</i>	9	0,08 %
<i>efter</i>	1.184	25,9 %
<i>efter at</i>	3.393	74,1 %

Tabel 1. Fordelingen af ledsætninger indledt af hhv. *fordi/før/efter* og *fordi at/før at/efter at* i avis-korpusset Korp.

Som man kan se i tabel 1, er *fordi at* og *før at* nærmest ikkeeksisterende i landsdækkende aviser.⁵ Omvendt er *efter at* hyppigere end *efter* uden *at*, selvom *efter* uden *at* også forekommer ret hyppigt, dvs. i lige knap 26 % af tilfældene. Vores hypotese 1, 2 og 3 er således understøttet, og vi kan konkludere at man i redigeret skriftsprog holder sig til de præskriptive forbud og påbud mht. brugen af *fordi* og *før* – og i vid udstrækning i brugen af *efter*. Med en chi-i-anden-*goodness of fit*-test (fordelingstest) kan vi godtgøre at der er en statistisk skæv fordeling mellem antallet af *efter*-ledsætninger og *efter at*-ledsætninger i avis-korpusset ($p < 0,0001^6$), under den nulhypotese at de er ligeligt fordelte. Det er altså næppe tilfældigt at vi finder så stor en overvægt af sætninger med *efter at* i vores data.

Vi antog at vi ville kunne underbygge hypotesen om at det er de præskriptive regler der spiller en rolle i skriftsproget hvis det viste sig at det forholdt sig omvendt i talesproget – eller at der i hvert fald her var en langt større andel af de normstridige ordforbindelser. Hvis *fordi at*, *før at* og *efter* uden *at* er mere almindeligt i det antageligt mere utvungne talesprog, må der ske en afretning af disse under overførslen til skrift, og det er plausibelt at det skyldes årtiers prædiken mod disse. Vi foretog således den samme søgning i LANCHART (med få modifikationer, jf. afsnit 3) og fik følgende resultater:

⁵ Det er væsentligt at bemærke at fem af de 11 forekomster af *fordi at* i avisdataene med sikkerhed stammer fra citater. Det samme gælder tre af de ni eksempler på *før at*. I de fleste af alle disse forekomster kan vi se inkvittet i vores søgeudtræk, fx det understregede her: [...] *fordi at han er godt kørende opad og har overskud*», *sagde Bjarne Riis efter etappen.*, og dermed slå fast at de stammer fra mundtlige interviews. Vi kan kun gisne om hvor mange *før at*'er og *fordi at*'er der er blevet bortredigeret fra sådant nedskrevet talt sprog.

⁶ Vi antager her at p skal være $= < 0,05$ for at vidne om statistisk signifikans.

Sætningsinitiale elementer	Antal forekomster	Antal forekomster i %
<i>fordi</i>	3.108	88,5 %
<i>fordi at</i>	402	11,5 %
<i>før</i>	308	95,7 %
<i>før at</i>	14	4,3 %
<i>efter</i>	146	86,4 %
<i>efter at</i>	23	13,6 %

Tabel 2. Fordelingen af ledsætninger indledt af hhv. *fordi/før/efter* og *fordi at/før at/efter at* i talesprogs-korpusset LANCHART.

Vi kan se af tabel 2 at der også i talesproget er mere *fordi* end *fordi at* og mere *før* end *før at*. Til gengæld er forholdet vendt på hovedet mht. *efter*: Der er langt mere *efter* end *efter at* i talesprogs-korpusset. Resultatet for *efter (at)* er det mest slående, og det ser vi lidt nærmere på nedenfor. I alle tre sætningsindlederpar er fordelingen også statistisk set skæv ($p < 0,0001$), dvs. det er statistisk underbygget at man i talesproget bruger mere *fordi*, *før* og *efter* end varianterne med *at*.

Ser man nærmere på resultaterne for *fordi (at)* og *før (at)* i lyset af vores hypotese om forholdet mellem skrift- og talesprog, fremgår det at andelen af forekomster af *fordi at* er langt højere i talesprog end i avissproget (11,5 % vs. 0,04 %), og andelen af forekomster af *før at* er noget højere i talesprog end i avissproget (4,3 % vs. 0,08 %). En chi-i-anden-uafhængighedstest af fordelingen af hhv. *fordi (at)* og *før (at)* i skrift- og talesprog viser at den observerede forskel også er statistisk signifikant (med en nulhypotese om at fordelingen i skrift- og talesprog for hhv. *fordi (at)* og *før (at)* skulle være ens). I begge tilfælde er $p < 0,001$.

Der er altså statistisk belæg for at der er mere *fordi at* og *før at* i talesproget end i avis-skriftsproget. Dette resultat er i sig selv ikke overraskende, og det stemmer overens med vores opstillede hypoteser. A priori ville vi især forvente langt mere *fordi at* i talesprog end i avis-skriftsproget. Vi forventede også at der ville være mere *før at* i talesproget end i avis-skriftsproget (men at forskellen ville være langt mindre end for *fordi at*). Det hænger sammen med at vi ikke på forhånd havde nogen særlig fornemmelse af at *før at* i det hele taget var særlig udbredt, hverken i talt eller skrevet dansk.

Også fordelingen af *efter* og *efter at* i de to korpusser stemmer overens med hvad vi forventede: at der er mere *efter* uden *at* i tale end på skrift. Det overraskende er at der er så overvældende meget mere *efter* uden *at* i talesproget, nemlig godt 86 % overfor knap 26 % i avissproget. Forholdet er således vendt på hovedet.

6. Diskussion

Hvad angår hypotesen om at det er præskriptive regler der styrer brugen af de tre ledsætnings-initiale elementer i skriftsproget, er det mest overbevisende at det er sådan fat mht. *efter (at)*. Vi kan se i vores data at det endog er meget 'naturligt' at indlede en ledsætning med *efter* uden *at* i talesproget, og at denne 'naturlige tilbøjelighed' sandsynligvis i meget høj grad bliver undertrykt i skriftsproget. Vi mener at det er en absolut plausibel tolkning at den skriftlige sprogbrug er styret af præskriptive regler mht. brugen af *efter* og *efter at*. Det samme gør sig muligvis gældende for *fordi (at)* og i mindre tydelig grad for *før (at)* – her er forholdet mellem skrift og tale blot ikke vendt på hovedet (i modsætning til *efter (at)*).

Før at er i det hele taget ikke særlig brugt, hverken i talesprog eller i skriftsprog. Vi kan foreløbig kun gisne om hvad dette skyldes, fx lydlige forhold, strukturelle forhold, tilfældigheder? Der er i hvert fald én grammatisk kontekst hvor *efter at* er gangbart, men ikke *før at*, nemlig *at*-infinitivkonstruktioner, fx *han var meget træt efter at have optrådt hele aftenen*, men ikke **han var allerede meget træt før at have optrådt*. Det er tænkeligt at denne begrænsning i brugen af *før at* har en afsmittende effekt på forekomsten af *før + at*-ledsætning, og omvendt: at det faktum at *efter* kan indlede en *at*-infinitivkonstruktion, har en afsmittende effekt på forekomsten af *efter + at*-ledsætning.

Vi ville på forhånd have troet at forbuddet mod *fordi at* var det der – af disse tre – var mest salient i hovedet på sprogbrugerne. Hvad der foregår i hovedet på sprogbrugerne, ville vores data ganske vist ikke kunne sige noget om, men tallene kunne have givet et fingerpeg. Der er imidlertid ikke noget i tallene der kan indikere at forbuddet mod *fordi at* er mere salient når vi sammenligner *fordi at* og *før at*. Som man kan se i tabel 1, er *fordi at* og *før at* jo omtrent lige sjældent i avisskriftsprog (0,04 % overfor 0,08 %), og som man kan se i talesprogsdataene (tabel 2), er *fordi at* en del hyppigere end *før at* (11,5 % overfor 4,3 %; også denne forskel er statistisk signifikant, jf. afsnit 5). Hvis sprogbrugerne var meget mere bevidste om forbuddet mod *fordi at*, kunne man måske have forventet en større forskel på *fordi at* og *før at*, sådan at der var væsentligt mindre *fordi at* end *før at* i både tale og skrift.

Imidlertid er det muligt at man bør lade *før (at)* lidt ude af betragtning og snarere sammenligne *fordi (at)* og *efter (at)*. Som vi var inde på ovenfor, har *før at* muligvis aldrig været lige så udbredt i dansk (bortset fra i visse dialekter) som *fordi at*, og det ville altså kunne være derfor at *før at* i det hele taget er sjældent. Hvis man derimod ser på *fordi (at)* overfor *efter (at)* i avisskriftsproget (som er dér hvor en præskriptiv bevidsthed har de bedste udfoldelsesmuligheder), kan man se at det ildesete *fordi at* er langt mere undertrykt end det ildesete *efter* uden *at*. Det kunne tyde på at forbuddet mod *fordi at* er salient for sprogbrugerne.

Samtidig ville vi på forhånd have troet at *fordi at* var langt mere anvendt end vores resultater peger på – især i talesproget, jf. bemærkningerne ovenfor fra Jacobsen & Jørgensen (2013) om ordforbindelsens udbredelse og almindelighed. Vi mener netop også selv at høre det ganske ofte. Imidlertid er det jo muligt at vi og Jacobsen & Jørgensen tror det er mere udbredt end det er, **netop** fordi *fordi at* afgjort er den mest berømte og berygtede af de tre præskriptivt forkerte ledsætningsindledere *fordi at*, *før at* og *efter* uden *at*. Denne berygtethed kan altså have en *primende* effekt i folks bevidsthed som i sig selv kan føre til dels færre forekomster i den aktive sprogbrug, dels en passiv fornemmelse af at det er langt hyppigere end det egentlig er.

For så vidt at resultaterne ovenfor er udtryk for at præskriptive regler stadig gør sig gældende (i skriftsproget), så formoder vi at de lever videre på det konkrete enkeltordsniveau, ikke på det abstrakte, generelle niveau. M.a.o. formoder vi at sprogbrugerne har reglerne: ”*fordi* uden *at*”, ”*før* uden *at*” og ”*efter* med *at*” i baghovedet når de skriver (eller redigerer/korrekturlæser), snarere end forbuddet mod to konjunktioner der efterfølger hinanden, og påbuddet om at en ledsætning skal indledes af et *at* når den står som styrelse for en præposition.

Grundlæggende må man fastslå at man ikke kan vide hvorfor en sprogbruger vælger den ene nærsynonyme sprogform fremfor den anden. Det kan skyldes at sprogbrugeren med sin umiddelbare intuition, slet og ret, foretrækker en form fremfor en anden. Det kan også skyldes tilfældigheder. Det blotte faktum at der ikke er noget der tilsiger at den ene form er mere grammatisk end den anden, kan imidlertid pege på at de præskriptive regler i det mindste spiller en rolle når sprogbrugeren anvender en given sprogform og ikke en anden. Og denne konklusion er endnu mere sandsynlig hvis den ene sprogform er altovervejende mere frekvent end den alternative – på tværs af sprogbrugere.

7. Sammenfatning og perspektivering

I redigeret moderne dansk skriftsprog i de landsdækkende aviser følger man (bevidst eller ubevidst) i meget høj grad de gamle præskriptive regler om ikke at bruge det indre *at*, i hvert fald hvad angår de ledsætningsinitiale elementer vi har set på: *før* og *fordi*. Samtidig følger man i vid udstrækning reglen om at en ledsætning der kommer efter *efter*, skal indledes af *at*. Kombinationerne *før at* og *fordi at* er ligefrem næsten helt fraværende i de landsdækkende aviser. *Efter* uden *at* som ledsætningsindleder bruges derimod i omkring en fjerdedel af tilfældene.

I moderne dansk talesprog (fra 1970'erne og frem, dog primært fra 2000'erne og 2010'erne) bruges der langt overvejende *fordi*, *før* og *efter* uden *at*. Mest overraskende er det umiddelbart at det præskriptivt set forkerte *efter* uden *at* er det afgjort dominerende i talesproget (med godt 86 % *efter* uden *at*).

Ser man på forholdet mellem skrift og tale, finder vi at der er meget mere *fordi at* og noget mere *før at* i tale end på skrift. Samtidig var det ventet at der ville være mere *efter* uden *at* i talesproget end i skriftsproget, blot ikke at forholdet ville være vendt på hovedet.

Ved hjælp af chi-i-anden-tests kunne vi godtgøre at de forskelle vi observerede også er statistisk signifikante, og det gælder både forskellene mellem varianterne internt i hhv. skriftsprog og talesprog og forskellene mellem varianterne på tværs af skrift- og talesprog. Alt i alt understøtter tallene vores hypoteser mht. andele af forekomster af de ledsætningsinitiale elementer med og uden *at*.

Vi havde endvidere den hypotese at en signifikant større andel af de præskriptivt set forkerte *fordi at*, *før at* og *efter* uden *at* i talesproget end i skriftsproget kunne pege på at man i skriftsproget er styret af netop de præskriptive regler: at man bortredigerer *at*'erne efter *fordi* og *før* og tilføjer *at* efter *efter* netop med baggrund i sådanne indlærte regler. Vi kan se i vores data at *før at* generelt er meget lavfrekvent i både skrift- og talesprog; det giver dermed ikke megen mening at sige at der er et 'naturligt' forekommende *før at* som avisskribenterne og -redaktørerne luger ud i fordi de ved at det anses for forkert. Derimod mener vi at det er endog meget plausibelt at *fordi at* og i allerhøjeste grad *efter* uden *at* bliver rettet til *fordi* og *efter at* i avisteksterne som følge af de præskriptive regler herom. Vi må dog understrege at vi ikke kan **vide** hvorfor den ene form er valgt frem for den anden, og hvilke overvejelser – hvis nogen overhovedet – der ligger bag sprogbrugerens valg.

I fremtidig forskning ville det være særlig interessant at se nærmere på om der er strukturelle og/eller diskursive mønstre i brugen af varianterne med og uden *at*, særligt *efter* overfor *efter at* i skrift- og i talesproget, men også *fordi (at)* og *før (at)* i talesproget. (Derimod er der ikke mange mønstre at granske mht. *fordi (at)* og *før (at)* i avisteksterne, hvor varianterne uden *at* er næsten eneherkende). Bl.a. ville det være interessant at se på om vi parallelt med Boye et al. (2012) ville kunne finde at proportionen af ledsætninger med *at* er desto større, jo vanskeligere sætningerne er at producere.

Litteratur

- Boye, Kasper, Mads Poulsen, Hannah Bruun Pedersen, Marie Herget Christensen, Line Dalberg & Nicoline Munck Vinther (2012) *At* eller ikke *at* i tale og skrift. I *NyS, Nydanske Sprogstudier* 42:41-61.
- Christensen, Tanya Karoli, Torben Juel Jensen & Marie Herget Christensen (2015) Adverbielle ledsætningers ledstilling i dansk talesprog. I Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.), *15. Møde om Udforskningen af Dansk Sprog*: 97-115.
- Diderichsen, Philip & Jørgen Schack (2015) Jagten på den gode og sikre sprogbruger. I *Nyt fra Sprognævnet* 2015/3:1-8.
- Gregersen, Sune (2019) Årsagssætninger med *fordi* i (formelt) dansk talesprog. I *NyS, Nydanske Sprogstudier* 57:81-112.
- Hansen, Erik (1975) *At* eller ikke *at*. I *Nyt fra Sprognævnet* 14:1-4.
- Hansen, Erik (1983) Det pleonastiske *at*. I *Danske studier* 78:61-80.
- Hansen, Erik & Lars Heltoft (2019) *Grammatik over det danske sprog* (2. udg.). Odense: Syddansk Universitetsforlag.
- Hartling, Anna Sofie & Michael Nguyen (2022) *Efter* eller *efter at*. I *Nyt fra Sprognævnet* 2022/1:18.
- Jacobsen, Henrik Galberg & Peter Stray Jørgensen (2013) *Politikens Håndbog i Nudansk*. 6. udg. Politikens Forlag.
- Jensen, Torben Juel (2011) Ordstilling i ledsætninger i moderne dansk talesprog. I *Ny forskning i grammatik* 18:123-150.
- Mikkelsen, Nicholas Hedegaard (2011) Betydningsforskelle ved brugen af helsætningsordstilling efter *fordi*. I Inger Schoonderbeek Hansen & Peter Widell (red.), *13. Møde om Udforskningen af Dansk Sprog*: 213-224.
- Nguyen, Michael & Anna Sofie Hartling (2022) Konjunktion eller præposition? Om nogle ledsætningsinitiale elementers ordklasse. I *Ny forskning i grammatik* 29:139-155. <https://doi.org/10.7146/nfg.v1i29.132905>
- Pedersen, Karen Margrethe (2009) Indre *at* i danske dialekter. I Rita Therkelsen & Eva Skafte Jensen (red.), *Dramatikken i grammatikken: Festskrift til Lars Heltoft*. Roskilde: Institut for Kultur og Identitet, Roskilde Universitet: 321-333.
- Steensig, Jakob (1998) Om *fordi* i forskellige sætningstyper i dansk talesprog. I *Selskab for Nordisk Filologi: Årsberetning* 97:179-192.

Diktaroriske befølelser

Om ord og uord i Det Centrale Ordregister

Peter Juel Henriksen

Dansk Sprognævn

To kendte russiske historikere Andronik Mirganjan og Igor Klamkin tror ikke, at Rusland kan udvikles uden en "jernnæve". De hævder, at Ruslands vej til demokrati går gennem diktatur. I en af deres artikler hedder det: "I et autoritært regime lagdel samfundet og forskellige interesser modnes. Og når deres repræsentanter er parate til at gå i struben på hinanden, så stopper en "jernnæve" det. På den måde skabes hele tiden betingelserne for en harmonisering af interesser og følgelig for demokratiske reformer".

Og de tilføjede: "Netop fordi vi end ikke har et kim af civilt samfund, var vi imod Folkekongressen. Bortset fra illusionen om demokrati kunne den ikke give noget reelt. Det kunne den ikke helt objektivt. Det var slet ikke på grund af apparatets lænker. Illusioner er farlige. De føder skuffelser og leder til sidst til en destabilisering af samfundet, hvad vi allerede har set og desværre vil komme til at se mere af." De mener, at Folkekongressen skal give præsidenten diktaroriske befølelser.

Det danske PAROLE-korpus (to første afsnit)¹

1. Det Centrale Ordregister – en introduktion

Det Centrale Ordregister (COR) er et register for det danske ordforråd. Hvert dansk ord har (eller kan få) et COR-nummer der fungerer som ordets leksikalske holdepunkt.

¹ PAROLE-korpusset består af små, ukomplette tekststykker og kan derfor citeres frit iflg. citatretten (se Link3 2023). Tekststykket er citeret fra en artikel i Det Fri Aktuelt d. 1/12 1992 (nærmere info herunder).

'To' / COR.01528.600.01
'kendte' / COR.18159.303.01
'russiske' / COR.22261.303.01
'historikere' / COR.58774.112.01

Her ses de første fire ord i det danske PAROLE-korpus annoteret med COR-numre (herfra kaldet *links*). Et COR-link består af præfikset "COR" efterfulgt af tre numeriske indekser.

COR . <lemmaindeks> . <bøjningsindeks> . <variantindeks>²

Lemmet *russisk* har altså lemmaindekset 22261 mens *historiker* har 58774. Bøjningsindekset viser ordets morfologiske klasse og bøjning, fx står 112 for bøjningen sb.fk.pl.ubest (i Retskrivningsordbogens notation) mens 303 står for adj.pl. Sidst i artiklen findes en tabel over de bøjningsindekser der forekommer her i artiklen (den fulde definition ses i ordregister.dk, både som tabel og som opslagsfunktion).

Såvel lemmaindekser som bøjningsindekser er entydige, dvs. to forskellige lemmer vil altid have to forskellige lemmaindekser, og tilsvarende har forskellige bøjningsformer (fx 'russiske'_{sg.best} og 'russiske'_{pl}) altid forskellige bøjningsindekser (hhv. 302 og 303). Tekst der er annoteret med COR-links, er altså fri for homografi og er dermed lettere at databehandle (for både mennesker og computere).

COR-annoterede tekster er gode korpusressourcer – bedre end de velkendte Part-of-Speech-taggede korpora. PoS-tagging er, som bekendt, en af de hyppigst anvendte metoder til forberedelse af inputtekst til maskinoversættelse, talesyntese, tekstbehandling og mange andre slags NLP (*Natural Language Processing*). Ulempen ved PoS-tagging er at selve tagdefinitionen, altså de tags som man klassificerer tekstens tokens med, varierer fra det ene projekt til det næste. En grov ordklassebestemmelse er fx nok til at støtte en talesyntese mens et tekstbehandlingsprogram med grammatikkontrol behøver en meget finere morfologisk klassifikation. To korpora der er annoteret med forskellig tagdefinition, er besværlige at lægge sammen, i praksis ofte umulige (Kirchmeier et al. 2020). En af fordelene ved COR's standardiserede bøjnings- og

² Variantindekset bruges i de tilfælde hvor ortografien tillader flere forskellige stavninger af den samme ordform (fx "ressource" eller "resurse", "allevegne" eller "alle vegne"). Lemmaindekser har 5 cifre, bøjningsindekser har 3, variantindekser har 2. Det faste format gør indekserne lette at genkende i artiklens eksempler. Formatet gælder i grundregisteret, emnet for denne artikel, men ikke for alle COR-ressourcer (omtales herunder).

lemmainformation er at COR-annoterede sprogressourcer lavet til projekt A og projekt B umiddelbart kan lægges sammen og bruges som korpus i projekt C, uanset hvilke forskelle der måtte være mellem A, B og C's formål og metoder.

COR-links er mere informationsrige end PoS-tags. En COR-annoteret tekst afbildes på en PoS-tagget tekst simpelthen ved at fjerne lemmaindekserne. Fjerner man dertil de to sidste cifre i bøjningsindekserne, står hovedordklasserne tilbage (1 for substantiver, 2 for verber, 3 for adjektiver osv.). Har man brug for en lemmatiseret tekst, fjernes i stedet bøjningsindekserne. På den måde erstatter COR-linking såvel PoS-tagging som lemmatizing.

Variantindekset spiller en mindre rolle, og vi udelader det i resten af artiklen for læselighedens skyld; men faktisk kan også dette indeks gøre nytte for eksempel som input til et tekstbehandlingsprogram med stilkontrol (forfatteren er fx tilbøjelig til at flakke mellem varianterne 'resurse' og 'ressource', og her ville en løftet pegefinger være rart).

I artiklens første del præsenterer vi COR's grundregister (COR1), som alle andre COR-ressourcer refererer til. Man kan frit downloade COR1 og andre COR-relaterede sprogressourcer (ordbøger, korpora, værktøj etc.). En af de nyeste tilføjelser er PARC, det danske PAROLE-korpus³ med COR-annotation, der har leveret teksteksemplerne til denne artikel. I artiklens sidste del giver vi et overblik over de vigtigste COR-ressourcer og -metoder, med omtale af bl.a. automatisk COR-linking og praktiske COR-anvendelser i sprogteknologi. Vi må erkende at det bliver et hastigt flyby, men vi håber at læseren får lyst til at følge referencerne til de uddybende publikationer.

COR-ressourcer kan downloades fra www.ordregister.dk, som også giver adgang til forskellige online-tjenester såsom leksikalske opslag i COR1.

2. Nøgen tekst er dårligt input

Computere har svært ved at arbejde med naturligt sprog. Sproget er fuldt af flertydigheder, som vi mennesker navigerer sikkert imellem takket være vores forståelse af andre menneskers intentioner; computerprogrammer, derimod, opfatter ethvert inpu-telement som et entydigt datapunkt uanset om det drejer sig om et ord, et ciffer, en lyd-sample eller en pixel. Når et NLP-program får galt fat på et ord, giver det tit anledning til følgeføj af en karakter som mennesker

³ Den oprindelige projektbeskrivelse for PAROLE indebærer at alle tokens skulle bevares præcis som de forekom i kildeteksterne. Korpusset er rigt på stavfejl og andre anomalier og derfor en god kilde til 'økologiske' tekstprøver. Læs om PAROLE i Bilgram et al. (1993, 1998), Keson (1999) og Henrichsen (2002).

aldrig ville begå, fejl uden en gnist af fornuft.

Når MS-Word's stavekontrol læser ordet 'hundehvalpenne', bliver det godkendt selv i konteksten "kattekillingerne og hundehvalpenne drak af samme vandskål". Ordet kan analyseres som et kompositum 'hunde-hval-penne' og er dermed et velstavet ord, pedantisk betragtet. En sætning som "killingerene og hvalpenne drag ad same skol" giver ikke anledning til en eneste kommentar fra funktionen Stave- og Grammatikkontrol. Alle ordene er jo velstavet (med 'skol' som imperativ). Der er ikke megen hjælp at hente for hverken dyslektikere eller andre.

Men stavetjekkerne er ikke værre end resten af NLP-familien. Hvis man siger til Siri: "Pæren i lampen er sprunget, hvor kan jeg købe en ny pære?", svarer hun "Jeg ved ikke hvor der er en grønhandler". Og hvis man serverer denne tekst til Google Translate:

"Jeg har lige været til to studenterfester. Lærkes var fed. Majs var dødssyg."

og beder om at få den oversat til engelsk, så vender Translate tilbage med denne tekstperle

"I just went to two student parties. Larks were fat. Maize was deathly ill."⁴

Translate har åbenbart forvekslet den intenderede form 'Majs'_{prop.gen} med dens homograf 'majs'_{sb.fk.sg.ubest} mens 'Lærkes'_{prop.gen} er forvekslet med 'lærkes'_{sb.fk.pl.ubest} (som ret beset er et uord).

Skal man pege på én grundlæggende udfordring som næsten alle slags sprogprogrammer lider under, uanset om de behandler tekst eller tale, så er det at de danske ordformer kun sjældent er leksikalsk entydige (ganske særligt de kortere ord). Det problem løser NLP-udvikleren, som før nævnt, ved at tage hvert teksttoken, fx morfologisk, prosodisk eller terminologisk, ud fra sit eget formål. Der er mange annotationsstandarder i brug i de danske NLP-værksteder i dag (Kirchmeier et al. 2019). Alt for mange. Det kan alle blive enige om. Derfor har den fælles og versatile annotationsstandard – lige relevant til alle NLP-formål – været et mål for datalingvistikken gennem mindst fire årtier. Desværre er vi ingen vegne kommet, og derfor står vi i dag med et pletora af inkompatible tekstkorpora. En nylig undersøgelse (Kirchmeier et al. 2020) konkluderede at, ud af 100+ store danske korpora skabt for offentlige midler gennem de seneste 30 år, har kun omkring fem relevans i dag. Resten er endt på datakirke-

⁴ Afprøvningen af MS-Word (aktuel version distribueret af Statens-IT), Google Translate (i Chrome) og Apple's Siri (afprøvet på iPad, ny model 2022) blev foretaget d. 5/1 2023.

gården, ikke mindst på grund af inkompatible annotationsformater.

COR-linking blev udviklet som et versatilt alternativ. Vi vender tilbage til hvordan COR-baseret tekstbehandling kan bruges mod homografblindhed (som i eksemplerne herover).

3. COR's grundregister

COR's grundregister, kaldet COR1, er afledt af Retskrivningsordbogen og tildeler hver af RO's godt 530.000 ordformer et unikt COR-indeks. RO dækker, ifølge sine redaktionsprincipper, det almene danske ordforråd forstået som de ord der (i) benyttes i hele samfundet (dvs. ikke er typiske for bestemte alders- eller samfundsgrupper), (ii) forekommer stabilt over tid og (iii) generelt ikke refererer til specifikke organisationer, sagsforhold, produkter og personer (Jervelund et al. 2012). Den offentlige sektor er forpligtet til at følge RO's ortografi i skriftligt arbejde, fra ministerierne til folkeskolen (iflg. Lov om Dansk Retskrivning, se ref. Link2 1997). RO er – ifølge sine selektionskriterier, informationstyper og status som ortografisk norm – relevant for næsten alle typer af NLP-applikationer og er dermed det naturlige fundament for Det Centrale Ordregister.

Her er nogle eksempler (fra PARC) der viser hvordan homografer kan adskilles med COR1-links, selv i tilfælde hvor der ikke er forskel på deres stavning og bøjningsparadigme. Hvis COR-linking skal slå an for alvor, er der imidlertid to spørgsmål som må besvares: Hvordan linker man et token som ikke forekommer i COR1? Og hvorfra skal NLP-applikationen hente de arter af leksikalsk information som ikke findes i RO (stavning, bøjning og sammensætningspotentialer)? Det ser vi nærmere på i de næste afsnit.

4. COR-linking af det ukendte token

Når man afprøver en almindelig PoS-tagger på en tekst, får hvert eneste token et tag, selv de særeste stavfejl. Det ukendte token er ikke et principielt problem for en tagger, den skal bare klassificere det ud fra dets materielle lighed med kendte tokens. Linkeren, derimod, er forpligtet til at finde et leksikalsk mål til hvert eneste token. Lad os tage nogle konkrete eksempler. I PAROLE-citatet først i artiklen er der tre tokens (foruden *proprierne*) som ikke findes i COR1: 'jernnæve', 'diktaroriske' og 'befølelser', og desuden også formen 'lagdel' der ikke synes at passe i sin syntaktiske sammenhæng. Disse fire anomale tokens kan annoteres på forskellig måde i COR, afhængigt af hvilken status man ønsker at tillægge teksten; er den fx kanonisk, et 'værk', så alle tokens *skal* tages bogstaveligt, eller er den et løbende input til et tekstbehandlingssystem som skal granskes og rettes?

Der er i alt fald seks måder at COR-linke det ukendte token på, som alle bidrager med forskellig teksttolkning og dermed tilgodeser forskellige typer af NLP. Tabel 1 kobler de forskellige varianter og NLP-typer til hinanden og giver eksempler på de enkelte varianters relevans.

<i>Variant</i>	'jernnæve'	'lagdel'	'diktatoriske'	'befølelser'	<i>NLP-relevans</i>
<i>afvisning</i>	00000.000	00000.000	00000.000	00000.000	–
<i>PoS-tagging</i>	00000.110	00000.204	00000.303	00000.112	tegnsætning
<i>korrektion</i>	–	37419.204	29293.303	77788.112	stavekontrol oversættelse
<i>abstraktion</i>	–	37419	–	–	emnesøgning resumering
<i>analyse</i>	[97002] [] [98252.110]	–	–	[be-] [30029] [-else] [-r]	talesyntese læremidler
<i>accept</i>	PARC.p.110	PARC.q.204	PARC.r.303	PARC.s.112	tekstkanon

Tabel 1. COR-linking af det ukendte token – seks forskellige tilgange. Id'er af typen PARC.p.110 er forkortelser af COR.PARC.p.110 (osv.). Symbolerne p, q, r og s er lemmaindekser i ordbogen COR.PARC. Den højre kolonne giver eksempler på relevante NLP-formål.

De seks tilgange til annotation af det ukendte token har forskellige fordele og ulemper. Her må vi nøjes med at give nogle stikord, men reelt afspejler de en grundlæggende holdningsforskel til teksten.

Afvisning. Blot at konstatere at et ukendt token er *out of vocabulary* (ofte kaldet OOV) er det mindst risikable, men også det mindst værdiskabende valg.

PoS-tagging. Hvis formålet er at forberede en tekst til automatisk tegnsætning (kommatering), er en klassisk PoS-tag et naturligt valg. I COR-termer svarer det til kun at instantiere bøjningsindekset og lade lemmaet forblive ukendt.

Korrektion. En automatisk oversætter er afhængig af en egentlig korrektur, altså en erstatning af det ukendte (eller fejlbojede) token med et kendt alternativ.

Abstraktion. Denne tilgang er komplementær til PoS-tagging: at identificere lemmaet, men

ikke bøjningen. Automatisk tekstresumering og emnesøgning refererer typisk kun til lemmaer, ikke bøjningsformer.

Analyse. En talesyntese omsætter tekst til lyd, med lydskrift som mellemed. Udtaleordbogen er ofte et 'morfemikon' med udtaleregler for morfemer snarere end fuldformer. Det ukendte token skal derfor analyseres i morfemer, så langt som muligt.

Accept. I de fem første tilfælde er det ukendte token tolket som en tekstfejl. Men hvad hvis inputteksten er 'kanonisk', altså fejlfri a priori? I det tilfælde skal der oprettes en ny COR-ordbog til supplement af COR1 (in casu ordbogen COR.PARC), og det ukendte token linkes til et helt nyt leksem.

5. Nye COR-ordbøger

COR's grundregister dækker naturligvis ikke det samlede danske ordforråd (hvordan det end bestemmes), og derfor er det nødvendigt også at kunne linke *ud af* COR1. Oplagte eksempler på leksemer der er svagt repræsenteret i COR1, er proprierne og numeralierne, som RO kun dækker nødtørfigt. COR1 mangler også fagord, fremmedord, neologismer, samt en ubegrænset mængde af komposita. Dertil kommer interpunktionstegnene, som sprogfolk ganske vist ikke regner for rigtige sprogtegn (med Victor Borge som en undtagelse), men som ikke desto mindre bærer syntaktisk information der er uhyre relevant for NLP. Hvis COR-linking skal være sit salt værd, må ethvert teksttoken derfor kunne få et link. I dette afsnit diskuterer vi hvordan nye ordbøger kan føjes til COR-systemet og give NLP-programmerne mulighed for at importere leksikalsk information fra mange sider.

Vi har allerede nævnt den nye ordbog COR.PARC (se tabel 1), der dækker de ordformer i PAROLE-korpusset som ikke findes i COR1. Korpus PARC kan downloades fra ordregister.dk, ligesom ordbøgerne COR.PARC og COR.XP, sidstnævnte med COR-indekser for alle de almindelige interpunktionstegn, fx

! / COR.XP.1.03

, / COR.XP.2.01

(/ COR.XP.8.02

For at gøre en eksisterende ordbog COR-kompatibel skal hver enkelt leksikalsk indgang blot

forsynes med et unikt COR-indeks efter visse formelle regler⁵. Når alle indgange er indekseret, kan man sikre at ordbogen i sin helhed er formelt COR-kompatibel ved hjælp af Sprognævnets COR-værktøj. Når en ny COR-kompatibel ordbog er klar, kan man vælge at distribuere den frit; men alternativt kan man nøjes med at udstille dens opslagsord og holde den leksikalske information skjult bag en licens- eller betalingsmur (se tabel 2). På den måde kan fx en salgsvirksomhed gøre opmærksom på sit leksikalske produkt, dets omfang og dækning, uden at offentliggøre indholdet.

Artikler	BETALINGSMUR	COR-indeks	COR1-afbildning
		COR.OPEN.SMOVS.702 50	COR.70250 [bouillon] ← COR.96878 [kapital]
		COR.OPEN.SMOVS.600 58	COR.22093 [adj. genert] COR.32735 [vb. undgå] COR.60058 [sb. gelé] ← COR.73562 [sb. vanddamp]

Tabel 2. Firmaet Smovs-i-Sovs har udgivet en COR-kompatibel ordbog. Tabellen viser to af ordbogens artikler ("fond" og "sky") med deres COR-indekser. Hvor det giver mening, kan indekser i ordbogen afbildes på COR1 (fx COR.OPEN.SMOVS.70250 ← COR.70250). Dermed kan firmaets sovse-app importere leksikalsk info fra RO uden at blive vildledt af de mange homografer til 'fond' og 'sky'. Bemærk at Smovs-i-Sovs (kun) har offentliggjort oplysningerne til højre for betalingsmuren; de fungerer som reklame for den nye ordbog. COR-annotation kan således også være en forretningsmodel.

Der findes, i skrivende stund, omkring 10 COR-kompatible ordbøger og korpora der er offentligt tilgængelige (heriblandt COR1 og PARC). Mange flere er på vej. Sprognævnet arbejder på at COR-annotere det største frit tilgængelige tekstkorpus i Danmark, nemlig DAGW (Danish GigaWord-corpus; Derczynski et al. 2021). Dette korpus, der tæller omkring 1 mia.

⁵ En ordbog er COR-kompatibel når hver indgang har et COR-indeks af formen COR.OPEN.NAVN.ix, hvor NAVN er navnet på ordbogen, og ix indekserne for de informationstyper der er defineret for ordbogen. Denne type indeksering kaldes niveau 3, og den er åben for alle. Professionelle sprogredaktioner (DSN's repræsentantskab iflg. Lov om Dansk Sprognævn, se ref. LINK1) har også mulighed for at indekserer i niveau 2 (med linkformat COR.NAVN.ix). COR-ordbøger i niv. 2 forventes at være kvalitetssikret i henhold til alm. leksikografiske principper (Widmann 2022; Dideriksen et al. 2022).

tokens, er hurtigt blevet et referencekorpus for dansk korpuslingvistik takket være sin størrelse og tilgængelighed. En COR-annoteret version vil opgradere DAGW betydeligt som sprogresource for både forskning og udvikling.

Også på den taleteknologiske front er der nyt i vente. Alexandra-instituttet er netop gået i gang med at udvikle Danmarks største frit tilgængelige talesprogs-korpus (projekt CoRaL). Over 1000 timers indtaling vil, efter planen, blive transskriberet og tidskodet i de kommende år og gjort tilgængelige for træning af akustiske modeller til bl.a. talegenkendelse. Alle transskriptioner bliver COR-annoteret.

Og sidst, men ikke mindst, er en afgørende vigtig COR-ordbog ved at være færdig-udviklet, nemlig en omfattende betydningsordbog udviklet af DSL (Det Danske Sprog- og Litteraturselskab) og CST (Center for Sprogteknologi, KU), se Pedersen et al. (2022). Når ordbogen står klar i slutningen af 2023, vil den sætte nye standarder for hvordan dansk tekst kan annoteres semantisk – og dermed for udviklingen af NLP i retning ad kunstig intelligens.

6. COR-linket tekst som udgangspunkt for NLP

COR-linking giver et forbedret grundlag for de fleste typer af automatisk tekstprocessering. Det gælder ikke mindst tekstbehandling. Med COR-links kan man give et tekstbehandlingsprogram med stave- og grammatikkontrol stærkere briller på, og det giver vi her et eksempel på.

”I et autoritært regime lagdel samfundet og forskellige interesser modnes.”

Set fra et grammatisk synspunkt kan denne tekststreng ikke umiddelbart analyseres som en velformet sætning på grund af det skæve 5. token, ’lagdel’. En rimelig hypotese, understøttet af en rundspørge blandt kolleger, er at ordet egentlig skulle have været ’lagdeles’. En ordentlig stavekontrol bør altså foreslå korrekturen ’lagdel’→’lagdeles’, med det bagvedliggende argument at sætningen dermed bliver normaliseret både syntaktisk og semantisk. Med den COR-linkede tekst som udgangspunkt kan man opstille det følgende formelle ræsonnement (som kan overlades til en computer).

- 1 Teksten kan ikke, som den står, analyseres som en sætning. Det antages at årsagen er en stavefejl. Der gives derfor licens til at ændre ét token; korrekturen skal dog være realistisk (begrænset til nogle få bogstaver og også på anden måde minimal). Denne *licence to change* kalder vi *LiC*.
- 2 Der er to sætningsled (forstået som kongruensdomæner) som kan identificeres

- sikkert vha. COR-annotationen: ['et' 'autoritært' 'regime'] og ['forskellige' 'interesser']
- 3 Tekstens 7. token er 'og'₉₇₀. En ₉₇₀'er (sideordningskonjunktion) knytter to kongruerende led sammen.
 - 4 Konjunktionsfrasen [*V* 'og' *H*] har tre mulige højreled: $H=[\text{'forskellige'}]_{AP}$, $H=[\text{'forskellige' 'interesser'}]_{NP}$ og $H=[\text{'forskellige' 'interesser' 'modnes'}]_s$. Den første mulighed kan ikke finde et kongruerende venstreled *V* (modulo *LiC*, altså selv efter én korrektur); den anden kan godt kongruere med et venstreled, $V=[\text{'samfundet'}]_{NP}$, men det åbner stadig ikke for en sætningsanalyse (mod. *LiC*). Det tredje *H*-led tillader derimod en fuldstændig sætningsanalyse (mod. *LiC*). Med korrektoren 'lagdel'₃₇₄₁₉→'lagdeles'_{37419,204} kan teksten analyseres som [['I..'regime' 'lagdeles' 'samfundet']_s 'og' ['forskellige','interesser','modnes']_s]_s.
 - 5 Der udskrives et forslag til brugeren: korrektoren 'lagdel'→'lagdeles', evt. med en motiverende note.

Ræsonnementet i 1-4 viser sig altså at føre til et velmotiveret korrekturforslag. I modsat fald kunne man have løst *LiC* og givet flere frihedgrader. Arten og graden af korrektur kunne endda præsenteres som et valg for brugeren af tekstbehandlingsprogrammet.

Også korrekturerne 'diktatoriske'→'diktatoriske' og 'befølelser'→'beføjelser' kan understøttes med COR-baserede ræsonnementer; den sidstnævnte vil dog forudsætte en COR-ordbog med ret sofistikeret betydningsinformation. Vi venter spændt på DSL og CST's magnum opus.

7. Automatisk COR-linking

Indtil nu har vi diskuteret *nytt*en af COR-links, men ikke fremstillingen. Hvordan COR-linker man en tekst i praksis? Mindre tekststykker til grammatisk analyse og undervisningsmaterialer kan man naturligvis håndlinke; men til korpusannotation og NLP-udvikling er automatiske metoder et must. Sprognævnet har derfor udviklet et program til tekstannotation i stor skala, en såkaldt *linker*. Linkeren fungerer som en input-output-automat; den tager en tekststreng som input og udskriver teksten med hvert token annoteret med et COR-link.

Rent teknisk sker det ved at hvert teksttoken, som en indledende øvelse, annoteres med alle de links som er mulige i COR1. For eksempel har 'sky' ikke mindre end ni alternative links (tjek selv i ordregister.dk). Når 'sky' forekommer som token i en sammenhængende tekst, er

otte af de mulige links altså forkerte og skal fjernes. Algoritmen bag COR-linkeren kan beskrives som et loop, en stadigt gentagen analyse af teksten in toto. For hvert gennemløb fjernes ét link, og dermed ændres også beslutningsgrundlaget for den følgende iteration. Bliver man ved længe nok, står teksten til sidst tilbage, færdiglinket.

Sprognævnets COR-linker har fået navnet CLINK (af oplagte grunde), og programmet er p.t. i version 1.0. Denne version bruger tre forskellige strategier, kaldet LSYN (lokalsyntaktisk analyse), CTXT (kontekstanalyse) og FREQ (frekvens). Hver af de tre strategier er implementeret som et selvstændigt modul. LSYN ræsonnerer efter klassiske syntaktiske principper (som i afsnit 6) hvorimod CTXT benytter semantiske associationer mellem lemmaer uden at skele til bøjningen. FREQ foretager bare simple opslag i en frekvensordbog. De tre moduler er seriekoblet, sådan at det ene moduls output er input til det næste. Man kan eksperimentere med at ændre på deres rækkefølge, og desuden kan man skyde nye moduler ind. For tiden arbejder en af Sprognævnets tekniske assistenter på at træne et neuralt net (baseret på *TensorFlow*) som også kan indgå som modul. CLINK-moduler er hyggelige at programmere, og vi opfordrer alle til at bidrage. Med hvert nyt, velskrevet modul bliver CLINK stærkere, til gavn for os alle. Læs mere om CLINK i ordregister.dk og referencer dér.

8. Det nye økosystem af danske sprogressourcer

Det Centrale Ordregister har været under udarbejdelse siden april 2020 (på en statslig bevilling administreret af Digitaliseringsstyrelsen); men selve projektideen er af langt ældre dato. Gennem de seneste 20 år er det blevet stadigt tydeligere at Danmark er et svært sted at overleve som NLP-udvikler (Pedersen et al. 2011, Kirchmeier et al. 2019). Vores sprogsamfund er lille, og vores talesprog er vanskeligt at processere som data – men det kan vi ikke gøre noget ved. Hvad vi *kan* gøre noget ved, er tilgængeligheden af grundlæggende sprogressourcer for det danske sprog. Staten har, ind til 2020, stort set ikke bidraget til at sørge for gode og frit tilgængelige sprogressourcer, sådan som det er sket i vores nabolande. Hver NLP-virksomhed har måttet skabe sine egne sprogdata – maskinlæsbare ordbøger, tekstkorpora, talemateriale, transskriptioner og sprogmodeller. Det kan ikke undre at et firma som har ofret tusinder af arbejdstimer på at skabe et talekorpus for at kunne træne en talegenkender, ikke føler trang til at dele sit korpus med andre. Derfor har danske taledata til brug for NLP stort set været låst inde i firmaerne i årtier. Hvilke chancer giver det et nyt lille firma for at teste en god ide til et sprogprodukt og føre det til marked? Ikke mange. Det er for dyrt at komme gang når man begynder uden data. Som følge heraf har udviklingen af ny sprogteknologi til det danske sprog været begrænset til mellemstore danske virksomheder og internationale kæmper. Sverige har i

modsatning til Danmark i snesevis af mikrofirmaer med egen NLP-udvikling der drager nytte af en række offentlige sproressourcer, ligesom Norge, Finland, Island og Færøerne (Kirchmeier et al. 2019; Simonsen et al. 2022).

Men nu er der grøde i luften herhjemme. Staten bevilgede i 2020 omkring 40 mio. kroner til udvikling af danske sproressourcer. Nye materialer til NLP er under udarbejdelse og bliver løbende frigivet open-source. Hold øje med Digitaliseringsstyrelsens hjemmeside www.sprogteknologi.dk, hvor historien om Statens initiativ bliver fortalt og de nye sproressourcer præsenteret.

Vi i COR-projektet (DSN, DSL og CST) mener at man som dansker har ret til sit sprog, også som materiale til NLP. Vi ser Det Centrale Ordregister som vores bidrag til den nye demokratisering og opfordrer alle korpusejere og ordbogsredaktioner til at gøre deres data frit tilgængelige – cum COR.

Litteratur

- Bilgram, Thomas; Hans Arndt (1993) Automatisk morfologisk analyse af danske tekster. I Mette Kunøe og Erik Vive Larsen (red.). *4. Møde om Udforskning af Dansk Sprog*, Aarhus Universitet.
- Bilgram, Thomas; Britt Keson (1998) The Construction of a Danish Tagged Corpus. I *NODALIDA '98 Proceedings*.
- Derczynski, Leon et al. (2021) The Danish Gigaword Corpus. *Proceed. of NODALIDA-23*. Linköping Electronic Conference Proceedings 178 (2021).
- Diderichsen, Christina; Peter Juel Henriksen; Thomas Widmann. (2022) Det Centrale Ordregister. I *Nyt Fra Sprognævnet*. Oktober 2022.
- Henriksen, Peter Juel (2023) Det Centrale Ordregister – et indeks for det danske ordforråd – en gave til dansk sprogteknologi. *Proceedings NFL-2022*, Lund Universitet.
- Henriksen, Peter Juel (2002) Sidste Års Aviser. *LAMBDA 27*, Institut For Datalingvistik (KU).
- Keson, Britt (1999) *Vejledning til det Danske Morfosyntaktisk Taggede PAROLE-korpus*. DSL Press.
- Kirchmeier, Sabine; Peter Juel Henriksen; Philip Diderichsen (2019) Dansk Sprogteknologi i Verdensklasse. *Rapport fra sprogteknologiudvalget under Dansk Sprognævn nedsat af Kulturministeriet*.
- Kirchmeier, Sabine; Bolette Sandford Petersen; Peter Juel Henriksen; Sanni Nimb; Philip Diderichsen (2020) World Class Language Technology – Developing a Language Technology Strategy for Danish. *Proceed. of LREC 2020*.
- Pedersen, Bolette; Nathalie Carmen Hau Sørensen; Sanni Nimb; Sussi Olsen; Ida Flørke; Thomas Troelsgård (2022) Compiling a Suitable Level of Sense Granularity in a Lexicon for AI Purposes: The Open Source COR Lexicon *Proceed. of LREC2022*.
- Pedersen, Bolette Sandford; Wedekind, J.; Bøhm-Andersen, S.; Hoffensets-Andersen, S.; Juel Henriksen, P.; Kirchmeier-Andersen, S.; Kjærum, J. O.; Larsen, L. B., Nimb, S., Rasmussen, J., Revsbech, P. & Thomsen, H. (2011) Languages in the European Information Society – Danish. *META-NET White Paper Series*.

Budapest, Ungarn. (37 sider).

Jervelund, Anita, Schack, Jørgen, Jensen, Jørgen Nørby & Andersen, Margrethe Heidemann (red.) (2012) *Retskrivningsordbogen*. 4. udgave. Dansk Sprognævn.

Simonsen, Annika; Sandra Saxov Lamhauge, Iben Nyholm Debess; Peter Juel Henriksen (2022) Creating a basic language resource kit for Faroese. *Proceedings of LREC2022*.

Widmann, Thomas (2022) Det Centrale Ordregister og dets leksikografiske anvendelser. *Proceedings of NLF 2022*.

Link

Link1 (1997) Lov om Dansk Sprognævn (LOV nr. 320 af 14/05/1997).

<https://www.retsinformation.dk/eli/lta/1997/320>.

<https://kum.dk/ministeriet/organisation-og-institutioner/bestyrelser-raad-naevn-og-udvalg/dansk-sproгнаevn-repraesentantskab>.

Link2 (1997) Lov om dansk retskrivning (LOV nr 332 af 14/05/1997)

<https://www.retsinformation.dk/eli/lta/1997/332>.

Link3 (2023) Det Danske Sprog og Litteraturselskabs' præsentation af PAROLE-projektet (såvel de pan-europæiske som de specifikt danske aktiviteter): <https://korpus.dsl.dk/resources/details/parole.html>.

Appendiks

1. Tabel for bøjningsindeks og PoS (komplet tabel i ordregister.dk)

Bøjningsindeks (COR1)	Part-of-speech (Retskrivningsordbogen)
100	sb.sg.ubest
102	sb.pl.ubest
110	sb.fk.sg.ubest
111	sb.fk.sg.best
112	sb.fk.pl.ubest
114	sb.fk.sg.ubest.gen
120	sb.itk.sg.ubest
122	sb.itk.pl.ubest
201	vb.inf.pass
203	vb.præs.akt
204	vb.præs.pass
214	vb.perf.part.sg.best

300	adj.sg.ubest.fk
302	adj.sg.best.
303	adj.pl
500	prop
600	talord
880	præp
900	adv
910	art.best.sg.fk
915	art.ubest.sg.fk
970	konj

2. PAROLE med COR-links (første 22 tokens)

To/COR.01528.600.01
kendte/COR.18159.303.01
russiske/COR.22261.303.01
historikere/COR.58774.112.01
Andronik/COR.PARC.110001.500
Mirganjan/COR.PARC.110002.500
og/COR.00099.970.01
Igor/COR.PARC.110003.500
Klamkin/COR.PARC.110004.500
tror/COR.30425.203.01
ikke/COR.10161.900.01
./COR.XP.2.01
at/COR.00145.970.01
Rusland/COR.05764.500.01
kan/COR.30580.203.01
udvikles/COR.30794.201.01



En sprogholdningsundersøgelse blandt fans af Rikke Thomsen

Maja Agermose Hinrichsen

Aarhus Universitet

1. Indledning

På trods af Danmarks størrelse har vi tidligere haft en rig dialektal variation landet over (Arboe 2012). Især Sønderjylland er kendt som et område, hvor dialekt traditionelt har stået stærkt, og hvor dialekt fortsat bruges aktivt af en del af befolkningen. De senere år har vi set eksempler på alt fra film og reklamer til bøger og musik på sønderjysk. Den sønderjyske sangskriver Rikke Thomsen er et eksempel herpå. I 2019 udgav hun pladen *Omve 'n hjemve*, der indeholder sange på sønderjysk. I 2022 udgav hun opfølgeren *Opland* – igen 100 % på sønderjysk. For Thomsen var det en selvfølge at synge på sønderjysk, som hun i én af sine sange kalder sin ”moues stem” (’mors stemme’, dvs. sit modersmål). Albummet fra 2019 blev anledning til, at Rikke Thomsen senere samme år modtog Modermåls-Prisen som den yngste modtager nogensinde for sin indsats for sønderjysk dialekt (jf. Hansen 2019:39 f.).

Sange på dialekt kan være med til at bekræfte et sprogligt fællesskab og til at markere en identitet. Karen Margrethe Pedersen (2010:248 f.) peger på, at sange på sønderjysk kan fungere som en identitetsmarkør, der viser et tilhørsforhold til Sønderjylland og kan bidrage til en følelse af fællesskab. Imidlertid er dialekterne i Danmark under pres. I takt med øget social og geografisk mobilitet går udviklingen mere og mere i retning af et standardsprog, og de yngre generationer taler mindre og mindre dialekt (Maegaard 2020). Tiltag som Rikke Thomsens kan være med til at skabe og fastholde positiv interesse for dialekterne, men hvordan ser billedet ud i lidt flere detaljer, når man udforsker holdninger til sprogbrugen i Rikke Thomsens sange systematisk?

I denne artikel søger jeg at tegne et mere tydeligt billede af, hvordan Rikke Thomsens sange opleves og vurderes af en respondentgruppe bestående af hendes fans. Jeg anvender nogle teknikker inspireret af sprogholdningsforskningen i dansk sociolingvistik. Artiklen tager afsæt i en spørgeskemaundersøgelse, jeg har foretaget på Rikke Thomsens fanside på Facebook (”Rikke Thomsen Support”). Spørgeskemaet omfattede en sprogholdningsundersøgelse inspireret af Tore Kristiansens (fx 2009) klassiske dialekthitliste- og masketestundersøgelserdesign

samt en række eksplicite spørgsmål til Rikke Thomsens sange og sprogbrug. De stemmeprøver, jeg har taget udgangspunkt i, er fem lydclip fra det digitale formidlingsprojekt *SomMansiger* (www.sommansiger.nu, tilgået 6. juni 2022). Disse anvender jeg i den masketestinspirerede undersøgelse. Altså undersøger jeg sprogholdninger med brug af forskellige, mere og mindre direkte teknikker. Mit oprindelige formål var at undersøge bevidste og underbevidste sprogholdninger hos respondenterne, inspireret af Tore Kristiansens undersøgelsesdesign, men da denne metode ikke tidligere er blevet brugt i en online undersøgelse, er jeg efterfølgende blevet opmærksom på nogle problematikker ved dette format i forhold til at afdække underbevidste sprogholdninger, som jeg vil uddybe nærmere i mit metodeafsnit. Artiklen er baseret på mit sprogbrygsanalyseprojekt i foråret 2022 på Aarhus Universitet.¹

Afsnit 2 præsenterer teorigrundlaget for min undersøgelse. Afsnit 3 indeholder en gennemgang af mine metodiske overvejelser i projektet. I afsnit 4 præsenterer jeg resultaterne af min undersøgelse. Afsnit 5 præsenterer nogle afsluttende bemærkninger.

2. Teorigrundlag: Danmarks dialekter og tidligere sprogholdningsundersøgelser

Der har sandsynligvis været dialekter i Danmark lige så længe, som der har været mennesker (Pedersen 2002:39 f.). I takt med at samfundet har ændret sig, og der er kommet mere mobilitet, er dialekterne dog blevet mindre fremtrædende. Gennem de sidste par århundreder er der generelt sket en ændring hen imod en langt større brug af standardsprog, eller det man også kunne kalde *rigsdansk* (Arboe 2012:18 f.). Dette begreb er dog langt fra neutralt, men kan bl.a. forstås som en modsætning til dialektalt sprog, og er forbundet med en række værdiladninger (Kristiansen 2021:185-186 f.). Inger Schoonderbeek Hansen (2022) peger desuden på, at det lader til, at forståelsen af begrebet rigsdansk er under forandring. De jyske sprogbrugere i undersøgelsen forstod overvejende rigsdansk som et ”korrekt, forståeligt og ulokaliserbart sprog (...); en art regional standardvarietet, *jysk standard*” (Hansen 2022:112 f.).

En undersøgelse af Inger Schoonderbeek Hansen og Yonatan Goldshtein (2021) tyder desuden på, at der er ved at opstå sproglige normcentre uden for København. Her blev lokale og regionale stemmer generelt vurderet højere end stemmer fra København; lignende er kun blevet observeret i Sønderjylland (jf. Maegaard & Quist 2020).

Sønderjysk er på mange måder en særegen dialekt, som for mange sønderjyder er forbundet med en identitetsfølelse. Hvis man tager Sønderjyllands historie i betragtning, er dette ikke så underligt. I lang tid var det netop ved at tale sønderjysk, at danskere i det Sønderjylland, der

¹ Tak til min underviser Kristoffer Friis Bøegh for vejledning undervejs, samt Inger Schoonderbeek Hansen. Tak til fagfællebedømmeren for gode kommentarer og til publikum ved mit oplæg på MUDS 19.

i 1864 til 1920 var under tysk styre, kunne vise deres tilhørsforhold til Danmark. Dialekten var af denne grund også skærmet fra indflydelse fra rigsdansk. En undersøgelse af unge bornholmeres, nordjyders og sønderjyders sprogbrug viser, at det kun er unge i Sønderjylland, der i dag bruger dialekt som deres umarkerede hverdagssprog (Maegaard 2020).

Siden 1980'erne har sprogforskere været interesserede i at undersøge sprogspsykologiske drivkræfter bag sprogforandring og dialektudtynding, og dette har man bl.a. gjort ved at undersøge befolkningens sprogholdninger. Især Tore Kristiansen har lavet en lang række sprogholdningsundersøgelser i Danmark. Han skelner mellem bevidste og underbevidste sprogholdninger, dvs. de sprogholdninger folk eksplicit tilkendegiver, og de sprogholdninger, der fremkommer i sammenhænge, hvor man ikke er klar over, at der spørges til holdninger til sprog (Jensen et al. 2015). Når man undersøger de bevidste sprogholdninger, benytter man direkte metoder. Her kan man fx benytte *dialekthitlisten*, hvor man beder deltagerne om at rangordne de dialekter, de bedst kan lide, næstbedst kan lide osv. (Jensen et al. 2015:58 f.). Disse undersøgelser har vist, at folk generelt placerer deres egen dialekt øverst, herefter følger sprogborgen i den nærmeste lokale storby, og på tredjepladsen placeres rigsdansk, forstået som et dansk fælles kommunikationsmiddel (Jensen et al. 2015:58-59 f.). Københavnsk placeres oftest relativt langt nede på skalaen.

Dette holdningsmønster svarer dog dårligt til den sprogforandring, der reelt finder sted i Danmark, hvor de lokale dialekter mange steder bliver erstattet af træk fra københavnsk. Dette altså på trods af, at mange har negative bevidste holdninger over for københavnsk sprogbrug (Jensen et al. 2015:59 f.). Det er derfor også nødvendigt at undersøge de underbevidste sprogholdninger. Disse undersøges vha. en *masketest* (Jensen et al. 2015). Inden respondenterne laver dialekthitlisten, bliver de bedt om at lytte til en række lydindspillede stemmeprøver og vurdere dem i forhold til otte forskellige personlighedstræk. Masketesten foretages inden dialekthitlisten for at sørge for, at respondenterne ikke kender til undersøgelsens egentlige formål. De otte personlighedstræk repræsenterer to forskellige dimensioner: superioritet og dynamik. Superioritet vurderes på skalaerne klog-dum, seriøs-ligeglad, målrettet-sløv og til at stole på-ikke til at stole på. Dynamik vurderes på skalaerne selvsikker-usikker, spændende-kedelig, tjekket-utjekket og flink-usympatisk. De forskellige stemmer på optagelserne repræsenterer hver især tre forskellige måder at tale på, som bliver kaldt *aksanger*: KONSERVATIV, som svarer til betegnelsen rigsdansk, MODERNE, som svarer til københavnsk og LOKAL, som svarer til de lokale dialektnavne, der også blev brugt i dialekthitlisten. LOKAL er typisk ikke traditionel dialekt, men derimod en regionalstandard, fx standardsprog med en lokal farvning. Generelt viser resultater af sådanne undersøgelser, at den lokalt farvede sprogbrug nedvurderes i sam-

menligning med den konservative og den moderne aksang. Dette er interessant, da det står i modsætning til de bevidst tilkendegivne sprogholdninger, hvor det modsatte var tilfældet (Jensen et al. 2015:60 ff.). Der er altså en tendens til at nedvurdere de regionale varianter til fordel for københavnsk og rigsdansk, hvilket også svarer bedre til det generelle billede af sprogforandring i Danmark, hvor rigsdansk og københavnsk mange steder erstatter dialekterne.

Marie Maegaard lavede i 2002 samme type undersøgelse blandt en række sønderjyder for at se, om resultaterne ser anderledes ud her. Her viste sig dog de samme resultater: Når deltagerne skulle vurdere dialekterne bevidst, placerede de sønderjysk højest, men når det kom til de underbevidste sprogholdninger, blev den sønderjyske dialekt vurderet mere negativt end københavnsk og rigsdansk (Maegaard 2002:165 f.). En ny undersøgelse af Marie Maegaard og Pia Quist fra 2020 af sprogholdninger blandt unge i den sønderjyske by Bylderup viser dog en opvurdering af sønderjysk på det underbevidste plan, hvilket ikke tidligere er blevet observeret (Maegaard & Quist 2019:158 f.). Malene Monka, Maria Maegaard og Nicolai Pharaos har dog siden foretaget en lignende undersøgelse i Aabenraa, som blev præsenteret på MUDS 19. Denne undersøgelse kunne vise, at samme opvurdering ikke fandt sted i Aabenraa, hvilket tyder på, det har noget med den specifikke egn Bylderup at gøre (Monka et al. 2023). Generelt indikerer resultaterne dog, at der er noget andet på spil i Sønderjylland.

3. Metode

Jeg ønskede at undersøge sprogholdningerne hos Rikke Thomsens fans for at se, hvilket mønster der tegner sig, når man undersøgte respondenternes holdninger med afsæt i mere og mindre direkte analyseteknikker. For at undersøge sprogholdningerne udarbejdede jeg et spørgeskema, som jeg offentliggjorde på facebookside "Rikke Thomsen Support". Spørgeskemaet blev præsenteret på en sådan måde, at de potentielle respondenter ikke på forhånd skulle tænke, at spørgeskemaets hovedfokus ville ligge specifikt på holdninger til Rikke Thomsens sprogbrug og sønderjysk, men handlede om hendes virke som sanger og musiker samt forskellige stemmer mere generelt. Selve spørgeskemaet havde titlen "Rikke Thomsen" og på opslaget skrev jeg, at det bl.a. handlede om "hvad I synes om Rikke Thomsens sange" (figur 1). Alligevel kan det ikke udelukkes, at respondenterne har haft et fokus på sprog i baghovedet ved besvarelsen af spørgeskemaet, da konteksten (en sønderjysk sangerindes fanside) muligvis kan have ført deres tanker i dén retning. Da der var tale om en online undersøgelse, var det heller ikke muligt at være i dialog med respondenterne omkring deres tanker om undersøgelsen undervejs eller efterfølgende.

Hej Rikke Thomsen Support!

I forbindelse med et eksamensprojekt på Aarhus Universitet har jeg fået lov at dele et spørgeskema med jer her på siden. Det handler bl.a. om, hvad I synes om Rikke Thomsens sange 🎵 Det ville gøre mig meget glad, hvis I havde lyst til at besvare det - det tager kun ca 4 min og ville være en kæmpe hjælp 😊

SURVEY.AU.DK
survey.au.dk



Figur 1: Opfordring til at deltage i spørgeskema på Facebook.

Efter respondenterne havde svaret på en række baggrundsspørgsmål om sig selv, foretog jeg en undersøgelse, der var inspireret af en klassisk masketest, jf. Kristiansens undersøgelsesdesign (se fx Kristiansen 2009). Jeg tog udgangspunkt i fem lydklip fra sommansiger.nu. Grundet spørgeskemaets format måtte jeg lave visse praktisk motiverede afgrænsninger i mit undersøgelsesdesign. Klippene var af relativ kort varighed (ca. 15 sekunder), samt vurderinger af disse ud fra fire skalaer i stedet for alle otte, Kristiansen anvender. De fire skalaer var fordelt, så de repræsenterede både Kristiansens superioritets- og dynamikdimension. På superioritetsdimensionen udvalgte jeg skalaerne klog-dum og seriøs-ligeglad, mens jeg på dynamikdimensionen valgte selvsikker-usikker og flink-usympatisk. I min udvælgelse af de fem lydklip, der repræsenterer regionale varianter af henholdsvis nordjysk, sydsjællandsk, sønderjysk, aarhusiansk og københavnsk (se figur 6), forsøgte jeg at kontrollere for sociolingvistiske variabler ved at vælge personer, der var så ens som muligt på faktorer som alder, køn og uddannelse. Personerne på de udvalgte lydklip er derfor alle kvinder mellem 35-51 år med en lang videregående uddannelse. Jeg tog ikke højde for om de talte om samme emne; dette kan have haft en effekt på respondenternes besvarelser (jf. fx Hansen 2022).

Efterfølgende bad jeg respondenterne om at placere forskellige sproglige varieteter på en dialektthitliste (jf. Jensen et al. 2015:58 f.). Til slut stillede jeg en række eksplicitte spørgsmål til Rikke Thomsens sange samt dét faktum, at hun synger på dialekt. Rækkefølgen gik altså fra anvendelse af en mindre direkte/indirekte metode til en direkte metode til at undersøge sprogholdninger, og til sidst helt eksplicitte spørgsmål om sprogbrug. Her er en ulempe ved det online format, at deltagerne i princippet kunne gå tilbage i spørgeskemaet og ændre deres svar. I det fysiske format kan man kontrollere dette i højere grad ved at indsamle de fysiske skemaer efter hver test og uddele nye, så respondenterne ikke ændrer deres svar.

Jeg er opmærksom på, at min fremgangsmåde har nogle begrænsninger ud over de allerede nævnte forbehold, bl.a. at alle skalaer i den klassiske masketestundersøgelse ikke blev bragt i spil, og at de forskellige regionalsprog blot var repræsenteret af en enkel taler hver. Derudover

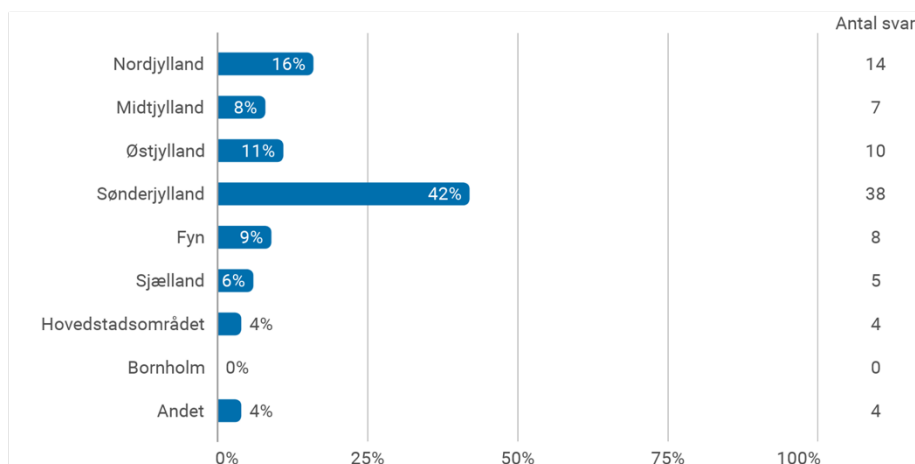
kommer respondenterne fra forskellige områder og har forskellige aldre, hvilket ikke er tilfældet i klassiske sprogholdningsundersøgelser, hvor man oftest undersøger unge fra samme geografiske område. Det medfører en række forbehold, da det kan være problematisk at generalisere på tværs af aldersgrupper og geografisk placering. I modsætning til tidligere undersøgelser er konteksten i nærværende undersøgelse desuden ikke neutral, da der netop er tale om en Facebookgruppe med et bestemt interesseområde. Derfor er det vanskeligt at konkludere, at der er tale om underbevidste sprogholdninger. Jeg har dog gjort, hvad jeg kunne for at kontrollere, at respondenterne ikke var opmærksomme på konteksten, jf. fx den neutrale deling af spørgeskemaet på gruppen. Endelig skal jeg gøre opmærksom på, at undersøgelsen må ses som et pilotstudie, altså bør resultaterne vurderes ud fra det mindre antal respondenter.

4. Resultater

Jeg fik i alt 92 besvarelser på spørgeskemaet. 41 af disse var kun delvise besvarelser, mens 51 var komplette, dvs. hele spørgeskemaet blev gennemført. Jeg tager udgangspunkt i de komplette besvarelser. Der var tale om en jævn kønsfordeling, med 49 % kvindelige respondenter og 51 % mandlige. Respondenterne havde alderen 21-77 år, så der er tale om et bredt aldersspænd. Gennemsnitsalderen var 66 år, så der er altså tale om en ældre informantgruppe, end man almindeligvis har i danske sprogholdningsundersøgelser.

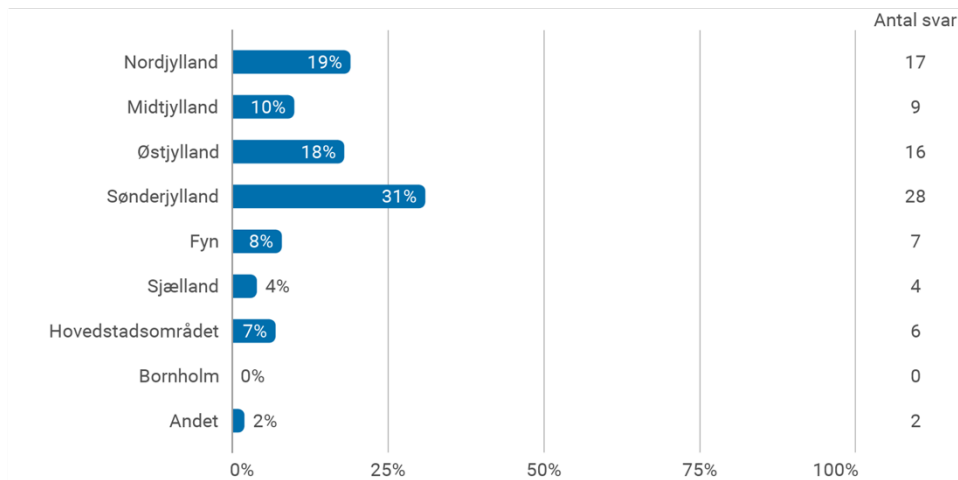
Geografisk kom størstedelen af respondenterne fra Sønderjylland (42 %), herefter var næstflest fra Nordjylland (16 %). Derudover var den geografiske distribution relativt spredt, men der var dog ingen respondenter fra Bornholm (jf. figur 2). 31 % af respondenterne bor i Sønderjylland i dag, mens 19 % bor i Nordjylland (jf. figur 3). Man kan altså ikke sige, at respondenterne er fra ét bestemt sted i Danmark.

Hvor er du oprindeligt fra?



Figur 2: Geografi.

Hvor bor du nu?

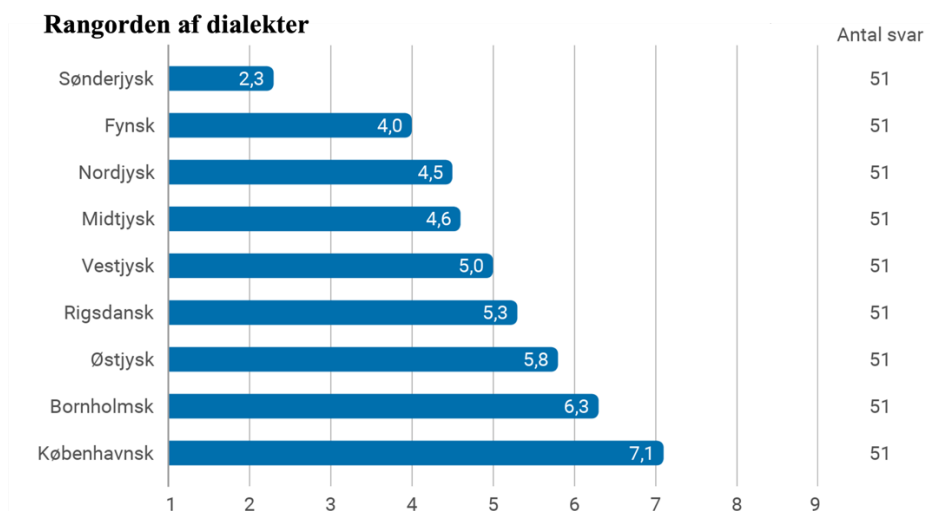


Figur 3: Bopæl.

I det følgende præsenterer jeg først resultaterne af dialekthitlisterne, dernæst den maske-testinspirerede undersøgelse og til sidst svarene på de eksplicite spørgsmål om Rikke Thomsens musik. Altså er præsentationen af resultaterne anderledes end rækkefølgen i mit spørgeskema (jf. beskrivelsen i afsnit 3).

4.1 Dialekthitlisteundersøgelsen

I dialekthitlisteundersøgelsen blev respondenterne bedt om at rangliste en række dialekter, efter hvilken de bedst kunne lide. Et var den bedste vurdering og ni den værste (se figur 4). Ikke overraskende fik sønderjysk den bedste vurdering med et gennemsnit på 2,3, mens københavnsk fik den laveste vurdering på gennemsnitligt 7,2 (jf. figur 4). Det svarer meget godt til de tendenser, der har vist sig i tidligere sprogholdningsundersøgelser, hvor respondenterne typisk opvurderer egne lokale dialekter og nedvurderer københavnsk. Derudover er det ikke overraskende, at sønderjysk bliver vurderet relativt højt hos alle, da deltagerne i lige præcis denne kontekst – med fokus på Rikke Thomsen – sandsynligvis i særlig grad vil have sønderjysk i baghovedet.



Figur 4: Dialekthitliste – gennemsnit

Ift. afvigelser fra tidligere undersøgelser ligger rigsdansk relativt langt nede på listen med en gennemsnitlig vurdering på 5,3 (jf. figur 4). Typisk bliver rigsdansk placeret omkring toppen, ca. tredje bedst (Gregersen et al. 2015), men her er det nede som nummer seks. Igen skyldes det måske konteksten. Da det ikke er blevet konkretiseret præcis, hvad der forstås ved rigsdansk som koncept, er det sandsynligt, at deltagerne i denne kontekst primært forstår det som noget, der står i modsætning til den dialekt, som Rikke Thomsen synger på. Det kan forklare, at rigsdansk her vurderes lidt lavere, end når det typisk forstås som et fællesdansk standardsprog. En alternativ forklaring er, at deltagerne bare bedre kan lide dialekt end rigsdansk.

Når man kigger på ranglisten fordelt på det område, personerne selv oprindeligt kommer fra (jf. figur 5), kan det bemærkes, at hverken østjyderne, københavnerne eller sjællænderne vurderer deres egen dialekt højest. Sønderjysk bliver derimod igen vurderet højt hos alle, hvilket er forventeligt, da det netop er Rikke Thomsens fans, der undersøges. Midtjysk bliver også vurderet relativt højt, og dette skyldes måske, at denne ofte ses som en relativ neutral dialekt (jf. fx Svenstrup & Thøgersen 2009). Sjællænderne placerer rigsdansk øverst på dialekthitlisten, skarpt efterfulgt af sønderjysk. Også respondenterne fra hovedstadsområdet og Nordjylland har sønderjysk som nr. 2 på listen. I Østjylland og i selve Sønderjylland placeres sønderjysk øverst på dialekthitlisten.

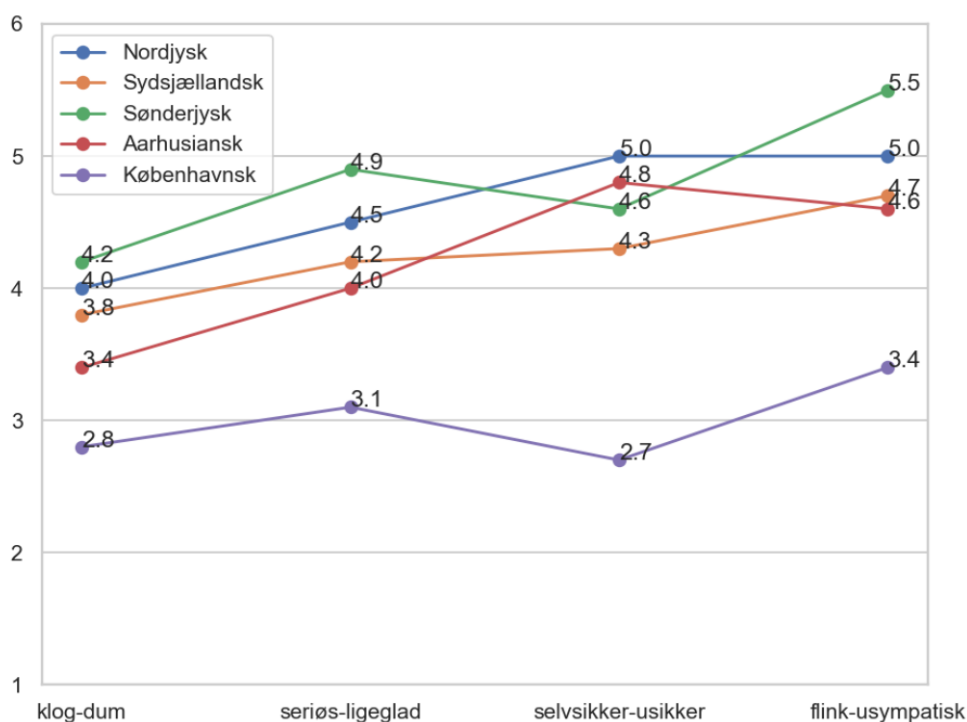
Hovedstadsområdet	Sjælland	Sønderjylland	Østjylland	Nordjylland
Midtjysk 2	Rigsdansk 3	Sønderjysk 1,6	Sønderjysk 2	Nordjysk 2,63
Sønderjysk 2,5	Sønderjysk 3,33	Fynsk 4,1	Fynsk 2,75	Sønderjysk 3,25
Østjysk 3	Midtjysk 4	Midtjysk 4,5	Nordjysk 3,25	Fynsk 4,25
Vestjysk 4,5	Vestjysk 4,67	Vestjysk 4,9	Midtjysk 4,75	Vestjysk 5
Fynsk 4,5	Fynsk 5,33	Nordjysk 5,2	Vestjysk 5,5	Rigsdansk 5,25
Bornholmsk 5,5	Nordjysk 5,67	Rigsdansk 5,6	Østjysk 5,75	Bornholmsk 5,25
Nordjysk 6	Københavnsk 5,67	Østjysk 5,6	Københavnsk 5,75	Midtjysk 5,88
Rigsdansk 8,5	Østjysk 6	Bornholmsk 6,4	Rigsdansk 7	Østjysk 5,88
Københavnsk 8,5	Bornholmsk 7,33	Københavnsk 7,1	Bornholmsk 8,25	Københavnsk 7,63

Figur 5: Rangorden fordelt på område, personerne oprindeligt er fra. Sønderjysk er markeret med fed skrift, rigsdansk med blå og københavnsk med rød.

Når respondenterne blev bedt direkte om at placere forskellige dialekter i forhold til hinanden, ser vi altså, at sønderjysk – som Rikke Thomsen bruger i sine sange – vurderes positivt. Resultatet af dialekthitlisteøvelsen må altså tolkes, som om respondenterne generelt har positive holdninger til sønderjysk. Opvurderingen af sønderjysk er desuden ikke begrænset til de sønderjyske respondenter, men gælder også de af Rikke Thomsens fans, der selv kommer fra og er bosat i andre dele af Danmark.

4.2 Den masketestinspirerede undersøgelse

Resultaterne fra undersøgelsen, der tog afsæt i det klassiske masketestdesign og altså er den mindre direkte metode til at undersøge sprogholdninger ud fra en række vurderingsskalaer, kan ses i figur 6. Her kan der argumenteres for, at disse sprogholdninger, selv hvis de ikke er direkte underbevidste, dog kan give udtryk for *mindre bevidste* sprogholdninger, eller mindre direkte eliciterede, end det er tilfældet ved dialekthitlisten. På x-aksen ses de fire vurderingsskalaer, mens y-aksen viser gennemsnitsværdierne på en 6-trinsskala, hvor en høj værdi svarer til en positiv vurdering. Her kan vi igen som i forrige afsnit se, at den københavnske taler bliver vurderet lavere end de andre talere, hvorimod den sønderjyske taler bliver vurderet højt. På skalaerne klog-dum, seriøs-ligeglad og flink-usympatisk ligger sønderjysk højest: Man kan altså ifølge respondenterne både være mest klog, seriøs og flink, hvis man taler sønderjysk.



X-aksen = de fire vurderingsskalaer

Y-aksen = gennemsnitsværdier på 6-trinsskala (høj værdi = positiv vurdering)

Figur 6: Resultaterne fra den masketestinspirerede undersøgelse.

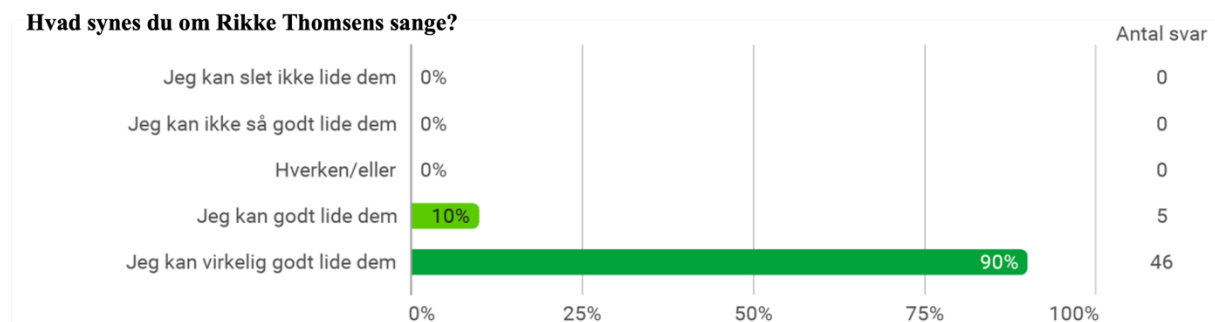
Det er i denne test, at mine resultater umiddelbart adskiller sig mest fra tidligere sprog-holdningsundersøgelser, og man kan overveje, hvad årsagen til forskellen er. Ofte peger sprog-forskere på medierne som primær årsag til opvurdering af københavnsk, fordi det er den varie-tet, man bliver mest eksponeret for gennem medierne (Jensen et al. 2015). Netop denne under-søgelses målgruppe, givet deres engagement i Rikke Thomsens sønderjyske musik, er i højere grad eksponeret for sønderjysk, og det er muligvis årsagen til, at de opvurderer sønderjysk ift. københavnsk. De ser også mere positivt på regionale varieteter; fx er forskellen på vurderingen af sønderjysk og nordjysk osv. ikke stor (og nordjysk og aarhusiansk vurderes desuden hhv. højest og næsthøjest på skalaen selvsikker-usikker). Resultaterne tegner her et billede af en respondentgruppe, der på tværs af skalaer vurderer regionale varieteter af dansk positivt, og især sønderjysk dialekt (jf. også Pedersen 2010).

Disse resultater ville kunne tolkes som et udtryk for respondenternes underbevidste hold-ninger til de forskellige dialekter, og ved første øjekast var jeg tilbøjelig til at sige, at vi her så en effekt af en befolkningsgruppe, som er mere eksponeret for sønderjysk og også mere positiv

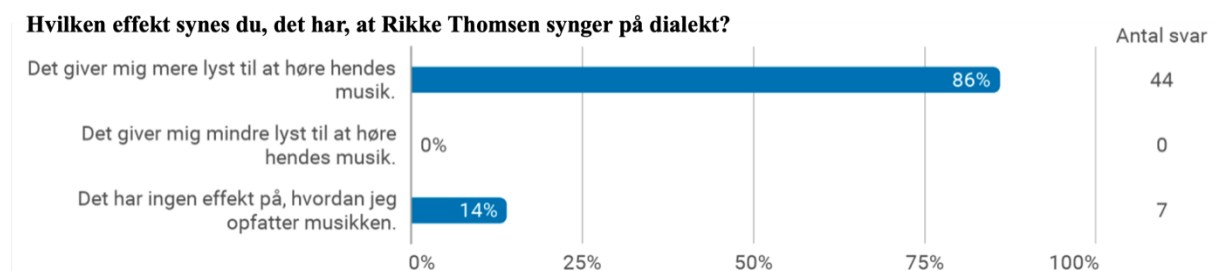
over for denne. Imidlertid er der dog, som også tidligere nævnt, ikke tilstrækkeligt grundlag for at fastslå, at der faktisk er tale om underbevidste holdninger. Det er også værd at påpege, som allerede nævnt, at denne undersøgelsesgruppe er en del ældre, end dem man typisk har undersøgt i sprogholdningsundersøgelser. Dette kan også påvirke resultaterne, hvorfor denne undersøgelse ikke kan sammenlignes direkte med tidligere undersøgelser. Men ligesom ved resultaterne af opgaven med dialekthitlisten kan vi forventeligt fastslå, at respondenterne generelt opvurderer sønderjysk, også ved benyttelse af mindre direkte metoder til undersøgelse af deres sprogholdninger.

4.3 Eksplicit tilkendegivne holdninger til Rikke Thomsens sange og sprogbrug

Til sidst i spørgeskemaet spurgte jeg direkte til respondenternes holdninger til Rikke Thomsens sange og sprogbrug. Disse resultater viste, meget forventeligt, at respondenterne generelt var meget positive over for Thomsens sange og det faktum, at hun synger på dialekt. 90 % svarede, at de virkelig godt kunne lide sangene, mens 10 % svarede, at de godt kunne lide dem (jf. figur 7). 86 % svarede desuden, at det gav dem mere lyst til at høre Rikke Thomsens musik, at hun sang på dialekt. Ingen svarede, at det, at sangene var på dialekt, gav dem mindre lyst til at høre hendes musik (jf. figur 8). 80 % svarede direkte, at de ”i meget høj grad” kunne lide, at Rikke Thomsen synger på dialekt, mens 20 % svarede ”i høj grad”. Igen svarede ingen, at det ikke kunne lide det (jf. figur 9).



Figur 7: Eksplicitte holdninger til Rikke Thomsens sange.

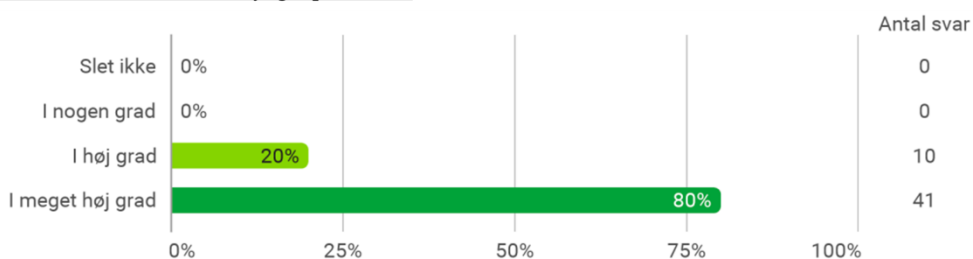


Figur 8: Eksplicitte holdninger til brugen af dialekt.

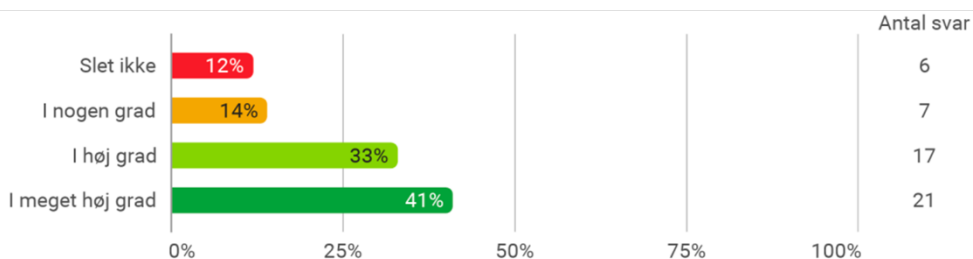
Til slut skulle deltagerne svare på, i hvilken grad de følte, de kunne relatere mere til teksterne, end når der synges på rigsdansk. Her svarede 21% ”i meget høj grad”, mens 33 % svarede ”i høj grad” (jf. figur 9). Disse resultater tyder på, at der er en forbindelse mellem dialekt og identitet, siden en stor del af respondenterne selvrapporterer at kunne relatere mere til teksterne pga. dialekten. Dette fandt Pia Quist (2017) også i sin undersøgelse, der bl.a. peger på, at sprogbrugerens tilknytning til et bestemt ”sted” bør undersøges som ”sociale handlinger der har med social orientering og identitet at gøre” (Quist 2017:11 f.). Dette passer også godt sammen med resultaterne i den tidligere nævnte undersøgelse af Pedersen (2010).

Hvor enig er du i følgende udsagn?

Jeg kan godt lide, at Rikke Thomsen synger på dialekt



Jeg føler jeg kan relatere mere til teksterne, end når der synges på rigsdansk



Figur 9: Spørgsmål om dialekt og identitet.

Denne sociale orientering og forbindelse mellem dialekt og identitet bliver understreget af de kvalitative data til sidst i spørgeskemaet, hvor jeg bad deltagerne ”skrive lidt om, hvorfor du lytter til Rikke Thomsens musik”. Her er der bl.a. én, der skriver: ”Hun synger godt og jeg elsker den sønderjyske dialekt, da det får mig til at føle mig hjemme.” En anden respondent skriver: ”Jeg kommer til at tænke på min egen barndom og ungdom”. Begge disse indikerer, at dialekten minder dem om deres tilhørsforhold til hjemstavnen. Nogle peger endda (ligesom Rikke Thomsen selv) på sønderjysk som værende deres modersmål. En respondent skriver: ”Sønderjysk (Alsisk) er mit modersmål, rigsdansk er mit første fremmedsprog.” Andre respondenter peger også på en form for modstilling mellem dialekt og rigsdansk. Der er en, der skriver: ”Fordi der er nuancer i dialekten som ikke kommer frem på rigsdansk”. En anden skriver: ”Hun griber noget omkring det at være fra provinsen i modsætning til det tit opbasunerede køben-

hav[n]eri.” Svarene viser, hvordan der for nogle af respondenterne knytter sig en identitetsfølelse og en positiv holdning til brugen af dialekt i Rikke Thomsens sange. Dette er blot et udvalg af respondenternes besvarelser, men vurderes som repræsentative for materialet som helhed.

5. Afslutning

I denne artikel har jeg søgt at tegne et mere tydeligt billede af, hvordan Rikke Thomsens sange opleves og vurderes af en respondentgruppe bestående af hendes fans ved at undersøge deres sprogholdninger. Resultaterne viser, at respondenterne vurderer sønderjysk højt og københavnsk lavt, både når det gælder de bevidste og de mindre bevidste sprogholdninger. Derudover peger mine resultater på, hvorfor brugen af dialekt er vigtig for repræsentativitet i kunst og kultur, da dét, at Rikke Thomsen synger på sønderjysk, gør, at mange af respondenterne føler, de kan spejle sig i teksterne.

Videre forskning kunne foretage en lignende undersøgelse med en større gruppe respondenter end min forholdsvis lille informantgruppe. Derudover ville det, som jeg allerede har været inde på, være interessant at undersøge og evt. sammenligne med en yngre målgruppe. Til slut ville det være interessant at undersøge, om lignende tendenser gør sig gældende ved en anden dialekt end sønderjysk, fx kunne man undersøge sprogholdninger hos fans af sangere som Allan Olsen eller Niels Hausgaard, der synger på nordjysk regionalsprog, eller hos helt andre.

Derudover kan denne pilotundersøgelse åbne op for en metodisk diskussion om muligheden for at foretage sprogholdningsundersøgelser i et online format. Det online format medfører en række problematikker, bl.a. diskussionen om, hvorvidt de sprogholdninger, jeg oprindeligt ønskede at undersøge, virkelig kan siges at være et udtryk for faktisk underbevidste sprogholdninger, eller om konteksten medførte, at holdningerne var mindre underbevidste end først antaget. Jeg havde dog også, som artiklen har vist, en række positive erfaringer ved det online format, bl.a. at jeg fik besvarelser med en jævn kønsfordeling, samt havde mulighed for at undersøge en gruppe mennesker indenfor et bestemt interesseområde uden at skulle have dem samlet på en enkel fysisk lokation.

Litteratur

- Arboe, Torben (2012) Dialekter og standardsprog – synspunkter og undersøgelser siden 1960. I *Ord & sag* 32. Peter Skautrup Centret for Jysk Dialektforskning, Aarhus Universitet:38-45.
- Gregersen, Frans & Tore Kristiansen (2015) *Hvad ved vi nu – om danske talesprog?* (1. oplag). Kbh: Sprogforandringscentret, Københavns Universitet.
- Hansen, Inger Schoonderbeek (2022) Findes der en særlig jysk sprogidentitet? I *Nydanske Sprogstudier NyS*, 61. Peter Skautrup Centret for Jysk Dialektforskning, Aarhus Universitet:73-125.
- Hansen, Inger Schoonderbeek & Yonatan Goldshtein (2021) Popular geopolitics and the conceptualization of linguistic norm centres. I *Journal of Postcolonial Linguistics*, 5:204-225.
- Hansen, Inger Schoonderbeek (2019) Dialektalt skriftsprog – Rikke Thomsens Mojn Når Vi Komme. I *Ord & sag*, 39:38-45.
- Jensen, Torben Juel, Tore Kristiansen & Marie Maegaard (2015) Sprogforandring og sprogholdninger. I Frans Gregersen & Tore Kristiansen (red.). *Hvad ved vi nu – om danske talesprog?* (1. oplag.) København: Sprogforandringscentret, Københavns Universitet:49-65.
- Kristiansen, Tore (2009) Rigsdansk. I Henrik Hovmark, Iben Stampe Sletten & Asgerd Gudiksen (red.). *I mund og bog: 25 artikler om sprog tilegnet Inge Lise Pedersen på 70-årsdagen d. 5. juni 2009*. København: Museum Tusulanum:179-191.
- Kristiansen, Tore & Marie Maegaard (2021) ”Rigsdansk og dialekterne” i fremtidsperspektiv. I Tore Kristiansen & Anne Holmen (red.). *Sprogs status i Rigsfællesskabet 2031. Københavnerstudier i Tosprogethed* (Vol. C 13). København: Det Humanistiske Fakultet, Københavns Universitet:185-196.
- Maegaard, Marie (2002) Sønderjysk i dag – holdninger og forestillinger. I *Dialekter – sidste udkald?* København: Modersmål-Selskabet: C. A. Reitzels Forlag A/S:163-167.
- Maegaard, Marie (2020) Introduction - Standardization as Sociolinguistic Change. I Kristine Køhler Mortensen, Monka Malene, Maegaard Marie, & Stæhr Andreas Candefors (red.). *Standardization as Sociolinguistic Change: A Transversal Study of Three Traditional Dialect Areas*. Taylor and Francis:1-26.
- Maegaard, Marie, & Pia Quist (2020) Perception, Recognition, and Indexicality: Experimental Investigations of Variation in Northern Jutland, Southern Jutland, and on Bornholm. I Kristine Køhler Mortensen, Monka Malene, Maegaard Marie, & Stæhr Andreas Candefors (red.). *Standardization as Sociolinguistic Change: A Transversal Study of Three Traditional Dialect Areas*. Taylor and Francis:145-168.
- Monka, Malene (2020) Sprogændring gennem tre generationer – fra sønderjysk til mere rigsdansk. I Ulla Weinreich & Michael Bach Ipsen (red.). *Sprog på grænsen : sprog o æ græns = Grenzsprachen* (1. udgave. red.). København: Modersmål-Selskabet:159-169.
- Monka, Malene, Marie Maegaard & Nikolaj Pharao (2023) *Sprogholdninger i center og periferi i Nordjylland og Sønderjylland*. Endnu ikke udgivet, men resumé kan findes på: <https://projekter.au.dk/muds/muds-2022/program-muds-19/resumeer-muds-19>. MUDS19.
- Pedersen, Inge Lise (2002) Hovedtrækkene i de danske dialekter. I *Dialekter : sidste udkald?* København: Modersmål-Selskabet: C.A. Reitzels forlag:39-80.
- Pedersen, Karen Margrethe (2010) Sønderjysk sang og rap. I Widell, Peter & Hansen, Inger Schoonderbeek. *13. Møde om Udforskningen af Dansk Sprog: Aarhus Universitet 14.-15. oktober 2010*. Nordisk Institut, Aarhus Universitet:237-250.

- Quist, Pia (2017) Sprog-krop-sted: Dialektsamfundets meningskabende orden. *Nordica Helsingiensia*, 48:57-71.
- SomManSiger (2022). Lyt til Danmark. [Lydklip] Hentet den 06.06.2022 fra <https://sommansiger.nu/>.
- Svenstrup, Christoph Hare & Jacob Thøgersen (2009) On the discursive construction of “Rigsdansk”, “Jysk”, “Københavnsk” and other varie-ties among adolescents in western Jutland. I Marie Maegaard, Tore Kristiansen, Marie Maegaard, Frans Gregersen, Pia Quist, & J. Normann Jørgensen (red.). *Language attitudes, standardization and language change : perspectives on themes raised by Tore Kristiansen on the occasion of his 60th birthday*. Oslo: Novus Press:181–210.

Hvad kan nonsensord fortælle os om stødets fonologi?

Præsentation af et ph.d.-projekt

Annika Høeg

Stockholms Universitet

1. Indledning

Stødet er et prosodisk fænomen i dansk, som forekommer på visse stavelser i visse ord, fx i ordene *bil* ['bi:ʔ] og *april* [a'pʁi:ʔ]. For at et ord kan have stød, antages det, at den trykstærke stavelse skal have stødbasis, hvilket vil sige, at den enten har en lang vokal eller en kort vokal fulgt af en sonor konsonant (Grønnum 2005:62). At et ord kan have stød, betyder dog ikke nødvendigvis, at det har det. Stødet har en kompleks distribution, men der findes nogle generelle mønstre; ord med tryk på sidste stavelse har ofte stød, som *bil* ['bi:ʔ] og *april* [a'pʁi:ʔ], og ord med tryk på næstsidste stavelse har sjældent stød, som *måned* ['mɔ:nəð] og *visé* ['vi:sə] (vb.). Denne generalisering indgår i Basbølls ikke-stødprincip, som forudsiger forekomsten af stød ud fra trykkets placering og endelsers produktivitet (Basbøll 2005). Alle undtagelser til dette antages at have en leksikalsk markering for +stød, fx i *hummer* ['hɔmʔʌ], –stød, fx i *vært* ['væʔd], eller en ekstraprosodisk sonorant¹, fx i *ven* ['vɛn]. En anden generalisering er, at enstavede ord som slutter på lang vokal altid har stød, fx *bi* ['bi:ʔ], hvorimod det varierer mere ved enstavede ord med kort vokal fulgt af en sonor konsonant, fx *hal* ['halʔ] og *øl* ['øl] (Grønnum 2005:221-223).

I denne artikel vil jeg præsentere mit igangværende ph.d.-projekt, som handler om generaliseringer af stødets forekomst i dansk og deres produktivitet. I projektet vil jeg se på, hvornår talere udtaler ikkeeksisterende ord med stød, såkaldte nonsensord (eller pseudoord). Dette gøres for at kunne sige noget om, hvilke generaliseringer fra eksisterende danske ord, som talere overfører til nonsensord. Eftersom projektet er i et tidligt stadie, og der endnu ikke findes noget resultat, vil fokus for denne artikel være en overgribende præsentation af projektet samt nonsensordene og fremgangsmåden i det første eksperiment.

¹En ekstraprosodisk sonorant kan, kort forklaret, siges at ”svæve frit” i slutningen af ordet og er derfor ikke med til at give ordet stødbasis. Et eksempel på dette er sonoranten /n/ i ordet *ven*. Dette er en leksikalsk undtagelse ligesom –stød, men med en anden forklaring til undtagelsen.

2. Formål

Formålet med projektet er at undersøge, hvilke generaliseringer for stødets forekomst som talere overfører til nonsensord, for at undersøge stødets produktivitet. Dette bliver gjort ved at sammenligne generaliseringer ud fra eksisterende ord med taleres udtale af forskellige nonsensord. Disse nonsensord er mulige danske ord som ikke betyder noget. Det antages, at nonsensord kan sige noget om stødets produktivitet, eftersom de per definition ikke har nogen leksikalsk information. En forventning er, at dette bl.a. vil kunne give empirisk støtte for, hvorvidt det kun er trykkets placering der afgør om et ord har stød, eller om der er andre faktorer som kan påvirke.

Projektet kommer til at bestå af tre eksperimenter, der alle indebærer at informanternes udtale af forskellige nonsensord bliver spillet ind på en lydoptagelse. Nonsensordene i eksperimenterne bliver udformet ud fra hvilken generalisering som skal testes. I eksperiment 1 er det den første del af Basbølls ikke-stødprincip som er i fokus: ”The penultimate syllable of the min-word has non-stød” (Basbøll 2005:380).² Dvs. at trykstærke stavelser (med stødbasis), som er tredjesidst eller sidst i et ord, har stød, og trykstærke stavelser, som er næstsidst i et ord, har ikke stød. Alle undtagelser til dette er leksikalske. Eftersom nonsensord ikke har leksikalsk information, burde stødforekomsten i nonsensord følge denne generalisering uden undtagelser. I eksperiment 1 er det derfor nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse, som vil blive testet. Formålet med eksperiment 2 er at undersøge forekomsten af stød i monosyllabiske ord, ud fra hvilken sonorant der er i codaen. Derfor vil det være monosyllabiske nonsensord med et rim³ bestående af en kort vokal fulgt af en sonor konsonant, som bliver testet. I eksperiment 3 er fokus på forekomsten af stød ud fra morfologiske alternationer. Her vil både bøjnings- og afledningsmorfemer blive kombineret med forskellige nonsensord for at se nærmere på, hvordan det påvirker forekomsten eller fraværet af stød. En del af eksperiment 3 vil også være at teste den anden del af Basbølls ikke-stødprincip: ”A monosyllabic stem has non-stød before a syllable (domain: Basic Word)” (Basbøll 2005:380).⁴ I denne del af ikke-stødprincippet afhænger stødets forekomst af antallet stavelser i stammen, og hvorvidt endelsen inkluderes i et vist domæne i ordstrukturen. Også her er der leksikalske undtagelser til reglen (se Grønnum 2005:228–236 for en gennemgang).

² Lidt forsimplet kan man sige, at ”min-word” er monomorfematiske ord.

³ Rim er stavelers kerne og coda, dvs. vokalen og efterfølgende konsonanter.

⁴ Se Basbøll (2005:377–379) for en gennemgang af domænerne *min-word* og *Basic Word*.

3. Nonsensord som metode

I dette afsnit vil jeg gå nærmere ind på, hvad der menes med nonsensord, og hvordan de kan anvendes til at teste forskellige fonologiske generaliseringer.

Kort sagt kan man sige, at nonsensord er opdigtede ord, som til formen (mere eller mindre) ligner rigtige ord, men ikke har nogen betydning eller funktion i det aktuelle sprog.

Nonsensord kan udformes på forskellige måder og med forskellige grader af nonsens, alt efter hvad det er, man vil teste. Det kan fx være nonsenssammensætninger, hvor eksisterende ord kombineres til en ny sammensætning, som fx *taxigris* i Bruces (1974) undersøgelse af toneaccent i sammensatte ord i sydsvensk. Det kan også være sammensætninger, hvor et eksisterende ord kombineres med et nonsensord, som fx *korstamse* i Grønnum m.fl.s (2020) undersøgelse af tilkomst af stød i sammensætninger. Et andet eksempel er nonsensord, eller nonsensrødder, som testes i en eksisterende morfologisk kontekst, hvilket gøres i Bennett og Bravers (2020) undersøgelse af palatalisering i Xhosa med nonsensrødder som *-fam-* i konteksten *iya-fam-ba* og *iya-fam-wa*. Endvidere kan nonsensord også bestå af fuldstændigt nonsens, forstået som at der ikke indgår nogle eksisterende morfemer i ordet. Et eksempel på fuldstændigt nonsens er *galeaft* fra Turcsan og Herments (2015) undersøgelse af tildeling af tryk i engelsk. I mit projekt består nonsensordene i det første eksperiment, som er i fokus i denne artikel, også af fuldstændigt nonsens.

Selv når nonsensord udelukkende består af nonsenselementer, er de for det meste mulige ord i det aktuelle sprog, hvilket vil sige, at de følger sprogets fonotaks. I det følgende, og resten af artiklen, vil udgangspunktet være danske nonsensord. Selvom nonsensord skal være mulige eller "tilladte" ord i sproget, kan de have en struktur, som påmindrer om fremmedord.⁵ Fx vil polysyllabiske nonsensord med mere end én fuldvokal, som *gnimpov* eller *suvulga*, virke mindre hjemlige, eller måske til og med mindre danske, end monosyllabiske nonsensord og polysyllabiske nonsensord med schwa, som *fril* eller *slærbe*. Det kan dog næppe påstås, at to (eller flere) fuldvokaler kun forekommer i usædvanlige eller eksotiske ord i dansk, jf. fx *tomat*, *ananas*, *sofa* og *cola*. Desuden kan stødet sagtens forekomme i låneord, som ikke ligner danske ord, fx *gymnasium* [gym'næ:ʔɔm]. At nonsensordene har en struktur, som ikke forekommer i danske arveord, er ikke nødvendigvis et problem, så længe informanterne udtaler ordene som om det er dansk. En anden grund til at have "mindre danske" nonsensord med i undersøgelsen er, at ord med tryk på før-næstsidste stavelse, som *kamera*, er mindre sædvanlige i dansk, og denne gruppe er ganske lille blandt eksisterende ord. Derfor ser ingen af nonsensordene med

⁵ Alt efter hvad det er, man vil undersøge. Nonsensord kan fx også bruges til at undersøge, hvor grænsen går for, hvad talere af et vist sprog bedømmer som acceptabelt.

tryk på den før-næstsidste stavelse særligt danske ud, men de er nødvendige for at kunne sige noget om, hvorvidt nonsensord med tryk på før-næstsidste stavelse altid udtales med stød, som det antages i Basbølls ikke-stødprincip.

Pointen med nonsensord er, at de er nye for taleren, og dermed ikke er memorerede med nogen leksikalsk information. De kan derfor bruges til at sige noget om sprogsystemet, i det her tilfælde fonologien, fx gennem at undersøge, hvordan informanter reagerer på, opfatter eller vurderer forskellige nonsensord, som de præsenteres for auditivt eller visuelt. I mit projekt anvendes nonsensordene til at undersøge produktion (udtale), hvor informanter bliver præsenteret for nonsensord i skrift og læser dem højt.

I det informanterne bliver præsenteret for nonsensordene i skrift, giver ortografien visse begrænsninger. Eksempelvis markeres vokallængde eller trykstærke stavelser generelt ikke i dansk ortografi. Hvis formålet fx er at teste nonsensord, hvor rimet består af en kort vokal fulgt af en sonor konsonant, bliver det et problem, hvis informanterne i stedet udtaler nonsensordene med en lang vokal. En anden udfordring er, hvordan man får informanterne til at lægge trykket på en specifik stavelse i polysyllabiske ord, men samtidigt undgår at de udtaler stavelsen med emfase. En mulighed er at skrive den trykstærke stavelse med versaler, som det gøres i en studie med nonsensord af Braver (2014). Jeg vil komme nærmere ind på hvordan det bliver gjort i det aktuelle projekt i næste afsnit.

Nonsensord er altså potentielle ord, som ikke har nogen betydning, og kan bruges til at teste forskellige generaliseringer, når der ikke er nogen leksikalsk information. I de følgende afsnit vil jeg præsentere projektets første eksperiment, med fokus på forsøgsdesignet.

4. Eksperiment 1 – trykkets placering

I dette eksperiment er formålet at undersøge, hvilken udstrækning taleres udtale af forskellige nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse stemmer overens med Basbølls ikke-stødprincip. For at teste dette har jeg genereret tre kategorier af nonsensord, ud fra hvilken stavelse der er trykstærk. I tabel 1 kan der ses eksempler på forskellige nonsensord, som testes i eksperimentet. Fremgangsmåden for genereringen og udvalget af nonsensordene vil blive uddybet nærmere i det næste afsnit.

Nonsensord med tryk på sidste stavelse	Nonsensord med tryk på næstsidste stavelse	Nonsensord med tryk på før-næstsidste stavelse
SPLU	PRULbe	NILdefe
PRILP	BÅpa	VÅNlupe
fluNIL	saKRÆLje	HÆJgopi
stjeKRÆL	fyKVILru	FImefa

Tabel 1. Eksempler på nonsensord som testes i eksperiment 1.

I eksperimentet læser informanterne i alt 336 nonsensord højt. Nonsensordene er fordelt på fire blokke med 84 ord i hver. Rækkefølgen på blokkene varierer fra gang til gang, og rækkefølgen på nonsensordene i hver blok er randomiseret. Inden oplæsningen får informanterne bl.a. følgende instruktion: ”Det her går som sagt ud på, at du skal læse forskellige ord højt på dansk. Det er ord som du ikke har set før. Du kan forestille dig, at du spiller et nyt spil, og at det her er ting som du kan købe i den verden.” Dette gøres for at få informanterne til at tænke på ordene som substantiver, men i øvrigt er der ikke nogen information, som peger mod at ordene har en bestemt ordklasse. Informanterne ser ordene et efter et i midten af en grå skærm og trykker videre, når de har læst ordet højt.



Figur 1. Eksempel fra dataindsamlingen til eksperiment 1.

I Figur 1 kan der ses et eksempel på, hvordan nonsensordene vises for informanterne.⁶ Som det kan ses i figuren, er den næstsidste stavelse skrevet med versaler. Dette anvendes for at angive, hvilken stavelse som trykket skal lægges på (jf. Braver 2014). Inden oplæsningen instrueres informanterne i at lægge trykket på den stavelse, som har store bogstaver. De ser forskellige minimale par på skærmen, som fx skrives FORMel og forMEL, og derefter ord hvor trykket flyttes, som CARdigan og carDIGan, for at sikre at de har forstået, hvordan de skal gøre, og at de kan flytte trykket. Denne markering af den trykstærke stavelse behøves for at få alle informanterne til at lægge trykket på samme stavelse, og fordi det ellers er svært at få dem til

⁶ Eksperimentet er bygget i programmet PsychoPy (se Peirce m.fl. 2019).

at lægge trykket på den før-næstsidste stavelse. Jeg har udført pilotundersøgelser, hvor nonsensordene enten blev vist med versaler i den trykstærke stavelse eller uden nogen markering af tryk, for at sikre at markeringen ikke medførte, at informanterne udtalte flere ord med stød. Versalerne virkede ikke til at få informanterne i pilotundersøgelserne til at udtale flere af nonsensordene med stød.

5. Generering af nonsensord til eksperiment 1

Under arbejdet med udformningen af nonsensordene er der meget at tage hensyn til. I det følgende vil jeg redegøre for, hvilke overvejelser jeg har gjort om udformningen af nonsensordene, deres struktur, hvilke specifikke kombinationsmuligheder der kan indgå, hvordan de er blevet genereret samt hvordan udvælgelsen blandt de genererede nonsensord er gået til.

I eksperimentet er trykkets placering den uafhængige variabel, som bliver manipuleret for at undersøge den afhængige variabel, som er forekomst eller fravær af stød. Nonsensordene udformes dermed så trykkets placering varierer. Dog er trykkets placering ikke det eneste, som varierer; bl.a. vokalkvalitet og antal stavelser i nonsensordene varierer også. Dette er variabler, som ikke skal testes, men som eventuelt kan påvirke resultatet. Denne type af uvedkommende variabler (eng. *extraneous variables*) kan bl.a. håndteres med hjælp af randomisering eller afbalancering (Barr m.fl. 2013:257). De uvedkommende variabler kan balanceres på tværs af den uafhængige variabel, så at der fx er lige mange af en bestemt vokal eller sonorant i nonsensordene.

De uvedkommende variabler, som jeg stræber efter at kontrollere i udformningen af nonsensordene, er stavelsesantal, vokalkvalitet i den toniske (trykstærke) stavelse, codaens struktur i den toniske stavelse samt rimets struktur i den posttoniske stavelse for nonsensordene med tryk på næstsidste stavelse. Disse variabler kontrolleres gennem en balancering af dem i de tre kategorier af nonsensord, således at der ikke er en specifik vokal eller sonorant, som dominerer i en af kategorierne, fx at en stor andel af nonsensordene med tryk på næstsidste stavelse har en åben trykstærk stavelse (dvs. ingen coda), hvorimod en stor andel af nonsensordene med tryk på sidste stavelse har sonoranten /n/ i codaen i den trykstærke stavelse. For at mindske sandsynligheden for at en sådan forskel kan påvirke, hvor ofte nonsensordene i kategorien får eller ikke får stød, har jeg forsøgt at balancere de ovennævnte uvedkommende variabler i hver kategori af nonsensord.

For den uvedkommende variabel, *stavelsesantal*, er det hvorvidt nonsensordet har en prætonisk stavelse eller ej, som jeg vil kontrollere, snarere end selve stavelsesantallet. Dette er kun relevant for nonsensordene med tryk på sidste eller næstsidste stavelse, eftersom mono-

morfematiske nonsensord med fire stavelser bliver for unaturlige i dansk tale.⁷ I kategorien med tryk på sidste stavelse er der altså en underkategori med monosyllabiske nonsensord og en med bisyllabiske nonsensord, ligesom der i kategorien med tryk på næstsidste stavelse er en underkategori med bisyllabiske nonsensord og en med trisyllabiske nonsensord.

Den uvedkommende variabel *vokalkvalitet* kontrolleres kun i den trykstærke stavelse. For vokalkvalitet har jeg forsøgt at få så stor spredning i vokaldiagrammet som muligt, men samtidigt begrænset det til fire vokaler for at kunne balancere dem på tværs af kategorierne uden at behøve for mange nonsensord. Disse fire vokaler er <i>, <æ>, <u> og <å>. Eftersom nonsensordene præsenteres i normal skrift, begrænses kontrollen over vokalkvaliteten, og dermed balanceringen, af ortografien. Jeg er vidende om, at den ortografiske repræsentation ikke nødvendigvis vil blive udtalt med samme fonem i alle kontekster, eller endda kan variere fra gang til gang, men jeg håber, at de fire grafemer kan give en nogenlunde spredning og afbalancering af vokalkvaliteten.

Den næste uvedkommende variabel er *codaens struktur i den trykstærke stavelse*. Eftersom det er forekomsten af stød, som undersøges, skal nonsensordene kunne have stød, dvs. udtales med stødbasis. I alle tre kategorier af nonsensord er konsonanten direkte efter vokalen derfor en af sonoranterne /m/, /n/, /l/, /ð/, /ɹ/, /ʁ/ eller /ɣ/.⁸ Codaen kan også være tom, hvilket angives med Ø i genereringen af nonsensordene. Ved en tom coda, altså åbne stavelser, er forventningen, at informanterne udtaler vokalen som en lang vokal. Nonsensordene ville også have stødbasis, hvis det var en lang vokal fulgt af en obstruent, men da der er risiko for, at det bliver en kort vokal, og dermed ingen stødbasis, og for at begrænse mængden kombinationer, som skal testes, indgår obstruenterne ikke i codaen i nonsensordenes trykstærke stavelser (med mindre de foregås af en sonorant). I nonsensordene med tryk på sidste stavelse kan sonoranten enten være final eller fulgt af endnu en konsonant.⁹ I nonsensordene med tryk på næstsidste eller før-næstsidste stavelse består codaen af maksimalt én sonorant, eftersom konsonanten efter bliver ansat i den efterfølgende stavelse. Desuden vil for lange interludes¹⁰ medføre større risiko for, at nonsensordene opfattes som sammensætninger.

Den sidste uvedkommende variabel er *rimets struktur i den posttoniske stavelse* og er kun relevant for nonsensordene med tryk på næstsidste stavelse. Anledningen til dette er, at der i Grønnums inddeling (2005:222) er forskel på, hvorvidt forekomsten af stød er forudsigteligt i

⁷ Nonsensordene med tryk på før-næstsidste stavelse er allerede på (eller måske over) grænsen. Se foregående afsnit for en diskussion om dette.

⁸ Ortografisk repræsenteret som <m>, <n>, <l>, <d>, <j>, <v> og <r>.

⁹ Hvis vokalen er kort, falder disse under det Grønnum (2005:222) kalder kategori III respektive kategori IV.

¹⁰ Konsonanterne mellem to stavelser, dvs. både coda og ansats. Termen er hentet fra Basbøll (2005).

ord med tryk på næstsidste stavelse, ud fra hvorvidt den sidste stavelse indeholder en fuldvokal eller et schwa, og hvorvidt stavelsen med schwa er åben eller lukket. Derfor deles de bisyllabiske nonsensord med tryk på næstsidste stavelse ind i undergrupperne *åben eller lukket posttonisk stavelse med fuldvokal*, *åben posttonisk stavelse med schwa* og *lukket posttonisk stavelse med schwa*. De trisyllabiske nonsensord med tryk på næstsidste stavelse deles kun ind i undergrupperne *åben posttonisk stavelse med schwa* og *åben posttonisk stavelse med fuldvokal*.¹¹ At de to underkategorier ikke deles ind i lige mange undergrupper, resulterer i, at der indgår flere bisyllabiske nonsensord med tryk på næstsidste stavelse i eksperimentet end trisyllabiske.

Vokallængde er egentlig også en uvedkommende variabel, som burde balanceres, men da det er svært at styre, hvilken vokallængde informanterne udtaler vokaler med i stavelser, som har en enkelt sonorant i codaen, valgte jeg ikke at lave en opdeling baseret på dette. I stedet vil antallet lange og korte vokaler blive opregnet ved databehandlingen, for at eventuelt kunne tage hensyn til det i analysen.

Antallet nonsensord i hver af de tre kategorier, *tryk på sidste stavelse*, *tryk på næstsidste stavelse* og *tryk på før-næstsidste stavelse*, bestemmes af antallet kombinationer af de ovennævnte uvedkommende variabler. Nonsensord med tryk på sidste stavelse har to muligheder for stavelsesantal, fire muligheder for vokal og 14 muligheder for coda¹². Dette giver 112 kombinationsmuligheder. I eksperimentet testes ét nonsensord for hver kombinationsmulighed og dermed 112 nonsensord med tryk på sidste stavelse (56 monosyllabiske og 56 bisyllabiske). For de bisyllabiske nonsensord med tryk på næstsidste stavelse er der fire muligheder for vokal, otte muligheder for coda og tre muligheder for den posttoniske stavelse, hvilket giver 96 kombinationsmuligheder. For trisyllabiske nonsensord med tryk på næstsidste stavelse er der fire muligheder for vokal, otte muligheder for coda og to muligheder for den posttoniske stavelse, hvilket giver 64 kombinationsmuligheder. I eksperimentet bliver der altså testet 160 nonsensord med tryk på næstsidste stavelse. For nonsensordene med tryk på før-næstsidste stavelse er der fire muligheder for vokal og otte muligheder for coda, hvilket giver 32 kombinationsmuligheder. For at der ikke skal blive for få nonsensord i denne kategori, indgår der to nonsensord for hver kombinationsmulighed, hvilket giver 64 nonsensord.

Øvrige dele af nonsensordene, dvs. ansatsen, vokalen i den prætoniske stavelse og inter-

¹¹ Mit indtryk er, at den posttoniske stavelse i trisyllabiske ord med tryk på næstsidste stavelse altid er åben. Dette er baseret på en gennemgang af en ordliste bestående af Retskrivningsordbogen (Dansk Sprognævn 2012) kombineret med transskriptioner fra NST-leksikonet (Andersen 2011).

¹² /ð/ forekommer kun finalt i disse nonsensord, eftersom dette segment ikke virker til at forekomme ordfinalt med efterfølgende konsonanter. Dette er baseret på en gennemgang af en ordliste bestående af Retskrivningsordbogen (Dansk Sprognævn 2012) kombineret med transskriptioner fra NST-leksikonet (Andersen 2011).

ludes som ikke er efter den toniske stavelse, antages ikke at påvirke forekomsten af stød. Disse vælges derfor tilfældigt ud fra mulige kombinationer. For dette har jeg gået ud fra beskrivelser af dansk fonotaks af Schachtenhaufen (2022) og Grønnum (2007), samt Retskrivningsordbogen som elektronisk ordliste (Dansk Sprognævn 2012) kombineret med fonetiske transskriptioner fra NST-leksikonet (Andersen 2011). I tabel 2 angives grafemerne, som kan forekomme i hver del af stavelsen i de bisyllabiske nonsensord med tryk på sidste stavelse. I den første kolonne angives hvilken del af stavelsen grafemerne kan udgøre, og i den anden kolonne angives grafemerne. Nogle af grafemerne angives i parentes, hvilket betyder at disse udgør en af de ovennævnte uvedkommende variabler. Her bliver der altså valgt tilfældigt inden for hver parentes.

At det er grafemer, og ikke fonemer, som angives i tabellen skyldes at nonsensordene er genereret ud fra ortografien. Jeg har udgået fra fonemerne i begrænsningerne for fonotaksen, og valgte derefter et grafem, som kan modsvare hvert fonem. Som det blev nævnt i forbindelse med vokalerne, kan dette resultere i, at ordene udtales anderledes end jeg havde tænkt. Dog er den eksakte kvalitet på fonemerne ikke afgørende, eftersom det er trykkets placering, som er i fokus, men det kan have betydning for balanceringen og mindsker kontrollen (Jf. Schmitz m.fl. 2021:582–583).

Del af stavelsen	Grafemer som kan indgå
Ansats, prætonisk	b t d k g f s sj h m n l v j Ø st sp sk tj spl spr spj str stj skr skj sv sm sn sl pl pr pj tr kn kl kr kj bl br bj dr dj gn gl gr gj fn fl fr fj vr
Kerne, prætonisk	a e i o u y
Interlude, prætonisk– tonisk	r b j k l m n t v s f br bl d fr fl kr kn kv p pl sj sp st sk sl sv tr bs bt pt ft gs gt sm sf
Kerne, tonisk	(i) (u) (æ) (å)
Coda, tonisk	(l) (lp lt lk lf ls) (m) (mp mt ms) (n) (nt nk ns) (j) (jn jl jt js) (v) (vn vl vs) (r) (rm rn rp rt rk rs) (d) (Ø)

Tabel 2. Mulige grafemer i de forskellige dele af stavelserne i bisyllabiske nonsensord med tryk på sidste stavelse. Grafemer, som angives i parentes, udgør en gruppe. Ø betyder, at pladsen kan være tom.

I tabellen angives altså de grafemer, som bliver anvendt til at generere nonsensordene, hvilket bliver gjort med hjælp af programmeringssproget Python, version 3.10.11 (van Rossum &

Drake 2009).^{13,14} Jeg har anvendt en kode for hver kategori af nonsensord og underkategorierne ud fra antal stavelser.¹⁵ I koden defineres de grafemer, som kan stå på de forskellige plader i stavelserne, og i hvilken rækkefølge de skal kombineres. Koden for de bisyllabiske nonsensord med tryk på sidste stavelse giver 56 grupper, hvor hver gruppe er en liste med alle kombinationsmuligheder.¹⁶ Længden på hver liste varierer alt efter, hvor mange forskellige grafemer der kan være i codaen (se tabel 2). De korteste lister består af 11448 nonsensord, og de længste af 68688 nonsensord. I alt genereres 1.465344 bisyllabiske nonsensord med tryk på sidste stavelse.¹⁷ Af disse er det dog kun ét ord fra hver liste, som skal testes i eksperimentet. For at gøre udvælgelsen af disse 56 ord mere overskuelig, og for at minimere bias i udvælgelsen, anvendes *random.seed*, som er en funktion i modulet *random*, til at vælge et tilfældigt nonsensord fra hver af de 56 lister. Med *random.seed* bliver det randomiserede udvalg desuden det samme, hver gang man kører koden.

Koden for de bisyllabiske nonsensord med tryk på sidste stavelse giver altså 56 tilfældigt udvalgte nonsensord, med ét ord fra hver gruppe. En manuel gennemgang er dog nødvendig, inden de kan testes i eksperimentet. Nogle af kombinationerne er allerede eksisterende ord eller homonymer, eller stavelser, som er identiske med eksisterende affikser. Nonsensordene kan også indeholde uønskede kombinationer, som fx <vr> + <u> i den trykstærke stavelse. For at forsøge at minimere bias i valget af nonsensord har jeg opsat kriterier for, hvilke ord der ikke vælges. Fx fravælges nonsensord med samme konsonant inden og efter vokalen, som *krirk*, og nonsensord med samme ansats to gange i træk, som *flefliv*. Det er altså det første tilfældige nonsensord, som ikke fravælges på grund af kriterier som de ovenstående, som testes, også selvom et senere nonsensord måske lyder bedre. De nonsensord som denne proces resulterer i, testes derefter i en pilotundersøgelse. I tilfælde hvor informanterne havde svært ved at læse ordene højt, hvor flere informanter tøvede eller hvor de ofte læste noget andet, valgtes nye nonsensord ud fra samme proces.

¹³ Se Hammond (2020:50–52), for et eksempel på hvordan monosyllabiske nonsensord kan genereres. Fremgangsmåden for polysyllabiske nonsensord er nogenlunde den samme, bortset fra at konsonanter og vokaler for hver del af stavelsen skal defineres.

¹⁴ For generering og randomiseret udvalg i Python anvendes pakken *numpy* (Harris m.fl. 2020), og det indbyggede modul *random* (van Rossum & Drake 2009).

¹⁵ Tak til Carl Hedman for hjælp med at skrive koden. Koden for generering, sortering og tilfældigt udvalg af nonsensordene vil blive gjort tilgængelig på GitHub. Kontakt forfatteren for at få koden, inden den bliver tilgængelig.

¹⁶ Hver gruppe udgøres af en kombination af de uvedkommende variabler *vokalkvalitet* (fire muligheder) og *codaens struktur i den trykstærke stavelse* (18 muligheder). Dvs. for de bisyllabiske nonsensord med tryk på sidste stavelse består en gruppe af alle kombinationer, som slutter på <il>, en anden gruppe består af alle kombinationer, som slutter på <ilp>, <ilt>, <ilk>, <ilf> eller <ils>, en tredje gruppe består af alle kombinationer, som slutter på osv.

¹⁷ Listerne indeholder også eksisterende ord, som indgår i det som her kaldes ”genererede nonsensord”.

At vælge ord fra som er 'for mærkelige' uden at vælge med bias, er en svær balancegang. Her har jeg vægtet at undgå bias tungere. Man kan diskutere, hvorvidt det er den bedste beslutning. Et udvalg hvor nonsensordene i stedet vælges ud fra intuition, ville formodentlig give mere 'naturligt danske' nonsensord, men det ville også øge risikoen for, at der ubevidst vælges nonsensord, som udtales med eller uden stød.

6. Afslutning

I denne artikel har jeg givet en præsentation af mit ph.d.-projekt, som har til formål at undersøge stødets produktivitet med hjælp af nonsensord. Artiklen har fokuseret på projektets første eksperiment, hvor 336 nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse skal læses højt af forskellige informanter. Ved eksperimentdesign er der mange hensyn at tage, og som Arunachalam (2013:221) pointerer; "designing a carefully controlled study is not trivial." I artiklen har jeg præsenteret nogle af de beslutninger, jeg har taget under udformningen af eksperimentet, fx i forhold til hvordan nonsensordene præsenteres for informanterne, og i forhold til udformningen af nonsensordene. Under arbejdet med nonsensordene har en svær balancegang været at få nonsensord, som virker "naturligt danske", men samtidigt holde kontrollen og validiteten så høj som muligt, og desuden undgå den bias det ville give at håndplukke de bedste (mest plausibelt danske) nonsensord.

Litteratur

- Andersen, Gisle (2011) *Leksikalsk database for dansk*. Technical report, Nasjonalbiblioteket.
<https://www.nb.no/sprakbanken/en/resource-catalogue/oai-nb-no-sbr-26/>.
- Arunachalam, Sudha (2013) Experimental Methods for Linguists. *Language and Linguistics Compass*, 7(4), 221–232. <https://doi.org/10.1111/lnc3.12021>
- Barr, Dale J., Levy, Roger, Scheepers, Christoph & Tily, Harry J. (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
<https://doi.org/10.1016/j.jml.2012.11.001>
- Basbøll, Hans (2005) *The phonology of Danish*. Oxford: Oxford University Press.
- Bennett, William & Braver, Aaron (2020) Different speakers, different grammars: Productivity and representation of Xhosa labial palatalization. *Phonological Data & Analysis*, 2(6), 1–29.
<https://doi.org/10.3765/pda.v2art6.9>
- Braver, Aaron (2014) Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua* (152), 24–44.
<https://doi.org/10.1016/j.lingua.2014.09.004>
- Bruce, Gösta (1974) Tonaccentregler för sammansatta ord i några sydsvenska stadsmål. I C. Platzack (red.). *Svenskans beskrivning* 8, 62–75.

- Dansk Sprognævn (2012) *Retskrivningsordbogen* (4. udg.) <https://dsn.dk/ordboeger/retskrivningsordbogen/ro-elektronisk-og-som-bog/>
- Grønnum, Nina (2005) *Fonetik og fonologi: almen og dansk*. 3 udg. København: Akademisk forlag.
- Grønnum, Nina (2007) *Rødgrød med fløde, en lille bog om dansk fonetik*. København: Akademisk Forlag.
- Grønnum, Nina, Pharaoh, Nicolai & Basbøll, Hans (2020) Stød in unexpected morphological contexts in Standard Danish: An experimental approach to sound change in progress. *Nordic Journal of Linguistics*, 43(2), 147–180. <https://doi.org/10.1017/S0332586520000074>
- Hammond, Michael (2020) *Python for Linguists*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108642408>
- Harris, Charles R., Millman, K. Jarrod, van der Walt, Stéfan J., Gommers, Ralf, Virtanen, Pauli, Cournapeau, David, Wieser, Eric, Taylor, Julian, Berg, Sebastian, Smith, Nathaniel J., Kern, Robert, Picus, Matti, Hoyer, Stephan, van Kerkwijk, Marten H., Brett, Matthew, Haldane, Allan, Fernández del Río, Jaime, Wiebe, Mark, Peterson, Pearu, Gérard-Marchant, Pierre, Sheppard, Kevin, Reddy, Tyler, Weckesser, Warren, Abbasi, Hameer, Gohlke, Christoph, & Oliphant, Travis. E., (2020) Array programming with NumPy. *Nature*, 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Peirce, Jonathan, Gray, Jeremy R., Simpson, Sol, MacAskill, Michael, Höchenberger, Richard, Sogo, Hiroyuki, Kastman, Erik & Lindeløv, Jonas (2019) PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Schachtenhaufen, Ruben (2022) *Ny dansk fonetik*. Hellerup: BoD.
- Schmitz, Dominic, Baer-Henney, Dinah & Plag, Ingo (2021) The duration of word-final /s/ differs across morphological categories in English: evidence from pseudowords. *Phonetica*, 78(5-6), 571-616. <https://doi.org/10.1515/phon-2021-2013>
- Turcsan, Gabor & Herment, Sophie (2015) Making Sense of Nonce Word Stress in English. I Mompean, J. A., & Fouz-González, J. (red.). *Investigating english pronunciation: Trends and directions*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137509437_2
- van Rossum, Guido, & Drake, Fred L., (2009) *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Hvordan vurderes jyske participiumsformer?

Torben Juel Jensen, Marie Maegaard & Nicolai Pharao
Københavns Universitet

1. Indledning

Selvom der er der sket en voldsom standardisering i det danske sprogsamfund i det 20. århundrede, er der åbenlyst stadig geografisk betingede forskelle i sprogbrugen. En af forskellene ses ved perfektum participium-bøjningen af stærke verber, hvor den form der slutter på -en, jf. eksempel 1-2, anses for at være et af de mest resistente dialekttræk i Jylland (Jensen & Maegaard 2010; Monka 2013).

- 1) den er bleven noget større siden
- 2) så er æ tid ellers bare gåen

En tidligere undersøgelse baseret på LANCHART-korpusset har dog vist at denne form i de seneste årtier har været i kraftig tilbagegang og erstattes af den standarddanske form et-form (udtalt [əð]). Spørgsmålet som vi stiller i denne artikel, er hvad der ligger til grund for denne sprogforandringsproces. Vores hypotese har været at forandringen bunder i varianternes sociale betydninger, og vi har derfor gennemført en sprogholdningsundersøgelse med henblik på at afdække disse.

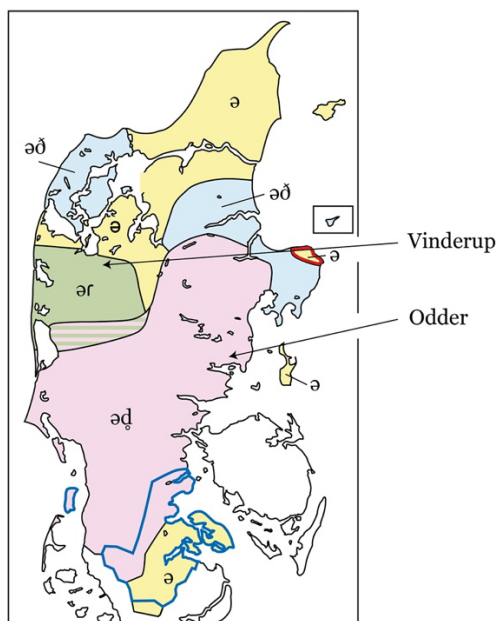
I artiklen vil vi først redegøre nærmere for variabelen og de hypoteser der ligger til grund for undersøgelsen. Herefter vil undersøgelsens design blive gennemgået efterfulgt af en præsentation og diskussion af resultaterne samt en kort konklusion.

2. Baggrund og hypoteser

Indledningsvis skal det bemærkes at der i eksempel 1 og 2 ovenfor i begge tilfælde er tale om infinitte former med verbal funktion (som del af perifrastisk perfektum eller passiv), altså det der i bl.a. Grammatik over det danske sprog (Hansen & Heltoft 2011:668) kaldes supinum, hvilket vi også vil gøre i det følgende. Det er kun former med denne funktion der vil blive behandlet i denne artikel.

I de sjællandske dialekter og i standarddansk anvendes supinum-fleksivet *-et* ved såvel svage som stærke verber (fx *blevet* og *gået*). I størstedelen af de traditionelle jyske dialekter anvendes derimod som nævnt fleksivet *-en* i forbindelse med stærke verber mens *-et*-formen kun anvendes ved svage verber, og verber der er overgået til svag bøjning i forhold til supinum-formen (I. L. Pedersen 2019; K. M. Pedersen 2019).

Ud over variationen mellem *-en* og *-et* er der geografisk variation i udtalen af *-et*-fleksivet. Figur 1 viser udtalen i de traditionelle jyske dialekter.



Figur 1: Kort over udtalen af *-et* i jyske dialekter (Jysk Ordbog: Kort 6.1; lydskrift ændret til IPA).

I standarddansk udtales *-et*-suffikset med det såkaldte bløde (approksimative) *d*, dvs. som [əð], men i et stort centralt og østjysk område dominerer udtalen med hårdt (klusilt) *d*, dvs. som [ɛ̃d̥]; i perifere områder af Jylland ses mere reducerede former, især [ə]. Udtalen med klusilt *d* tolkes derfor generelt som en novation, muligvis opstået ud fra en læseudtale, med udgangspunkt i en mere reduceret form, og den har frem til i hvert fald begyndelsen af det 20. århundrede spredt sig i Jylland og endda til Fyn (Pedersen 1987, 2002). Vi henviser til Jensen og Maegaard (2012) for en mere uddybet beskrivelse af den historiske udvikling.

Ovenstående beskriver de traditionelle dialekter, men vi har i en tidligere undersøgelse baseret på LANCHART-korpusset undersøgt brugen af supinum-former af stærke verber i de to byer Vinderup og Odder (jf. figur 1) i perioden fra slutningen af 1970'erne til 2005-07 (Jensen & Maegaard 2010, 2012). Her undersøgte vi for det første udviklingen i brugen af [ən]-formen de to steder, og for det andet hvilke former der tager over når [ən]-formen forsvinder som udtryk for den generelle afdialektisering der ses i anden halvdel af det 20. århundrede

(Maegaard et al. 2013). Vi var her særligt interesserede i om det er den østjyske [əɖ]-form der tager over i en videreførelse af den regionalisering der har fundet sted tidligere i det 20. århundrede, eller om det er den standarddanske [əð]-form. Resultaterne viste for det første at [ən]-formerne var i kraftig tilbagegang i perioden; i starten især i Odder, men senere også i Vinderup. For det andet viste resultaterne at der ikke er tale om regionalisering, men om standardisering: Det var meget tydeligt [əð]-formen der vandt frem på bekostning af såvel [ən]- som [əɖ]-formerne.

Tidligere undersøgelser har vist at sprogbrugsforandringer både i Danmark og andre steder finder sted i retning af den sprogbrug som vurderes positivt når talere vurderes uden at dem der vurderer, er klar over at det er sproglig variation der undersøges (Kristiansen 2009; Maegaard et al. 2013; Grondelaers & Kristiansen 2013). En oplagt hypotese er derfor at de forandringer i brugen af supinum-former der er sket gennem de seneste årtier, er drevet af formernes sociale betydninger. Formålet med den undersøgelse der vil blive præsenteret i det følgende, er således at afdække forskelle i varianternes sociale betydninger.

Til det brugte vi et såkaldt *matched guise*-design, hvor forsøgspersoner præsenteres for lydoptagelser og bliver bedt om at vurdere den talende i forhold til ikke-sproglige parametre som alder, uddannelse og personlighedstræk, sidstnævnte typisk i forhold til skalaerne superioritet, attraktivitet og dynamik (Zahn & Hopper 1985; Garrett 2010). Lydklippene afviger dog kun fra hinanden mht. sproglige træk, men dette skjules så vidt muligt for forsøgspersonen. På denne måde eliciteres underbevidste holdninger og stereotype forestillinger om de sproglige varieteter der undersøges.

Resultater af sådanne undersøgelser har som nævnt vist sig at korrelere med sprogforandringsprocesser, men de er dog i stort set alle tilfælde udført i forhold til sproglige varieteter og ikke enkelttræk. En del af formålet med vores undersøgelse har derfor været at teste om denne metode kan anvendes i forhold til et enkelttræk som supinum-former.

Vi betragter supinum-formerne af stærke verber som en sociolingvistiske variabel med varianterne [ən], [əɖ] og [əð], og vi havde på forhånd en række hypoteser om den sociale betydning af de forskellige varianter: For det første forventede vi at [ən] forbindes med dialekt og jysk tilhørsforhold, mens [əð] forbindes med standard-/rigsdansk – og muligvis også københavnsk eller i hvert fald østdansk tilhørsforhold.

Da [ən]-varianten er knyttet til traditionelle dialekter, forventede vi at den ville føre til en vurdering af taleren som kommende fra landet, mens [əð] omvendt ville trække i retning af en vurdering af at taleren kommer fra byen.

I forhold til vurderingsskalaen superioritet, som omhandler kompetence og social status, ville vi ud fra de holdninger vi normalt ser i forhold til regional præget sprog og standarddansk (fx Kristiansen 2009), forvente at [ən] fører til en lav vurdering, mens [əð] fører til en høj vurdering. Samme forventning havde vi i forhold til dynamik- og attraktivitetsskalaerne. Dynamik handler om selvtillid, aktivitet og entusiasme, mens attraktivitet relaterer sig til solidaritet, troværdighed, venlighed. Det generelle sprogholdningsbillede i Danmark har været at lyttere nedvurderer lokal sprogbrug i alle tre dimensioner (se fx Kristiansen 2009 for en sammenstilling af fire forskellige steder), men der er også undersøgelser der viser et mere komplekst billede hvor lokalsprog opvurderes, enten generelt eller mht. specifikke skalaer (se Maegaard & Quist 2020; Maegaard et al. 2020; Hansen & Goldshtein 2021).

Da lokalt og regionalt præget sprog i Danmark både generelt og i forhold til supinum-varianter i højere grad ses hos ældre end yngre sprogbrugere (Jensen & Maegaard 2012:182, 187), forventede vi også at [ən] ville føre til en vurdering af taleren som ældre end ved [əð]. Såvel i dette tilfælde som i forhold til de tidligere nævnte skalaer var vores forventning at [əɗ] vurderingsmæssigt ville ligge imellem [ən] og [əð].

Endelig forventede vi at der kunne være forskelle i vurderingen af de tre varianter afhængigt af om taleren er mand eller kvinde, fordi det er et generelt mønster at variation opfattes forskelligt alt efter talerens køn, og mere specifikt at kvindelige talere vurderes lavere end mandlige talere når de benytter lokale dialekttræk (se fx Maegaard 2005). Dette hænger igen sammen med den sociale betydning af lokale træk idet dialektale og regionale træk ofte knyttes til betydninger som har at gøre med egenskaber som opfattes mere positivt hos mænd end hos kvinder (fx ligefremhed, praktisk arbejde, lokal forankring og traditionelle værdier). Samtidig forventede vi at der ville være forskelle i vurderingerne afhængigt af hvor lytteren kommer fra, og af vedkommendes øvrige baggrundsvariable som fx køn og alder.

3. Data og metode

Til undersøgelsen brugte vi lydclip med supinum-varianter af to forskellige lemmaer, *hjælpe* og *skrive*. Vi valgte disse lemmaer fordi de i sprogbrugsundersøgelsen var blandt dem der forekom med alle tre suffiksvarianter hos talerne fra Odder i Østjylland.

I et forsøg på at sikre den økologiske validitet af stimuli til undersøgelsen, og derved sørge for at de ligger så tæt som muligt på faktisk forekommende sprogbrug, udvalgte vi ytringer fra LANCHART-korpusset og forsøgte at klippe de forskellige varianter af suffikset ind i sætningerne. Det blev desværre aldrig med et tilfredsstillende resultat, især ikke for varianten [əð] (formentlig fordi den indeholder en approksimant). Derfor valgte vi i stedet at finde

autentiske eksempler fra korpusset og konstruere vores stimuli ud fra dem. Vi valgte syv sætninger som enten indeholdt *skrive* eller *hjælpe*. Sætningerne gengives her, og de variable suffikser er fremhævet med kursiv:

1. Det har været dejligt at mine store brødre har *hjulpet* mig meget
2. Det har da *hjulpet* meget at hun kan komme ud og få nogle nye friske input
3. Vi har mange venner og kammerater som har *hjulpet* os rigtig meget
4. Jeg har lige *skrevet* på en sag om magtanvendelse i dag her på arbejdet
5. Jeg har jo *skrevet* i min ansøgning at jeg er indstillet på at videreudanne mig
6. Det er den eneste gang mine lærere har *skrevet* i min kontaktbog
7. Jeg har ikke hørt fra dem endnu men jeg har *skrevet* over til dem

Disse syv sætninger fik vi indtalt af fire forskellige talere fra Østjylland i tre forskellige versioner, en for hver variant af suffikset. Alle fire talere var fra Østjylland og havde boet der det meste af deres liv. Vi fandt talere der alle var omkring 50 år gamle for at sikre at eventuelle spor af deres alder i udtalen af sætningerne var forenelig med alle tre varianter af suffikset. Da den forudgående sprogbrugsundersøgelse havde vist at særligt [ən] men også [əɗ] kun sjældent (hvis overhovedet) forekom hos yngre talere, virkede det mere rimeligt at bruge midaldrende stemmer til indtalingen. Til gengæld ville vi gerne vide om der ville være en effekt af talerens køn (som i tidligere undersøgelser), og derfor brugte vi to mænd og to kvinder til indtalingerne.

Indtalingerne blev foretaget via Zoom under ledelse af en af os, og indtalerne blev bedt om at gentage en sætning hvis de tøvede eller afbrød sig selv i løbet af indtalingen, eller hvis de brugte en anden variant af suffikset end den de var blevet bedt om. Alle fire talere nævnte at det ikke faldt dem lige naturligt at bruge hver variant i hver sætning (men alle mente at de selv kunne finde på at bruge en hvilken som helst af varianterne med begge lemmer). Det kan ikke udelukkes at stimuli bærer et vist præg af oplæsning. Indtalerne blev bedt om ikke at fremhæve enkelte ord i sætningen, men blot sige den stille og roligt og naturligt (hvilket så var op til dem selv at fortolke). Det har medvirket til et meget ensartet intonationsforløb i hver indtaling, ligesom vi heller ikke har bemærket variation i trykgruppemønstre hverken hvad angår timingen eller omfanget af det tonale forløb. Sætningerne indeholdt i øvrigt ikke segmentelle træk der er kendt som stereotypiske for østjysk, bortset fra at en enkelt af talerne i to af sætningerne anvendte den regionale form [mɑjəɗ] i stedet for det standarddanske [mɑ:ð]. Da han gjorde det konsekvent uanset hvilken supinum-variant sætningerne blev indtalt med, anså vi dog ikke dette for et problem i vores design. Det kan ikke udelukkes at det kan have påvirket vurderingen af

stemmen, men da vi er interesserede i vurderingen af varianterne af suffikset, og eftersom alle indtalere var repræsenteret med alle varianter, så bør dette særtræk ved den ene mandlige stemme ikke kunne påvirke de resultater vi præsenterer her.

På denne måde fik vi konstrueret syv sætninger i tre versioner hver indtalt af fire talere, dvs. 84 stimuli i alt. Det ville være alt for meget for en enkelt lytter at skulle forholde sig til, og de ville formentlig tabe koncentrationen hvis de skulle deltage i et forsøg med så mange stimuli. Vi ønskede desuden ikke at deltagerne skulle høre hver stemme mere end én gang, idet vi så ville risikere at deltagerne blev bevidste om den specifikke sproglige variation vi ville undersøge. Derfor anvendte vi et *between subject*-design til selve eksperimentet. Hver respondent hørte således kun hver taler én gang, og reaktionerne på de enkelte stimuli blev så sammenlignet på tværs af respondenterne.

Eksperimentet blev gennemført online (blandt andet på grund af corona-nedlukningerne i perioden for dataindsamling) via SoSci Survey (www.soscisurvey.de). Deltagere blev rekrutteret via vores egne Facebook-profiler og netværk i øvrigt. Deltagere fulgte et link og blev i første omgang tildelt en af 21 forskellige lister med fire stimuli. De hørte altså allesammen fire stimuli hver, to med lemmaet *skrive* og to med lemmaet *hjælpe*. Alle tre varianter af supinum og alle fire talere var repræsenteret på hver liste, så hver deltager blev eksponeret for den fulde variation.

Forsøget bestod af to dele: I den første del skulle deltagerne vurdere taleren bag det lyd-klip de lige havde hørt, på de tre skalaer Intelligent, Flink og Interessant (én skala fra hver af dimensionerne *superioritet*, *attraktivitet* og *dynamik*). Vi valgte at benytte få skalaer, men én fra hver dimension, fordi vi derigennem får vurderinger fra hver dimension, men samtidig bliver det ikke uoverskueligt langt for respondenterne. Vurderingerne gik fra 0 til 4, hvor 0 betød 'slet ikke' og 4 betød 'meget' (disse værdier blev omkodet til regressionsanalyserne, så de i stedet gik fra 1 til 5). Efter vurderingen på skalaerne skulle deltagerne give deres bud på hvilken uddannelse taleren senest havde taget (blandt seks valgmuligheder), og hvor gammel de mente taleren var. Efter at have vurderet den første taler på denne måde blev hver deltager præsenteret for en ny taler som skulle bedømmes på samme måde. Når en deltager havde vurderet fire lydklip, gik forsøget over til den anden del. Her fik deltagerne at vide at de nu skulle høre hvert lydclip en gang til og for hver af dem angive i hvor høj grad taleren talte rigsdansk, hvilken region i Danmark de troede taleren kom fra, og om taleren boede på landet eller i byen. Vi lavede denne inddeling i forsøget for at undgå at deltagerne under vurderingen på personlighedstrækkene blev opmærksomme på den specifikke sproglige variation vi ønskede at under-

søge. Forsøget sluttede med et åbent spørgsmål til deltagerne om hvad de troede undersøgelsen handlede om. Spørgeskemaet kan ses som bilag til artiklen.

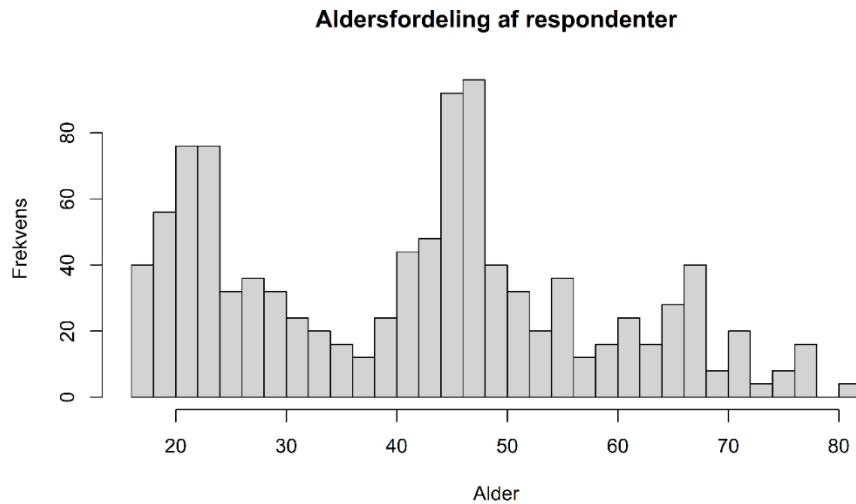
Undersøgelsen blev gennemført i begyndelsen af 2022, og deltagere blev som sagt rekrutteret gennem private netværk (dog blev sprogvidenskabelige kolleger bedt om ikke at deltage). Vi frasorterede alle ikke-fuldførte besvarelser og endte med i alt 262 respondenter fra hele Danmark i alderen 17-81 år. 64 % af respondenterne var kvinder, 31 % var mænd og 5 % angav ”andet” eller ønskede ikke at svare på dette spørgsmål. Fordelingen angående uddannelsesniveau og geografi ses i tabel 1 og 2, mens distributionen i forhold til alder fremgår af figur 2. Vi havde kun i meget begrænset omfang kontrol over hvem som deltog, og respondenter er således heller ikke ligeligt fordelt i forhold til køn, alder, uddannelsesniveauer og geografi. Der er således overrepræsentation af kvinder og respondenter med enten gymnasial uddannelse (overvejende studerende) eller mellemlange og lange videregående uddannelser, og aldersfordelingen er ikke jævn hvorfor vi valgte at opdele respondenterne i tre aldersgrupper: 17-34, 35-59 og 60-81 år. Vi sørgede i forbindelse med rekrutteringen for at få god repræsentation af respondenter med østjysk baggrund, og i modsætning til i de fleste tidligere danske sprogholdningsundersøgelser er der en god aldersspredning, dvs. der er ikke kun unge respondenter. Baggrundsvariablenes eventuelle indflydelse på respondenternes vurderinger blev inddraget i analyserne.

Uddannelsesniveau	%
9. klasse	6
Gymnasiet	18
Erhvervsuddannelse	3
Lang videregående	42
Mellemlang videregående	26
Kortere videregående	4
Ved ikke	1

Tabel 1: Respondenternes uddannelsesniveau.

Fødeegn	%
Østjylland	21
Øvrige Jylland	21
Øvrige Danmark	58

Tabel 2: Respondenternes geografiske baggrund.



Figur 2: Respondenternes alder.

4. Resultater

Vi brugte ordinal logistisk regression med blandede effekter (Baayen 2008; Christensen 2019) til at analysere dataene. Den vigtigste uafhængige variabel var naturligvis supinum-varianten da vi jo især var interesserede i hvordan den påvirker vurderingen af taleren, og denne faktor blev derfor inkluderet som fast effekt. Derudover inddrog vi følgende faktorer som faste effekter: hvilket lemma supinum-suffikset forekom sammen med, hvilket køn taleren havde, respondentens fødested, køn, aldersgruppe og uddannelse. For at kontrollere for potentiel indflydelse fra de enkelte sætninger eller de enkelte talere (eller en kombination) på vurderingerne inddrog vi (som tilfældige effekter) også hvilken af de syv stimuli-sætninger bedømmelsen skete i forhold til, og hvilken af de fire talere der havde produceret den enkelte stimulus. Herved undgår vi at resultaterne påvirkes af forskelle i indhold mellem sætningerne eller eventuelle idiosyncrasier hos indtalerne (jf. at kun den ene mandlige indtaler brugte den østjyske form af ordet 'meget').

I forbindelse med responsvariablene Intelligent, Interessant, Taleralder, Sted (land/by) og Region blev bedømmelsen af taleren ikke påvirket statistisk signifikant af hvilken variant af supinum der blev anvendt. Bedømmelsen blev i alle tilfælde påvirket af andre af de inddragede faktorer, men da vi primært var interesserede i hvordan bedømmelsen af taleren påvirkes af supinum-varianten, vil disse resultater, bortset fra variabelen Region, ikke blive beskrevet nærmere.

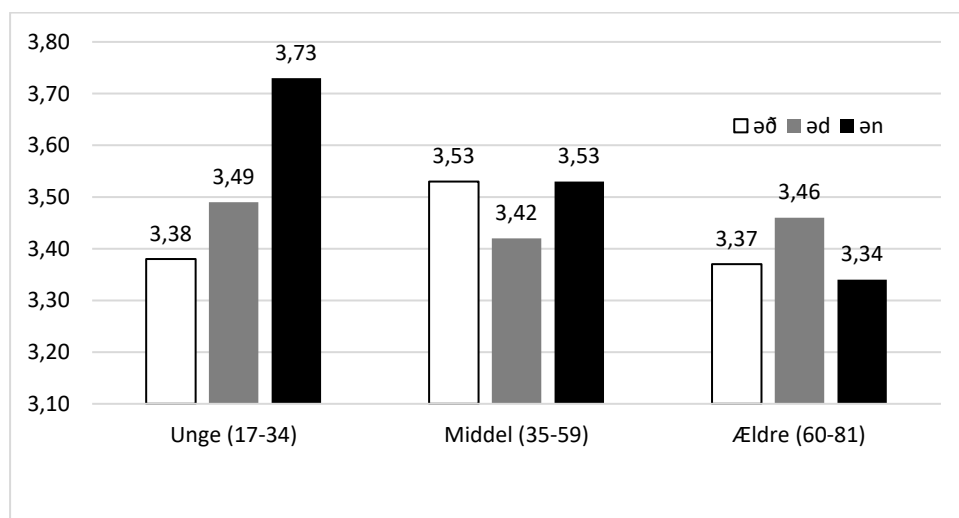
Resultaterne viser derimod at vurderingen af hvor Flink taleren er, og hvor Rigsdansk/standarddansk vedkommende taler, blev påvirket statistisk signifikant af supinum-varianten.

4.1. Flink

Figur 3 viser den statistiske model der bedst forklarer variationen i vores datamateriale mht. hvor Flink taleren vurderes til at være. Med 'bedste model' menes den model som indeholder alle de faktorer (og interaktioner mellem faktorer) som har en statistisk signifikant effekt på den afhængige variabel, og som samtidig hver især gør modellen signifikant bedre til at forklare variationen end en model uden denne faktor. Ud over supinum-variablen har også respondentens alder og uddannelse signifikant betydning for vurderingen. Analysen viser desuden at supinum-variablen og respondent-alderen interagerer med hinanden i forhold til hvordan de påvirker vurderingen, og deres effekt er derfor svær at gennemskue alene ved at se på tallene i modellen; det ses tydeligere ved at se på figur 4, som viser de gennemsnitlige bedømmelser. Vi kommer her ikke nærmere ind på faktoren uddannelse (ListenerEducation) da den ikke interagerer med supinum-variablen, dvs. den påvirker bedømmelsen af taleren men uafhængigt af hvilken supinum-variant der anvendes, hvilket gør den relativt uinteressant i forhold til formålet med vores undersøgelse.

Random effects:					
Groups	Name	Variance	Std.Dev.		
Speaker	(Intercept)	0.2199	0.4689		
Number of groups: Speaker 4					
Coefficients:					
		Estimate	Std. Error	z value	Pr(> z)
AgeGroupMiddle		-0.006936	0.247635	-0.028	0.97765
AgeGroupOld		-0.424439	0.312783	-1.357	0.17479
Variant əđ		0.241097	0.236564	1.019	0.30813
Variant ən		0.751457	0.234849	3.200	0.00138 **
ListenerEducationVocShortMedium		0.549625	0.196028	2.804	0.00505 **
ListenerEducationLong		0.346650	0.187680	1.847	0.06474 .
AgeGroupMiddle:Variant əđ		-0.468409	0.316075	-1.482	0.13835
AgeGroupOld:Variant əđ		-0.054168	0.408655	-0.133	0.89455
AgeGroupMiddle:Variant ən		-0.719346	0.314051	-2.291	0.02199 *

Figur 3: Bedste model for responsvariablen Flink (*treatment*-kodning med følgende default/reference-niveauer: Variant: æð, AgeGroup: Young, ListenerEducation: Other).



Figur 4: Gennemsnitlige bedømmelser for responsvariablen Flink.

Som nævnt blev respondenterne opdelt i tre aldersgrupper, og det er kun inden for den yngste aldersgruppe (17-34 år) at supinum-varianterne fører til signifikant forskellige bedømmelser. Resultaterne viser at [æn]-varianten, dvs. den mest dialektale variant, hos de unge modsat vores hypotese fører til at taleren vurderes som *mere* flink end både den regionale [æð]-variant ($p=0,029$) og den standarddanske [æð]-variant ($p=0,001$). Der er derimod ingen signifikant forskel mellem [æð] og [æð] mht. hvordan de påvirker bedømmelsen.

4.2. Rigsdansk

Figur 5 viser den statistiske model der bedst forklarer variationen i vores datamateriale i forhold til hvor Rigsdansk taleren vurderes at tale. Som det fremgår, er der her kun supinum-variablen der har statistisk signifikant effekt på vurderingen, og den har den effekt man ville forvente ud fra sprogbrugen: Talere der anvender [æn]- og [æð]-varianterne, vurderes som mindre rigsdansktalende end talere der anvender [æð]-varianten. Vurderingen er uafhængig af respondent-variablene, altså af hvor respondenterne kommer fra i landet, samt deres alder, uddannelse og køn. Figur 6 viser gennemsnitsvurderingerne.

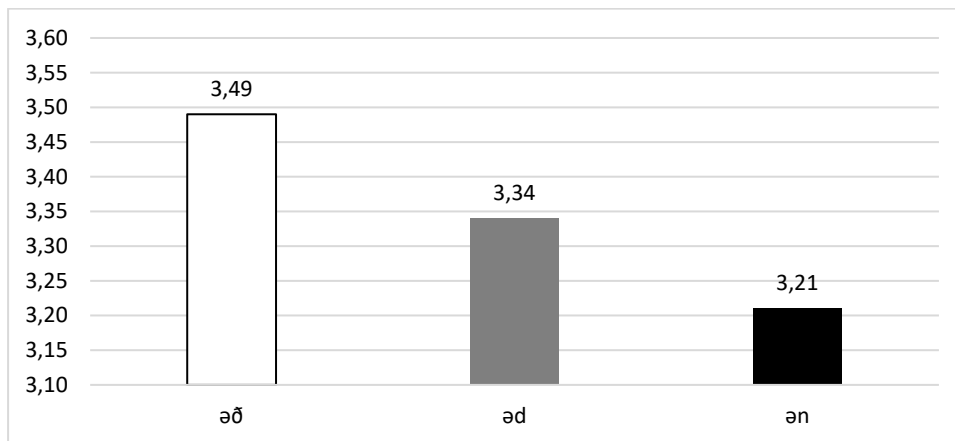
```

Random effects:

Groups   Name              Variance Std.Dev.
Sentence (Intercept) 0.02002  0.1415
Speaker  (Intercept) 0.12004  0.3465
Number of groups:  Sentence 7,  Speaker 4

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
Variant æð  -0.2814     0.1407  -2.000  0.045496 *
Variant æn  -0.4737     0.1396  -3.393  0.000691 ***
    
```

Figur 5: Bedste model for responsvariablen Rigsdansk (*treatment*-kodning med default/reference-niveauet æð for Variant).



Figur 6: Gennemsnitlige bedømmelser for responsvariablen Rigsdansk.

4.3. Region

Alle fire talere kom som nævnt fra Østjylland, og de taler for os at høre med tydelig østjysk intonation. I 36 % af vurderingerne placeres de også i Østjylland, men i størstedelen af tilfældene placeres de altså andre steder i landet. I 25 % af tilfældene placeres de dog i andre dele af Jylland, så samlet set placeres de i de fleste tilfælde (61 %) i Jylland, men i 27 % af tilfældene placeres de faktisk på Sjælland. Som nævnt betyder supinum-varianten, modsat vores hypotese, ikke noget for denne vurdering, men det gør til gengæld respondenternes baggrund: Respondenter fra Jylland og især Østjylland har størst tendens til at placere talerne i Østjylland. Desuden viser resultaterne at ældre respondenter har større tendens til at placere taleren i Østjylland end de yngre. Vi henviser til Jensen et al. (2023) for en nærmere beskrivelse af disse resultater.

5. Diskussion

Alt i alt må man sige at vores undersøgelse kun har kunnet påvise en temmelig afgrænset effekt af supinum-variablen, og resultaterne kan derfor kun i beskedent omfang siges at understøtte vores hypotese om at ændringerne i sprogbrugen drives af sprogholdninger.

Det kan der være rent metodiske grunde til: For det første kan det være at stimuli lyder unaturlige fordi der er tale om oplæste sætninger. Vi forsøgte naturligvis at få talerne til at sige sætningerne så naturligt som muligt, men man må alligevel sige at de i flere tilfælde lyder lidt oplæste, og at det kan være at det forstyrrer vurderingerne set i forhold til hvis det havde været spontan tale.

For det andet kan det være at stimuli har været for korte til at respondenterne har følt det rimeligt at skulle vurdere talerne, og at de derfor har givet middelbedømmelser til alle. I de fleste sprogholdningsundersøgelser anvendes der længere lydclip, men det har vi afholdt os fra fordi vi var bekymrede for at en effekt af en enkelt supinum-variant så ville drukne i støj.

For det tredje kan det være at respondenterne var bevidste om at det er supinum-suffikset der undersøges, og det derfor er bevidste snarere end underbevidste sprogholdninger udtrykkes – måske fordi der er tale om korte og lidt unaturlige lydclip, måske fordi grammatiske variable tiltrækker sig mere opmærksomhed end fonetiske. Dette kan ikke afvises, men ud fra respondenternes egne svar er der ikke meget der tyder på det. Som nævnt blev alle respondenter spurgt om hvad de troede at undersøgelsen handlede om, men kun ganske få angav supinum, eller grammatik i det hele taget. De fleste angav derimod at undersøgelsen handlede om generelle holdninger til dialekter, hvilket ikke er overraskende da de sidste spørgsmål jo omhandler rigsdansk og geografisk tilhørsforhold.

Endelig kan det være at variabelen primært har social betydning *i kombination* med andre variable, altså som del af et register eller en stil. I modsætning til de fleste tidligere sprogholdningsundersøgelser retter vores undersøgelse sig mod en enkelt variabel og ikke en varietet, og det kan måske forklare de anderledes resultater.

Designet har dog kunnet vise effekter af supinum-variablen, og afgørende er det her at [əð]-varianten fører til en vurdering af taleren som mere rigsdansk end både [əɖ]- og [ən]-varianterne, også selvom respondenterne tilsyneladende ikke bemærker at det alene er supinum-variablen som varieres i forsøget. Supinum-variablen har derimod ikke betydning for hvorvidt taleren placeres i (Øst)Jylland eller ej, hvilket er overraskende eftersom [əɖ]- og [ən]-varianterne sprogbrugsmæssigt både traditionelt og i LANCHART-korpusset i modsætning til [əð]-varianten er knyttet til Jylland. På den måde peger resultaterne på at [əð]-varianten opfattes som

en rigsdansk variant, som på trods af sit oprindeligt østdanske ophav ikke stedsbestemmes til et særligt sted, men som nu virker stedsneutral (se Maegaard, Quist & Pharao 2020 for en lignende diskussion i forbindelse med hvad der kan opfattes som hhv. københavnsk og bornholmsk).

Den effekt at [əð]-varianten opfattes som mere rigsdansk, er forventet i forhold til de konstaterede forandringer i sprogbrugen; derimod er det uventet at vurderingen mht. rigsdansk ikke synes at blive knyttet til de evalueringsdimensioner som normalt knytter sig til standarddansk versus lokalt præget sprogbrug. Vi ser således ikke en positiv effekt af den standarddanske [əð]-variant i forhold til superioritet, dynamik, attraktivitet og urbanitet. Derimod ser vi at den mest lokale variant, [ən], fører til en *højere* vurdering af taleren i forhold til attraktivitetsdimensionen (Flink).

Internationalt set er det ikke specielt overraskende at der kan være en positiv effekt af den lokale, ikke-standardvariant i forhold til attraktivitetsdimensionen (jf. fx Garrett 2010:55 ff.), men i en dansk kontekst er det faktisk overraskende da mange tidligere sprogholdningsundersøgelser har vist at den lokale varietet nedvurderes på alle skalaer set i forhold til standarddansk (Maegaard 2005; Kristiansen 2009). Dog har visse nyere undersøgelser i Jylland, som nævnt i afsnit 2, netop vist et mere komplekst billede hvor lokal sprogbrug opvurderes enten generelt eller på visse skalaer (Maegaard et al. 2020; Maegaard & Quist 2020; Hansen & Goldshtein 2021). Det er også værd at bemærke at det kun er hos den unge del af respondenterne denne effekt ses. Som nævnt er de talere som vurderes, omkring 50 år gamle, og den positive vurdering kan således skyldes at respondenterne vurderer talere der er betydeligt ældre end dem selv. I de fleste tidligere danske undersøgelser har respondenterne vurderet talere som er omtrent jævnaldrende med dem selv. Det er således muligt at den lokale, 'gammeldags' sprogbrug har en nostalgisk positiv effekt når den høres hos talere der er ældre end dem selv (i dette tilfælde svarende til forældregenerationen), selvom den ikke ville have det hvis respondenterne hørte den hos jævnaldrende. Dette ville svare til resultater fra nyere undersøgelser af sproglig kommodificering som viser at ældre, lokale sprogformer opfattes som autentiske og sympatiske samtidig med at de forsvinder fra hverdagsprogbrug (Kelly-Holmes & Pietikäinen 2014; Jaffe 2019; Maegaard & Karrebæk 2019).

Resultaterne er også interessante mht. til forholdet mellem bevidsthed og social betydning. Den undersøgte supinum-variation er tilsyneladende en variation som der er en høj grad af bevidsthed om. Den benyttes fx i parodier, og både [əð]- og [ən]-varianterne kan opfattes som *stereotypes* med Labovs terminologi (Labov 2001:196). Samtidig kan vi også se af vurderingerne på *rigsdansk*-skalaen at lytterne registrerer variationen og reagerer på den. Alligevel

vurderer de ikke talerne ret forskelligt på de tre personlighedstræk og mht. uddannelsesniveau. Det kan tolkes som udtryk for at et træk på trods af en høj grad af saliens ikke nødvendigvis behøver at have stor betydning for sociale vurderinger. Vi henviser til Jensen et al. (2023) for yderligere diskussion af undersøgelsens resultater.

6. Konklusion

Formålet med undersøgelsen var at afdække den sociale betydning af tre supinum-varianter i en østjysk sprogbrugskontekst. Resultaterne har kun bekræftet én af vores hypoteser om den sociale betydning, nemlig at [əð]-varianten opfattes som mere rigs-/standarddansk end både [əɖ]- og [ən]-varianterne.

Denne vurdering følges uventet ikke af de vurderinger som i tidligere undersøgelser ses knyttet til standarddansk vs. lokalt præget sprogbrug; derimod ses det at den mest lokale variant, [ən], hos de yngre respondenter fører til vurdering af taleren som mere ”flink” end de øvrige varianter.

Resultatet er overraskende set i lyset af den forandring i sprogbrugen der har fundet sted i sidste del af de 20. århundrede, hvor [əð]-varianten har vundet voldsomt frem på bekostning af de lokale varianter. Den begrænsede effekt af supinum-variablen i forsøget kan være en indikation af at afvigelser fra standardnormen på dette område ikke (længere?) tilskrives samme negative sociale betydninger som andre lokale varianter. Dette kan igen danne grundlag for den hypotese at de lokale varianter af supinum-suffikset kun vil forsvinde langsomt fra sprogbrugen i Østjylland.

Litteratur

- Jysk Ordbog*. Århus: Peter Skautrup Centret for Jysk Dialektforskning, Aarhus Universitet. <http://jyskordbog.dk>.
- Baayen, R. Harald (2008) *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Christensen, Rune Haubo Bojesen (2019) *ordinal - Regression Models for Ordinal Data. R package version 2019.12-10*. www.r-project.org
- Garrett, Peter (2010) *Attitudes to language*. Cambridge: Cambridge University Press.
- Grondelaers, Stefan & Tore Kristiansen (2013) On the need to access deep evaluations when searching for the motor of standard language change. Tore Kristiansen & Stefan Grondelaers (red.). *Language (De)standardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus:9–52.
- Hansen, Erik & Lars Heltoft (2011) *Grammatik over det Danske Sprog*. København: Det Danske Sprog- og Litteraturselskab.

- Hansen, Inger Schoonderbeek & Yonatan Goldshtein (2021) Er København (virkelig) eneste sproglige normcenter? I *MUDS* 18:189-210.
- Jaffe, Alexandra (2019) Poeticizing the economy: The Corsican language in a nexus of pride and profit. I *Multilingua* 38(1):9-27.
- Jensen, Torben Juel & Marie Maegaard (2010) Hvor er bleven blevet af? Præteritum participium af stærke verber i Jylland. I *Danske talesprog* 10:34-56.
- Jensen, Torben Juel & Marie Maegaard (2012) Past participles of strong verbs in Jutland Danish: A real-time study of regionalization and standardization. I *Nordic Journal of Linguistics* 35(2):169-195.
- Jensen, Torben Juel, Marie Maegaard & Nicolai Pharao (2023) Investigating indexicality of regional morphological variation and change. Hanna Fischer & Stefan Rabanus (red.). *Morphologische und syntaktische Variation in den deutschen Regionalsprachen: Impulse für die Erforschung der sprachlichen Vertikale*. Hildesheim/Zürich/New York: Georg Olms Verlag:241-272.
- Kelly-Holmes, Helen & Sari Pietikäinen (2014) Commodifying Sámi Culture in an Indigenous Tourism Site. I *Journal of Sociolinguistics* 18(4):518–538.
- Kristiansen, Tore (2009) The macro-level social meaning of late-modern Danish accents. I *Acta Linguistica Hafniensia* 41:167-192.
- Labov, William (2001) *Principles of linguistic change - Social factors*. Malden/Oxford: Wiley-Blackwell.
- Maegaard, Marie (2005) Language attitudes, norm and gender a presentation of the method and results from a language attitude study. I *Acta Linguistica Hafniensia* 37(1):55-80.
- Maegaard, Marie, Torben Juel Jensen, Tore Kristiansen & Jens Normann Jørgensen (2013) Diffusion of language change: accommodation to a moving target. I *Journal of Sociolinguistics* 17(1):3-36.
- Maegaard, Marie & Martha Sif Karrebæk (2019) 'The illusion or the truth?' – backstage constructions of authenticity in an up-market restaurant. I *Language & Communication* 69:54-66.
- Maegaard, Marie & Pia Quist (2020) Perception, Recognition, and Indexicality. Experimental Investigations of Variation in Northern Jutland, Southern Jutland, and on Bornholm. Marie Maegaard, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (red.). *Standardization as Sociolinguistic Change. A Transversal Study of Three Traditional Dialect Areas*. New York: Routledge:145-168.
- Maegaard, Marie, Pia Quist & Nicolai Pharao (2020) Genkendelse, sprogholdninger og standardisering: Nye fund i Bylderup, Hirtshals og Nexø. I *Danske talesprog* 20:143-174.
- Monka, Malene (2013) Sted og sprogforandring : en undersøgelse af sprogforandring i virkelig tid hos mobile og bofaste informanter fra Odder, Vinderup og Tinglev. I *Danske talesprog* 13:1-336.
- Pedersen, Inge Lise (1987) Fynsk mellem øst og vest. I *MUDS* 1:177-189.
- Pedersen, Inge Lise (2002) The impact of internal or contact-induced change on weak preterites in -et in Danish dialects with an outlook to Norway and Sweden. Jan Berns & Jaap van Marle (red.). *Present-day dialectology. Problems and findings*. Berlin: Mouton de Gruyter:269-282.
- Pedersen, Inge Lise (2019) Sociolekter – bøjning. Ebba Hjorth, Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). *Dansk sproghistorie*. Bd. 3. København: Det Danske Sprog- og Litteraturselskab:383-389.

Hvordan vurderes jyske participiumsformer?

- Pedersen, Karen Margrethe (2019) Dialekter - bøjning. Ebba Hjorth, Henrik Galberg Jacobsen, Bent Jørgensen, Birgitte Jacobsen, Merete Korvenius Jørgensen & Laurids Kristian Fahl (red.). *Dansk sproghistorie*. Bd. 3. København: Det danske sprog- og litteraturselskab:277-301.
- Zahn, Christopher J. & Robert Hopper (1985) Measuring Language Attitudes: The Speech Evaluation Instrument. *I Journal of Language and Social Psychology* 4(2):113-123.

Bilag

Spørgeskemaet

Del 1 – holdningsspørgsmål

Velkommen

Velkommen til vores undersøgelse og mange tak fordi du vil bruge tid på at deltage i den.

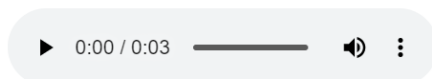
Du skal nu høre 4 korte lydclip sagt af 4 forskellige personer. Du skal selv trykke på play-knappen øverst på hver side for at høre et lydclip. For hvert lydclip vil vi bede dig om at svare på en række spørgsmål.

Til sidst vil vi bede dig om at give os lidt oplysninger om dig selv. Disse oplysninger bevares sikkert og fortroligt og din besvarelse vil ikke kunne spores tilbage til dig.

Undersøgelsen tager ca. 5 minutter at gennemføre.

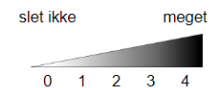
Klik på knappen Næste for at gå videre og begynde undersøgelsen!

Næste



1. Personlighed

Her kan du høre et kort klip med en mand. Lyt til klippet én gang og giv så din vurdering af ham på skalaerne nedenfor.



Intelligent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Flink	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interessant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Hvor gammel tror du han er?

Gæt alderen på ham der har indtalt lydklippet

3. Hvilken uddannelse tror du han har?

Sæt kryds ved den uddannelse du tror manden senest har færdiggjort.

- Grundskole
 - Gymnasieuddannelse
 - Erhvervsuddannelse
 - Kort videregående
 - Mellemlang videregående
 - Lang videregående
-
- Ved ikke

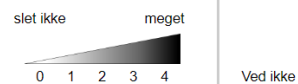
Næste

Denne del blev gentaget for i alt 4 stemmer

Del 2 – genkendelsestest

13. Rigsdansk/Standarddansk

Nu skal du høre den første person een gang til. Lyt til klippet igen og angiv i hvor høj grad du synes han taler rigsdansk/standarddansk



Rigsdansk

14. Region

Sæt kryds ud fra hvor du tror manden kommer fra

- Nordjylland
- Vestjylland
- Østjylland
- Sønderjylland
- Fyn (og omkringliggende øer)
- Sjælland
- Lolland-Falster
- Bornholm
- Andet:

15. Land eller by?

Tror du manden bor på landet eller i byen?

- Byen
- Landet

Næste

Denne del blev gentaget for i alt 4 stemmer.

Del 3 – baggrundsplysninger om deltageren

25. Baggrundsplysninger

Vi vil gerne vide lidt om dig til brug for vores analyser af de svar vi indsamler. Alle dine svar opbevares på en sikker server og komplet anonymiserede.

Hvor gammel er du?

Sæt kryds

Kvinde

Mand

Andet:

Ønsker ikke at svare

26. Hvor er du født?

Sæt kryds

Nordjylland

Vestjylland

Østjylland

Sønderjylland

Fyn (og øer)

Nordsjælland (inkl. København)

Sydsjælland

Lolland-Falster

Bornholm

Andet:

27. Hvor bor du?

Sæt kryds

Nordjylland

Vestjylland

Østjylland

Sønderjylland

Fyn (og øer)

Nordsjælland (inkl. København)

Sydsjælland

Lolland-Falster

Bornholm

Andet:

28. Baggrundsplysninger

Vi vil gerne vide lidt om dig til brug for vores analyser af de svar vi indsamler. Alle dine svar opbevares på en sikker server og komplet anonymiserede.

Hvad hedder den by,
du bor i?

29. Hvilken uddannelse har du senest afsluttet?

Sæt kryds

Grundskole

Gymnasieuddannelse

Erhvervsuddannelse

Kort videregående

Mellemlang videregående

Lang videregående

Ved ikke

30. Beskæftigelse

Hvad er dit arbejde?

31. Hvad tror du undersøgelsen handler om?

Her vil vi gerne høre dit bud på hvad formålet med undersøgelsen er. Var der noget bestemt du lagde mærke til i de lydclip, du hørte, som du tror er vigtigt for undersøgelsen?

***Eftertid*. En grammatikterms historie og historiens relevans for nyere dansk verbalteori**

Viggo Bank Jensen

Københavns Universitet

Peter Skautrup (1953:431) giver Rasmus Rask æren for mere end 50 danske grammatiske termer, herunder bl.a. *Eftertid*. Paul Diderichsen (1960:92 ff.) gør opmærksom på at Rask i mange af de tilfælde som Skautrup nævner, har overtaget betegnelserne fra andre, hvilket svarer til hvad Rask selv understregede; det drejer sig fx om ”nutid” og ”fremtid”. Diderichsen fremhæver imidlertid at Rask med rette skal have æren når det gælder termerne for ’måderne’, i al fald ”fremstættende” og ”betingende” måde. Diderichsen nævner ikke *Eftertid*, en term som interessant nok viser sig at have en lidt speciel placering mellem tempus og modus, hvilket man for eksempel kan se ved at følge Rasks diskussion af modusindplaceringen i hans første introduktion af termen. I *Ordbog over det danske Sprog (ODS)*, bind 4, spalte 194 (1922) beskrives termen således:

Eftertid: [...] (dannet af Rask; grammatisk betegnelse, lidet brugt) den af verbets tider, der betegner noget i forhold til fortid (præteritum) fremtidigt: fortids (ell. datids) fremtid (futurum in præterito ell. 1. konditionalis), fx “jeg vilde (ell. skulde) elske”. Rask.FS.Fortalen 32. Mikkels.Sprogl. 203.¹

Kristian Mikkelsens *Dansk Sproglære*, som der henvises til, er fra 1894; *FS* er Rasks frisiske sproglære fra 1825 hvor han kort skriver (s. XXXII) om *Eftertid*: ”futurum perfectum v. conditionale” og senere giver et kort eksempel (s. 87): ”skolde stéra (skulde styre)”. Som det ses, er der i *ODS*-opslaget i en parentes en noget modificerende parafrase af Rasks tekst: ”futurum perfectum” (der normalt betyder futurum exactum/førfremtid) er blevet erstattet af ”futurum in præterito”; der er derudover blevet indført et ”1.” ved konditionalis, en praksis der vist først for alvor begynder at vinde hævd i Danmark i tyskgrammatikker efter Rasks død i 1832.

¹ Jeg har udskrevet de forkortelser der ikke er umiddelbart forståelige.

Ud fra opslaget i *ODS* kunne man fristes til at tro at termen blev introduceret via beskrivelse af frisisk, altså et med dansk nært beslægtet sprog, Rask havde imidlertid allerede indført termen i 1824 i forbindelse med analysen af et romansk sprog, i sin spanske sproglære, for den verbalform der typisk omtales som konditionalis (*el condicional*). Rasks grundigste og mest diskuterende behandlinger af *Eftertid* finder man i den spanske sproglære og i Rask (1827) om italiensk. Rask publicerede aldrig en analyse af dansk hvor han bruger betegnelsen *Eftertid*. Som det vil fremgå af afsnit 2 om Rask, kan man dog godt tale om at han også introducerer termen i forhold til dansk grammatik.

Kristian Mikkelsen (1894) er måske den eneste der benytter tempustermen *Eftertid* i en publiceret, systematisk analyse af dansk grammatik; han har også en lang række eksempler. I 1895 udsendes et cirkulære om anbefalede danske grammatiske betegnelser til brug i almueskoler og på seminarier (Diderichsen 1961:60), og her er *Eftertid* ikke med, men i stedet ”fortids fremtid”. *Eftertid* indgår i diskussionen mellem Wiwel (1901:197) og Mikkelsen (1902:43). Hansen og Heltoft (2011:721) inddrager Wiwels ’afstandsbegreb’ i deres analyse, men har på linje med Mikkelsen otte tempora, herunder ”futurum in præterito”.

Inden vi når frem til tesen i min artikel, er her en kort introduktion til spansk konditionalis med udgangspunkt i en dansk universitetsgrammatik i spansk:²

1) Konditionalis’ grundbetydning er at angive at en situation er fremtidig i forhold til en situation i fortiden: fortids-fremtid:

Belén le contó que iría a estudiar a otro país =

Belén fortalte ham at hun ville tage til et andet land for at læse

2) Konditionalis om en forestillet situation. I en hovedsætning kan konditionalis angive en forestillet situation i ytringssituationen. Hvis hovedsætningen har en hypotetisk ledsætning knyttet til sig, står denne oftest i imperfektum konjunktiv [...]:

Si fuera poeta, yo te escribiría un poema =

Hvis jeg var digter, ville jeg skrive et digt til dig

Eksempel 2 viser at man i hypotetiske konstruktioner kan tale om en form for arbejdsdeling mellem konjunktiv (i protasis, i dansk grammatik ofte kaldt betingelsessætning³) og konditionalis (i apodosis, her er det sværere at pege på en typisk betegnelse i dansk grammatik⁴). Jeg

² Dam Jensen & Dam (2018): 241-42.

³ Fx i Hansen og Heltoft (2011: 686, 1552).

⁴ Mikkelsen (1894: 380) taler om den ”betingede (overordnede) sætning”.

vælger at bruge betegnelserne protasis og apodosis (ofte brugt i romansksprogede grammatikker) når jeg taler om fænomenerne generelt uafhængigt af Rask eller andre grammatikere. Selve ordet 'konditionalis'⁵ er noget tvetydigt i forhold til om man mener betingende eller betinget, engelsk 'conditional' kan således også oversættes til begge dele. Det er meget vigtigt for Rask at undersøge forskellen mellem disse to betydninger. Rask (1824, 1827) præciserer således forskellen mellem den betingende sætning (altså protasis) hvor man har den betingende måde (konjunktiv), og den betingede sætning (altså apodosis) hvor man har den betingede verbalform (eftertid). Denne præcise skelnen mellem betingende og betinget var ikke spor oplagt, hverken i Rasks værker indtil den spanske sproglære i 1824 eller i den grammatiske beskrivelse af dansk i den pågældende periode.

Tese: Det er min tese at Rask udvikler sit *Eftertids*begreb i forbindelse med analyse af romanske sprog og herved når frem til en opfattelse af fænomenet i dansk der på flere punkter er nyskabende i forhold til hans forgængere blandt danske grammatikere, og at han herved foregriber vigtige træk i Mikkelsens fremstilling vedrørende *Eftertid* og i Hansen og Heltofts analyse af "futurum in præterito". Undersøgelsen af denne tese udgør kernen i min artikel.

I det følgende vil jeg: (1) redegøre for nogle fremstillinger af fænomenet i dansk grammatikhistorie før Rask; (2) analysere Rasks behandling af *Eftertid* fra den første publicering i den spanske sproglære i 1824 frem til Rasks korte men prægnante analyse af fænomenet i dansk i *Grammar of the Danish Language for the Use of Englishmen* (1830); (3) redegøre for *Eftertids* rolle i Mikkelsen (1894) og påpege parallellerne (og de vigtigste forskelle) i forhold til Rasks fremstilling; (4) undersøge relevansen af Rasks *Eftertids*begreb i grammatisk analyse i dag ved inddragelse af Hansen og Heltoft (2011). Rask indfører også en *Föreftertid*. For at fastholde den historiske linje i min tekst fokuserer jeg på *Eftertid* og behandler kun sporadisk *För*-formen.

1. Forudsætningerne i forhold til Rasks behandling af *Eftertid*

Jens Høysgaard (1747) indfører et *futurum secundum*: "Futurum secundum er (a) altid sammensat af Imperfectum skulde eller vilde og af Præsens Infinitivi til det rette Verbum", fx skulde eller vilde elske. "(b) Dette Futurum seer tilbage paa det forbigangne, og betegner uden vilkaar,

⁵ Om 'condizionale' på italiensk: En 'proposizione subordinata condizionale' (underordnet betingende sætning) er den del af en betingelseskonstruktion, *protasi*, hvor man typisk bruger konjunktiv ('congiuntivo', se fx Dardano & Trifone 1997:412) og altså netop ikke verbalformen 'condizionale' (der bruges i hovedsætningen, *apodosi*). Brugen af selve ordet 'condizionale' er således ret forvirrende.

at noget vist skulde skee, førend i den nærværende tid, efter noget, som var tilforn, saasom: Om mandagen snakkede jeg med ham i Kjøbenhavn, dagen efter skulde han rejse til Falster” (Høysgaard 1747:68). Efterfølgende har Høysgaard en længere argumentation for at formen er indikativisk fordi den udtrykker noget ”for vist”. *Futurum secundum* bruges også når direkte tale omskrives til indirekte tale og *futurum primum* (*vil det gaae dem ilde*) ændres til (*vilde det gaae dem ilde*) (ibid:74). Høysgaard behandler forholdet mellem *Antecedens* og *Conseqvens* (ibid:152 f.) og er opmærksom på de relevante konjunktioner i den førstnævnte, men diskuterer ikke verbalformen i den sidstnævnte. Han gør opmærksom på at *Conseqvens* godt kan stå før *Antecedens* og dokumenterer det med velvalgte eksempler (dog ikke med *futurum secundum*). Vi kan kalde hans opfattelse logisk-grammatisk. Høysgaards *futurum secundum* kan meget vel have været en vis inspiration for Rask.

Høysgaards *futurum secundum* får en hård medfart af Jacob Baden, den toneangivende grammatiker i Danmark i slutningen af 1700-tallet. Baden (1785:219) bruger 25 linjer på at afvise behovet for et nyt futurum à la Høysgaards. Jeg undlader at referere Badens lange argumentation, der går via latin, men indlægget illustrerer at periodens grammatikere på ingen måde er enige om relevansen af at inkludere et sådant tempus i det danske verbalsystem.

Matthias Bøgh (1800:18 ff.) opererer med otte tempora i indikativ, heraf fire futurale, bl.a. ”betinget futurum” (*skulde elske*) og ”betinget futurum præteritum” (*skulde have elsket*). Som flere af sine samtidige prøver Bøgh (1800:47 ff.) at bevare konjunktiv i dansk, og han skriver således:

Det danske Sprog mangler ikke Konjunktiv, skjønt denne Modus ikke er forskjellig fra Indikativ [...] I konjunktiviske Sætninger bruges ofte Imperfektum for Præsens; saasom: vidste han det, saa kom han strax.

Efterfølgende omtaler han også brug af omskrevne former med ’vilde’/’skulde’, men uden at præcisere hvor de bruges i betingelseskonstruktionen. Hans analyse er muligvis påvirket af tysk (jf. Diderichsen 1960:93), hvor der kan være konjunktiv i både den betingende og den betingede sætning.

Niels Lang Nissen (1808) holder sig i verbalparadigmerne til de seks syntetiske latinske tempora, og behandler derfor ikke *Eftertid*. Til gengæld giver han præcise analyser af *betingende sætninger*, hvilket i flg. Paul Diderichsen (1960:95) har banet vejen for Rasks *betingende Måde* som betegnelse for konjunktiv. Ét sted bruger Nissen (1808:214) betegnelsen *betingende sætninger* når han skriver: ”Dog står verbet først: [...] naar i betingende sætninger konjunktio-

nen udelades, f.Ex: “vil du gjøre det, saa skal du have tak””, og i en paragraf om *betingende* konjunktioner har han fire præcist formulerede betingelseskonstruktioner, dog ingen med *skulde, vilde* + infinitiv i apodosis og alle med protasis først (Nissen 1808:194). Andetsteds skriver Nissen (1808:47):

Betingede Sætninger, der kunde betragtes som indbyrdes afhængige, hinanden underordnede, udtrykkes i Dansken ved Imperfecter i uvis og næsten præsentisk Betydning, f.Ex: “Hvis du var klog, fulgte du mit råd”.

To kommentarer til dette eksempel: 1) her får Nissen ødelagt sin systematik ved at gøre begge sætninger ’betingede’, altså også betingelsessætningen (protasis). I citatet ovenfor (1808:214) kalder han den tilsvarende ledsætning for *betingende*. Det er blandt andet uklarer af denne type som Rask sætter sig for at afklare i de to romanske grammatikker; 2) eksemplet viser også hvorfor det er svært at finde *Eftertid* på dansk. Nissen vælger jo et velformet eksempel fra dansk – med *imperfectum* i begge dele af konstruktionen⁶ - og illustrerer således at det ikke nødvendigvis er oplagt at vælge *vilde, skulde* + infinitiv i apodosis, samtidig med at der ikke er en konjunktivform man kan bruge som kendetegn i protasis. Nissen (1808:226 ff.) bruger begreberne *Forsætning* og *Eftersætning* på linje med Bloch (se nedenfor).

Søren Bloch (1817:206) har ”Imperfectfuturet” med i sin paradigmatiske opstilling, med *vilde, skulde elske*, og han har en ganske præcis paragraf om denne temporale side af *Eftertid*. Det er værd at bemærke at hans analyse ikke tager udgangspunkt i formen, men i funktionen, og at han også inddrager brugen af *Imperfectum*:

Fortids-Fremtiden, Imperfectfuturet, naar Fremtiden eller den tilkommende Handling tænkes som følgende efter en forbigangen Handling eller Tilstand. Dette Futurs form er *enten* dannet ved det imperfecte Hjelpeord *vilde* og *skulde* med Verbets Infinitiv; f. Ex. jeg *haabede* [...], at det *vilde blive forandret* [...] *eller* eens med Imperfectet. f. Ex. jeg *troede*, at han *blev* Student om to Aar (Bloch 1817:192).

Bloch (1817:41-42) skriver andetsteds om den ”toleddede periode” med *Forsætningen* (der kendes på, ”at der foran Samme enten staar eller kan staa et Betingelsesord, saasom: naar, da, efterdi, efterat, imedens, saalængesom, hvis, ifald, dersom, om endskjönt o.a.”) og *Eftersæt-*

⁶ På italiensk eksisterer en nogenlunde tilsvarende konstruktion, men den regnes ikke for en velformet sætning af grammatikerne, så det er ikke den de tager udgangspunkt i (se yderligere kommentar i afslutningsafsnittet).

ningen (der kendes på ”et foranstaaende eller forantænkt saa eller da”), og bringer eksemplet: ”Snakkede man stedse med Overlæg, vilde man ey saa ofte fortryde sine Ord (*dersom* man stedse snakkede med Overlæg, *da* vilde man o.s.v.)” (fremhævelserne er Blochs). Han skriver også

[...] at den første eller *Forsætningen* (antecedens [...]) fastsætter den Betingelse eller Omstændighed, under hvilken den sidste eller *Eftersætningen* (consequens [...]) bekræftes eller nægtes (Bloch 1817:41).

I modsætning til Høysgaards logisk-grammatiske opfattelse (se ovenfor) fremgår det af Bloch-citaterne ovenfor og af Blochs konkrete eksempler i bogen at han har en topologisk-grammatisk opfattelse som vedrørende *Eftersætningen* svarer til ODS's 2. definition: *hovedsætning* (*overordnet sætning*), som følger efter sin *bisætning* (ODS, bind 4, spalte 190). Dermed mangler han et begreb for alle de steder hvor apodosis står først og dermed et begreb der kan sammenfatte de steder hvor man møder *Eftertid* i betingelseskonstruktioner. Nissen har samme opfattelse som Bloch af antecedent, forsætning osv., og Nissens *betingende sætning* er derfor altid placeret i *Forsætningen* og har dermed et uhensigtsmæssigt topologisk træk på linje med problemerne ved Blochs (og Nissens) begreb om *Eftersætning*. Da Nissen og Bloch er to af Rasks vigtigste forudsætninger, er det muligt at man kan se Rasks analyse som en implicit diskussion af anvendelsen af deres *Eftersætnings*-begreb.

2. Rask og *Eftertid*

I Rasks værker om islandsk (1811), om angelsaksisk (1817) samt prisskriftet (1818) behandles *Eftertid* ikke, heller ikke under andre betegnelser. Rask (1811) kalder konjunktiv for ”betingende måde”. Dermed får han frigjort betegnelsen *betingende* fra det uhensigtsmæssige topologiske træk det havde i Nissens anvendelse som *betingende sætning* (se afsnit 1). Rask vakler dog i denne periode lidt mellem ”betingende måde” (1811, 1818) og ”betinget måde” (1817). I 1823 kommer han hjem fra sin mere end 6 år lange rejse, hvor han bl.a. har været i Indien. På tilbagerejsen til Europa afslutter han sine udkast til den franske, italienske og spanske verbalteori (Petersen 1870:307). Der er store forventninger om at han skal udgive noget om sanskrit, men til skuffelse for mange udgiver han i 1824 i stedet en spansk sproglære, der ikke får så megen bevågenhed. Det er her han introducerer termen *Eftertid*, der fx ikke findes i terminologilisten i Rask (1817).

Den ovenfor omtalte spanskgrammatik (Dam Jensen og Dam 2018) indplacerer konditionalis under indikativ. Det er stadig i vore dage et omdiskuteret punkt om konditionalis kan indplaceres under indikativ, en diskussion der vedrører spansk, italiensk og andre romanske sprog. På Rasks tid var det en nærmest ikke-eksisterende idé for spansk og italiensk, og ofte valgte grammatikerne at indplacere konditionalis under konjunktiv, sandsynligvis fordi den nye romanske form, konditionalis, havde overtaget nogle af de funktioner som konjunktiv havde i latin. Rask (1824:XXII) bringer en liste over grammatiktermer, og her skriver han at formen *Eftertiden/betinget Fremtid* svarer til formen ”*prétérito imperfecto de subjuntivo*” i det spanske akademis sproglære på Rasks tid. I akademiets sproglære placeres formen altså under konjunktiv, mens Rask indplacerer den under ”fremsettende måde” (indikativ).

Rask definerer den således: *Eftertiden* henføres ”til en anden Tidspunkt helst forbigangen, som om man vilde sige *Tiden derefter*” (Rask 1824:22). I den italienske formlære (1827:29, jf. også 1824:91) præciserer han betegnelsen: Han skriver således at ”Fremtid” og ”Eftertid” egentlig mere præcist burde kaldes ”Efternutid” og ”Efterdatid” da de forholder sig på samme måde til hinanden som ”Nutid” og ”Datid”, men holder dog fast i at kalde dem ”Fremtid” og ”Eftertid”. Men hvorfor vælger Rask at indplacere *Eftertid* (konditionalis) under indikativ? Her skriver han vedrørende spansk:

At skjelne den fremsættende Måde fra den betingende [konjunktiv, vbj] er ikke så vanskeligt, når man først klart adskiller den Sætning, der udtrykker en Betingelse fra den, som derved betinges, den første er *betingende* den sidste *fremsettende*, skönt den ofte kan være betinget [...] Den betingede Sætning står således i fremsættende Måde, og på denne Grund tror jeg med rette at have henregnet *Eftertid* [...], eller om man heller vil: den betingede Fremtid [...], til fremsættende Måde; f.E. [...] *yo querría ir á Sevilla*, jeg kunde have Lyst at gå til Sevilla (ikke: *yo quisese ir á Sevilla* (Rask 1824:171).⁷

Han skelner altså præcist mellem den *betingende* sætning i konjunktiv og den *betingede* sætning i indikativformen *Eftertid*. Efterfølgende har Rask (1824:171-72) et mere komplekst eksempel med en *hvis*-sætning i konjunktiv som protasis og en overordnet sætning i *Eftertid* som apodosis. Om italiensk bemærker Rask (1827:62) at *Eftertiden* ”svare aldeles til vore Omskrivninger med [...] *skulde, vilde*⁸ [om] hvad der vilde skje i Eftertiden (dvs. efter Datiden, derefter)”. Rask gør altså opmærksom på hvordan *Eftertid* i betydningen ”fortids fremtid” udtrykkes på

⁷ *querría* er ’condicional’, mens *quisese* er konjunktiv.

⁸ I Rask (1824: 91) nævnes kun *skulde*.

dansk. Rask har imidlertid hverken i den spanske sproglære eller den italienske formlære et typisk eksempel på brug af de to sprogs konditionalis som rent temporal ”fortids fremtid”.⁹

Rask (1827:62) behandler også betingelseskonstruktioner i den italienske formlære: ”Eftertid bruges også i betingede sætninger, der afhænge af betingende”, med bl.a. dette eksempel: ”*che sarebbe di me, s'io non potessi vederla piu!* – hvad vilde der blive af mig, dersom jeg ikke kunde se hende mere!” (Rasks egen oversættelse). ”Man kunde kalde dette en *betinget* Fremtid, når man blot ikke forblande det med *betingende* [konjunktiv, vbj] Måde; men man ser, at den også i disse Tilfælde svarer til vor Omskrivning med *vilde, skulde*.”

Jeg ved ikke om det er tilsigtet, men med forstavelsen ”Efter-” får Rask konstrueret en term der kan referere såvel til den temporale betydning ”fortids fremtid” som til den logiske betydning, dvs. betinget, altså logisk efter betingelsen.

På baggrund af Rasks to romanske grammatikker kan man sige at han har givet elementer til en forståelse af hvordan *Eftertid* kunne indgå i en dansk grammatik. Formmæssigt skriver han at der er tale om omskrivninger med præteritumsformerne *vilde* og *skulde* (*vilde* er dog først med fra 1827), funktionsmæssigt peger han på anvendelsen af *vilde, skulde*-omskrivningerne såvel rent temporalt (”fortids fremtid”) som i apodosis i betingelseskonstruktioner. Men har Rask så faktisk prøvet direkte at beskrive *Eftertid* i dansk? Ja, i al fald to steder.

I Rask (1825ms, med sidepaginering 28), et manuskript om dansk grammatik, optræder *Eftertid* ved siden af *Fremtid* i en paradigmatiske opstilling, med eksemplet *skulde, vilde høre*. Manuskriptet er sandsynligvis udarbejdet i forbindelse med undervisning af Rasmus Rasks fætter (Bjerrum 1959:137-38, 269). Dele af det danner basis for Rask (1830) og Rask (1837), men betegnelsen *Eftertid* og den tilhørende paradigmatiske opstilling er ikke med i disse to udgivelser.

Rask (1830:59) behandler imidlertid kort eksempler på *brug* af *Eftertids* former. Her omtales brugen af *skulde* og *vilde* til at angive ”a futurity relative to some other time; they are prefixed to the infinitive” med eksemplet ”*Han sagde jeg skulle komme, he said (that) I should come o: told me to come*”. Rask tilføjer efterfølgende: ”The past of these auxiliaries also expresses the conditional future in French; f.i. *jeg skulde nok skrive, hvis jeg havde noget at skrive om, I would write (to him) indeed, if I had any thing to write about*”. Rask angiver som nævnt ikke nogen term for det tempus han analyserer. I broren H.K. Rasks danske oversættelse (1837) er det pågældende afsnit udeladt, og *Eftertid* optræder ikke i bogens liste over Rasks termer.

⁹ For Rasks analyse af verbalsystemet i spansk (Kuhlmann Madsen 1994) og i italiensk (Bank Jensen: 2017).

Som det ses, gør Rask her ligesom i analysen af de romanske sprogs konditionalis/*Eftertid*: Også i analysen af dansk tager han udgangspunkt i *verbalformen* (uden dog at give den omskrevne form et navn) og giver først et eksempel på ”fortids fremtid” og umiddelbart efter et eksempel på brug i betingelseskonstruktion (med henvisning til fransk konditionalis). Ved at tage udgangspunkt i verbalformen i stedet for i indholdet og ved at henvise til et romansk sprog opnår han altså en forståelse der forener de to anvendelser som optræder adskilt hos Bloch. Henvisningen til fransk er en effektiv og kortfattet måde at introducere det danske system på. For nøjere at kende Rasks begrebsapparat og hans overvejelser bag henvisningen til fransk konditionalis er man dog nødt til at gå til hans analyser vedrørende betingende og betingede sætninger i den spanske og den italienske grammatik. Under alle omstændigheder bekræfter henvisningen til fransk konditionalis at Rask medtænker de romanske sprog, også når han skriver om disse fænomener i dansk.¹⁰

Delkonklusion 1: Som det fremgår af min oversigt i afsnit 1, er det kun få af Rasks forgængere (Høysgaard, Bøgh) der analyserer et futurum in præterito-tempus med udgangspunkt i en bestemt verbalform; Høysgaard analyserer ikke brugen af denne verbalform i hypotetiske konstruktioner, og Bøgh får ikke differentieret verbalanvendelserne i protasis og adoposis. Hverken Baden eller Nissen inkluderer et futurum in præterito. Søren Bloch har en gedigen analyse af temporal ”fortids fremtid”; den er dog helt adskilt fra undersøgelsen af betingelseskonstruktioner, og denne er præget af et utilstrækkeligt begrebsapparat. Rask tilføjer: 1) en ny term der kan bruges som overbegreb for de to diskuterede funktioner; 2) et adækvat begrebsapparat med den præcise skelnen mellem betingende og betingede sætninger hvor de sidstnævnte udgør de steder i teksten hvor man ofte finder *Eftertid* når det ikke er temporalt; 3) en koncentreret præsentation vedrørende dansk af *Eftertids* to centrale funktioner: a) fortids fremtid, b) betinget fremtid i betingelseskonstruktioner (Rask 1830). Der er to vigtige forudsætninger for Rasks tre resultater: 1) Han tager udgangspunkt i en veldefineret arbejdsdeling i spansk og italiensk mellem konjunktiv og konditionalis i (velformulerede) hypotetiske konstruktioner; 2) han fokuserer i dansk på verbalformen med *skulde*, *vilde* + infinitiv og ser bort fra den udbredte anvendelse af alternativet præteritum.

Man kunne opstille Rasks begreber på denne måde:

¹⁰ Rask benytter termen *Eftertid* i sin engelske grammatik og henviser også her til fransk: ”Ligeledes bruges *should* og *would* med Navneform, til at udtrykke vor Omskrivning med *skulde*, *vilde* eller den franske betingede Fremtid (*Future conditionel*) hvad vi kunde kalde *Eftertiden*” (Rask 1832a: 102).

Figur 1	Rasks udspecificeringer	En mulig fortolkning
Eftertid	Efter-datid	Tidsligt efter
	Betinget Fremtid	Logisk efter

Med *Eftertid* og *betinget sætning* får Rask flyttet forgængernes *Efter* i *Eftersætning* fra et sætningsbegreb der her er utilstrækkeligt, til en mere fleksibel term for en verbalform.

3. Kristian Mikkelsen og *Eftertid* i dansk grammatik

Det er vist kun i Kristian Mikkelsens *Dansk Sproglære* (1894) at tempusbetegnelsen *Eftertid* bliver brugt på en systematisk og omfattende måde vedrørende dansk grammatik.¹¹ Mikkelsen har to større afsnit om verberne ('udsagnsordene' hos Mikkelsen): det ene er et underafsnit under "Ordene og deres former" (s. 200-232), det andet er et underafsnit (s. 349-383) under "Ordføjningen", først og fremmest om *brugen* af udsagnsordenes former. På de allersidste 4 sider af det andet verbalafsnit (s. 379-383) anfører han at indikativformerne kan bruges modalt, i delafsnittet "En blot tænkt Forestilling, udtrykt ved en forskydning af Tiden". Mikkelsens rækkefølge i fremstillingen er i høj grad med til at fremme en opfattelse af at tempus i dansk er en klar temporal bøjningskategori på linje med latin.

Mikkelsen opstiller dette bøjningsmønster for fremsættende måde, aktiv ("handleform"), her i min forenkede fremstilling:

Fremsættende måde:	
Nutid	læser
Datid	læste
Förnutid	har læst
Fördatid	havde læst
Fremtid	skal læse – vil læse

¹¹ Mikkelsen (1902: 43) henviser til at han har optaget *Eftertid* og *Föreftertid* i sit tempusskema fra Wimmers store oldnordiske Formlære § 159 og § 160 (Wimmer 1870: 79, 126ff). Rask (1832b) havde selv vedrørende oldnordisk indført *Eftertid* i paradigmerne for omskrevne verbalformer i *Kortfattet Vejledning til det oldnordiske eller gamle islandske Sprog*, til forskel fra situationen i sit ungdomsværk (1811). Der er ingen teksteksempler.

Eftertid	skulde læse – vilde læse
Förfremtid	skal have læst – vil have læst – faar læst
Förefertid	skulde have læst – vilde have læst – fik læst

Som det ses, bruger Mikkelsen betegnelserne *Eftertid* (og *Förefertid*). Mikkelsen introducerer *Fremtid* og *Eftertid* sammen og forklarer *Eftertid* på følgende måde:

Til at betegne, hvad der skal henføres til de efter [Datid] liggende Tidspunkter, bruges Hjælpeudsagnsordene *skulle* og *ville* i Forbindelse med Nutids Navnemaade. [Denne Tid] kaldes: [...] **Eftertid** eller Datids Fremtid (Futurum in præterito): *jeg lovede, at jeg **skulde komme** den næste Dag; jeg formodede, at han **vilde komme** den næste Dag* (Mikkelsen 1894:203, fremhævelserne og kursiveringerne er Mikkelsens).

I *Eftertid* ”bruges som Hjælpeudsagnsord *ville*, der oprindeligt betegner en Vilje, og *skulle*, der oprindeligt betegner en Pligt” (Mikkelsen 1894:372, kursiveringerne er Mikkelsens). Mikkelsen skriver yderligere i sit afsnit om tidernes brug:

[...] **Eftertid** bruges om hvad der er tilkommende [...] i Forhold til [...] det forbigangne, f. Eks. [...] *Jeg troede, at det vilde være for sent den næste Dag. Da han havde sagt, at han skulde fortælle mig det en anden Gang, gik han hen til det øvrige Selskab. Jeg har altid været bange for, at det skulde gå ham galt* (Mikkelsen 1894:376, fremhævelserne og kursiveringerne er Mikkelsens).

Et par sider længere fremme skriver han:

Brugen af den forestillende Maade [altså konjunktiv, vbj] er i Dansk meget indskrænket, da den ikke alene mangler en særlig Datidsform, men ogsaa saa godt som udelukkende forekommer i Hovedsætninger [...]. Imidlertid har Sproget et andet Middel [end forestillende måde] til at udtrykke en blot tænkt Forestilling, idet en saadan ofte betegnes ved en Forskydning af Tiderne, saaledes at: [...] **Eftertid** bruges i Steden for **Fremtid**, f. Eks. *Han taler, som om han skulde dø i Morgen;* (Mikkelsen 1894:379, fremhævelserne og kursiveringerne er Mikkelsens).

Dette uddybes af Mikkelsen, der skriver at forskydningen finder sted ved *antagelse mod virkeligheden*, herunder

i *Betingelsesudtryk* saavel i den betingede (overordnede) Sætning som i den betingende (underordnede) Sætning, f. Eks. *Hvis han var rask, blev han ikke borte* [...] Da det, der er betinget, kan tænkes at være tilkommende i Forhold til Betingelsen, bruges ofte Eftertid og Föreftertid i den betingede Sætning i steden for Datid og Fördatid, f. Eks. *Hvis han var rask, vilde han nu være her. Hvis han havde vidst det, vilde han have sagt det* (Mikkelsen 1894:380 f., kursiveringerne er Mikkelsens).

Delkonklusion 2: Som det ses af det foregående, er Mikkelsen i forhold til *Eftertid* på linje med Rask med hensyn til: 1) tempusterm, 2) valgte hjælpeudsagnsord, 3) temporal anvendelse og anvendelse i betingelsesudtryk, herunder også præciseringen af forskellen mellem betingende og betinget sætning (Mikkelsen nævner også anvendelsen til at angive høflighed, beskedenhed). Det sidstnævnte citat ovenfor viser dog også at Mikkelsen gør opmærksom på at datid/præteritum lige så vel som *Eftertid* kan tage pladsen i den betingede sætning, noget som Rask ikke går ind på.

4. Eftertid og Hansen og Heltofts 'eftertid' i *Grammatik over det danske sprog* (2011)

Tempusbetegnelsen *Eftertid* optræder ikke som sådan hos Hansen og Heltoft. De to forfattere giver imidlertid Mikkelsen medhold i ideen om otte tempora i dansk, herunder et tempus med betegnelsen 'futurum in præterito' med *vilde* som hjælpeverbum. Samtidig påpeger de i forlængelse af Wiwel (1901) at der i dansk er klare tilfælde hvor præteritumsformen ikke udtrykker noget fortidigt, men etablerer forestillede verdener af forskellig slags (647-48), og at præteritum således ikke blot kan forstås som tidslig afstand, men også som modal afstand i realitet, non-realis (683-684). Hansen og Heltoft understreger at dette afstandsbegreb bør inddrages ved opstilling af systemet.

Vedrørende 'futurum in præterito' har forfatterne to mindre afsnit: et om tidslig 'futurum in præterito' (703-04) og et om 'modal futurum in præterito' (708-09); det sidstnævnte drejer sig primært om betingelseskonstruktioner. Vi kan således genfinde den type systematisering Rask introducerede først via sine grundige romanske grammatikker og siden koncentreret i den korte danske.

Hansen og Heltoft (2011:682) benytter som sagt ikke tempusbetegnelsen *Eftertid*, men bruger udtrykket 'eftertid' som modsætning til 'neutral' til at indholdsbeskrive kategori-mod-

sætningsparret 'futurum vs. ikke-futurum'. Lidt længere fremme (s. 685-86) påpeger de at futurums hjælpeverb *ville* har betydningen 'følge' enten i den tidslige betydning 'følge efter', 'eftertid' – eller i betydningen 'følge af', 'logisk følge'. Selv om Hansen og Heltoft giver ordet 'eftertid' en lidt anden plads i deres systembeskrivelse end Rask, så har deres overvejelser lighedspunkter med Rasks betragtninger over *Eftertids* dobbeltnatur: som fremtid i forhold til noget datidigt eller som noget 'betinget' fremtidigt i betingelseskonstruktioner (jf. min figur 1).

Hermed ikke sagt at de to analyser af dansk er sammenfaldende. Rask tager i en vis forstand verbalformen for givet og abstraherer fra brugen af præteritum som alternativ til *Eftertid*. Hansen og Heltoft tager i deres analyse udgangspunkt i Wiwels "afstandsbegreb", hvor den omfattende anvendelse af den syntetiske form præteritum er et centralt omdrejningspunkt. Det betyder dels at de selvfølgelig ikke i den samlede analyse kan abstrahere fra brugen af syntetisk præteritum på den plads hvor futurum in præterito optræder, dels at de føler sig forpligtet til at argumentere grundigt for at *ville* er røget ud af "modalverbssystemets centrum" og i moderne dansk er "flyttet til tempussystemet som futurisk hjælpeverb" (Hansen og Heltoft 2011: 796).

5. Konklusion og afsluttende overvejelser

Med udgangspunkt i en diskuterende analyse af den veldefinerede arbejdsdeling i spansk og italiensk mellem konjunktivformerne og konditionalisformerne i (velformulerede) hypotetiske konstruktioner når Rask frem til en skarp analyse af forholdet mellem betingende og betingede sætninger og får dermed udpeget stedet hvor man finder *Eftertid* i hypotetiske konstruktioner. Denne analyse finder man hos Rask kun i de to grammatikker om romanske sprog, men et koncentrat af resultatet af analysen overføres til dansk (og engelsk) når han henviser til fransk betinget fremtid (*Future conditionnel*). Rask vælger på dansk, ligesom på spansk og italiensk, at lade analysen tage udgangspunkt i *verbalformen* (*skulde, vilde* + infinitiv), og det hjælper ham til i en syntese at abstrahere fra alternative måder at udtrykke apodosis på og til at se fællesskabet mellem temporal brug som fortids fremtid og modal brug i apodosis. Hans term *Eftertid* kan opfattes som et forsøg på at sammenfatte betydningerne i de to anvendelser.

Mikkelsen er på linje med Rask med hensyn til termen *Eftertid*, de valgte hjælpeudsagnsord, skellet mellem betingende og betingede sætninger samt de to anvendelser af *Eftertid* som henholdsvis "fortids fremtid" og apodosis i betingelsesudtryk. I kontrast til Rask behandler Mikkelsen dog de to anvendelser helt adskilt. Det hænger sammen med at Mikkelsen vedrørende rollen i apodosis primært tager udgangspunkt i funktionen og ikke i verbalformen og derfor bringer lige så mange eksempler med datid som med *Eftertid*.

Hansen og Heltoft accepterer kun *ville* som hjælpeverbum i ”futurum in preterito”, men har vedrørende anvendelserne en todelt struktur der svarer til Rasks korte indslag i den danske grammatik. De har desuden en brug af ordet ’eftertid’ der ligesom Rasks tempusterm *Eftertid* synes at kunne sammenfatte den tidslige og den logiske følge. Det ses således at Rasks begreb om *Eftertid* stadig kan have relevans i grammatisk analyse.

På flere romanske sprog, jeg tager italiensk som eksempel, kan man udtrykke betingelseskonstruktionen ved (modal) brug af *imperfetto* i både den betingende og den betingede sætning. De italienske grammatikere har dog generelt været ret enige om at det er en mere uformel, nogle vil kalde den ukorrekt, sprogbrug, og derfor er det ukontroversielt at Rask ser bort fra denne sprogbrug når han analyserer italiensk. På dansk er grammatikerne enige om at både præteritum og futurum in præterito er gangbare i apodosis. Det kunne rejse spørgsmålet: Er det først og fremmest en stilistisk eller en indholdsmæssig markør når man på dansk vælger mellem præteritum og futurum in præterito i apodosis?

Rasks introduktion af termen *Eftertid* i forhold til italiensk (og spansk) giver mulighed for at tolke formen både som tidsligt efter og som logisk følge og dermed fange at formen hører hjemme i apodosis i betingelseskonstruktioner, et element jeg har kunnet bruge i undervisning af italienskstuderende som supplement til den noget tvetydige italienske *condizionale*-betegnelse (se Bank Jensen 2018). Som sprogunderviser er det efter min mening relevant at kende forskellige terminologier, såvel internationale som danske, og her kan man sommetider finde inspiration i sprogvidenskabens historie.

Litteratur

- Baden, Jacob (1785) *Forelæsninger over det Danske Sprog eller Resonneret Dansk Grammatik*. Kjøbenhavn: Horrebow.
- Bank Jensen, Viggo (2018) Den italienske *condizionale* ifølge Rasmus Rask. I *Mål og Mæle* 39, 1: 25-31.
- Bank Jensen, Viggo (2017) Rasmus Rask og det italienske verbalsystem. I *Ny Forskning i Grammatik* 24: 4-20.
- Bjerrum, Marie (1959) *Rasmus Rasks afhandlinger om det danske sprog*. København: Dansk Videnskabs Forlag.
- Bloch, Søren (1817) *Fuldstændig Dansk Sproglære: til Brug i lærde Skoler og ved Selvunderviisning. Første Deel, om Ordenes Dannelse og Retskrivning*, 2. aldeles omarbejdede Oplag. Odense: Hempel.
- Bøgh, Matthias (1800²) *Udtog af den danske Sproglære til brug ved Underviisning*. Kjøbenhavn: Gyldendal.
- Dam Jensen, Helle & Dam, Lotte (2018) *Spansk universitetsgrammatik. Former og kontekst*. Århus: Aarhus Universitetsforlag.
- Dardano, Maurizio & Trifone, Pietro (1997) *La nuova grammatica della lingua italiana*. Bologna: Zanichelli.
- Diderichsen, Paul (1960) *Rasmus Rask og den grammatiske tradition*. København: Ejnar Munksgaard.

- Diderichsen, Paul (1961) Dansk grammatisk terminologi. I *Undervisningsvejledning for folkeskolen II. Betænkning nr. 297*: 58-84.
- Diderichsen, Paul (1966, 3.udg.) 51. Verbets sammensatte eller omskrevne Former. I *Elementær dansk grammatik*. København: Gyldendal: 127-138.
- Diderichsen, Paul (1968) *Sprogsyn og sproglig opdragelse, historisk baggrund og aktuelle problemer*. København: Nyt Nordisk Forlag.
- Hansen, Erik & Heltoft, Lars (2011) *Grammatik over det danske sprog I-III*. København: Det Danske Sprog- og Litteraturselskab.
- Høysgaard (1747) *Accentuered og Raissonered Grammatica som viser Det Danske Sprog i sin naturlige Skikkelse*. Kjøbenhavn: Groth.
- Kuhlmann Madsen, John (1994) Rask, Bello og Det spanske Akademis nye grammatik. I *Vindue mod den romanske verden* (red. Skytte, G. et. al.). København: Museum Tusulanum: 193-215.
- Mikkelsen, Kristian (1894) *Dansk sproglære med sproghistoriske tillæg. Haandbog for lærere og viderekomne*. København: Lehmann og Stages Forlag.
- Mikkelsen, Kristian (1902) *Om det ny System i den danske Sproglære*. København.
- Nissen, Niels Lang (1808) *Grundtræk af Dansk Sproglære*. Kjøbenhavn: Schultz.
- Ordbog over det danske Sprog (ODS)*, Bind 4 (1922), ved Verner Dahlerup. København: Det Danske Sprog- og Litteraturselskab. Sidst tilgået 02.01.2023: <http://ordnet.dk/ods>
- Petersen, N.M. (1870) Bidrag til Rasmus Kristian Rasks Levned. I *Samlede Afhandlinger I*. København: Samfundet til den Danske Litteratur Fremme: 217-343.
- Rask, Ramus (1811) *Vejledning til det islandske eller gamle nordiske Sprog*. Kjøbenhavn: Schuboths Forlag
- Rask, Rasmus (1818) (Prisskriftet) *Undersøgelse om det gamle Nordiske eller Islandske Sprogs Oprindelse*. Kjøbenhavn: Gyldendalske Boghandel.
- Rask, Rasmus (1824) *Spansk Sproglære efter en ny Plan udarbejdet af Prof. R. Rask*. København: Beekens Forlag.
- Rask, Rasmus (1825) *Frisisk Sproglære udarbejdet efter samme Plan som den islandske og angelsaksiske*. København: Beekens Forlag.
- Rask, Rasmus (1825 ms) *Udkastoptegnelser til en Dansk Sproglære*, Ny kgl. Samling 149 c⁵ 4^o (jf. også Bjerrum 1959: 137-38).
- Rask, Rasmus (1827) *Italiensk Formlære udarbejdet efter samme Plan som den spanske Sproglære*. København: Schultz.
- Rask, Rasmus (1830) *Grammar of the Danish Language for the Use of Englishmen*.
- Rask, Rasmus (1832a) *Engelsk Formlære, udarb. efter en ny Plan*. København: Gyldendal.
- Rask, Rasmus (1832b) *Kortfattet Vejledning til det oldnordiske eller gamle islandske Sprog*. København: Reitzel.
- Rask, Rasmus (1837) *Dansk Sproglære* (fra engelsk oversat af H.K. Rask). København: Popske Offisin.
- Skautrup, Peter (1953) *Det danske sprogs historie*, III. København: Gyldendalske Boghandel/Nordisk Forlag.
- Wimmer, Ludvig (1870) *Oldnordisk formlære til brug ved undervisning og selvstudium*. København: Chr. Steen og Söns Forlag.
- Wiwel, H.G. (1901) *Synspunkter for dansk sproglære*. København: Det Nordiske Forlag.

Interaktion på noder

Julia Jørgensen & Maria Jørgensen

Aarhus Universitet

1. Indledning

Sproget er musikalsk. Musikalske virkemidler som rytme, tempo, lydstyrke og tonehøjde kan fungere som ressourcer i organisationen af samtaler og konstruktionen af sociale handlinger (fx Auer m.fl. 1999a; Local & Walker 2004).

På trods af dette mangler vi viden om prosodiens betydning i dansk interaktion. Der er dokumentation for, at nogle intonationskonturer bruges til interaktionelle formål på dansk (fx Sørensen 2021) – men hvad med andre musikalske virkemidler? Hvilke interaktionelle funktioner har de på dansk, og hvad er sammenspillet mellem dem? Hvordan kan dette undersøges for naturligt forekommende interaktion?

I denne artikel udforsker vi¹, om nogle af disse spørgsmål kan besvares ved at notere samtaler på noder – og om sådan et eksperiment overhovedet kan lade sig gøre. Vores tilgang er eksplorativ, og undersøgelsen er forankret i konversationsanalysens (Sidnell & Stivers 2013) metoder, forståelse af sprog og tradition for datasessioner. Vores indgangsvinkel er en *single case*-analyse (Sidnell 2013) af et kort samtaleuddrag fra det offentligt tilgængelige danske samtalekorpus Samtalebank (MacWhinney & Wagner 2010).

2. Baggrund

Idéen om at notere tale på noder er ikke helt ny. I ældre dansk litteratur om fonetik (fx Bo 1933) gengives stiliserede udtalemåder af sætninger nogle gange på noder, som reflekterer både tonehøjde, intonationskonturer og rytme². Nodenotationen er dog begrænset til enkeltsætninger, og selvom analyseobjektet er det talte sprog, lader eksemplerne til at være tænkte.

I denne artikel forsøger vi at notere et længere uddrag af en naturligt forekommende samtale på noder, og i kombination med konversationsanalyse (Sidnell & Stivers 2013) bruger vi noder som analyseredskab til at undersøge interaktionelle systematikker i prosodien i uddraget.

¹ Julia Jørgensen er bachelor i musikvidenskab fra AU. Maria Jørgensen er ph.d.-studerende ved Lingvistik, AU.

² Tak til Yonatan Goldshtein for at gøre os opmærksomme på dette.

Denne kombination af nodenotation og konversationsanalyse er meningsfuld på flere planer. Før vi gennemgår disse vil vi dog først argumentere for, hvorfor det overhovedet er interessant at undersøge prosodi fra et interaktionelt perspektiv.

Vi ved fra interaktionsforskningen, at musikalske elementer i ”samtalesproget” (Brøcker m.fl. 2012) kan have interaktionel relevans. Bl.a. kan øget tempo ved turovergangsrelevante steder bruges til turfastholdelse (Local & Walker 2004), tilpasning til eller brud på en forudgående turs rytme kan markere tilslutning eller fraslutning (Auer m.fl. 1999b), og bestemte intonationskonturer kan markere affektive standpunkter som fx overraskelse (Sørensen 2021). Disse detaljer er blevet opdaget med konversationsanalysens mikroanalytiske perspektiv og detaljerede Jeffersonske transskriptionssystem (Jefferson 2004; Hepburn & Bolden 2013), som tillader notering af visse prosodiske kvaliteter, herunder relativ lydstyrke og trykfordeling. Det er dog begrænset, hvor præcist det kan blive.

Andre transskriptionssystemer går mere i detaljen: Det tyske *Gesprächsanalytisches Transskriptionssystem* (GAT2; Selting m.fl. 2009) tager udgangspunkt i Jeffersons system, men inkorporerer ord lånt fra musikken, bl.a. *staccato*, og impressionistiske tegninger af intonationskonturer. Auer m.fl. (1999a:ix-xi) tager også udgangspunkt i Jefferson, men inkluderer flere detaljer, lydstyrkenotering som *forte* og *piano*, og tekstopstilling, der tydeligt viser beatet i samtalen. Både GAT2 og Auer m.fl.s systemer er dog stadig meget begrænsede, og vi står tilbage med spørgsmålet om, hvorvidt det er muligt at indfange prosodien endnu mere præcist.

Naturligvis kan man nærstudere tale i computerprogrammer, fx *Praat* (Boersma & Weenink 2022), og det bruges da også i konversationsanalyser (fx i Sørensen 2021). At analysere en hel samtale i *Praat* ville dog både blive uoverskueligt og kompromittere konversationsanalysens emiske perspektiv: De detaljer, analysen afdækker, skal være beviseligt betydningsbærende for deltagerne i samtalen (fx Steensig 2015). Analytikerens egen opfattelse af, hvad der siges, er afgørende for at indfange netop dette.

Det leder os til nodenotation – en oplagt mulighed, hvis man vil opnå høj præcision og detaljegrad i sin prosodiske transskription. Her kan der tages højde for stort set alle tænkelige prosodiske detaljer: tonehøjde, -længde og intonationskonturer, pauser, rytme, lydstyrke og tempo (inklusive kvaliteter som *rallentando*), samt stavelseskvaliteter som *staccato*. Også stemmekvaliteter, fx knirkende stemme, kan noteres ved hjælp af fremførelsesangivelser (se afsnit 5). Desuden gør systemet det muligt at notere både absolutte og relative prosodiske detaljer.

Nodenotation passer også godt til konversationsanalytisk tradition. Transskription er den vigtigste analytiske indgang i konversationsanalysen, og noder er blot en anden måde at transskribere på. De må derfor, på samme måde som en Jeffersonsk transskription, kunne bruges som redskab til at synliggøre detaljer i data, både i transskriptionsprocessen og når data genses, fx i forbindelse med en datasession.

Konversationsanalytikeren bruger desuden, ligesom musikeren der lytter et stykke musik af, sit øre til efter bedste evne at gengive det, hun hører. En nodetransskription af et samtaleuddrag er altså kvalitativ og bevarer et emisk perspektiv. Dertil kan nævnes, at nogle konversationsanalytikere, der bruger *Praat* til prosodisk analyse (fx Sørensen 2021), foretrækker at konvertere forskellen i tonehøjderne i ytringer fra Hz til semitoner, da det i højere grad svarer til, hvordan øret opfatter lyde (S. Sørensen, personlig kommunikation). Nodenotation visualiserer semitonerne uden at skulle konvertere fra Hz.

Det er selvfølgelig vigtigt at nævne, at sprog ikke *er* musik – en pointe, vi er bevidste om og vender tilbage til i diskussionen. Vi mener ikke, at det gør vores foretagende mindre oplagt eller interessant, men det er klart, at det på nuværende stadie blot er et eksperiment; et forsøg på at låne indsigter fra én disciplin til at forfine observationer i en anden.

3. Metode, data og fremgangsmåde

Som sagt er undersøgelsen konversationsanalytisk (Sidnell & Stivers 2013), både i metode og fremgangsmåde, og i vores grundlæggende antagelser om samtaleproget.

Samtalenproget er organiseret ned til mindste detalje (Sacks 1984). Pauser, syntaktiske og leksikalske valg, gestik og mimik, prosodiske variationer osv. er alle potentielt interaktionelt motiverede. Disse detaljer kan nemlig bruges som ressourcer til at få interaktion til at lykkes, dvs. skabe intersubjektiv forståelse (Schegloff 1992) og få samtalens projekter i mål ved at gøre sociale handlinger genkendelige (Levinson 2013). Disse grundantagelser motiverer konversationsanalysens brug af optagelser af naturligt forekommende samtaler og detaljerede transskriptionssystem (Hepburn & Bolden 2013).

Vi begyndte undersøgelsen med to forskningsspørgsmål: Kan naturligt forekommende interaktion skrives ud på noder? Og afslører dette interaktionelt motiverede regelmæssigheder i det prosodiske design af interaktion? For at besvare disse spørgsmål foretager vi en tredelt *single case*-analyse (Sidnell 2013) af et samtaleuddrag på ca. 10 sekunder. Uddraget er fra en videooptaget samtale fra Samtalebanken (McWhinney & Wagner 2010) og blev valgt for sin

praktiske og analytiske tilgængelighed: data kan tilgås online og var derfor nemt at dele forfatterne imellem, og uddraget optræder desuden i en kollektion i forbindelse med en af forfatternes andre undersøgelser (Jørgensen 2021). Deltagerne i uddraget er anonymiserede.

Første led i vores fremgangsmåde var en klassisk konversationsanalyse. Vi lavede først en Jeffersonsk transskription af samtaleuddraget. Med udgangspunkt i transskriptionen og gentagne afspilninger af data undersøgte vi derefter, hvordan samtaledeltagerne skaber intersubjektiv forståelse – med andre ord, hvilke sociale handlinger deltagerne udfører, hvordan de bruger lingvistiske, prosodiske og kropslige ressourcer til dette, og hvilke interaktionelle forhold der motiverer brugen af netop disse ressourcer.

Andet led var at forsøge at lave en nodetransskription af uddraget i nodebehandlingsprogrammet Sibelius (AVID 2023). Vi afspillede uddraget gentagne gange, både i normalt og nedsat tempo, gentog selv, hvad vi hørte i sang og tale, og brugte gehør og klaver til at fastslå de omtrentlige toner samt rytmen i uddraget. I Sibelius noterede vi noderne og afspillede og redigerede dem løbende, så de kom så tæt på, hvad vi hørte, som muligt. Det sidste led i fremgangsmåden var at sammenligne den konversationsanalytiske transskription, nodetransskriptionen og den interaktionelle redegørelse i en datasession, hvor vi sammen lavede observationer over de potentielle interaktionelle regelmæssigheder i uddraget.

De næste tre afsnit præsenterer resultaterne af de tre analytiske led.

4. Konversationsanalyse af data

I dette afsnit præsenterer vi vores samtaleuddrag, inklusive den Jeffersonske transskription.

Uddraget er fra en samtale mellem Anne og Beate, to kvinder i tyverne. Kvinderne diskuterer, hvad der gik galt, da de lørdagen forinden skulle mødes på diskotek Luux, men endte med at gå fejl af hinanden. De fortæller skiftevis om deres perspektiv på begivenhederne. Anne har netop givet en del af sin forklaring, som Beate kvitterer for i linje 1. Beate igangsætter herefter selv en historiefortælling (fx Couper-Kuhlen & Selting 2018), hvor hun giver sit perspektiv, i line 2.

Faldt i søvn [AnneOgBeate | SamtaleBank | 03:50-04:05]

01 BEA: NÅ: OKAY?=
02 BEA: =MEN ALTså så ring↑ed du.=
03 BEA: =å jeg sagde til dig rigtig mange gange at du skulle ↑komme_
04 (0.4)
05 BEA: på ↑Luux_
06 BEA: å det ville du ↑ikk_
07 (0.9)
08 BEA: å::: f:
09 BEA: (.)|
10 BEA: [ja.]
11 ANN: [så] [faldt #jeg i søvn#.
12 BEA: [°fts°
13 (0.7)
14 BEA: ↑gjorde du;↑
15 ANN: °*j[a*°]
16 BEA: [εokay] for jeg troede nemligε=
17 BEA: =εjeg tænkte at je:gf_
18 (.)
19 BEA: ε#ø:#::h var helt vildt led nuε
20 BEA: εfordi at så glemte jegε(h) H:: [HAH]
21 ANN: [hh]
22 BEA: .H:: εså glemte jeg at du havde haft ringet?

Beates kvittering (l. 1) virker kort for hovedet, og hun lader til at skynde sig til sit eget perspektiv – der er ingen pause mellem hendes første og anden turkonstruktionsenhed (Steensig 2001) i hhv. line 1 og 2. Grunden lader til at være skyld: hun tror hun har såret Anne ved at glemme hende (jf. ll. 16-21) og vil forsvare sig. Det ser vi i Beates formuleringer af de næste par linjer. Hun påpeger, at hun sagde *rigtig mange gange* (l. 3), at Anne skulle komme, men at *det ville [Anne] ikk* (l. 6) – med andre ord, Anne er potentielt medskyldig i Beates foreteelse, og Beate selv gjorde alt, hvad hun kunne. Dette forsvar bygges op over flere inkremitter (ll. 3-10), der danner én lang sætning. Det sidste inkrement indledes med et langt *og* (l. 8), der hænger lidt i luften inden det afsluttes med *ja* (l. 10). Den faldende intonation lyder opgivende og bidrager til Beates argument: Hvad der skete, var beklageligt, men hun gjorde, hvad hun kunne. Beate konstruerer altså sig selv som skyldig, men undgår at eksplicitere sin skyld. Hun understreger at hun gjorde, hvad hun kunne, og gør Anne til medskyldig.

Forløsningen på skylden kommer, da Anne tilføjer en konklusion til Beates historie med *så faldt jeg i søvn* (l. 11), et nyt inkrement. Efter en pause kvitterer Beate med et proform-spørgsmål, *gjorde du*, i et højt toneleje (l. 14; om proform-spørgsmål, se Jørgensen 2021). Spørgsmålet viser, at linje 11 er mod Beates forventninger, og det hjælper hende med at indlede en forklaring på, hvorfor hun har opført sig, som hun har i samtalen: Siden Anne *ikke* blev såret (fordi hun sov og ikke stod og følte sig glemt et sted), er Beate skyldfri og kan derfor frit fortælle om sin bekymring (ll. 16-21). Hun er ikke længere i forsvar.

5. Samtaleuddraget på noder

Vi vil nu forsøge at besvare vores første forskningsspørgsmål: Kan dette samtaleuddrag sættes på noder? Først præsenterer vi dog lynintroduktion til nodelæsning (se mere i Rushby 2020).

En enkelt 'stemme' i en nodetransskription kaldes et nodesystem. Nodesystemer er organiseret i linjer, som er inddelt i mindre enheder – takter. Tonehøjde læses vertikalt: her kan toner placeres både på og mellem linjerne med enten en hel eller en halv tones interval mellem hver placeringsmulighed. Toner kan desuden markeres som hævede eller sænkede med fortegnene # (en halv tone højere) og b (en halv tone lavere). Horisontalt i systemet læses det melodiske forløb, dvs. tonernes forhold til hinanden i tid. Stilhed angives med pausetegn af forskellig varighed. De enkelte noders varighed angives bl.a. med antallet af faner og bjælker – jo flere bjælker noden har, jo kortere varer den. Desuden er det muligt at angive det generelle tempo (*beats per minute*, BPM) over en given takts øverste venstre hjørne. En nodetransskription kan indeholde flere nodesystemer ('stemmer') oven over hinanden, og disse skal læses samtidigt.

Nedenfor præsenterer vi resultatet af vores forsøg på en nodetransskription af uddragets linje 01-14. Det har været udfordrende at notere, men svaret på forskningsspørgsmålet er ja: Det kan lade sig gøre at notere samtaleuddraget på noder, men kun hvis man tager sig visse kreative friheder.

Kirstine Boas, Inger Schoonderbeek Hansen, Tina Thode Hougaard
& Ea Lindhardt Overgaard (udg.):
19. Møde om Udforskningen af Dansk Sprog, Århus 2022

ANNE

BEATE

♩ = 110

nå o - kay men alt - så så ring - ed du og jeg

2 ♩ = 130 ♩ = 100

sag til dig rig-tig mang gang at du sku kom på Luux og det ville du ik

4 *knirkende stemme*

så faldt jeg i søvn

å ja gjorde du

Uddraget er ikke noteret i en toneart, hvilket forklarer de mange løse fortegn. Almindeligvis indikeres en toneart i begyndelsen af nodesystemer med dertilhørende faste fortegn, men da det ikke er muligt at bestemme en egentlig toneart for tale, har vi valgt at arbejde uden fast toneart. Et lignende problem har været angivelse af taktart – hvor mange slag hver takt er værd. Der er ikke noteret en taktart, men vi har arbejdet ud fra en standard-takt på 4/4, fordi det skaber overblik og gør noden mere læsevenlig. Det er også vigtigt at pointere, at toneangivelserne i et nodesystem er en anelse stiliserede. Klaveret er stemt i tolv toner med fastlagte Hz-værdier, som kan vises i nodesystemet. Frekvenserne i tale har flere nuancer, og i bestemmelsen af tonerne i Anne og Beates tale har vi nogle gange været nødsaget til at 'runde op og ned' til nærmeste tone på klaveret. Beates *og jeg* i første linje ligger fx på en tone mellem C og C#, som i vores transskription er rundet op til nærmeste halve tone, C#.

Andre ting fungerer uden problemer i noteringen. Anne og Beate har hvert sit nodesystem (hver sin 'stemme') i noteringen, hvilket gør det muligt at vise overlap. Samtidig viser det både tonehøjden i deres ytringer, og forholdet mellem de to taleres tonehøjder. Det er også muligt at

vise varierende hastighed med løbende angivelser af BPM. At angive kvaliteter som knirkende stemme er muligt med en kombination af noder med kryds-nodehoved og fremførelsesangivelser som i Annes *jeg i søvn*. Særligt glidende toneforløb kan angives med *glissando* som i Beates *komme og gjorde du*.

6. Afslører nodeanalysen interaktionelle systematikker i prosodien?

Svaret på vores andet forskningsspørgsmål – afslører nodeanalysen systematikker i prosodien – er også ja. Nodeanalysen tydeliggør detaljer, der nuancerer den interaktionelle analyse af uddraget. Observationer underbygges og udvides, og i nogle tilfælde giver nodeanalysen anledning til at se tingene i et nyt lys. Vi fokuserer her på de ting, vi mener er særligt interessante.

Beates hastige afvikling af kvitteringen for Annes forudgående tur og introduktion af eget projekt gøres særligt tydelig i nodeanalysen. Her kan vi se, at den rytme og det tempo, hun starter ytringen med (*nå okay*), fortsætter ubrudt ind i hendes historiefortælling (*men altså så ringed du*). De to elementer kobles tæt sammen. Samtidig starter historiefortællingen (på *men*) på en lavere tone end den forudgående stavelse *-kay* – vi mener, at det gør det hørbart, at der, trods den tætte rytmiske forbindelse, stadig er tale om to særskilte turkonstruktionsenheder.



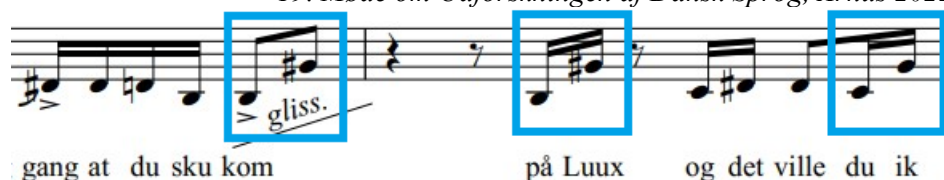
Figur 1: Én ubrudt rytme over to turkonstruktionsenheder.

Nodeanalysen underbygger således, hvad der også står klart i den Jeffersonske transskription. De to turkonstruktionsenheder er prosodisk meget tæt forbundne, og det giver Beate en lyd af at skynde sig frem til sit eget perspektiv og kun anerkende Annes side af sagen *proforma*. Hun har travlt med at komme frem til sit.

I Beates efterfølgende ytring argumenterer hun for, at hun har gjort sit for at få Anne til at komme ned på Luux. Ytringen når en første mulig afslutning efter *at du sku kom*. På det sidste ord glider Beate op fra et H til et G#. Herefter følger en pause, og Beate tilføjer derefter to inkremitter, der også mellem sig har en kort pause. Det første inkrement, *på Luux*, rammer præcis samme toner som Beates første afslutning – H og G#. Det næste inkrements afslutning, *du ikk*, bruger igen næsten samme toner. Denne gang starter Beate en halv tone højere, på C, men sluttonen er stadig G#.

Kirstine Boas, Inger Schoonderbeek Hansen, Tina Thode Hougaard
& Ea Lindhardt Overgaard (udg.):

19. Møde om Udforskningen af Dansk Sprog, Århus 2022



Figur 2: Melodiske kopier.

Nodetransskriptionen afslører, at Beate ikke bare afslutter med stigende intonation tre gange i træk, som der står i den Jeffersonske transskription, men at de tre afslutninger er overraskende præcise melodiske 'kopier' af hinanden. Det minder om den *prosodiske parallelisme* man ser i listeintonation (Selting 2007: 506) – intonationen hjælper med at tydeliggøre, at inkrementerne er delelementer af et hele, der endnu ikke er afsluttet. Denne type intonation gør Beates forsvarsargument og forsøg på at gøre Anne til medskyldig stærkere. At nævne Annes fejltagelser som en liste af begivenheder, der følger hinanden, får dem til at lyde mere problematiske: hun har ikke kun begået én, men gentagne fejl, og det efter flere bestræbelser fra Beates side. At nå op på den samme høje sluttone hver gang skaber forbindelse mellem og kronologi i Beates inkremitter og giver muligvis også en indigneret lyd. Det kunne være interessant at undersøge flere listekonstruktioner med nodetransskription for at se, om konturen bliver lige så præcist kopieret, eller om det er særligt for bestemte lister – og i så fald hvilken effekt det har.

Omkring nodetransskriptionens afslutning ser vi en række interessante detaljer. Først er det værd at bemærke, at Beate tydeligt markerer både sin historie og listen i den som afsluttet ved et markant fald på ordet *ja*, som når den dybeste tone hidtil. Denne ytring indledes med *å*, som først lyder som oplægget til endnu et inkrement i stil med de forudgående. Det hører vi ikke kun grammatisk (*og* kan ikke stå alene – det peger på, at der kommer mere), men også tonalt: *å* ligger nemlig på et H, samme starttone som i de to første melodiske kopier beskrevet ovenfor. Det lyder, som om Beate er i gang med endnu en kopi. Efter en pause kommer dog et afsluttende fald³ på *ja*.



Figur 3: Potentielt ny kopi og dybeste tone hidtil.

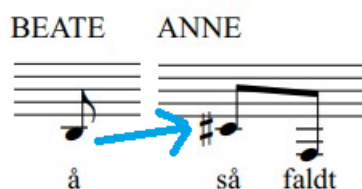
³ Interessant nok falder Beate til G#, samme sluttone som i kopierne – dog en oktav dybere.

Anne tilføjer herefter sit eget inkrement til turen. Grammatisk kobler det sig nemt på Beates tur (... og det ville du ikke, og så faldt jeg i søvn.), men Anne bruger også tone til at skabe sammenkoblingen. Hun starter sit inkrement på et C#, en hel tone over Beates å. Som vi kan se i node-transskriptionen, starter Beates turkonstruktionsenheder med en lidt lavere tone end den næste betonede stavelse. Ser vi fx på turkonstruktionsenheden, der starter *og jeg sagde til dig...*, ligger *og jeg* på et C#, mens den næste betonede stavelse, *sagde*, ligger på et D, en halv tone højere.



Figur 4: Ny turkonstruktionsenhed stiger på første betonede stavelse.

Ved at starte sin ytring lidt højere end Beate i nederste linje får Anne de to ytringer til at lyde som én enhed: hun gør dem melodisk til én sætning.



Figur 5: Anne stiger efter Beates indledende *og*.

Det er også interessant, at de to talere har en fælles fornemmelse af, hvilket leje den afsluttende tone bør ligge i. Anne giver også sin tilføjelse afsluttende, faldende intonation, og hun falder til en tone meget nær Beates.



Figur 6: Fælles fald.

Til sidst vil vi nævne, at Beates proform-spørgsmål, *gjorde du*, er særligt markeret, både i kropssprog – Beate hæver øjenbrynene og skubber hovedet let frem⁴ – og i prosodien. Beate foretager et stort tonalt spring og når uddragets tone. Denne markerethed hjælper med at udtrykke, at det forudgående er mod hendes forventninger (Jørgensen 2021).



Figur 7: Beates højeste tone.

7. Diskussion

En nodeanalyse af samtaleuddrag kan altså bruges som led i en interaktionel analyse. Den afslører prosodiske detaljer, der understøtter, supplerer og forfiner konversationsanalysen af uddraget.

Der er dog nogle åbenlyse faldgruber ved anvendelsen af nodetranskription i denne sammenhæng. Selvom sproget er musikalsk, og selvom vi kan udlede toner og rytme af samtaler, så *er* sprog naturligvis ikke musik. Undervejs i analysen har vi overvejet, om vi er kommet til at overmusikalisere analysematerialet. Hører man først noget som musik, er det svært at lade være, og man risikerer at begynde at lede efter rytmer og tonale sammenhænge, der ikke nødvendigvis svarer til virkeligheden. Et andet kritikpunkt er, at nodenotation ikke er designet til samtalesprog. Som beskrevet i afsnit 5 har vi været nødt til at være kreative og gå på kompromis med visse dele af systematikken i nodenotationen for at få uddraget til at passe i et nodesystem.

Vi mener dog ikke, at disse faldgruber behøver at være hæmsko for videre forsøg med analysemetoden. Var nodenotering vores primære måde at tilgå data, ville der være grund til kritik og til i højere grad at have øje for både potentiel overmusikalisering og de friheder, vi bliver nødt til at tage os med nodesystemet. I vores undersøgelse har nodeanalysen dog kun fungeret supplerende og som et værktøj, der får bestemte detaljer til at stå tydeligere frem. Disse detaljer kan senere undersøges mere systematisk, for eksempel i kollektionsanalyser – mere om det nedenfor.

Det er interessant, at der kommer så mange forskellige observationer ud af analysen, og vi mener, der er grundlag for at arbejde videre med nodenotation af samtaler. Måske er analyse-

⁴ Det er en fast bestanddel af denne type spørgsmål, at de akkompagneres af hoved og/eller bevægelser omkring øjne og øjenbryn (Steensig et al. 2023).

metoden dog mest anvendelig, hvis der fokuseres på et afgrænset materialegrundlag, fx i undersøgelser af kollektioner af mere begrænsede praksisser (fx listeintonation eller proform-spørgsmål). Her ville analyseredskabet kunne bruges til at belyse prosodiske kvaliteter med interaktionel betydning, der er gennemgående på tværs af eksempler, snarere end enkeltstående detaljer i en *single case*.

Det kunne dog også være interessant at se mere på nodeanalyser af hele samtalebidder for at drage nytte af analysemetodens kvalitet som kreativt opdagelsesredskab⁵. Før vi noterede uddraget på noder, havde vi ikke bemærket, at Beates listeintonation indeholder så præcise melodiske kopier, eller at Anne kobler sit inkrement på Beates tur både grammatisk og tonalt. At undersøge data med et kreativt perspektiv kan belyse detaljer, der ellers ikke var blevet opdaget, og som bagefter kan udforskes nøjere med andre metoder.

8. Konklusion

I denne artikel spørger vi, om man kan notere naturligt forekommende samtaler på noder, og om dette kan være analytisk brugbart i interaktionelt forankrede undersøgelser af dansk samtaleprosodi. Svaret på begge spørgsmål er ja – med visse forbehold.

Nodenotation er ikke skabt til sprog, men med nogle bevidste, kreative modifikationer og en åben tilgang til metoden kan det lade sig gøre at skrive hele stykker af samtale ned med relativt stor præcision. Både det kronologiske samtaleforløb, inklusive taleroverlap og pauser, samt prosodiske kvaliteter som bl.a. toner, intonationskonturer, rytme, betoning og temposkift kan gengives. Analysen afslører detaljer om samtaleuddraget, som i høj grad beriger den interaktionelle analyse. Datagrundlaget er dog stadig småt, og det kunne være interessant at lave flere undersøgelser med samme metode for at se, om det positive resultat kan repliceres, og om nogle af de fundne detaljer dukker op i andre interaktioner.

Selvom det er vigtigt at huske, at sprog ikke er musik, så er der meget at arbejde videre med i denne måde at undersøge prosodi i interaktion. Klarest står måske nodenotationens brugbarhed som kreativt opdagelsesredskab i prosodiske undersøgelser samt dens potentiale i udforskningen af kollektioner af begrænsede praksisser.

⁵ Tak til Yonathan Goldshtein og Liv Moeslund Ahlgren for denne pointe.

9. Litteratur

- Auer, Peter, Elizabeth Couper-Kuhlen & Frank Müller (1999a) *Language in Time: The Rhythm and Tempo of Spoken Interaction*. New York: Oxford University Press.
- Auer, Peter, Elizabeth Couper-Kuhlen & Frank Müller (1999b) Rhythm and Preference Organization in English. I Peter Auer, Elizabeth Couper-Kuhlen & Frank Müller (reds.). *Language in Time: The Rhythm and Tempo of Spoken Interaction*. New York: Oxford University Press: 92-115.
- AVID (2023) = <https://www.avid.com/sibelius> (tilgået 04/05-23).
- Bo, Alf (1933) *Tonegangen i Dansk Rigsmaal*. København: Povl Branner.
- Boersma, Paul & David Weenink (2022) *Praat: doing phonetics by computer [Computerprogram]*. Version 6.2.06. <https://www.praat.org> (tilgået 04/05-23).
- Brøcker, Karen Kiil, Magnus Glenvad Tind Hamann, Maria Jørgensen, Simon Bjerring Lange, Nicholas Hede-gaard Mikkelsen & Jakob Steensig (2012) Samtalesprogets grammatik. Fire fænomener og nogle metodiske overvejelser. I *NyS, Nydanske Sprogstudier* 42:10-40.
- Couper-Kuhlen, Elizabeth & Margret Selting (2018) *Interactional Linguistics*. Cambridge: Cambridge University Press.
- Hepburn, Alexa & Galina B. Bolden (2013) The Conversation Analytic Approach to Transcription. I Jack Sidnell & Tanya Stivers (reds.). *The Handbook of Conversation Analysis*. Malden, MA: Wiley-Blackwell: 57-76.
- Jefferson, Gail (2004) Glossary of transcript symbols with an introduction. I Gene Lerner (red.). *Conversation Analysis: Studies from the first generation*. Amsterdam: John Benjamins:13-31
- Jørgensen, Maria (2021) Er *har du det* et reelt spørgsmål – og hvilken forskel gør ‘*det*’? En interaktionel analyse. I Yonatan U. Goldstein, Inger S. Hansen & Tina T. Hougaard (reds.). *18. Møde Om Udforskningen Af Dansk Sprog*. Aarhus: Aarhus Universitet:337-58.
- Levinson, Steven (2013) Action Ascription and Formation. I Jack Sidnell & Tanya Stivers (reds.). *The Handbook of Conversation Analysis*. Malden, MA: Wiley-Blackwell:103-130
- Local, John & Gareth Walker (2004) Abrupt-Joins as a Resource for the Production of Multi-Unit, Multi-Action Turns. I *Journal of Pragmatics* 36 (8): 1375-1403.
- MacWhinney, Brian og Johannes Wagner (2010) Transcribing, searching and data sharing: The CLAN software and the TalkBank data repository. I *Gesprächsforschung* 11:154-173.
- Rushby, Simon (2020) *Discovering Music Theory: The ABRSM Grade 1 Workbook*. London: ABRSM.
- Sacks, Harvey (1984) Notes on Methodology. I J. M. Atkinson og John Heritage (reds.). *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press: 2-27.
- Schegloff, Emanuel (1992) Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. I *American Journal of Sociology* 97 (5): 1295-1345.
- Selting, Margret (2007) Lists as embedded structures and the prosody of list construction as an interactional resource. I *Journal of Pragmatics* 39(3):483-526.
- Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, m. fl. (2009) Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung* 10: 353-402.
- Sidnell, Jack (2013) Basic Conversation Analytic Methods. I Jack Sidnell & Tanya Stivers (reds.). *The Handbook of Conversation Analysis*. Oxford: WileyBlackwell:77-99.

- Sidnell, Jack & Tanya Stivers (2013) *The Handbook of Conversation Analysis*. Oxford: Wiley-Blackwell.
- Steensig, Jakob (2001) *Sprog i virkeligheden: bidrag til en interaktionel lingvistik*. Aarhus: Aarhus Universitetsforlag
- Steensig, Jakob (2015) Konversationsanalyse. I S. Brinkmann & L. Tanggaard (reds.). *Kvalitative metoder og tilgange – en grundbog*. København: Hans Reitzels Forlag:321-47
- Steensig, Jakob, Maria Jørgensen, Nicholas Mikkelsen, Karita Suomalainen & Søren S. Sørensen (2023) Towards a Grammar of Danish Talk-in-Interaction: From Action Formation to Grammatical Description. *Research on Language and Social Interaction* 56 (2):116-140
- Sørensen, Søren Sandager (2021) Affiliating in Second Position: Response Tokens with Rising Pitch in Danish. *Research on Language and Social Interaction* 54 (1):101-2.

Resumé:

Om svar på specificerende HV-spørgsmål i dansk

Maria Jørgensen

Aarhus Universitet

I interaktionsforskningen kategoriseres HV-spørgsmål som hhv. specificerende eller fortællende (Fox & Thompson 2010). Hvor fortællende HV-spørgsmål efterspørger længere beretninger, beder specificerende HV-spørgsmål i stedet om meget afgrænset information, og de kan, til forskel fra fortællende spørgsmål, besvares med et enkelt syntagma. Det specificerende spørgsmål ”hvornår skal vi spise” kan for eksempel besvares med ”nu” eller ”om to timer”. Det kan til gengæld også besvares med en fuld sætning: ”det skal vi nu” eller ”det skal vi om to timer”.

I denne præsentation argumenterer jeg for, at der er en interaktionelt motiveret funktionel forskel på, om man på dansk besvarer specificerende HV-spørgsmål med en fuld sætning eller blot med et syntagma. Det viser jeg gennem en analyse af to repræsentative eksempler, et på hver responstype. Min metode er konversationsanalyse (Sidnell & Stivers 2013) og interaktionel lingvistik (Couper-Kuhlen & Selting 2018). Analysen er baseret på en kollektionsanalyse af 79 svar på specificerende spørgsmål, identificeret og indsamlet fra 7 timer og 42 minutters videooptagelser af naturligt forekommende dansk interaktion. Data stammer fra Aarhus Universitets samtalekorpus (AULing) samt fra den offentligt tilgængelige Samtalebank.

Fuldsætnings svar er på dansk den hyppigste type respons på specificerende spørgsmål. Denne type svar indekserer en høj grad af uafhængighed fra, hvad der ellers er sket i interaktionen, og svartypen følger ofte efter spørgsmål, der indleder nye emner eller retninger for samtalen. Uafhængigheden i fuldsætnings svar lader taleren hævde høj epistemisk autoritet over svaret. Syntagmesvar indekserer derimod en høj grad af integreret i, hvad der foregår i samtalen netop nu. Denne type svar følger ofte efter spørgsmål, der ikke ændrer samtals retning, men i stedet bygger videre på samtaledeltagernes nuværende projekt i udvidelsessekvenser. Integreret i syntagmesvar lader desuden taleren hævde en mere ligevægtig epistemisk status mellem spørger og taler.

De to svartyper kan bruges af samtaledeleter til bl.a. at dele deres umiddelbare forståelse af, hvor i samtalen de befinder sig, og til at orientere sig mod egne og andre samtaledeleteres vidensrettigheder og -pligter (Stivers et al. 2011). Desuden er svartypernes forskellighed et udtryk for, at grammatik er positionssensitiv (Schegloff 1996) og derfor også bør udforskes fra et interaktionelt perspektiv.

Litteratur

- Couper-Kuhlen, Elizabeth & Margret Selting. 2018. *Interactional Linguistics: Studying Language in Social Interaction*. Cambridge: Cambridge University Press.
- Fox, Barbara A. & Sandra A. Thompson. 2010. Responses to Wh-Questions in English Conversation. *Research on Language & Social Interaction* 43 (2): 133–56.
- Schegloff, Emanuel A. 1996. Turn Organization: One Intersection of Grammar and Interaction. I: E. Ochs, E. A. Schegloff & S. Thompson (red.): *Interaction and Grammar*, 52–133. Cambridge: Cambridge University Press.
- Sidnell, Jack & Tanya Stivers (red.). 2013. *The Handbook of Conversation Analysis*. Malden, MA: Wiley-Blackwell.
- Stivers, Tanya, Lorenza Mondada & Jakob Steensig (red.). 2011. *The Morality of Knowledge in Conversation*. Cambridge: Cambridge University Press.

Altså Bertel Haarder han ku ikk li noget med universiteter

Deltagelse i historiefortælling i en dansk hverdagssamtale

Nicholas Mikkelsen

Aarhus Universitet

1. Indledning

Denne artikel handler om historiefortælling i hverdagssamtaler. Tidligere forskning inden for konversationsanalysen har bl.a. studeret, hvordan historier indledes (Jefferson 1978), hvordan de afsluttes (Sacks 1974) og hvad der sker, efter de er afsluttede (Kjærbeck & Asmuß 2005). Men der er ikke så meget forskning, der ud fra et interaktionelt perspektiv har undersøgt, hvad der sker mellem starten og afslutningen på en historie, og det er nogle af de perspektiver, jeg vil prøve at undersøge i denne artikel. Mere konkret vil jeg gøre det med udgangspunkt i følgende spørgsmål:

- Hvilke lingvistiske og multimodale ressourcer bruger fortælleren til at invitere til deltagelse fra andenparten i samtalen?
- Hvordan og med hvilke ressourcer bidrager andenparten til at forme talerens handlinger?

Artiklen er en del af projektet Grammar in Everyday Life (GEL)¹, hvor vi med udgangspunkt i interaktionel lingvistik (Couper-Kuhlen & Selting 2018) søger at undersøge forskellige aspekter af hverdagssamtaler med fokus på handlinger. Min del af projektet bærer titlen “How people negotiate participation in Danish everyday interaction” – deraf også mit fokus på hvordan samtaledeltagere forhandler deltagelse i løbet af større ture som historiefortælling. Konkret er min forskning inspireret af blandt andre Charles og Marjorie Harness Goodwins arbejde om netop deltagelse (2004) og Per Linells tanker om dialogisme (2009). Det kommer til udtryk

¹ Projektet er ledet af Jakob Steensig (Aarhus Universitet) og har tilknyttet to ph.d.-studerende (Maria Jørgensen og undertegnede) foruden en postdoc (Søren Sandager Sørensen). Det er finansieret af midler fra Danmarks Frie Forskningsfond. Projektperioden er september 2019 til august 2023. Projektets fulde titel er *Grammar in Everyday life (GEL) – How people ask questions, negotiate participation and tell others what to do in Danish everyday interaction*.

ved, at jeg ikke blot fokuserer på, hvordan en taler producerer en historie, men hvordan deltagerne i samtalen bidrager til at samskabe historien via deres bidrag eller hvordan manglende bidrag fra andenparten får taleren til at justere udformningen af sin tur. Dette er modsat, hvad Linell omtaler som monologiske beskrivelser, hvor sådanne interaktionelle perspektiver er fraværende (se fx Møller 1993). Det er derfor også væsentligt at inddrage samspillet mellem krop og grammatik (jf. fx Keevallik 2018) i sin beskrivelse, da blikretning, nik, hovedryst og andre kropslige udtryk er en væsentlig del af, hvordan man forhandler deltagelse og hvordan deltagere i en samtale bidrager til den.

I denne artikel vil jeg arbejde med en såkaldt *single case* analyse, som jeg vil bruge til at svare på de ovenstående spørgsmål og på den baggrund lægge op til, hvordan vi kan studere historier fra et dialogisk perspektiv med udgangspunkt i en større kollektion af historier (Sidnell 2012:78). Intentionen er altså at gøre observationer i dette specifikke excerpt for i et evt. senere studie at efterprøve, om de systematisk finder anvendelse på tværs af samtaledeltagere i forskellige samtaler (Sidnell 2012:88).

Data til artiklen er en samtale fra det offentligt tilgængelige korpus Samtalebank (tilgængeligt via <https://samtalebank.talkbank.org/>), som er transskriberet ud fra Hepburn og Boldens konventioner (2012). For at beskrive de multimodale aspekter af samtalen gør jeg også brug af Mondadas konventioner for multimodal transskription (2019).

I afsnit 2 vil jeg redegøre for tidligere forskning på området og vise, hvordan min tilgang til historier vil kunne bidrage til den viden, vi allerede har. I afsnit 3 følger en analyse af én udvalgt historie, og i afsnit 4 vil jeg samle op på, hvilke konklusioner vi kan drage af analysen, og diskutere, hvordan disse vil kunne efterprøves i en større kollektion af data.

2. Tidligere forskning

Med udgangspunkt i konversationsanalytisk metode har flere forskere over de seneste mere end 40 år bidraget til en interaktionel forståelse af historiefortælling som fænomen. *Interaktionel* er her et nøgleord, hvormed konversationsanalysen adskiller sig fra tidligere forskning i narrativer² som Labov og Waletzky (1967). Narrativer hos dem er dog eliciteret som en del af sociolingvistiske interviews (jf. Schegloff 1997). På dansk har Møller (1993) også arbejdet med fortællinger. Disse er også til dels indsamlet under interviewsituationer, men han behandler

² Jeg bruger her begrebet "narrativer," da det er det ord, forskerne selv bruger til at beskrive deres genstandsfelt. Når jeg i den øvrige tekst skriver om "historiefortælling" eller "fortællinger", er det med udgangspunkt i, at jeg undersøger aktiviteten at fortælle en historie som en lokalt situeret praksis. Til det formål er "historiefortælling" og "fortællinger" bedre dækkende, og det er desuden også sådan, fænomenet normalt omtales inden for den konversationsanalytiske litteratur.

også gruppesamtaler (1993:282-337). Dog forholder han sig kun i ringe grad til, hvordan andre end fortælleren bidrager. Konversationsanalytiske studier af historier fokuserer mere på den lokale forhandling og sekventielle forankring af historier i hverdagssamtaler. Jefferson (1978) beskriver fx, hvordan historier opstår i samtaler og hvordan de afsluttes ved at vende tilbage til en *pointe*, der netop antydes i begyndelsen. På samme vis har Sacks (1974) også beskrevet, hvordan en joke fortælles ved at klimakset foregribes i begyndelsen, hvilket senere er med til at vise, hvornår den har nået sit klimaks eller *punchline*. Ifølge Kjærbeck og Asmuß (2005) følges en histories klimaks oftest af en såkaldt *post punchline sequence*, hvor man først adresserer historiens modalitet (trist, sjov osv.) og dernæst vurderer selve historien.

På den baggrund har forskningen altså fint afdækket, hvordan historier startes og afsluttes. Fokus er her på den sekventielle indlejring og implikativitet ved historier – altså på hvilken baggrund historien initieres i samtalen og hvilken efterfølgende retning for samtalen, de kan siges at udstikke. Foruden den sekventielle organisering af historier foregår der under historiefortælling også en række handlinger, der snarere er simultant organiserede (Goodwin 2018:97). Goodwin beskriver fx brugen af fortsættelsesmarkører som broer mellem enheder i fortællinger (1986:207) og brugen af vurderinger fra tilhørerne. Her orienterer taleren sig mod, at vurderingerne gøres færdige, inden de fortsætter historien (Goodwin 1986:210). Eriksson har beskrevet den samme brug af fortsættelsesmarkører og vurderinger i svenske samtaler (1997:234-245). En væsentlig *pointe* er her, at tilhørere også er med til at forme historien gennem deres bidrag (Duranti 1986:243). Historier fortælles altså i en vekselvirkning mellem fortæller og tilhører, hvor bidrag fra tilhørerne påvirker talerens handlinger og omvendt (Goodwin 1986:215).

Ikke kun det verbale er en vigtig del af, hvordan en historie forløber – også kropslige handlinger spiller en væsentlig rolle. Fx kan blikretning hos tilhørerne være med til at forme talerens ytringer (Goodwin 1979), ligesom talerens blikretning også har betydning i forhold til turfastholdelse og -overdragelse (Kendon 1967:42). Ifølge Stivers (2008) bruges *nik* fra modtageren af en historie til at vise tilslutning til den holdning, der kommer til udtryk igennem fortællingen. *Nik* kan dog også bruges af taleren og altså ikke kun som respons (Aoki 2011). Inden for japansk interaktionsforskning har man arbejdet meget med netop *nik* i samtaler (Oshima 2014:32). Det er her væsentligt at være opmærksom på, at der på tværs af sprogsamfund også kan være forskel på brugen af kropslige resurser i samtaler foruden de rent verbale.

Særligt nyere forskning i interaktion er præget af, hvad Nevile har kaldt *the embodied turn* (2015). Når vi interagerer, bruger vi også vores kroppe, og på den måde er kroppen også relevant i grammatisk beskrivelse (Keevallik 2018). Min analyse af historien i nærværende

artikel vil således også fokusere på, hvordan deltagerne bruger blikretning, gestik m.m. til at udføre nogle af de handlinger, der ikke entydigt kan beskrives sekventielt (Linell 2009:184-86).

3. Analyse

Historien, min analyse vil fokusere på, er hentet fra en samtale mellem tre kvinder: Lis, Kirsten og Ann. De sidder i nogle bløde stole omkring et lavt bord, imens de drikker kaffe, spiser småkager og snakker. For overskuelighedens skyld præsenterer jeg det her i seks mindre bidder (1A-F). I 1A nedenfor indleder Lis en historie om en tidligere kollega:

1A. ”arbejdsløs” [Samtalebank | omfodbold | 3:17]

01 *LIS: €som en a:f af mine€ Δg[amle kolle]ger

lis €hh peger bagud €

ann Δh albue på v hånd--->

02 *ANN: [uhhuh]

03 *LIS: ovre på søgård_

04 *LIS: hun var- \$ung å ±hun var \$lige €færdig\$uddannet

lis \$nik

ann ±cigaret til mund--->

lis \$kigger på KIR

lis €hæver og sænker hænderne

lis \$kigger på ANN

kir \$nik

I linje 1 til 3 etablerer Lis sin historie. Med *som en a:f mine gamle kolleger* laver hun en anknøytning til den forudgående tale. I samme linje indtager Ann en position som modtager af historien ved at lade højre albue hvile på venstre hånd. I sin højre hånd holder hun samtidig en tændt cigaret, og med denne positur er hun genkendelig som værende i gang med at ryge snarere end en, der vil tage ordet. Hun kigger samtidig på Lis.

I linje 4 begynder Lis sin egentlige historie. Det er her værd at bemærke hvordan hun fremhæver ordet *ung* ved også at nikke ligesom hun skifter blikretning fra at have kigget mod Ann til at kigge på Kirsten i stedet. Ordet *færdiguddannet* fremhæves også ved, at hun hæver og sænker hænderne. I løbet af ordet kigger hun tilbage mod Ann, imens Kirsten nikker. Kirstens nik ligner en spejling af Lis nik, der akkompagnerede *ung*, og orienterer sig mod, at Lis har færdiggjort en enhed.

I linje 5 begynder hun på en afvigelse fra den historie, hun har igangsat i de forløbne linjer.

1B. ”arbejdsløs” [Samtalebank | omfodbold | 3:21]

05 *LIS: €hun- \$(.) å hun ↓↓eller\$\$ lige å lige↓↓_

lis €hæver hænderne

lis \$lukker øjnene \$åbner øjnene, kigger op

kir \$nik

06 *LIS: .hh\$ hun\$± havde været €arbejdsøløs et stykke tid.

lis \$kigger på KIR

lis \$kigger på ANN

ann ---->±cigaret fra mund

lis €vip m. hh, sænker hænder

ann €nik opad

07 *LIS: €°hun \$havde to *små børn*°.

lis €hæver og sænker hænder

lis \$kigger lige frem

Afvigelsen indledes med en afsnupning ved ordet *hun* efterfulgt af en mikropause, hvorefter hun genstarter med *å hun*. Inden mikropausen har hun også lukket øjnene. Hun fortsætter ytringen i et lavere stemmeleje med *eller lige å lige*. Det sænkede stemmeleje bidrager til at vise, at her er tale om noget, der er underordnet den forudgående tale. Samtidig viser ordet *eller*, at der er tale om en afvigelse fra det forløb, der hidtil har været projiceret. Ved *eller* åbner hun samtidig øjnene igen og kigger op. Samtidig nikker Kirsten og viser dermed en sekventiel tilpasning – at hun tilslutter sig den afvigelse, der projiceres med *eller*. Tidligere har Lindström og Linell også beskrevet, hvordan formatet *x-och-x* kan bruges som en reparerende handling i svenske samtaler (2007:79). Noget af det samme synes at være på spil her, hvor Lis skal reparere noget af sin historie, før hun kan fortsætte. I løbet af linje 5 kigger hun desuden op snarere end på modtagerne, hvormed hun ikke udvælger dem til at bidrage til den igangværende ytring.

I linje 6 fortsætter hun med en kort indånding, hvorefter hun udvider sin tur om den tidligere kollega: *hun havde været arbejdsløs et stykke tid*. Her kigger hun først på Kirsten og dernæst på Ann. Imens hun siger ordet *arbejdsløs*, vipper hun med højre hånd og sænker hænderne igen. Ved samme ord nikker Ann opad. Med sit nik viser Ann genkendelse af den yderligere information om kollegaen, hvilket indpakning af ordet med gestus og tryk på første stavelse også lægger op til.

I linje 7 fortsætter Lis ved at tilføje yderligere information om kollegaen: *°hun havde to små børn*°*. Dette er pakket ind som noget parentetisk, ved at hun taler med en sagte stemme. Tilsvarende skifter hun blikretning efter ordet *hun* og kigger lige frem og ikke på nogle af de to øvrige samtaledeeltagere.

I linje 8 følger en længere indånding, hvorefter hun fortsætter: *å så:: sku vi lukke*.

1C. ”arbejdsløs” [Samtalebank | omfodbold | 3:25]

- 08 *LIS: .hhhh å \$så:: sku vi ¥lukke_€
lis \$kigger på ANN--->
lis ¥hovedryst
lis€hæver hh
- 09 *LIS: °altså° ¥det €hele □sku§ lukkes
lis ¥hovedryst--->
lis €vip og peg med hh
ann □nik
kir §nikker--->
- 10 *LIS: □der €engang i fir↓sern ¥ikk↑ås_€§
ann □nik
lis ----> ¥
lis €vip hh til højre
lis€hæver vh
kir ---->§
- 11 *ANN: □hm\$::_
ann □nik
lis ---->\$kigger til højre

Med *å så::* knytter hun denne enhed an til det tidligere og forbinder den til hovedhistorien. Der er dog også noget i indpakningen, der kunne tyde på, at der er et problem. Det ser vi ved tøven på ordet *så::*, hvormed hun foregriber, at Ann ikke nødvendigvis kender til omstændighederne omkring, at de skulle lukke. Under ordet *lukke* ryster Lis også på hovedet, men hverken Ann eller Kirsten bidrager med verbal eller nonverbal respons på dette tidspunkt. De øvrige samtale-deltagere demonstrerer altså ingen form for kendskab til de omstændigheder, Lis refererer til.

I linje 9 fortsætter Lis således med et svagt udtalt *altså*, der viser en afvigelse fra den igangværende tale (Heinemann & Steensig 2018). Det følges af *det hele sku lukkes*, hvor både *hele* og *lukkes* er markeret med tryk. Hun ryster ligeledes på hovedet og vipper og peger med højre hånd, samtidig med at hun udtaler *hele*. Både Ann og Kirsten responderer med nik. Kirstens nik fortsætter i løbet af den efterfølgende linje.

I linje 10 udvider Lis sin ytring fra linje 9 ved at placere den i tid med *der engang i fir↓sern ikk↑ås*. Det afsluttende *ikk↑ås* er her endnu en ressource, Lis anvender til at forfølge respons fra de andre samtaledeeltagere (Jensen et al. 2019), selvom både Ann og Kirsten nikker i løbet af linje 10. I linje 11 tilføjer Ann et *hm::* simultant med endnu et nik. Foruden den

nonverbale respons i form af nik bidrager hun altså også med en verbal responsenhed. Her kigger Lis væk fra Ann, inden hun fortsætter i linje 12.

1D. ”arbejdsløs” [Samtalebank | omfodbold | 3:17]

12 *LIS: >as< \$bertel haarder han ku ikk □ lide noget med univer€si*teter*.

lis \$kigger på ANN

ann □hovedryst

lis €sænker hænder

13 *LIS: [ba\$:re] luk det;

14 *ANN: [ng\$]_]

lis \$kigger til højre

I linje 12 udvider hun igen initieret med et noget forkortet *altså*, >as<, der igen viser en afvigelse fra det projicerede forløb. Der sker her et skift i aktivitet, hvor Lis introducerer endnu en karakter til sin historie – nemlig Bertel Haarder. Han tilskrives holdningen, at han ikke kunne lide noget med universiteter, efterfulgt af hvad der kan forstås som en rapport af hans holdning (*ba:re luk det* i linje 13).

Efter at have kigget til højre kigger Lis igen på Ann i starten af linje 12. Undervejs i ytringen responderer Ann til dels med hovedryst og med et *nej* i overlap med Lis' *ba:re* linje 13. Det forekommer tydeligt, at linje 12 til 13 om Bertel Haarder er designet til at få en anden respons end den, som Ann bidrager med. Holdningen, Bertel Haarder tilskrives, er formuleret som det Pomerantz kalder en *extreme case formulation* (1986) – i en mere moderat tone havde hun måske foreslået, at han var kritisk over for universiteterne. Sådanne bruges ofte i sekvenser, hvor en taler brokker sig over noget for at legitimere en anke (Pomerantz 1986:228), men Anns bidrag (hovedryst i linje 12 og *nej* i linje 14) flugter ikke med en sådan aktivitet, som der implicit lægges an til. Hendes *nej* og hovedryst forekommer at være en simpel kvittering for modtagelse af information, og hun tilslutter sig altså ikke den holdning, Lis giver udtryk for gennem sin rapport af Bertel Haarder. Efter at have nået konklusionen på denne del af historien med sit *ba:re luk det*, vender Lis således tilbage til sin egentlige historie i linje 15.

1E. ”arbejdsløs” [Samtalebank | omfodbold | 3:32]

15 *LIS: \$.hhh[hh tsk (.) å:: €±jah siger hun så tjl mig_

lis \$kigger ml KIR og ANN

lis €nik

ann ±cigaret til mund--->

16 *ANN: [ja_

17 *LIS: .h \$det (jo) nok synd \$for jgr hvis i bliver \$arbejdsløse.

- lis \$kigger på ANN \$kigger ml ANN og KIR \$kigger på KIR og
ml KIR og ANN
- 18 (0.4)±(0.4)
ann ---->±fjerner cigaret fra mund
- 19 *LIS: \$*ja* det sy\$nes \$jeg jo osse
lis \$kigger på ANN
lis \$blik til h og tilbage på ANN
kir \$drejer hovedet til h
- 20 *LIS: men hvorfør:r □±hvorfo\$\$:r-
ann □nik opad
ann ±tipper cigaret opad
lis \$kigger ned
kir \$drejer hovedet t v, kigger på LIS
- 21 (0.3)
- 22 *LIS: \$hvorfor er det ik synd \$for dig;
lis \$kigger på ANN \$kigger ned

Det sker først med en længere indånding samtidig med, at hun kigger i rummet mellem Kirsten og Ann. Med sit *å::* knytter hun igen i linje 15 an til hovedhistorien, og Ann tager efterfølgende igen sin cigaret op til munden. Brugen af cigaretten kommer igen på et tidspunkt, hvor det bliver gjort relevant at vise deltagelse som lyttende ved, at Lis har signaleret, at hun vender tilbage til historien. Med *jah* begynder Lis på noget, der er genkendeligt som rapporteret tale. Udåndingen på *jah* gør det hørbart som sagt med en anden stemme, og samtidig efterfølges det af en rapporterende ramme med *siger hun så til mig* (Bolden 2004:1073). Hun fortsætter rapporten i linje 17 med *det (jo) nok synd for jer hvis i bliver arbejdsløse*. I løbet af ytringen i linje 17 kigger hun på tværs af de andre deltagere – først på Ann, dernæst i rummet mellem Ann og Kirsten, på Kirsten og igen i rummet mellem Ann og Kirsten. Når hun ved ytringen afslutning ikke kigger på nogen af deltagerne, signalerer hun samtidig, at hun fortsætter sin tur, hvilket hun også gør efter en pause på 0,8 sekunder i linje 19 med **ja* det synes jeg jo osse*. Linje 19 indledes med knirkende stemme (**ja**), hvormed hun viser, at det er en anden aktør i historien, der taler. Det er da også tydeligt ud fra konteksten, hvor **ja** fungerer som svar på ytringen i linje 17, og brugen af pronomenet *jeg* henviser til fortælleren – altså Lis. I løbet af linje 19 kigger hun til dels på Ann, til højre og så tilbage på Ann igen. Men det er også væsentligt, at Kirsten drejer hovedet til højre og altså kigger væk fra Lis. Dette har implikationer for det videre forløb, hvor Lis i den efterfølgende linje 20 gøre brug af hvad C. Goodwin tidligere har beskrevet som *phrasal breaks* (1979, 1980). Han beskriver, hvordan en taler kan genstarte sin tur for at tiltrække sig blikket fra de andre samtaledeltagere. Det ser også ud til at være tilfældet her, hvor Lis efter at have gentaget *hvorfor* to gange, igen får Kirsten til at kigge på sig. Og

efter en pause på 0,3 sekunder fortsætter hun sin tur med at rapportere sin egen respons i linje 22: *hvorfor er det ik synd for dig*. Denne linje 22 er et rapporteret spørgsmål, der fungerer som setup til selve konklusionen på historien, som kommer i linje 23 til 25, hvor hun rapporterer kollegaens respons som en punchline på historien.

1F. "arbejdsløs" [Samtalebank | omfodbold | 3:42]

23 *LIS: .hhh \$ne:j for jeg jo vant til å leve af ingenting.

lis \$kigger på ANN, ned og på ANN

24 *LIS: >°siger \$ hun så°<=

lis \$drejer hovedet og kigger til højre

25 *LIS: =D(h)gt \$er j j(h)o ±j(h)kke.¥

lis \$kigger på ANN

ann ±læner sig tilbage, hoved let på skrå

lis ¥smiler--->

Hvordan man konstruerer klimakset på en historie, er allerede beskrevet andetsteds, og det vil jeg ikke gå dybere ind i her, hvor det primære fokus er de ting, der sker i løbet af historiefortællingsforløbet, hvilket jeg vil samle op på nedenfor.

4. Konklusion / diskussion

I det følgende vil jeg sammenfatte resultatet af analysen ovenfor for at kunne udpege hvilke lingvistiske og multimodale ressourcer både fortælleren (Lis) og andenparten (Kirsten og Ann) gør brug af. Efterfølgende vil jeg udpege nogle perspektiver af analysen, man med fordel vil kunne arbejde videre med i en større undersøgelse.

Det er her væsentligt at have både fortælleren og tilhørernes respektive perspektiver for øje. Lis bryder løbende ud af det projicerede forløb med *eller* (linje 5) og *altså* (linje 9 og 12). Samtidig gør hun også brug af et lavere register og en mere sagte stemme til at indikere underordning (linje 5-7 og linje 9). Ved sin tilbagevenden til det projicerede historieforløb gør hun i et par tilfælde brug af indånding og forlænget vokal på det efterfølgende *så* (linje 8) og *å* (linje 15). De øvrige deltagere i samtalen responderer her med fx *nik* på tidspunkter, hvor færdiggørelsen af en underordning kan forudsiges (fx linje 4 og 9-10), men gør også brug af verbale responsenheder (linje 11, 14 og 16). I linje 4 og 6 gør Lis også brug af en kombination af tryk og gestik på ordene *færdiguddannet* og *arbejdsløs*, hvilket inviterer til genkendelse fra andenparten. Deltagerne gør også brug af kroppen at indtage en positur, der viser, at de agerer lyttere til historien, hvilket Ann aktivt gør ad et par omgange. Tilsvarende orienterer Lis sig

også mod Kirstens manglende deltagelse i linje 19 til 20, og det er med til at forme hendes turs videre forløb.

Deltagelsen i løbet af historiefortællingen mellem Lis på den ene side og Ann og Kirsten på den anden foregår altså i en form for vekselvirkning mellem, at Lis producerer nogle enheder, der på forskellig vis er designet som dele af hovedhistorien eller som underordnede. Kirsten og Ann forholder sig til disse enheder ved på forskellig vis at respondere på dem. Således spiller de som samtaledeltagere også en vigtig rolle i forhold til, hvordan historien konstrueres.

På baggrund af min analyse er der en række fænomener, som kunne behandles i en større undersøgelse. Det er nærliggende at arbejde videre med, hvordan man i danske hverdags-samtaler forhandler underordninger eller afvigelser fra det projicerede forløb i historier, som det sker i løbet af Lis' historie (se Mikkelsen 2023). Som jeg indledningsvist var inde på, er nik og hovedrysten også en vigtig del af samtaler, og her kunne det bl.a. være interessant at undersøge deres rolle i forhold til verbale responsenheder. Kommer non-verbale responser fx tidligere i en ytrings produktion end de verbale? Det er i hvert fald, hvad ovenstående historie kunne indikere, men om det er systematisk på tværs af samtaler, er værd at undersøge. Tilsvarende er det også værd at bemærke, at det ikke kun er modtagerne af historien, der nikker eller ryster på hovedet undervejs. Også Lis nikker (linje 4 og 15) og ryster på hovedet (linje 8 og 9) i løbet af sin historie, og det kunne også være interessant i en dansk kontekst at undersøge, hvad det gør at nikke under sin egen tur – altså ikke som respons til andres (jf. Aoki 2011).

Litteratur

- Aoki, H. (2011) Some functions of speaker head nods. I C. Goodwin, C. LeBaron, & J. Streeck (red.). *Embodied interaction: Language and body in the material world*. Cambridge: Cambridge University Press:93-105.
- Bolden, G. (2004) The quote and beyond: defining boundaries of reported speech in conversational Russian. *Journal of Pragmatics* 36:1071-1118.
- Couper-Kuhlen, E., og M. Selting (2018) *Interactional Linguistics: Studying Language in Social Interaction*. Cambridge New York, NY Melbourne Singapore: Cambridge University Press.
- Duranti, A. (1986) The Audience as Co-author: An Introduction. I: *Text & Talk – Interdisciplinary Journal for the Study of Discourse*. De Gruyter Mouton.
- Eriksson, M. (1997) *Ungdomars berättande: En studie i struktur och interaktion*. Uppsala: Institutionen för nordiska språk vid Uppsala Universitet.
- Goodwin, C. (1979) The Interactive Construction of a Sentence in Natural Conversation. I G. Psathas (red.). *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers:97-121.

Altså Bertal han ku ikk li noget med universiteter. Deltagelse i historiefortælling i en dansk hverdags samtale

- Goodwin, C. (1980) Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry* 50(3-4):272-302. <https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>.
- Goodwin, C. (1984) Notes on story structure and the organization of participation. I J. M. Atkinson & J. Heritage (red.). *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press:225-246.
- Goodwin, C. (1986) Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies* 9(2-3):205-217. <https://doi.org/10.1007/BF00148127>.
- Goodwin, C. (2018) *Co-operative action*. Cambridge University Press.
- Goodwin, C., og M.H. Goodwin (2004) Participation. I *A Companion to Linguistic Anthropology*. Oxford: Blackwell:222-44. <https://doi.org/10.1002/9780470996522.ch10>.
- Hepburn, A., og G.B. Bolden (2012) The Conversation Analytic Approach to Transcription. I Jack Sidnell og Tanya Stivers (red.). *The Handbook of Conversation Analysis*. John Wiley & Sons, Ltd.:57-76.
- Heinemann, T., og Steensig, J. (2018) Justifying departures from progressivity: The Danish turn-initial particle altså. I J. Heritage & M.-L. Sorjonen (red.). *Between Turn and Sequence: Turn-Initial Particles Across Languages*. John Benjamins:445-476. <https://doi.org/10.1075/slsi.31.15hei>.
- Jefferson, G. (1978) Sequential aspects of story telling in conversation. I J. Schenkein (red.) *Studies in the organization of conversational interaction*. Academic Press:213-248.
- Jensen, A. M., Clausen, O. H. & N. Pedersen (2019) Et overblik over graden af responselicitering for det efterstillede påhæng ikk og formerne deraf (ikke, ikkå, ikkås og ing). I: *Skrifter om samtalegrammatik* 6 (1):1-37.
- Keevallik, L. (2018) What Does Embodied Interaction Tell Us About Grammar? *Research on Language and Social Interaction* 51(1):1-21. <https://doi.org/10.1080/08351813.2018.1413887>.
- Kendon, A. (1967) Some functions of gaze-direction in social interaction. *Acta Psychologica* 26:22-63. [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4).
- Kjærbeck, S., og B. Asmuß (2005) Negotiating meaning in narratives: An investigation of the interactional construction of the punchline and the post punchline sequence. *Narrative Inquiry* 15(1):1-24.
- Labov, W., og J. Waletzky (1967) Narrative Analysis. I J. Helm (red.). *Essays on the Verbal and Visual Arts*. Washington: University of Washington Press:12-44.
- Lindström, J., og P. Linell (2007) Roli å roli: x-och-x som samtalspraktik och grammatisk konstruktion. I E. Engdahl og A.-M. Londen (red.). *Interaktion och kontext: Nio studier av svenska samtal*. Lund: Studentlitteratur:19-89.
- Linell, P. (2009) *Rethinking language, mind, and world dialogically: interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Pub.
- Mikkelsen, N. (2023) Parenthesis in storytelling in Danish talk-in-interaction. Artikel under forberedelse.
- Mondada, L. (2019) *Conventions for multimodal transcription*. <https://www.lorenzamondada.net/multimodal-transcription>.
- Møller, E. (1993) *Mundtlig fortælling: Fortællingens struktur og funktion i uformel tale*. København: C.A. Reitzels Forlag.
- Nevile, M. (2015) The Embodied Turn in Research on Language and Social Interaction. *Research on Language and Social Interaction* 48(2):121-151. <https://doi.org/10.1080/08351813.2015.1025499>.

- Oshima, S. (2014) Consensus through professionalized head nods: The role of nodding in service encounters in Japan. *Journal of Business Communication* 51(1):31-57. <https://doi.org/10.1177/2329488413516207>.
- Pomerantz, A. (1986) Extreme case formulations: A way of legitimizing claims. *Human Studies* 9(2-3):219-30. <https://doi.org/10.1007/BF00148128>.
- Sacks, H. (1974) An analysis of the course of a joke's telling in conversation. I J. Sherzer & R. Bauman (red.). *Explorations in the Ethnography of Speaking*. London: Cambridge University Press:337-353.
- SamtaleBank. (u.å.) Hentet 5. juni 2017, fra <http://samtalebank.talkbank.org/>.
- Sidnell, J. (2012) Basic conversation analytic methods. I J. Sidnell og T. Stivers (red.). *The Handbook of Conversation Analysis*. Oxford: Wiley-Blackwell:77-99.
- Stivers, T. (2008) Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language and Social Interaction* 41(1):31-57. <https://doi.org/10.1080/08351810701691123>.

Resumé:

“Vi er nogle der bruger det!” – Et historisk perspektiv på danske kønsinklusive pronomener fra 1942 til 2021

Ehm Miltersen

Aarhus Universitet

Kønsinklusive (eller kønsneutrale) pronomener har fået en del opmærksomhed fra det offentlige øje de seneste år, både i Danmark og i udlandet, men er som akademisk emne forholdsvis underbelyst. Der findes enkelte studier af kønsinklusive pronomener på forskellige sprog, såsom dansk (Miltersen 2018; Miltersen 2020), svensk (Sendén et al. 2021), spansk (Papadopoulos 2022), fransk (Knisely 2020) og italiensk (Scotto di Carlo 2020), samt en større mængde litteratur om engelske pronomener, inklusiv sproghistoriske studier (eksempelvis Baron 2020). Dette oplæg tager udgangspunkt i mit speciale fra foråret 2022, der undersøgte den historiske udvikling af danske kønsinklusive pronomener. Jeg definerer her kønsinklusive pronomener som personlige pronomener i 3. person ental, der, når de tager personreference, enten ikke angiver referentens konceptuelle køn (McConnell-Ginet 2014; Ackerman 2019; jf. *sexus* i Hansen og Heltoft 2011) eller angiver referentens konceptuelle køn som noget andet end mand eller kvinde, på sprog hvor pronomener i samme paradigme ellers angiver konceptuelt køn. På nutidigt dansk drejer dette sig om pronomenerne *hen*, singularært *de* og *den*, i paradigme med *han* og *hun* (Miltersen 2018), omend en række andre pronomener er blevet foreslået, diskuteret og i begrænset omfang brugt gennem tiden, eksempelvis *hyn*, *høn*, *hæn*, *huan* og *hin*.

Specialet bruger en multimetodologisk tilgang til at analysere en kombination af dokumentdata (Bowen 2009) og data fra kvalitative interviews med personer fra nogle af de *speech communities / communities of practice* (Labov 1972; Eckert og McConnell-Ginet 1998), der i dag bruger pronomenerne. Data er blevet indsamlet og analyseret med metoder fra korpuslingvistik, lingvistisk etnografi og diskursanalyse. Dertil har jeg suppleret med et autoetnografisk perspektiv og inddraget indsigter fra min egen tilegnelse af de relevante pronomener. På baggrund af 242 eksempler på pronomener, indhentet fra diverse korpora og andre kilder, samt personlige beretninger om interviewpersonernes (og mine egne) første møder og erfa-

ringer med pronomenerne, konstruerer specialet en tidslinje over danske kønsinklusive pronomeners opståen og udvikling, der strækker sig fra 1942 til 2021. Ud over den kronologiske beskrivelse identificerer jeg også en række relevante diskurser om sprog og køn og disses udvikling.

Tidslinjen indikerer, at denne udvikling kan karakteriseres af to generelle 'bølger', der følger andre samfundsmæssige udviklinger i forhold til køn(sopfattelser): Først giver fremskridt inden for kvinderettigheder anledning til diskussioner om kønsneutralt sprog i perioden 1969-1983; senere, efter den gradvise udvikling af queerteori i 90'erne og 00'erne, fører en øget synlighed af trans og nonbinære personer til større opmærksomhed på kønnet og kønsinkluderende sprog. Afslutningsvis vil jeg i oplægget kort diskutere nogle sprogpoltiske implikationer af studiet.

Litteratur

- Ackerman, Lauren (2019) Syntactic and cognitive issues in investigating gendered coreference. I *Glossa: A Journal of General Linguistics*, 4(1). 117.
- Baron, Dennis (2020) *What's your pronoun? Beyond he & she*. New York: Liveright.
- Bowen, Glenn A. (2009) Document Analysis as a Qualitative Research Method. *Qualitative Research Annual* 9(2):27-40.
- Eckert, Penelope og Sally McConnell-Ginet (1998) *Language and Gender*. Cambridge: Cambridge University Press.
- Hansen, Erik og Lars Heltoft (2011) *Grammatik over Det Danske Sprog*. Odense: Syddansk Universitetsforlag.
- Knisely, Kris (2020) Le français non-binaire: linguistic forms used by non-binary speakers of French. *Foreign Language Annals*, 53(4):850-876.
- Labov, William (1972) On the Mechanism of Linguistic Change. I John Gumperz and Dell Hymes (red.) *Directions in Sociolinguistics: The Ethnography of Communication*. John Wiley & Sons:513-536.
- McConnell-Ginet, Sally (2014) Meaning-Making and Ideologies of Gender and Sexuality. I Susan Ehrlich, Miriam Meyerhoff, og Janet Holmes (red.) *The handbook of language, gender, and sexuality*. John Wiley & Sons:316-334.
- Miltersen, Ehm (2018) De, den, hen, and the rest. *Journal of Language Works – Sprogvidenskabeligt Studenter-tidsskrift* 3(1):31-42.
- Miltersen, Ehm (2020) Singular de and its referential use in talk-in-interaction. *Scandinavian Studies in Language* 11(2):22-61.
- Papadouopilos, Ben (2022) A brief history of gender-inclusive Spanish/Una breve historia del español no binario. *Deportate, esuli, profughe*, 48(1):31-48.

“Vi er nogle der bruger det!” – Et historisk perspektiv på danske kønsinklusive pronomener fra 1942 til 2021

- Scotto di Carlo, Genna (2020) An analysis of the use of inclusive language among Italian nonbinary individuals: a survey transcending binary thinking. Beyond binary thinking through inclusiveness: interdisciplinary reflections and perspectives. *I-LanD Journal: Identity, Language and Diversity*, 2:70-89.
- Sendén, Marie Gustafsson, Emma A. Renström og Anna Lindqvist. (2021) Pronouns beyond the binary: The change of attitudes and use over time. *Gender & Society* 35(4):588-615.

Resumé:

”Har du set *Nor*, efter *de* har fået et nyt job?”

Et acceptabilitetsstudie af singulært *de*

Maja Mittag & Line Burholt Kristensen

Københavns Universitet

Analogt med engelsk *they* (Bjorkman 2017) er der i de senere år kommet en ny anvendelse af *de* med singularisreference. Singulært *de* er et kønsneutralt alternativ til pronomenerne *han* eller *hun*, der bl.a. kan bruges om mennesker med en nonbinær kønsidentitet. Singulært *de* er det mest populære alternativ til *han* og *hun* – mere end dobbelt så mange personer foretrækker at blive omtalt med *de* fremfor *hen* (Miltersen 2018).

Den nye anvendelse af *de* kan give anledning til nye former for tvetydige sætninger, hvor det er uklart om *de* refererer til en entals- eller flertalsreferent. Vi undersøger her om den nye brug har negative konsekvenser for sætningsprocesseringsmål. Oplægget præsenterer resultaterne fra et todelt læsetids- og acceptabilitetsstudie med sætninger, der indeholder singulært *de* overfor tilsvarende sætninger med *han* eller *hun*, fx (1).

- (1) Efter jeres familiemedlem er blevet opereret, skal *de/han/hun* forholde sig i ro i et døgn.

I to eksperimenter (et online og et on-site med hhv. 52 og 37 deltagere i aldersgruppen 18-35 år) blev deltagerne bedt om at læse en række sætninger på tid og dernæst bedømme dem på en kontinuert skala fra “Slet ikke acceptabel” til “Helt acceptabel”.

Foertsch og Gernsbacher (1997) har vist at korrelatets specificitet påvirker processeringen af engelske sætninger med singulært *they*. Derfor undersøgte eksperimenterne også hvordan korrelatets grad af specificitet påvirker en sætnings acceptabilitet. Hver stimulisætning indeholdt en af tre korrelattyper med varierende grad af specificitet, fx *en patient, jeres familiemedlem* og *din kusine* (fra mindst til mest specifik), samt et af tre personlige pronomener i singularis (*han, hun* eller *de*).

Overordnet set blev sætningerne med singulært *de* anset som mindre acceptable end tilsvarende med *han* eller *hun*. Som forventet var forskellen mellem acceptabilitetsbedømmelserne af *han/hun* og singulært *de* større, des mere specifikt korrelateret var. I begge eksperimenter blev deltagernes læsetider også registreret, men her var, mod forventningen, ingen effekt af pronomenvalg.

Oplægget vil diskutere hvordan korrelatets grad af specificitet og deltagernes kendskab til LGBT+-miljøet påvirker acceptabilitetsbedømmelserne. Desuden diskuteres det om singulært *de* styrer kongruens i flertal eller ental.

Litteratur

- Bjorkman, Bronwyn M. (2017) Singular they and the syntactic representation of gender in English. *Glossa: A Journal of General Linguistics*, 2(1), artikel 80.
- Foertsch, Julie & Gernsbacher, Morton Ann (1997) In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he? *Psychological Science*, 8(2):106-111.
- Miltersen, Ehm (2018) De, den, hen, and the rest: a pilot study of the use of gender-neutral and nonbinary/genderqueer pronouns in Danish. *Journal of Language Works*, 3(1):31-42.

Resumé:

Sprogholdninger i center og periferi i Nordjylland og Sønderjylland

Malene Monka, Marie Maegaard & Nikolaj Pharao
Københavns Universitet

Et nyere studie viste forskellige mønstre for sprogbrug og sprogholdninger blandt unge i Hirtshals i Nordjylland og Bylderup i Sønderjylland (Maegaard et al. 2020). De unge i Nordjylland brugte ingen traditionelle dialekttræk i deres hverdagsprog; det gjorde en del af de unge sønderjyder (Maegaard & Monka 2020; Monka 2020). Et sprogmasketeksperiment med 12 stemmer (to drenge og to piger, der repræsenterer hhv. lokalt, regionalt og københavnsk talesprog) viser, at de unge fra Nordjylland ikke genkender lokale stemmer som nordjyske, og de sproglige forskelle på tværs af stemmerne har også begrænset indflydelse på bedømmelsen. En lignende masketest i Sønderjylland med 12 stemmer (to drenge og to piger i hhv. en dialekt- og en standardnær maske, samt to københavnske drenge- og pigestemmer) viste, at de unge genkendte dialektmaskerne som lokale og tilskrev de regionale masker til den lokale købstad Aabenraa. I forhold til tidligere danske sprogholdningsstudier er det bemærkelsesværdigt, at de unge sønderjyder overordnet set bedømmer de københavnske stemmer mere negativt end de lokale dialekttalende stemmer (Maegaard & Quist 2020). Her sammenlignes resultaterne af ovennævnte sprogholdningsundersøgelser med sprogholdningsundersøgelser fra de regionale centre Aalborg og Aabenraa. Ved at bruge samme design og stemmer kan vi undersøge forskelle mht. orientering mod københavnsk mellem de to regioner. Resultaterne fra Aalborg adskiller sig ikke nævneværdigt fra resultaterne fra Hirtshals. I Aabenraa afviger resultaterne til gengæld fra resultaterne fra Bylderup. De unge i Aabenraa nedvurderer de lokale stemmer og opvurderer de københavnske. Resultaterne afspejler forskelle i lokalhistorien i de to områder og de forskellige typer periferi, man finder i hhv. Nordjylland og Sønderjylland. Dette fører til, at den lokale sprogbrug tilskrives forskellig værdi. Resultaterne tolkes inden for rammen *sociolinguistic change* (Androutsopoulos 2016, Coupland 2014, 2016, Maegaard m.fl. 2020).

Litteratur

- Androutsopoulos, Jannis (red.) 2014. *Mediatization and Sociolinguistic Change*. Berlin: De Gruyter.
- Coupland, Nikolas. 2014. Sociolinguistic Change, Vernacularization and Broadcast British Media. *Mediatization and Sociolinguistic Change*. Jannis Androutsopoulos (red.) Berlin/Boston: De Gruyter. 67-96.
- Coupland, Nikolas. 2016 Labov, Vernacularity and Sociolinguistic Change. *Journal of Sociolinguistics* 20: 4. 409-30.
- Maegaard, Marie & Malene Monka. 2020. Pattern of Dialect Use: Language Standardization at Different Rates. *Standardization as Sociolinguistic Change: A Transversal Study of Three Traditional Dialect Areas*. Marie Maegaard, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (red.). New York: Routledge. 27-46.
- Maegaard, Marie, Malene Monka, Kristiane Køhler Mortensen & Andreas Candefors Stæhr (red.). 2020. *Standardization as Sociolinguistic Change*. Routledge Studies in Language Change. New York: Routledge.
- Maegaard, Marie & Pia Quist. 2020. Perception, Recognition, and Indexicality: Experimental Investigations of Variation in Northern Jutland, Southern Jutland, and on Bornholm. *Standardization as Sociolinguistic Change: A Transversal Study of Three Traditional Dialect Areas*, edited by Marie Maegaard, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (red.) New York: Routledge. 145-68.
- Monka, Malene. 2020. Southern Jutland; Language Ideology as a Means to Slow Down Standardization. *Standardization as Sociolinguistic Change: A Transversal Study of Three Traditional Dialect Areas*. Marie Maegaard, Malene Monka, Kristine Køhler Mortensen & Andreas Candefors Stæhr (red.) New York: Routledge. 70-118.

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”

Når performansen halter

Anne Mette Nyvad og Ken Ramshøj Christensen

Aarhus Universitet

1. Indledning¹

På dansk kongruerer subjektspredikativet i genus og numerus med subjektet, når førstnævnte er et adjektiv (fx *Han er wild*), men dette mønster brydes tilsyneladende hos nogle sprogbrugere i uformelle varianter af (talt og skrevet) dansk, når det prædikative adjektiv tager en præpositionsfrase (PP) med et præpositionsobjekt (P-Obj) som komplement. Et prædikativt adjektiv kan være bøjet i pluralis, selvom subjektet er i singularis, hvis P-Obj er i pluralis, og selv hvis P-Obj er i akkusativ, som i (1) og (2)a. Efter alt at dømme kongruerer det prædikative adjektiv i stedet med P-Obj, da strukturelt parallelle konstruktioner, hvor det prædikative adjektiv hverken kongruerer med subjektet eller P-Obj, ikke synes at forekomme (eller i hvert fald er meget sjældne, jf. også korpussøgningen i Engberg-Pedersen & Poulsen 2010), (5)):

(1) **Kongruens med P-Obj i pluralis**

- a. Han_{sg} er **vilde**_{pl} med dem_{pl}.
- b. Dem_{pl} er han_{sg} **vilde**_{pl} med.

- (2) a. Jeg undskylder for meget tørre neglebånd -
jeg har ikke været **gode** ved dem det sidste stykke tid ;) [appeal4.dk/blogs]
- b. De her er jeg **vilde** med ♥ en skøn dag med babyspam og tøsefnidder.
[instagram]

¹ Tak til Michael Nguyen for konstruktiv kritik på en tidligere udgave af denne artikel, og tak til tilhørerne på MUDES 19 (14. oktober 2022) for spørgsmål og kommentarer. Forskningen blev delvist støttet af Det Frie Forskningsråd (Bevillings-ID: DFF-9062-00047B).

(3) **Kongruens med P-Obj i singularis**

- a. Vi_{pl} er **vild**_{sg} med den_{sg}.
- b. Den_{sg} er vi_{pl} **vild**_{sg} med.

- (4) a. [H]an kalder sig ”amatør fotograf” – Vi er **ligeglåd** med det. [babyklar.dk]
- b. Danmark vandt EM i fodbold [...],
men det er vi **ligeglåd** med. [bilmagasinet.dk]

(5) **Ingen kongruens**

- a. *Han_{sg} er **vilde**_{pl} med den_{sg}
- b. *Den_{sg} er han_{sg} **vilde**_{pl} med

Som det ses i (1)-(2) og (3)-(4), er der på det syntaktiske plan overordnet to måder, hvorpå den atypiske kongruens kan komme til udtryk: I a-eksemplerne i (1)-(4) står styrelsen (P-Obj) på sin vante plads efter præpositionen, mens det i b-eksemplerne er flyttet frem i sætningen. Den atypiske kongruens kan også blive udløst i sætninger, hvor P-Obj er flyttet på tværs af en sætningsgrænse, som i (6), og faktisk er det ikke kun et P-Obj, der kan konkurrere om kongruensen, se fx (7):

- (6) Rigtig søde øringer og armbånd, dem er jeg **sikre** på de bliver rigtig glade for
[bryllupsklar.dk/forum]

- (7) Det er ikke dem der holder Putin **vågne** om natten.
[DR TV Avisen, 29.04.2022, ca. 18:35]

Men kan fænomenet betragtes som en del af det grammatiske system i dansk (dvs. er der evidens, der taler for det?), eller er det en performansrelateret fejl, der opstår som følge af manglende opmærksomhed, forvirring eller belastning af arbejdshukommelsen?

2. Tidligere behandlinger af fænomenet

En korpusundersøgelse af Engberg-Pedersen & Poulsens (2010) fandt en forekomst på gennemsnitligt 2,4 % og 5,5 % for atypisk kongruens med et P-Obj i henholdsvis singularis og pluralis (sammenlignet med utvetydig standardkongruens, dvs. tilfælde hvor subjektskongruens er tyde-

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

lig, da P-Obj har et andet tal end subjektet). Adjektivernes ubøjede form er generelt langt mere frekvent end den bøjede, men Engberg-Pedersen & Poulsens korpussøgning viste desuden, at når det kommer til forekomster af atypisk kongruens, er den *bøjede* form af det prædikative adjektiv mere frekvent, og således synes det ikke muligt at forklare fænomenet med henvisning til frekvens.

I et læsetidsforsøg fandt Engberg-Pedersen og Poulsen (2010), at sætninger uden kongruens af typen i (8)d blev læst langsommere end de sætninger, hvor der var kongruens mellem det prædikative adjektiv og P-Obj, (8)c. Måske mest overraskende er dog, at der ikke var nogen signifikant forskel i læsetid mellem sætninger med den atypiske kongruens, (8)c, og dem med standardkongruens (med subjektet), (8)b (eksempler fra Engberg-Pedersen & Poulsen 2010: 223):

- (8) a. **Fuld kongruens:**
Den bog var hun helt **vild** med for et år siden
- b. **Subjektskongruens:**
De bøger var hun helt **vild** med for et år siden
- c. **Kongruens med P-Obj:**
De bøger var hun helt **vilde** med for et år siden
- d. **Ingen kongruens:**
Den bog var hun helt **vilde** med for et år siden

I Christensen & Nyvad (2019) ville vi se, (i) om vi kunne elicitere den atypiske kongruens i et sætningsfuldførelsesstudie, og (ii) om numerus spiller en rolle i udløsningen af fænomenet. Vi havde en formodning om, at et fremflyttet P-Obj i pluralis, (9)a, i højere grad end et i singularis, (9)b, ville være i stand til at tiltrække den atypiske kongruens, da den kan siges at være mere prominent, markeret og svær (se Franck m.fl. 2002; se også Kandel m.fl. 2022, Staub 2009):

- (9) a. Jeg er helt vild med kager. Og dem her, dem er jeg specielt ___ med.
SVAR: _____
- b. Vi er helt vilde med kager. Og den her, den er vi specielt ___ med.
SVAR: _____

Vi testede 14 forskellige prædikative adjektiver: *flov over*, *ligeglad med*, *stolt af*, *tilfreds med*, *tryk ved*, *træt af*, *vild med*, *enig med*, *glad for*, *god ved*, *ked af*, *sikker på*, *sur på*, *vred på*.

Vores resultater viste imidlertid den modsatte tendens: Standardformen blev brugt signifikant oftere med et P-Obj i pluralis (83 %) end den alternative form, hvorimod standardformen og den alternative form af det prædikative adjektiv blev anvendt stort set lige ofte, når P-Obj var i singularis, fx (9)b. Med andre ord blev den *ubøjede* form brugt oftere end den *bøjede* form i sætninger med atypisk kongruens.

Vi forklarede disse data med henvisning til en kombination af flere forskellige faktorer, der hver især belaster arbejdshukommelsen og dermed øger risikoen for performansrelaterede fejl uafhængigt af P-Obj's numerus: For det første overtræder sætninger som i (9) præferencehierarkierne for ordstilling i (10) (se Newmeyer 2005: 207), da første lineære element, P-Obj, konsekvent var 3. person, mens subjektet enten var 1. eller 2. person.

- (10) a. Person: ({1., 2.} > 3. person)
b. Relation: Subjekt > Objekt > Oblik

For det andet optræder subjektet oftere i fundamentfeltet end nogen anden syntaktisk funktion (61 % i de skriftlige kilder, (61 % i de skriftlige kilder, Mikkelsen 2015 undersøgte, og meget sammenligneligt 62 % i tale ifølge Puggaard 2019), og når man processerer en sætning, hvor det den syntaktiske funktion af det første nominale led er tvetydigt, bliver det først tolket som subjekt (Friederici m.fl. 2001). Fremflytningen af P-Obj, som i b-eksempler i (1)-(4), fører således muligvis til en reanalyse af sætningen, når man umiddelbart efter det finite verbum kommer til subjektet, men fejlfortolkningen glemmes sandsynligvis ikke (jf. Christianson m.fl. 2001), og det kan påvirke konkurrencen til fordel for kongruens med det prædikative adjektiv. Sidst men ikke mindst kan en del af tilfældene muligvis tilskrives bortfald af endelser (fx apokope) på adjektivet i dansk.²

Ingen af disse tidligere studier har gransket effekten af fremflytning i relation til forekomsten af den atypiske kongruens. I et andet studie, Christensen & Nyvad (2021), testede vi de samme sætninger som i Christensen & Nyvad (2019), men uden fremflytning af P-Obj:

² I Christensen & Nyvad (2019, 2021) kaldte vi fænomenet ”alternativ kongruens”. Men fordi vi ikke fandt belæg for, at fænomenet er grammatikaliseret topic- eller fokuskongruens, og dermed ikke et reelt alternativ til standard subjektkongruens, refererer vi her til fænomenet som atypisk kongruens.

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

- (11) a. Jeg er helt vild med kager, især dem her. Jeg er specielt ____ med dem her.

SVAR: _____

- b. Vi er helt vilde med kager, især den her. Vi er specielt ____ med den her.

SVAR: _____

I sætninger af typen i (11) var der stort set ingen atypisk kongruens, mens et P-Obj i singularis tiltrak kongruensen i 14 % af tilfældene. I dette stimulusæt var der på grund af den kanoniske ordrækkefølge ingen overtrædelse af præferencehierarki eller mulighed for fejlfortolkning af det første nominale led, så vi forklarede mønstret i data med henvisning til bortfald af endelser på adjektivet. Data fra Christensen & Nyvad (2019, 2021) viser således, at den atypiske kongruens overordnet er mere sandsynlig, når der er fremflytning af P-Obj, end når det står *in situ* (dvs. på sin vante plads efter præpositionen).

En anden af teserne i Christensen & Nyvad (2019, 2021) opstod på baggrund af den type fora, hvor vi oftest fandt eksempler på atypisk kongruens, nemlig i blogosfæren: Vi forventede derfor, at den atypiske kongruens muligvis er mere udbredt blandt yngre kvinder, eller måske at fænomenet kunne være knyttet til faktorer som dialekt eller uddannelseslængde. Ingen af de sociolingvistiske faktorer, vi inkluderede (dvs. køn, alder, uddannelsesniveau, region), viste sig at være korreleret med forekomsten af fænomenet, hvilket er kompatibelt med formodningen om, at den atypiske kongruens er en performansrelateret fejl. Yderligere støtte til denne antagelse kommer fra det faktum, at vi fandt en positiv korrelation mellem verbet *er* (i præsensformen) + adjektivformens frekvens i KorpusDK og standardkongruens:³ Øget hyppighed fører givetvis til mindre belastning af arbejdshukommelsen, som i sidste ende nedsætter risikoen for performansrelaterede fejl og dermed forekomsten af atypisk kongruens.

3. Bedømmelser af sætninger med atypisk kongruens

Hvilke konstruktioner eller udtryk folk bruger, hænger ikke altid sammen med, hvad de finder er acceptable former (jf. fx udtalen af *håndklæder* som *hångklæder* eller brugen af *som at der* eller diverse sætningsknuder) (Christensen & Nyvad 2023). Hverken Engberg-Pedersen & Poulsens (2010) studier (der granskede læsetid og udbredelse i korpora) eller vores elicitationsstudier fortæller os, hvad de danske sprogbrugere rent faktisk synes om sætninger med atypisk

³ Tilføjelsen af *er* sikrer, at der er tale om et prædikativt, og ikke et attributivt, adjektiv.

kongruens, men det er det spørgsmål, som vi søgte et svar på i nærværende acceptabilitetsstudie. Vores forudsigelser var baseret på fund fra de tidligere studier: Givet at standardkongruensen er langt mere frekvent i korpusser end den atypiske kongruens, havde vi en formodning om, at den også ville blive bedømt til at være signifikant mere acceptabel end den atypiske kongruens. Derudover forventede vi, at der ville være en effekt i bedømmelsesniveauerne af numerus og fremflytning, sådan som vi også så det i elicitationsstudierne, således at (i) et P-Obj i singularis og (ii) et fremflyttet P-Obj i højere grad var i stand til at tiltrække kongruensen.

3.1 Design

Vi ville gerne finde ud af, hvad sprogbrugerne synes om naturligt forekommende eksempler, så vi fandt eksempler på sætninger med kontekst på Google⁴, som vi kun modificerede i det omfang, det var nødvendigt for at udvikle et stimulussæt med minimalt forskellige par, der tog højde for forskellige lingvistiske faktorer, såsom person og tal, stavfejl, osv.

Det endelige stimulussæt bestod af 96 målsætninger: 12 forskellige prædikative adjektiver (*enig med, glad for, god ved, ligeglad med, sikker på, stolt af, sur på, tilfreds med, træt af, tryk ved, vild med og vred på*), der optrådte med og uden standardkongruens, samt uden og med fremflytning af P-Obj. Numerus varierede systematisk på tværs af disse sætningstyper.

- (12) a. Disse tasker er fundet i en genbrugsbutik i Berlin.
Jeg er vild/e med dem/Dem er jeg vild/e med, fordi der er så mange farver, og pailletterne er for skønne.
- b. Denne taske er fundet i en genbrugsbutik i Berlin.
Vi er vild/e med den/Den er vi vild/e med, fordi der er så mange farver, og pailletterne er for skønne.

Derudover bestod stimulussættet af 17 kontrol- eller fyldsætninger (*fillers*) af tre forskellige typer: Filler 1 var klart grammatiske sætninger, se (13), filler 2 var klart ugrammatiske sætninger, der overtrådte *The Coordinate Structure Constraint* (Ross 1967: 161), der udelukker fremflytning fra en koordination, (14), mens filler 3 bestod af sætninger, hvor der manglede nutids-*r*, (15):

⁴ Vi søgte på Google.dk og med "site:dk" for at sikre at det kun var danske hjemmesider, der med en vis sandsynlighed var skrevet af danskere, og ikke maskinoversatte sider fra fx *Facebook* eller *Traveladvisor*.

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

- (13) Filler 1: Det er klart, at vi blev glade for, at gæsterne faktisk nød både pindemadderne og vinen til arrangementet.
- (14) Filler 2: *Det er vinen, som vi blev glade for, at gæsterne faktisk nød både pindemadderne og til arrangementet.
- (15) Filler 3: *Jeg forstår virkelig ikke hvordan man **parallelparkere** (men jeg skal nok lære det på et tidspunkt).

Disse sætninger blev fordelt i et Latin-square-design og i randomiseret rækkefølge på 8 lister med hver især 12 målsætninger og 13 fillers, og sat ind i et spørgeskema på Google Forms, hvor forsøgspersoner blev bedt om at vurdere sætningerne på en skala fra 1 (”helt uacceptabelt”) til 7 (”helt acceptabelt”). Deltagerne blev bedt om at basere deres bedømmelser på deres egen intuition, ikke på hvad de evt. mente var ’korrekt’ eller ’standard’. De blev også bedt om at se bort fra tegnsætning eller kommatering. Forsøgspersonerne blev rekrutteret på sociale medier.

3.2 Resultater

197 forsøgspersoner deltog (196 kvinder, 1 mand).⁵ Gennemsnitsalderen var 32 (18-72, SD = 10) og deres gennemsnitlige uddannelseslængde var 17 år (5-25, SD = 3).

Bedømmelsesdataene blev z-transformeret per deltager for at kontrollere for idiosynkratisk brug af skalaen⁶, og der blev kørt en *mixed-effects*-model, hvor de z-transformerede bedømmelsesdata blev modelleret som en funktion af kongruens, fremflytning og numerus, med *random slopes* per deltager for trevejsinteraktion mellem kongruens, fremflytning og numerus og *random intercepts* for *item*. Resultaterne ses i Tabel 1:

⁵ Det er en meget kraftig kønsbalance, men eftersom vi ikke har nogen grund til at forvente en kønsforskel i den ene eller anden retning i acceptabilitetsbedømmelserne, ser vi ikke det som et problem. Vi har desuden tidligere ikke fundet nogen signifikant effekt af køn ved elicitation af kongruens (Christensen & Nyvad 2019, 2021).

⁶ Den z-transformerede bedømmelse angiver antallet af standardafvigelser den rå (ikke-transformerede) bedømmelse er fra deltagerens gennemsnitsbedømmelse af den type sætning. Z-transformation er en måde at kontrollere for potentiel individuel bias i bedømmingsskalaen, fx hvis nogen kun bruger den ene ende af skalaen eller en mindre del af den. (Sprouse m.fl. 2012, Nyvad m.fl. 2022).

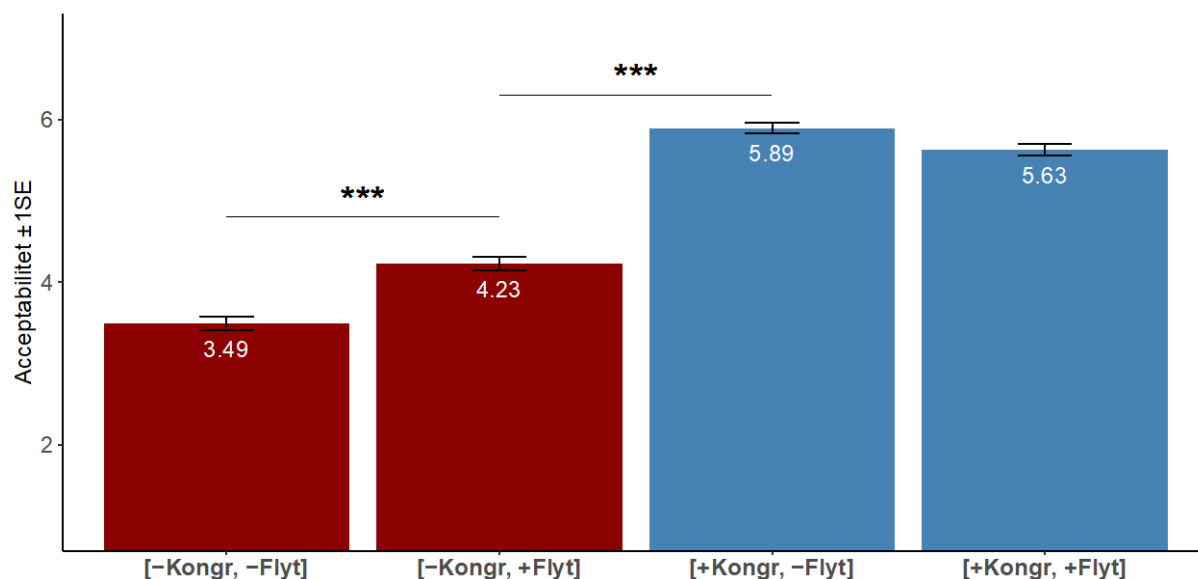
	estimat	SE	t	df	p	
(Intercept)	-0,32	0,08	-4,05	91,85	0,000	***
Kongruens [+Kongruens]	1,03	0,11	8,98	104,9	0,000	***
Fremflytning [+Fremflytning]	0,25	0,11	2,32	85,38	0,022	*
Numerus [Singularis]	-0,04	0,11	-0,37	85,94	0,716	
Kongruens × Fremflytning	-0,27	0,15	-1,76	85,36	0,082	
Kongruens × Numerus	0,06	0,15	0,36	86,92	0,721	
Fremflytning × Numerus	0,15	0,15	0,99	84,82	0,324	
Kongruens × Fremflytning × Numerus	-0,32	0,22	-1,48	84,82	0,143	

Tabel 1. Opsummering af den statistiske model. (***) betyder $p < 0,001$, * $p < 0,05$. SE = standard error, df = degrees of freedom.)

Der var ingen signifikante interaktioner ($p > 0,1$) (selvom interaktionen mellem kongruens og fremflytning var tæt på, $p = 0,082$). Mens der ikke var nogen signifikant effekt af numerus ($p > 0,7$), var der til gengæld en signifikant effekt af kongruens ($p < 0,001$) og af fremflytning ($p = 0,022$).

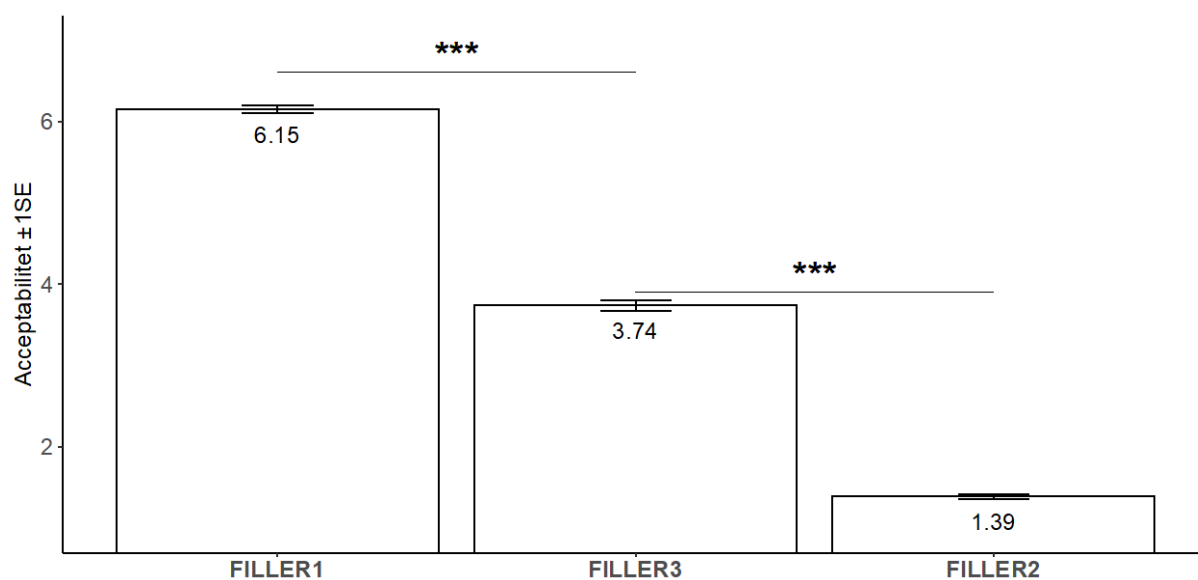
Helt som ventet er bedømmelserne af sætningerne med subjektkongruens højere end dem med atypisk kongruens (med P-Obj), sammenlign de blå med de røde søjler i Figur 1. Modsat vores forudsigelse var der dog ingen signifikant effekt af numerus, men der var en klar effekt af fremflytning: Sætningerne med den atypiske kongruens blev bedømt til at være markant bedre, når P-Obj var det første nominale led ($p < 0,001$), og det er helt i tråd med Christensen og Nyvad (2019, 2021), der bl.a. fandt, at en nærmest 50/50-fordeling mellem standardkongruens og atypisk kongruens i eliciteringsdata ved sætninger med fremflytning af et P-Obj i singularis. Det er bemærkelsesværdigt, at fremflytning gør den atypiske kongruens mere acceptabel (de røde søjler i figuren), når vi ved fra uafhængige studier, at fremflytning normalt reducerer acceptabiliteten (Christensen m.fl. 2013, Christensen & Nyvad 2014, Christensen & Nyvad 2023, Fanselow & Frisch 2006). Pudsigt nok forsvinder effekten af fremflytning i bedømmelserne, når sætningerne indeholder standardkongruens med subjektet (der er ikke signifikant forskel på højden af de to blå søjler i figuren).

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter



Figur 1: Gennemsnitsacceptabilitet per type for målsætninger. (***) betyder $p < 0,001$.)

Sammenholdt med bedømmelserne for den ugrammatiske filler 2, **Det er vinen, som vi blev glade for, at gæsterne faktisk nød både pindemadderne og til arrangementet*, (14), var bedømmelserne af sætningerne med atypisk kongruens relativt høje og (i hvert fald efter øjemål⁷) på niveau med sætningerne med manglende nutids-*r*, se Figur 2.



Figur 2: Gennemsnitsacceptabilitet per filler-type (sorteret efter faldende acceptabilitet).

⁷ Eftersom målsætningerne og de tre filler-typer ikke var minimalt forskellige, giver det ikke mening at sammenligne dem statistisk. Fx er målsætningen i (12)a og filler 3 i (15) forskellige fra hinanden på mange måder, både leksikalsk, semantisk og syntaktisk.

4. Diskussion

I studierne i Christensen og Nyvad (2019, 2021) fandt vi, at den atypiske kongruens mellem prædikativt adjektiv og P-Obj er *let at elicitere*, men at standardkongruens med subjekt generelt var væsentligt *lettere* at elicitere. Nærværende studie viste, at den atypiske kongruens er signifikant mindre acceptabel end standardkongruens (på niveau med manglende nutids-*r*), men klart bedre end ukontroversielt ugrammatiske sætninger. Fremflytning øger acceptabiliteten signifikant, men der er til gengæld overraskende nok ingen effekt af numerus. Disse fund er i tråd med antagelsen af en vid grad af produktivitet og systematik i udbredelsen af den atypiske kongruens, og at danske sprogbrugere ikke umiddelbart opfatter den atypiske kongruens som decideret uacceptabel.

Men hvordan kan fænomenet forklares? I det følgende vil vi forsøge at belyse dette spørgsmål ved at trække på fonetiske (Grønnum 2005), syntaktiske (Vikner 2001) (Baker 2008, Belletti 2001) og informationsstrukturelle (Comrie 2003) faktorer, som alle potentielt kunne spille en rolle i fænomenets opståen. Med Occams ragekniv i hånden må vi dog slutteligt erkende, at fænomenet sandsynligvis bedst kan forklares af simple, uafhængige faktorer.

4.1 Fonetiske faktorer

Ifølge Hansen (1967: 368–369) optræder den ubøjede form af prædikative adjektiver ofte i stedet for det pluralisbøjede ”i talesproget og talesprogsnært skriftsprog”. I både Christensen & Nyvad (2019, 2021) og nærværende studie af acceptabilitetsbedømmelser har vi i nogen grad forsøgt at imitere talesprog. Ord, der (ligesom singularis- og pluralisformen af prædikative adjektiver) udtales ens eller tæt på ens, bliver jævnligt forvekslet med hinanden i usikre sprogbrugeres skriftsprog, og det atypiske kongruensmønster kunne således være resultatet af ”en usikker omsættelse af tale til skrift” (Jensen 2004: 165). Det kan dog ikke være hele forklaringen, da der faktisk er en klar udtalemæssig forskel mellem eksempelvis den ubøjede form *vild* og den pluralisbøjede prædikative adjektiv *vilde* (Hansen 1990, Brink m.fl. 1991): Førstnævnte involverer stød, mens schwa i endelsen på pluralisformen typisk assimilerer med den forrige vokal og/eller sonorant, uden tab af stavelser (Grønnum (2005).

Engberg-Pedersen & Poulsen (2010: 230) argumenterer for, at den atypiske kongruens kan være opstået som følge af spontan fonetisk assimilation, der kumulativt har påvirket sprogbrugers efterfølgende sproglige bearbejdning af lignende sætninger. På sigt kan det have ført til, at disse performansudløste afvigelser er blevet optaget som en del af sprogsystemet. Imidlertid ville man i så fald forvente, at et tilsvarende atypiske bøjningsmønster ville optræde spontant, når der ikke er et nominal til stede, der kan konkurrere om kongruensen, fx når præposi-

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

tionsobjektet ikke er til stede, jf. (16) og når de pågældende adjektiver optræder attributivt, jf. (17), som påpeget i Nyvad (2007: 9):

- (16) a. Han er vild/*vilde
- b. Vi er vilde/*vild

- (17) a. En vild/*vilde fest
- b. Den vilde/*vild fest
- c. To vilde/*vild fester

Jensen (2004: 165) peger ligeledes på, at man hverken finder ”konstruktioner med entals-subjekter hvortil der er knyttet et ’nøgent’ prædikativ i flertal” (Jensen 2004: 165), se (16)a, og det sammenholdt med den observation, at komplekse nominaler med det atypiske bøjningsmønster, (17), ikke optræder, er en stærk indikation af, at tilstedeværelsen af præpositionsobjektet er afgørende for udløsningen af den atypiske kongruens, og at fonetiske faktorer ikke alene kan forklare fænomenet.

4.2 Syntaktiske analyser af fænomenet

Vigliocco & Nicol (1998) og Franck, Vigliocco & Nicol (2002) lavede psykolingvistiske forsøg i engelsk og fransk, hvor de forsøgte at elicitere fejl i kongruensen mellem subjekt og finit verbum ved at lade et nominalt led intervenere mellem subjektet og det finitte verbum, fx **The key to the cabinets are...* Sådanne afvigelser fra standardkongruensen er imidlertid helt grundlæggende forskellige fra den atypiske kongruens på dansk: Det intervenerende nominal mellem de kongruerende størrelser i den atypiske kongruens på dansk, er nemlig netop subjektet i sætningstypen der involverer fremflytning (*Dem er jeg vilde med*) og de engelske studiers interventionsprincip forudsiger således standardkongruensen og *ikke* den atypiske kongruens!

Kongruens med prædikative adjektiver findes ifølge Vikner (2001) på alle de romanske og skandinaviske sprog uafhængigt af, hvor rigt bøjningssystemet er. I den generative grammatik (se bl.a. Delsing 1993: 84, Kayne 1989) antages det, at kongruens er resultatet af en ”afstemning” af grammatisk person, tal og køn, i den generative litteratur tilsammen kaldet ϕ -træk (læses phi-træk). I standardkongruensen mellem et subjekt og et prædikativt adjektiv formodes det således, at subjektet er basisgenereret inde i adjektivfrasen, hvor det udgør en del af adjektivets argumentstruktur (Vikner 2001: 407–408). Det er igennem denne lokale relation, at kongruensen mellem subjektet og det prædikative adjektiv opstår. I analogi med Kaynes (1989)

syntaktiske analyse af alternativ kongruens med participiumsformer på fransk, kunne man således antage, at det fremflyttede P-Obj på sin vej mod sætningens første position indgår i en syntaktisk relation med det prædikative adjektiv, hvor det ”afstemmer” sine ϕ -træk. Det er dog ikke ligetil at forklare den *atypiske* kongruens med henvisning til syntaktisk derivation, da fænomenet forefindes uafhængigt af præpositionsobjektets syntaktiske placering (*in situ* og fremflyttet) (se Nyvad (2007) for en oversigt). Dette peger på, at hverken intervention eller fremflytning af nominelle led rummer forklaringen, der således muligvis kan findes i ekstra-syntaktiske faktorer (Christensen & Nyvad 2019, 2021).

4.3 Informationsstrukturelle forklaringsmodeller

Der er en tværsproglig tendens til, at subjekter tilhører baggrundsinformation i en sætning (Blake 2001: 131–132). I den traditionelle grammatik blev ”topic” oven i købet sidestillet med subjektet, der ifølge Hansen & Heltoft (2011: 1729) er ”antifokus”. Ifølge Lambrecht (1994: 199) er der en stærk tværsproglig tendens til, at et topikalt element indtager sætningens første position. Som Nyvad (2007: 23) pointerer, er det bemærkelsesværdigt, hvor ofte det prædikative adjektiv kongruerer med et P-Obj, der ser ud til at fungere som topic. Dette kunne tyde på, at det muligvis kongruerer med en pragmatisk størrelse (”topic”) fremfor et syntaktisk element (subjektet). Den alternative kongruens kunne på denne måde beskrives som overensstemmelse i ϕ -træk mellem et prædikativt adjektiv og dets P-Obj, når dette fungerer som topic (se Nyvad 2007 for detaljer), en hypotese, der kunne gå under navnet *topickongruens*, og som er kompatibel med Comrie (2003: 319 ff.), der hævder, at såkaldt *trigger-happy* kongruens (hvor der er ’konkurrence’ om kongruensen) er afhængig af den nominale frases *topicworthiness*, dvs. sandsynligheden for, at det pågældende element fungerer som topic, hvilket formodes at være relateret til den nominale frases bestemthed og *salience*, dvs. pragmatisk prominens (Comrie 2003: 329).

Overordnet synes den atypiske kongruens at være mest almindelig med pronominer og topikaliserede elementer (men det er stadig et empirisk spørgsmål), og det er muligt, at den alternative kongruens på de prædikative adjektiver i sin oprindelse startede med kun at optræde, når P-Obj var et topikaliseret pronomen. At det har kunnet tiltrække kongruensen, kunne antageligvis være forbundet med det faktum, at P-Obj ligesom subjektet er et valensbundet led i sætningen, og som sådan bærer ϕ -træk.

I så fald kunne den atypiske kongruens muligvis være opstået ved, at sprogbrugere har fortolket atypisk kongruens ved fremflytning som styret af informationsstruktur, altså en form for *topickongruens*. Efterfølgende kunne det så være blevet generaliseret til at omfatte syntak-

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

tiske konfigurationer, hvor P-Obj ikke er fremflyttet, muligvis understøttet af kontekstuelle faktorer og trykforhold. Problemet er dog, at det er svært at operationalisere begreberne topic og fokus i eksperimentelle undersøgelser, således at tesen vedrørende topickongruens kan efterprøves videnskabeligt (jf. Christensen & Nyvad 2021: 129–130).

5. Hvor er vi så nu?

Som overstående diskussion viser, er det svært at indkapsle fænomenet med atypisk kongruens ved hjælp af én forklaringsmodel. Vi bliver nødt til at ty til idéen om en sammensværgelse af flere forskellige faktorer: I Christensen & Nyvad (2019) fandt vi i vores eliciteringsdata noget, der lignede en 50/50-fordeling mellem standardkongruens og atypisk kongruens, når P-Obj var i singularis. Dette bemærkelsesværdige mønster kan dog tilskrives to velkendte faktorer, som vi ved, spiller en rolle uafhængigt af det grammatiske fænomen under lup her, nemlig dels bortfald af endelser på prædikative adjektiver på dansk, der ikke er et nyt fænomen (jf. Hansen 1967) og dels fejlparsing af sætninger, hvor første nominelle led viser sig ikke at være subjekt (jf. Friederici m.fl. 2001). Når disse to faktorer er sat ud af spil, er den atypiske kongruens nærmest ikke-eksisterende: I stimulussættet i Christensen og Nyvad (2021) var P-Obj *in situ*, og når sidstnævnte var i pluralis og subjektet i singularis, blev den bøjede form af det prædikative adjektiv kun brugt i 1 % af tilfældene!

Selvom vi ikke kan afvise, at fænomenet kunne afspejle en ny form for kongruenssystem, der blander sig med standard subjektkongruens, eller at der er tale om en gryende sprogforandring, så har vi imidlertid heller ikke nogen data, der understøtter det. Omvendt er performansfejl som regel ikke helt tilfældige, men ofte systematiske (se fx Garrett 2015), hvilket egentlig ikke er specielt overraskende, eftersom vi alle har det samme fysiologiske og kognitive apparatur, som vi bruger, når vi taler. Desuden er fænomenet med atypisk kongruens netop systematisk og produktivt, og af selvsamme årsager efter alt at dømme mere eller mindre acceptabelt hos danske sprogbrugere.

Litteratur

- Baker, Mark C. (2008) *The Syntax of Agreement & Concord*. Cambridge: Cambridge University Press.
- Belletti, Adriana. (2001) Agreement Projections. I Mark R. Baltin & Chris Collins (red.) *The Handbook of Contemporary Syntactic Theory*. Malden, MA: Blackwell: 482–510.
- Blake, Barry J. (2001) *Case*. 2. udg. Cambridge: Cambridge University Press.
- Brink, Lars, Jørn Lund, Steffen Heger & Jens Normann Jørgensen (1991) *Den Store Danske Udtaleordbog*. København: Munksgaard.
- Christensen, Ken Ramshøj, Johannes Kizach & Anne Mette Nyvad (2013) Escape from the Island: Grammaticality & (Reduced) Acceptability of *Wh*-Island Violations in Danish. *Journal of Psycholinguistic Research* 42:51–70.
- Christensen, Ken Ramshøj & Anne Mette Nyvad (2014) On the Nature of Escapable Relative Islands. *Nordic Journal of Linguistics* 37:29–45.
- Christensen, Ken Ramshøj & Anne Mette Nyvad (2019) De her sætninger er jeg vilde med!: Sproglig variation i prædikative adjektivers kongruensmønster. I Yonatan Goldshtein, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). *17. Møde om udforskningen af dansk sprog*. MUDS 17. Aarhus: Institut for Kommunikation og Kultur, Aarhus Universitet:95–113.
- Christensen, Ken Ramshøj & Anne Mette Nyvad (2021) Vi er (stadig) vild med den her sætning! I Yonatan Goldshtein, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). *18. Møde om udforskningen af dansk sprog*. MUDS 18. Aarhus: Institut for Kommunikation og Kultur, Aarhus Universitet: 169–86.
- Christensen, Ken Ramshøj & Anne Mette Nyvad (2023) Frequency, Animacy, Structural Complexity & Acceptability. (Ms. indsendt til bedømmelse.)
- Christianson, Kiel Andrew Hollingworth, John F. Halliwell & Fernanda Ferreira (2001) Thematic Roles Assigned along the Garden Path Linger. *Cognitive Psychology* 42:368–407.
- Comrie, Bernard (2003) When Agreement Gets Trigger-Happy. *Transactions of the Philological Society* 101:313–37.
- Delsing, Lars-Olof (1993) *The Internal Structure of Noun Phrases in the Scandinavian Languages*. Ph.d.-afhandling. Lund: Lund University.
- Engberg-Pedersen, Elisabeth & Mads Poulsen (2010) Alternative Agreement Controllers in Danish: Usage or Structure. I Kasper Boye & Elisabeth Engberg-Pedersen (red.). *Language Usage & Language Structure*. Berlin: Walter de Gruyter: 205–35.
- Fanselow, Gisbert & Stefan Frisch (2006) Effects of Processing Difficulty on Judgments of Acceptability. I Gisbert Fanselow, Catherine Fery & Matthias Schlesewsky (red.). *Gradience in Grammar: Generative Perspectives*. Oxford: Oxford University Press:291–316.
- Franck, Julie, Gabriella Vigliocco & Janet Nicol (2002) Subject-Verb Agreement Errors in French & English: The Role of Syntactic Hierarchy. *Language & Cognitive Processes* 17:371–404.
- Friederici, Angela D., Axel Mecklinger, Kevin M. Spencer, Karsten Steinhauser & Emanuel Donchin. 2001. Syntactic Parsing Preferences & Their On-Line Revisions: A Spatio-Temporal Analysis of Event-Related Brain Potentials. *Cognitive Brain Research* 11:305–23.
- Garrett, Merrill F. (2015) Speech Errors, Psychology Of. I *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam: Elsevier:229–34.

”Vi har fundet nogle sætninger på nettet, og dem bliver du vilde med!”. Når performansen halter

- Grønnum, Nina (2005) *Fonetik og fonologi: almen og dansk*. 3. udg. København: Akademisk.
- Hansen, Aage (1967) *Moderne Dansk. Bind II*. København: Grafisk Forlag.
- Hansen, Erik & Lars Heltøft (2011) *Grammatik over det Danske Sprog. bind I-III*. København: Det Danske Sprog- og Litteraturselskab.
- Hansen, Peter Molbæk (1990) *Udtaleordbog*. København: Gyldendal.
- Jensen, Jørgen Nørby (2004) Dem er jeg helt vilde med! Om fokuskongruens i moderne dansk. I Henrik Jørgensen & Peter Stray Jørgensen (red.). *På godt dansk. Festskrift til Henrik Galberg Jacobsen i anledning af hans 60 års fødselsdag den 4. februar 2004*. Aarhus: Wessel og Huitfeldt:163–71.
- Kandel, Margaret, Cassidy Rae Wyatt & Colin Phillips (2022) Agreement Attraction Error & Timing Profiles in Continuous Speech. *Glossa Psycholinguistics* 1.
- Kayne, Richard S. (1989) Facets of Romance Past Participle Agreement. I Paola Benincá (red.). *Dialect Variation & the Theory of Grammar*. Berlin: De Gruyter: 85–104.
- Lambrecht, Knud (1994) *Information Structure & Sentence Form: Topic, Focus & the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Mikkelsen, Line (2015) VP Anaphora & Verb-Second Order in Danish. *Journal of Linguistics* 51:595–643.
- Newmeyer, Frederick J. (2005) *Possible & Probable Languages: A Generative Perspective on Linguistic Typology*. New York: Oxford University Press.
- Nyvad, Anne Mette (2007) Revisiting Agreement Paradigms: Predicative Adjectives in Danish. MA-afhandling. London: University College London.
- Nyvad, Anne Mette, Christiane Müller & Ken Ramshøj Christensen (2022) Too True to Be Good? The Non-Uniformity of Extraction from Adjunct Clauses in English. *Languages* 7:244.
- Puggaard, Rasmus (2019) Fundamentfeltet i danske samtaler. I Yonatan Goldshtein, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). *17. Møde om udforskningen af dansk sprog*. Aarhus: Institut for Kommunikation og Kultur, Aarhus Universitet:395–414.
- Ross, John Robert (1967) Constraints on Variables in Syntax. Ph.d.-afhandling. Cambridge, MA: MIT.
- Sprouse, Jon, Matt Wagers & Colin Phillips (2012) A Test of the Relation between Working-Memory Capacity & Syntactic Island Effects. *Language* 88:82–123.
- Staub, Adrian (2009) On the Interpretation of the Number Attraction Effect: Response Time Evidence. *Journal of Memory & Language* 60:308.
- Vigliocco, Gabriella & Janet Nicol (1998) Separating Hierarchical Relations & Word Order in Language Production: Is Proximity Concord Syntactic or Linear? *Cognition* 68:B13–29.
- Vikner, Sten (2001) Predicative Adjective Agreement. Where German May Be ‘Easy’, but French & Danish Are Not ‘Easies.’ I Kirsten Adamzik & Helen Christen (red.). *Sprachkontakt, Sprachvergleich, Sprachvariation*. Tübingen: Max Niemeyer Verlag:399–414.

1500-talssalmer og danske sprogmodeller

Ea Lindhardt Overgaard

Aarhus Universitet

Indledning

Danske sprogmodeller til *Natural Language Processing* (NLP) er i de seneste år blevet forbedret i takt med at der er udviklet flere ressourcer inden for dansk sprogteknologi. Det viser sig blandt andet ved større og bedre korpora samt nye forbedrede sprogmodeller (Enevoldsen, Johansen m.fl. 2021, Strømberg-Derczynski, Ciosici m.fl. 2021, sprogteknologi.dk 2023). I takt med udviklingen af dansksprogede ressourcer er opmærksomheden på sprogmodellernes potentiale og inddragelsen af digitale værktøjer i undervisnings- og forskningsmiljøer også steget (Ondelli 2018). Trods den positive udvikling for danske sprogmodeller er der stadig et stykke vej endnu før modellerne uden videre kan benyttes på domænespecifikke områder, f.eks. salmelitteratur, og historiske datasæt, med tilfredsstillende resultater (Enevoldsen, Johansen m.fl. 2021). Det er et velkendt problem at sprogmodellerne præsterer dårligere på sådanne data fordi sprogmodellerne er udviklet og trænet på moderne dansk og hovedsageligt prosaisk tekstmateriale. Det er altså en udfordring at få sprogmodellerne til at opnå en høj nok grad af præcision når de benyttes på tekstsamlinger der afviger en moderne retskrivningsnorm.

For at undersøge hvordan man kan forbedre de danske sprogmodellers præcision på normafvigende korpora, præsenterer jeg i denne artikel forskellige strategier for at optimere sprogmodellernes præstation gennem normalisering af en lille samling danske reformationsalmer. Med afsæt i ti salmer¹ fra Thomissøns salmebog (1569) undersøger jeg hvorvidt normalisering af teksterne påvirker NLP-værktøjers præstation. Til at illustrere modellernes præstation tester jeg præcisionen af *Part of speech-tagging* (POS) for de to danske sprogmodeller *DaCy* og *SpaCy*, der begge er frit tilgængelige værktøjer til *Natural Language Processing* i programmeringssproget *Python* (Honnibal og Montani 2017, Enevoldsen, Johansen m.fl. 2021). Sprogmodellernes præstation testes på tre forskellige grader af normalisering af salmeteksterne. Denne undersøgelse indeholder kun 10 salmer, men hvis tallet steg bare til 100 sal-

¹ Titler i moderniseret form: *Aleneste Gud i Himmerig; Behold os Herre, ved dit ord; Det hellige kors, vor Herre selv bar; Et barn er født i Betlehem; Et lidet barn så lysteligt; Hvad kan os komme til for nød; Jeg vil mig Herren love; Krist stod op af døde; O Guds Lam Uskyldig og Vor Gud han er så fast en borg.*

mer, ville en manuel annotering være enormt ressourcekrævende. Af samme grund er det ofte nødvendigt at benytte digitale sprogmodeller når man arbejder med *big data*-analyser. Artiklen fungerer derfor som et eksempel på hvordan man kan håndtere normaliseringen og optimere brugen af sprogmodeller på tekster i meget større korpore.

I artiklen redegør jeg først for undersøgelsens overordnede design. Herefter vil jeg gennem et eksempel med POS-tagging argumentere for nødvendigheden af at normalisere sine tekstdata når man benytter sprogmodeller på ældre tekst. Efter eksemplet forklarer jeg hvordan undersøgelsens ti salmer er blevet normaliseret, og hvordan præcisionen af sprogmodellernes POS-tagging evalueres. Slutteligt diskuterer jeg normaliseringens indvirkning på sprogmodellernes POS-tagging af salmerne.

Undersøgelsens design

Undersøgelsens data består af digitale udgaver af de ti ovennævnte salmer fra Thomissøns salmebog 1569. Salmeteksterne er digitaliseret til et rent tekstformat (.txt-filer). I undersøgelsen bruger jeg programmeringssproget *Python* (Rossum og Drake 2009) hvor jeg, foruden de maskinlæringsbaserede sprogmodeller *SpaCy* og *DaCy*, benytter mig af basismoduler og basisfunktionalitet indbygget i *Python*. Jeg bruger særligt datatypen *dictionaries* til at opbygge en ordbogsressource for salmeteksterne. Her knytter jeg et opslagsord, *key*, med en eller flere variationsformer for ordet, betegnet som *values*. Det er muligt at knytte flere *values* til hver enkelt *key*. Nedenfor ses i eksempel (1) en dictionary med *key* *vederkvæger* og den tilhørende *value* *vederqueger*, samt i eksempel (2) en dictionary med *key* *hand* og de tilhørende *values* *han* og *hånd*.

```
(1) Ordbog_1 = {'vederkvæger': ['vederqueger']}
(2) Ordbog_2 = {'hand': ['han', 'hånd']}
```

Ud fra et opslagsord får jeg altså automatisk præsenteret dets mulige variationsformer, præcis som hvis jeg slog ordet op i en almindelig ordbog. Ordbøgerne fungerer begge veje så jeg i ovenstående eksempel kan udtrække både *vederkvæger* ved at søge *vederqueger* frem, og omvendt. For ord der i salmerne har en ældre stavemåde end den moderne retstavning foreskriver, oprettes indgangene *key: value* hvor den ældre stavemåde knyttes til en stavemåde der følger moderne retstavning. Ordbogen er på den måde et redskab til at normalisere de ældre salmetekster til en mere moderne retstavning. De moderne retstavningsformer er hovedsageligt baseret på Dansk Sprog- og Litteraturselskabs omfattende ordbogsressource knyttet til danske reformationssalmer. Ressourcen giver adgang til en lang række leksikografiske værker på én

gang. For ét opslag præsenteres man for opslagsartikler der matcher det søgte ord på tværs af forskellige ordbogsressourcer, og opslagene indeholder links til de originale opslagsartikler i de enkelte ordbøger (Det Danske Sprog- og Litteraturselskab 2023).

Normaliseringen af ordene i de ældre salmer er i mange tilfælde ikke entydigt bestemt. For ord med flere mulige *values* knyttet til hver *key* kræver det en manuel vurdering af hvilket ord der er det rette i den givne kontekst. Jeg har derfor opdelt normaliseringsprocessen i undersøgelsen i tre dele for at tage forbehold for de ord der kræver en manuel vurdering. I afsnittet *Normalisering af tekstdata* uddyber jeg denne opdeling samt hvordan de forskellige *dictionaries* bruges.

Som nævnt i indledningen tester jeg efter hver del af normaliseringsprocessen præcisionen af de danske sprogmodellers POS-tagging. Ved POS-tagging annoteres hvert enkelt ord i en tekst med et *part of speech*-tag, der angiver ordklassen for det givne ord. Ordklassenotationen følger de universelle POS-tags (Universal Dependencies (UD) 2014-2022). Forud for normaliseringsprocessen har jeg testet præcisionen for POS-tagging på de originale tekster. Til at evaluere præcisionen har jeg lavet en manuel POS-opmærkning for salmeteksterne, en såkaldt *golden standard*, som kan holdes op mod sprogmodellernes resultater. Dette er standard evalueringspraksis (Enevoldsen, Johansen m.fk. 2021). Præcisionen af sprogmodellerne måles ud fra hvor mange procent af POS-tagene sprogmodellerne har angivet korrekt.

Er normalisering nødvendig?

For at sprogmodellerne kan præstere godt nok på små domænespecifikke korpora til at man kan opnå brugbare resultater, er det nødvendigt at tilpasse sit korpus så det ligner modellens træningsdata mest muligt. Det er som nævnt et velkendt problem inden for sprogteknologi at sprogmodellernes præcision er for lav hvis man anvender dem på korpora der ikke følger moderne retskrivningsnormer (Bigi 2014). For at illustrere problematikken viser Tabel 1 et udsnit af en automatiseret POS-opmærkning af den originale version af salmen *Alleneste Gud i Himmerig* fra Thomissøns salmebog (1569) og af den moderne salme *Aleneste Gud i Himmerig* fra Den Danske salmebog (2003)². Fejlklassificerede POS-tags er markeret med grå, og POS-taggingen er lavet med *SpaCy*:

² En liste over de forskellige POS-tags findes til sidst i artiklen.

Tabel 1

Thomissøn 1569	POS	DDS 2002	POS
<u>Alleniste</u>	Propn	<u>Aleneste</u>	Propn
Gud	Noun	Gud	Noun
i	Adp	i	Adp
Himmerig	Noun	Himmerig	Noun
være	Aux	ske	Verb
<u>loff</u>	Adj	lov	Noun
<u>oc</u>	X	og	Konj
<u>priss</u>	Noun	pris	Noun
for	Adp	for	Adp
alle	Adj	sin	Det
sin	Det	nåde	Noun
<u>naade</u>	Noun	som	Pron
der	Pron	han	Pron
<u>hand</u>	X	har	Aux
<u>haffuer</u>	Propn	os	Pron
<u>giort</u>	Verb	skænket	Verb
i	Adp	faderlig	Adj
Jorderig	Noun	at	Part
J	Propn	fri	Adj
disse	Det	os	Pron
samme	Adj	af	Adp
<u>naadelige</u>	Adj	sydens	Noun
dage.	noun	våde	Adj

I skemaet kan man se at der forekommer flere fejlklassificeringer for det ældre teksteksempel end for eksemplet med moderne retstavning. Hvis man kigger på den procentvise fejlklassificering for salmeteksterne i deres fulde længde, har *Alleniste Gud i Himmerig* 44,8% forkerte POS-tags mens *Aleneste Gud i Himmerig* har 11,7%. Der er altså en tydelig forskel på sprogmodellernes præstation fra det ældre teksteksempel til det nyere. Det skal bemærkes at det moderne teksteksempel ikke er en normalisering af den ældre salme, men er den moderne udgave som den fremgår i den nugældende salmebog. Den moderne salme har altså, foruden at være tilpasset gældende retskrivningsnormer, været udsat for flere gendigtninger og revisioner gennem tiden. Jeg benytter begrebet *normalisering* om dét at tilpasse de enkelte ord til en gældende norm – i dette tilfælde retskrivningsnormen – i det omfang det er muligt uden at foretage drastiske indholdsmæssige ændringer ift. den originale tekst.

For salmen *Alleniste Gud i Himmerig* er fejltypene på de forkerte POS-tags varierende mellem proprier, adjektiver og X (andet). Disse fejltypen er hyppige gennem hele salmen, men ikke begrænset til disse ordklasser da der forekommer fejlklassificeringer for samtlige ordklas-

ser. Den hyppigste fejltipe er dog at SpaCy klassificerer noget som et proprium der ikke er det. Dette sker 30 gange bare i denne salme og udgør godt en tredjedel af alle fejlklassificeringerne.

I salmen *Aleneste Gud i Himmerig* ses samme fejltipe, som i det ældre eksempel, hvor SpaCy klassificerer noget som et proprium der ikke er det. Ellers er den hyppigste fejltipe her at ord klassificeres som substantiver og adjektiver når de ikke er det. Her er det særligt imperativformer af verber, som *fri* i eksemplet fra skemaet, samt interjektioner der bliver fejlklassificeret. Ordene *fri* og *våde* er begge homografer med et adjektiv hvor man kan forestille sig at den adjektiviske udgave er den hyppigste variant af ordet i normaliseret brug. Fejlklassificeringen af disse ord i eksemplet er her højst sandsynligt baseret på at sprogmodellens træningsdata har en højere frekvens af *fri* og *våde* som adjektiver end som verbum og substantiv.

Selvom der forekommer fejl i begge ovenstående eksempler, så er fejlene meget hyppigere for den ældre tekst end for den moderne. Fejlklassificeringerne synes oftest kategoriske for den moderne salme hvor de for den ældre salme både består af kategoriske fejl, men også har en høj grad af tilfældighed. Selvom fejlmarginen er betydeligt lavere for den moderne salme end for den ældre, så ligger mængden af fejlklassificeringer for den moderne salme på grænsen af hvad man kan acceptere som brugbart. De bedste danske sprogmodeller kan efterhånden klassificere med over 90 % nøjagtighed, og dette er inden for digital humaniora et præcisionsniveau der normalt accepteres (Enevoldsen, Johansen m.fl. 2021). En måde at reducere fejlmarginen for POS-taggingen af de ældre salmer er at normalisere salmeteksterne i en sådan grad at ordenes ortografi ligner det data, som sprogmodellen er trænet på, så meget som muligt. Dette behandles i følgende afsnit.

Normalisering af tekstdata

I arbejdet med tekst som data er man nødt til at klargøre sit datamateriale på flere niveauer for at gøre teksten maskinlæsbar (Bigi 2014, Ondelli 2018). Et af disse niveauer omhandler hvordan man skal opdele teksten i tokens, dvs. i enheder af ord, tegnsætning, mellemrum - små betydningsenheder man gerne vil kunne identificere maskinelt. I denne proces ensretter og normaliserer man ofte sit korpus så eksempelvis tal altid skrives ud i bogstaver, forkortelser skrives i deres fulde længde mm. Dette gør disse størrelser let genkendelige i korpus. I denne undersøgelse drejer spørgsmålet om normalisering sig om hvordan ældre ordformer og varianter gøres genkendelige for moderne sprogmodeller. Normaliseringsaspektet rejser for salmerne i Thomissøns salmebog flere filologiske spørgsmål, herunder hvor meget og hvordan teksterne skal normaliseres uden at normaliseringen skaber indholdsmæssige uoverensstemmelser mellem den originale og den normaliserede udgave af en salmetekst. Jeg ønsker at bringe

salmeteksterne så meget i overensstemmelse med den gældende retskrivningsnorm som muligt uden at ændre indholdsmæssigt på salmeteksterne. Hvis man eksempelvis vil undersøge forhold omkring et ord som *navn*, kan man så uproblematisk slå formerne *Naffn*, *Navn*, og *navn* sammen under én form: *navn*? Det er langt fra altid muligt at reducere et ords varianter til én form. Ord, der ikke entydigt henviser til samme lemma, er eksempelvis problematiske: Vil et ord som *ere* eksempelvis kunne udbyttes med *ære* eller *er*? Mange ord har flere mulige normaliseringsmuligheder, og man skal vurdere normaliseringen for disse ord i dets kontekst. I det følgende viser jeg et eksempel på hvordan de forskellige normaliseringsaspekter kan behandles.

Eksempel på normalisering af salmetekster

Normaliseringen af de ældre salmer er foretaget i tre trin. Første trin af processen er automatisk at ensrette de ord der uden problemer kan reduceres fra mange variantformer til én. Første trin ensretter altså ortografien for ord der gennem tiden har ændret stavemåder og samtidig ikke har sammenfald i stavemåde med andre ord. Eksempler herpå kunne være *nåde* > *naade* og *vilje* > *villie*. Denne ortografiske normaliseringsproces er automatiseret ved at jeg har lavet en *dictionary* hvor variationsformerne for et ord er knyttet sammen med den form af ordet som jeg vil normalisere med. Jeg har herefter lavet en funktion der for hver forekomst af variationsformerne i salmerne udskifter ordet med dens normaliserede form. Nedenfor i eksempel (3) ses eksempler på opslag i denne ordbog hvor den ønskede normaliserede form af ordet står først og den ældre variant af ordet følger efter:

```
(3)
ordbog = {'lignes': ['lignis'],
          'gjort': ['giort'],
          'nådelig': ['naadelig'],
          'yndest': ['yndist'],
          'vilje': ['villie'],
          'vederkvæger': ['vederqueger']}
```

I andet trin af normaliseringen har jeg lavet en funktion hvor homografer der senere har fået to forskellige stavemåder, identificeres i teksten, og den korrekte form for ordet udvælges manuelt ved at vælge mellem en række mulige normaliseringsmuligheder. Koden for funktionen er vist i nedenstående billede:

```
def moderne_2(tekst): #ordsplittet liste
    tekst_lav = tekst.lower()
    ordopdelt = rens_ord(tekst_lav)
    moderniseret = []
    for token in ordopdelt:
        val = 0
        for key, list_of_values in ordbog_2.items():

            if token == key:
                print('Opslag: ', key)
                print('list_of_values: ', list_of_values)
                valg = int(input('Skriv nummer på korrekt opslag (start ved indeks 0).'))

                moderniseret.append(list_of_values[valg])
                val = 1
                continue

            if val == 0:
                moderniseret.append(token)
                continue

    return " ".join(moderniseret)

moderne_2(tekst)
```

Opslag: loff

list_of_values: ['lov', 'løv', 'lo', 'løb', 'loft']

Skriv nummer på korrekt opslag (start ved indeks 0).

Her ses det at funktionen har identificeret *loff* som et ord der skal tages stilling til i sætningen ”være loff oc priss for alle sin naade”. Efter opslagsordet følger en *list_of_values* med en række ord der kan vælges imellem. I den sorte bjælke kan man skrive indeks for det ord man ønsker at udskifte *loff* med. Herefter retter funktionen automatisk ordet i salmeteksten til den valgte form. Dette trin er altså en semi-automatiseret proces der giver mulighed for en menneskelig vurdering af ordets betydning i dets kontekst.

Det sidste trin af normaliseringen omhandler brugen af versaler. Jeg har valgt at inddrage dette aspekt i normaliseringen af salmeteksterne da jeg fra eksemplet i Tabel 1 ved at mange ord bliver fejlklassificeret som *proprier*. Det sker ofte de steder hvor ord starter med versal hvor vi efter moderne retskrivning ikke ville bruge versaler. I 1569 herskede ingen officiel dansk retskrivning (Jacobsen 2010, Det Danske Sprog- og Litteraturselskab 2023), og brugen af versaler for substantiver benyttes inkonsekvent i Thomissøns salmebog. Den anderledes brug af versaler ift. den moderne retskrivningsnorm synes at påvirke sprogmodellernes præcision af POS-tagging i en sådan grad at det er væsentligt at inddrage dette aspekt i normaliseringen for at reducere fejlmarginen. Salmetekster har desuden en særlig brug af versaler da ord som *Gud*, *Herre* og *Himmerig* starter med versaler og kan anses som *proprier* i salmekonteksten. Jeg har derfor lavet en vurdering af hvilke ord der i salmekontekst skal bevare startversalerne ud fra et diakront blik på salmelitteraturen. Jeg bevarer derfor versalbrugen ved ord hvor denne versalbrug gør sig gældende gennem hele salmehistorien. Til dette tredje trin af normaliseringen har

jeg igen lavet en funktion der sørger for at bevare versaler for de ord jeg har gemt som versal-bevarende ord.

Nedenfor vises et eksempel fra salmen ”Det Hellige Kaarss” fra Thomissøns salmebog (1569), og de forskellige trin teksten har gennemgået i normaliseringsprocessen. Først vises den originale tekst a). Herefter er ændringerne for 1. modernisering markeret med fed i b), 2. modernisering er markeret med understregning i c) og 3. modernisering er markeret med gennemstregning i d):

(4)

a) Det hellige Kaarss vor Herre selff bar, met blodige vunder oc dødelig Saar, for oss hand plict oc bod fuldgjorde, at wi skuld oss der til forlade, wi vorder ej salig i andre maade.

b) Det hellige Kaarss vor Herre **selv** bar, met blodige vunder oc dødelig **sår**, for oss hand plict oc bod **fuldgjorde**, at wi **skulle** oss der til forlade, wi vorder ej salig i andre **måde**.

c) Det hellige kors vor Herre **selv** bar, med blodige vunder og dødelig **sår**, for os han pligt og bod **fuldgjorde**, at vi **skulle** os der til forlade, vi vorder ej salig i andre **måde**.

d) det hellige kors vor Herre **selv** bar, med blodige vunder og dødelig **sår**, for os han pligt og bod **fuldgjorde**, at vi **skulle** os der til forlade, vi vorder ej salig i andre **måde**.

Det er som nævnt den moderne retstavning der har styret graden af normalisering, samtidig med at der tages forbehold for at ændringerne har så lidt indflydelse på den oprindelige tekst som muligt. Syntaktiske mønstre er ikke ændret da moderne salmer også har en vekslende syntaks. Ældre ordformer (*vunder*) er heller ikke ændret til mere moderne synonymmer (*sår*) da sådanne ændringer ville lægge sig tæt op ad nyfortolkninger af salmen hvilket i højere grad ændrer betydningen af den oprindelige tekst. Teksten er gennem normaliseringsprocessen blevet forberedt til NLP-værktøjer der er trænet på moderne dansksproget materiale.

Resultater

De ti salmer i denne undersøgelse har alle gennemgået de tre niveauer af normalisering. Jeg evaluerer sprogmodellerne på alle de tre niveauer af normalisering ved at teste modellernes klassificeringsevner mod en manuel POS-opmærkning af salmerne. Forud for moderniserings-

processen har jeg målt den procentvise del af korrekte POS-klassificeringer, og efter hver del af normaliseringsprocessen har jeg igen målt den procentvise andel af korrekte klassificeringer. Det er på den måde muligt at vurdere det enkelte moderniseringstrin op mod sprogmodellens evne til at klassificere korrekt. I undersøgelsen testes både sprogmodellen SpaCy der, blandt mange mulige sprog, indeholder en model til dansk sprog (Honnibal og Montani 2017), og DaCy, en sprogmodel baseret på SpaCy som er blevet forbedret til dansksproget materiale. Blandt andet er DaCy trænet på en mængde data der kan mime slå- og stavfejl i tekst samt historisk ortografisk variation (Enevoldsen, Johansen m.fl. 2021). DaCy kan derfor antages at være mere robust over for variationer i korpora end SpaCy.

Nedenfor ses *Tabel 2* med resultater for undersøgelsen af hhv. spaCy og DaCy's POS-tagger. Der er for hver salme angivet de procentvise korrekte klassificeringer for hhv. spaCy og DaCy til hver moderniseringsgrad af salmerne.

Tabel 2

Normaliseringsgrad	Original		1. normalisering		2. normalisering		3. normalisering	
	spaCy	DaCy	spaCy	DaCy	spaCy	DaCy	spaCy	DaCy
<u>Aleneste Gud</u>	56,7	72,4	66,2	80,0	86,7	89,6	90,5	88,6
Behold os Herre	66,9	83,5	73,6	86,0	82,6	90,9	84,3	90,9
Det hellige kors	50,7	75,4	63,0	83,3	84,1	84,8	84,8	86,6
Et barn er født i Betlehem	53,5	58,0	64,3	74,5	80,3	84,1	84,7	86,6
Et lidet barn så lysteligt	46,4	63,1	51,4	70,3	74,8	77,5	74,8	79,7
Hvad kan os komme til for nød	57,1	75,8	67,1	83,9	84,5	89,7	84,5	89,7
Jeg vil mig Herren love	59,0	71,9	75,2	81,3	84,1	87,6	84,5	87,8
<u>Krist stod op af døde</u>	40,0	72,0	36,0	72,0	68,0	64,0	72,0	72,0
O Guds lam uskyldig	33,3	69,8	57,3	78,1	71,9	83,3	75,0	83,3
Vor Gud han er så fast en borg	46,8	68,9	61,2	81,1	81,6	87,9	83,2	88,4

For de originale salmetekster ligger andelen af korrekt klassificerede ord på mellem 33,3% og 66,9% for SpaCy og på mellem 58,0% og 83,5% for DaCy. Trods store udsving i begge modellers succesrater kan det konkluderes at DaCy præsterer væsentligt bedre på den originale tekst end SpaCy. Ingen af modellerne nærmer sig dog et niveau hvor opmærkningen vil kunne benyttes i et videre analytisk arbejde. Det er også værd at bemærke at de laveste succesrater for

SpaCy ikke nødvendigvis medfører en tilsvarende lav succesrate for DaCy. Det kan blandt andet ses ved ”O Guds lam uskyldig”, der har succesraterne 33,3% og 69,8%. Samtidig præsterer de to modeller næsten på samme niveau for salmen ”Et barn er født i Betlehem” med succesrater på 53,5% og 58,0%.

I tabellen kan man se at sprogmodellerne præsterer bedre i takt med at salmeteksterne moderniseres. Dette gælder i alle tilfælde bortset fra ”Krist stod op af døde” hvor både SpaCy og DaCy falder i succesrate fra hhv. originalteksten til udgaven med 1. modernisering for SpaCy og fra 1. modernisering til 2. modernisering med DaCy. Dette resultat skyldes formentlig at salmen kun indeholder 25 ord og én korrekt klassificering fra eller til derfor påvirker den procentvise succesrate med 4 %.

Den gennemsnitlige andel af korrekt klassificerede ord til de forskellige normaliseringsgrader ses i Tabel 3:

Tabel 3

Normaliseringsgrad	Original		1. normalisering		2. normalisering		3. normalisering	
	spaCy	DaCy	spaCy	DaCy	spaCy	DaCy	spaCy	DaCy
Gennemsnit	51,04	71,08	61,53	79,05	79,86	83,94	81,83	85,36

Her kan man se at SpaCy præsterer langt dårligere end DaCy jo lavere normaliseringsgrad af teksten man evaluerer modellen på. Dog sker der væsentlige forbedringer ved højere normaliseringsgrad og særligt 1. til 2. normaliseringsgrad har betydning for SpaCys evne til at klassificere korrekt. En grund til DaCys markant bedre evne til at klassificere korrekt for de lavere normaliseringsgrader kan skyldes at modellen netop er trænet på data, der gør modellen mere robust for mindre ortografiske forskelle.

Både 1. og 2. grad af normalisering har stor indflydelse på succesraten for sprogmodellerne hvor den 3. normaliseringsgrad påvirker resultatet i mindre grad. Det skyldes blandt andet at proprier ikke er en så udbredt ordklasse sammenlignet med andre ordklasser og en 100 % korrekt klassificering af proprier derfor stadig kun vil bidrage med en forholdsvis lav procentvis forbedring i den samlede succesrate af ordklassificeringer.

De normaliseringstrin der er benyttet i denne undersøgelse, har gennemsnitligt forbedret andelen af korrekt klassificerede ord fra 51,04 % til 81,83 % for SpaCy og fra 71,08 % til 85,36 % for DaCy. Det er, trods den store forbedring, stadig ikke nok til at komme over de gyldne 90 % som inden for digital humaniora kan antages som den accepterede grænse. Normaliseringen medvirker dog til at man når et langt stykke af vejen til at nå målet ved blot at normalisere

teksterne. Det diskuteres i følgende afsnit hvordan man videre vil kunne tilpasse sprogmodellerne for at nå et acceptabelt klassificeringsniveau.

Diskussion

Jeg har gennem artiklen præsenteret forskellige normaliseringsgrader for ældre salmetekster og evalueret to danske sprogmodellers succesrate for POS-tagging på hvert niveau af normaliseringsgraderne. Resultaterne viser at normalisering af ældre salmetekst forbedrer sprogmodellerne evne til POS-tagging væsentligt. Samtidig kan det konkluderes at normalisering af ældre salmetekst alene ikke kan medvirke til at sprogmodellernes procentvise succesrate kan nå et højt nok niveau til at de lever op til en tilfredsstillende standard. Ud fra denne mindre undersøgelse tyder det på at modellen DaCy virker mere robust for lavere normaliseringsgrader af dansksproget tekst end modellen SpaCy. Dette forhold udligner sig jo højere grad af normalisering teksten har gennemgået.

Selvom normalisering og ensretning af tekst altid er et spørgsmål af relevans i tilrettelæggelsen af et korpus, så skal man altid forholde sig til konsekvenserne ved en sådan ensretning af tekstkorpus. Er alle dele nødvendige? Og hvorfor? Og desuden, hvilke konsekvenser medfører det at modernisere versus hvis man ikke gør det? Det vil under alle omstændigheder altid være en konsekvens at videre undersøgelser af tekstdata i et korpus baserer sig på den måde teksten er repræsenteret i korpusset.

Domænespecifikke karakteristika i tekstkorpora kan også påvirke succesraten for sprogmodellernes klassificering. I en salmekontekst drejer det sig blandt andet om den særlige brug af versaler, der er blevet kommenteret tidligere. Desuden medvirker salmernes strofiske form og sangbarhed til at den syntaktiske struktur er afvigende fra en 'normal' sætnings syntaks, brugen af tegnsætning er friere end retskrivningens normer foreskriver – punktum afgrænser ikke nødvendigvis sætninger, og kommatering indikerer nogle gange linjeskift frem for syntaktisk placerede kommaer. Salmer har også en særlig brug af interjektioner og imperativer. Frekvensen af disse ordklasser i salmernes tekstdomæne kan altså variere meget fra den frekvens ordklasserne har i det træningsdatasæt som sprogmodellerne er trænet på, da salmer eller lignende tekstdata ikke er repræsenteret i tilstrækkelig grad i disse datasæt (Enevoldsen, Johansen m.fl. 2021).

En løsning på de domænespecifikke karakteristikas indvirken på sprogmodellernes klassificering af tekstdata kan være at træne sprogmodellerne yderligere på en mængde salmer der er manuelt annoterede med korrekte klassificeringer. Det er imidlertid et ressourcekrævende og omstændigt arbejde. Der er desuden også en begrænset mængde salmedata tilgængelig, og

det ville give dårligt mening at annotere en tilstrækkelig mængde salmedata til at optræne en model hvis der ikke ligger en tilsvarende stor mængde salmer tilbage som man kan benytte modellen på efterfølgende.

En måde at optimere sprogpakkerne til salmerne er at man manuelt går ind og retter sprogværktøjerne for specifikke situationer. Hyppige årsager til fejlklassificeringer i salmerne forekommer ved interjektionerne *O* og *Amen*. Disse optræder forholdsvis hyppigt i salmerne, og det vil være en mindre opgave at justere sprogpakkerne til altid at kategorisere disse ord som interjektioner frem for konjunktioner eller substantiver som de hyppigt fejlklassificeres som. En nærmere undersøgelse af strukturen for imperativer i salmerne vil på samme måde højst sandsynligt kunne afsløre strukturer for brugen af imperativer i salmerne. Hvis mønstre kan identificeres, kan de oversættes til regelbaserede rettelser af klassificeringen som kan implementeres i sprogmodellerne.

Litteratur

- Bigi, Brigitte (2014). A Multilingual Text Normalization Approach. *Human Language and Technology Challenges for Computer Science and Linguistics*, 8387:515-526.
- Den danske salmebog*. (2003). Det Kgl. Vajssenshus.
- Det Danske Sprog- og Litteraturselskab. (2023). *Vejledning: Sådan bruger du Danske Reformationssalmer. Melodier og tekster 1529-1573*. Det danske Sprog- og Litteraturselskab. Sidst tilgået den 29.01.2023 på <https://salmer.dsl.dk/guidelines>
- Enevoldsen, Kenneth; Johansen, Lasse & Nielbo, Kristoffer Laigaard (2021). DaCy: A Unified Framework for Danish NLP. *CEUR Workshop Proceedings*, 2989:206-216.
- Honnibal, Matthew & Montani, Ines (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. Sidst tilgået 02.02.2023 på <https://spacy.io/>
- Jacobsen, Henrik Galberg (2010). *Ret og Skrift: Officiel dansk retskrivning 1739-2005. Bind 1. Direktiver. Aktører. Normer*. Syddansk Universitetsforlag.
- Ondelli, Stefano (2018). Treat Texts as Data but Remember They Are Made of Words: Compiling and Preprocessing Corpora. I A. Tuzzi (Red.), *Tracing the Life Cycle of Ideas in the Humanities and Social Sciences* Springer International Publishing:133-150.
- Rossum, Guido van, & Drake, Fred L. (2009). Python 3 Reference Manual.
- sprogteknologi.dk. (2023). *sprogteknologi.dk*. Digitaliseringsstyrelsen. www.sprogteknologi.dk.
- Strømberg-Derczynski, Leon; Ciosici, Manuel R.; Christiansen, Morten H.; Baglini, Rebekah; Dalsgaard, Jacob Aarups; Fusaroli, Riccardo; Henrichsen, Peter Juel; Hvingelby, Rasmus; Kirkedal, Andreas & Kjeldsen, Alex Speed (2021). The Danish Gigawords Corpus. I *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*. Sverige. University Electronic Press.

1500-talssalmer og danske sprogmodeller

Thomissøn, Hans (1569). *Den danske Psalmebog / met mange Christelige Psalmer / Ordentlig tilsammenset formeret oc forbedret. Aff Hans Thomissøn. Prentet i Kiøbenhaffn / aff Laurentz Benedicht. Cum gratia et Privilegio Serenissimæ Regiæ Maiestatis.* Laurentz Benedicht.

Universal Dependencies (UD). (2014-2022). *Universal POS tags.* Universal Dependencies contributors. Sidst tilgået den 31.01.2023 på <https://universaldependencies.org/u/pos/all.html>

Resumé:

Den danske alveopalatale sibilant [ɶ]: Fonologi, ortografi og psykologisk realitet af et muligt /ɶ/ fonem

Sidsel Holm Rasmussen & Camilla Søballe Horslund

Aarhus Universitet

På dansk har vi to distinktive sibilanter: den alveolare frikativ [s] og den alveopalatale [ɶ]. Initialt er disse to kommutable lyde, altså betydningsadskillende i minimale par, som eksempelvis i sæl og sjæl. Alligevel analyseres den alveopalatale frikativ [ɶ] ofte som fusionen af fonemerne /s/ og /j/ (Grønnum 2005: 305-6; 2007: 118-121, Basbøll 2005: 152-3) med henvisning til fonotaks, fusioner på tværs af ordgrænser og i sammensatte ord, samt en ikke-udgivet undersøgelse (Beck 2004). Inden for anvendte discipliner, herunder førstesprogstilegnelse, andet-sprogstilegnelse og logopædi, behandles [ɶ] dog ofte som et fonem /ɶ/ (f.eks. Clausen og Fox-Boyer 2017, Bohn og Ellegaard 2019, Clausen og Fox-Boyer 2022). Men hvordan opfattes denne frikativ egentlig af danske sprogbrugere?

Fonetisk er [ɶ] utvetydigt en enkeltlyd, der hovedsageligt optræder i forlyd. Relativt nylige indlån, såsom *brunch* og *squash* udtales dog ofte med [ɶ] i udlyd (Andersen og Rathje 2007). Derudover optræder [ɶ] intervokalisk i ord som *mission* og *niche*.

Vi præsenterer et overblik over, hvor i det danske ordforråd [ɶ] optræder og kortlægger frekvensen af lydens forekomst på baggrund af korpussøgninger. Desuden undersøger vi vha. søgefunktionen Bogstavlyd (bogstavlyd.ku.dk), hvordan de ortografiske former af lyden optræder og fordeler sig i dansk retskrivning, og hvor forudsigeligt forholdet mellem [ɶ] og dets skriftlige repræsentation er i dansk. Med baggrund i ovenstående diskuterer vi argumenterne for og imod en monofonematisk tolkning af [ɶ] samt hvordan den psykologiske realitet af /ɶ/ versus /sj/ vil kunne testes eksperimentelt.

Litteratur

- Andersen, Margrethe Heidemann, og Marianne Rathje (2007) "Siger danskerne et fancy eller et fancyt bælte? Tilpasning af importord i dansk talesprog". *Moderne importord i språka i Norden, Stuntman og andre importord i Norden. Om udtale og bøjning*, 7:103–27.
- Basbøll, Hans (2005) *The phonology of Danish. The Phonology of the world's languages*. Oxford; New York: Oxford University Press.
- Beck, Jeppe Hjordt (2004) *Den psykologiske realitet af /ɛ/ – En psykolingvistisk undersøgelse*. Overbygningsopgave, Københavns Universitet.
- Bogstavlyd = <https://bogstavlyd.ku.dk>. Udviklet af Holger Juul i samarbejde med Philip Diderichsen.
- Bohn, Ocke-Schwen, & Anne Ellegaard (2019) Perceptual assimilation and graded discrimination as predictors of identification accuracy for learners differing in L2 experience: The case of Danish learners' perception of English initial fricatives. In *Proceedings of the 19th International Congress of Phonetic Sciences:2071-2074*.
- Clausen, Marit C. og Annette V. Fox-Boyer (2017) Phonological development of Danish-speaking children: A normative cross-sectional study. *Clinical Linguistics & Phonetics* 31 (6):440–458.
- Clausen, Marit C. og Annette V. Fox-Boyer. (2022) Diagnostic validity, accuracy and inter-rater reliability of a phonological assessment for Danish-speaking children. *Journal of Communication Disorders*, 95, 106168.
- Grønnum, Nina (2005) *Fonetik og fonologi: almen og dansk*. 3. udg. København: Akademisk Forlag.
- Grønnum, Nina (2007) *Rødgrød med fløde: en lille bog om dansk fonetik*. København: Akademisk Forlag.

Uden og med *-t*. Adverbialer dannet af adjektiver på *-ig* og *-lig*

Jørgen Schack & Eva Skafte Jensen

Dansk Sprognævn

1. Indledning

I denne artikel giver vi et historisk overblik over brugen, og til dels normeringen, af adverbialformer dannet af adjektiver, først og fremmest de adjektiver som ender på *-ig* og *-lig*, dvs. adjektiver som fx *rigtig* og *særlig*. Vi lægger ud med en kort oversigt over reglerne for brugen af former uden og med *-t* i de nuværende retskrivningsregler (Retskrivningsordbogen, 1. udg., 1986 – 4. udg., 2012) og i de foregående retskrivningsvejledninger (Retskrivningsordbog 1955, og Saabys Retskrivningsordbog 1918 ff.). Dernæst følger en oversigt over brugen af adverbialformer fra begyndelsen af den ældre nydanske periode og frem til slutningen af 1800-tallet. Vi har her særligt fokus på udviklingen i 1800-tallet og giver en række repræsentative eksempler på brugen af adverbialformer i forskellige syntaktiske kontekster. I den sidste del af artiklen vender vi tilbage til nutiden og giver her en række eksempler som skal illustrere moderne dansk sprogbrug i tilsvarende kontekster.

2. Hvad siger de nutidige retskrivningsregler?

De nutidige regler for brugen af *-t* på adverbialer dannet af adjektiver er beskrevet i Retskrivningsordbogen 1986-2012, § 36-39. Vi ser i det følgende bort fra adverbialer dannet af adjektiver på *-vis* (§ 39), som adskiller sig fra de øvrige adverbialer derved at adjektiverne på *-vis* for størstepartens vedkommende oprindeligt var rene adverbier, som senere (især i løbet af 1900-tallet) har udviklet en adjektivisk funktion (Galberg Jacobsen 2019:107). De har derfor i dag valgfrit *-t* i såvel adjektivisk som adverbial funktion, fx *et delvis(t) nederlag*; *De var kun delvis(t) uenige* (jf. Retskrivningsordbogen, § 27 og 39).

2.1. Hovedreglen

Adverbialer dannet af adjektiver ender som hovedregel på *-t*, hvis adjektivet kan have denne endelse (§ 36.2 om ”T-adverbialer”), fx *Hun synger smukt*; *Bygningen er rædsomt forfalden*;

Krisen var **hurtigt** forbi. Der er enkelte undtagelser fra reglen (udover de nedenfor nævnte systematiske undtagelser), fx *Jeg ser ham kun **sjælden(t)**; Og lad det nu gå lidt **tjep(t)**.*

2.2. Adverbialer dannet af adjektiver på **-ig** og **-lig**

De adverbialer som kan dannes af adjektiverne på **-ig** og **-lig**, kan eller skal skrives med **-t** afhængigt af deres syntaktiske funktion.

2.2.1. Måde

Mådesadverbialer (og andre adverbialer som specificerer verbalhandlingen) skal skrives med **-t** (§ 38.1), fx *Du må endelig ikke tage det **personligt**; Hun rejste sig **roligt** op; Bordplanen er **omhyggeligt** udtænkt.* Ved enkelte ord, nemlig *bitterlig* og *udtrykkelig*, er det dog valgfrit om man vil tilføje **-t**, fx *Jeg gjorde **udtrykkelig(t)** opmærksom på risikoen.*

2.2.3. Grad

Gradsadverbialer som står attributivt til et adjektivisk eller adverbialt led, har valgfrit **-t** (§ 38.2), fx *Jeg er ikke **særlig(t)** interesseret i fodbold; Bygningen er **skrækkelig(t)** forfalden; Det går **ualmindelig(t)** langsomt i dag.* Når adverbialet står som selvstændigt sætningsled, skal der imidlertid altid tilføjes **-t**, fx *Fodbold interesserer mig ikke **særligt**.*

2.2.4. ”Andre adverbialer”

Adverbialer dannet af adjektiver på **-ig** og **-lig** kan som bekendt have andre syntaktiske funktioner end de to der er beskrevet ovenfor (måde og grad). Disse øvrige funktioner er beskrevet samlet i Retskrivningsordbogens § 38.3 under overskriften ”Andre adverbialer.” Det drejer sig i hovedsagen om tids-, steds- og holdningsadverbialer. Sådanne adverbialer kan valgfrit skrives uden eller med **-t**, fx *Børnene vågnede **tidlig(t)** om morgenen (tid), Vinduerne skal males **indvendig(t)** (sted), **Personlig(t)** er jeg ikke i tvivl; Du kan **rolig(t)** blive hjemme (holdning).*

3. Den nære fortid

De ovenfor beskrevne regler for brugen af det adverbialle **-t** kan karakteriseres som en systematiseret, opstrammet og stærkt udvidet version af den tilsvarende regel i 1955-ordbogens retskrivningsvejledning (§ 30 om ”Biord dannet af tillægsord”):

I regelen bruges tillægsordenes intetkønsform som biord: *synge smukt, godt gjort*. I nogle tilfælde kan dog fælleskønsformen bruges som biord, nemlig når biordet står i uegentlig

betydning og betegner en grad: *fuldkommen tilfreds, grænseløs ulykkelig, rædsom grim*. Især er der mange biord, svarende til tillægsord på *-ig* og *-lig*, som det er almindeligt at skrive med *-t* eller uden *-t*, afhængende af om de står i deres egentlige, med tillægsordet stemmende betydning (betegnende en måde) eller de bruges for at angive en grad.

Som det fremgår, er udeladelse af *-t* her ikke begrænset til gradsadverbialer på *-ig* og *-lig* (jf. fx *rædsom grim*).

Den ovenfor citerede regel er, bortset fra en sproglig modernisering, identisk med den tilsvarende regel i 7. udgave af *Saabys Retskrivningsordbog* (1918), som var den første retskrivningsordbog der indeholdt en retskrivningsvejledning. Det skal dog bemærkes at 1918-reglen afsluttes med en ganske liberal passus vedrørende forskellen mellem 'egentlig' og 'uegentlig' betydning, som 1955-reglen udelader:

Det forlanges dog ikke, at denne Forskel skal overholdes, eller at der skal følges nogen bestemt Regel for, hvornaar man vil skrive disse af Tillægsord paa -ig dannede Biord med t og hvornaar uden t. (Saabys Retskrivningsordbog, § 45).

3.1. "Egentlig" vs. "uegentlig" betydning

Som det er fremgået, er den 'gamle' regel om brugen af adverbialt *-t* (Saaby 1918 ff. og Retskrivningsordbog 1955) en semantisk-syntaktisk motiveret regel, idet den skelner mellem de tilfælde hvor adjektivet bruges med 'uegentlig betydning', dvs. i funktionen som gradsadverbial (fx *rædsom grim*), og de tilfælde hvor adverbialet bibeholder adjektivets 'egentlige betydning', dvs. i funktionen som mådesadverbial (fx *Det lyder rædsomt*).

Denne regel skyldes skolemanden og sprogforskeren Axel Sørensen, som i en afhandling med den beskedne titel *En lille bemærkning om de danske biord* (1887) udleder et princip for brugen af *-t* på grundlag af en "ganske bestemt påvislig praxis, ikke blot i det dannede talesprog, men hos så godt som alle de bedste forfattere, i alt fald i dette århundrede" (Sørensen 1887:416). Sørensens regel for adverbialernes form lyder som følger:

- a) Så længe de er brugte som mådesbiord føjede til udsagnsord og i glosebetydning falder sammen med det tillægsord, hvoraf de er dannede, antager de efter hovedregelen tillægsordenes intetkønsform.

- b) Når de derimod står som gradsbiord tilføjede til andre ordklasser end udsagnsord og de tillige i betydning fjerner sig fra den oprindelige, dannes de af tillægsordenes fælleskønsform. (Sörensen 1887:429)

Sörensens afhandling indeholder en omfattende eksempelsamling, som skal påvise at reglen i udstrakt grad følges af 1800-tallets ”bedste forfattere”, bl.a. ”(en Person, der) mener det saa *inderlig* vel med hele Verden” (Kierkegaard) vs. ”– hun elskede jo dette Sted saa *inderligt*” (Thom. Lange).

4. Historisk tilbageblik

Adjektiverne på *-ig* og *-lig* har i dag kun to mulige adverbialformer, nemlig uden og med *-t*, fx *rigtig(t)*, *særlig(t)*, jf. ovenfor. Tidligere har andre former været almindelige, jf. fx Ruus (2019:89) om situationen i ældre nydansk (1500-1700):

Ganske mange adverbier er dannet af adjektiver afledt med derivativet *-lig*. Disse adverbier slutter i begyndelsen af den ældre nydanske periode næsten udelukkende på *-lige*, mens former på *-ligen* bliver almindelige hen i perioden.

Ruus (2019:89) giver en række illustrerende eksempler, bl.a.

*Oc beder ieg alle **kerlige** som samme bog læse* (C. Pedersen 1533)

’Og beder jeg kærligt alle som læser samme bog’

*Her maa mand nu **synderligen** giffue act paa tuende Stycker* (Kogebog 1616)

’Her må man nu især give agt på to ting’

4.1. Tidslinje

Den nedenstående tidslinje illustrerer udviklingen i brugen af adverbialformer dannet af adjektiver på *-ig* og *-lig* fra ca. 1500 til i dag. De nævnte former har været brugt i alle adverbielle funktioner. Tidslinjen er udfærdiget med udgangspunkt i beskrivelserne hos Falk & Torp (1900:110-113), Skautrup (1947:239-240, 383-384 og 1953:373-374), Ruus (2019:88-89), Galberg Jacobsen (2019:109-110) og viser i grove træk adverbialformernes kronologiske rækkefølge. Som det vil fremgå, sameksisterer nogle former i lange perioder, og visse former har i bestemte perioder været dominerende, mens andre (sameksisterende) former

har været sjældnere brugt. Det gælder bl.a. former på *-lig*, som har været i brug i hele den beskrevne periode (dvs. fra ca. 1500 til i dag). I ældre nydansk har man først og fremmest brugt former på *-lige* og *-ligen* (jf. Ruus 2019:89), men også former på *-lig* forekom. Fx anfører Skautrup (1947:240) om Chr. Pedersen (1480-1554) at han ”anvender alle endelser [dvs. *-lig*, *-lige*, *-ligen*], og nogen konsekvens er det vanskeligt at få øje på.” Om den yngste af formerne, nemlig formen på *-(l)igt*, som vandt frem i løbet af 1800-tallet, kan det bemærkes at den var uglest af de fleste af samtidens grammatikere, herunder Rask og N. M. Petersen, der opfattede den som en svecisme (jf. Skautrup 1953:373 f.).

-(l)ige Formerne på **-(l)ige** (fx *daglige*): rækker bagud til gammeldansk (dvs. før 1500). De går af brug i løbet af det 17. årh.

-(l)igen Formerne på **-(l)igen** (fx *dagligen*) er formentlig indlånt fra tysk, især middelnedertysk. De sameksisterer med former på *-(l)ige* i ældre nydansk. Endelsen *-(l)igen* ophører med at være produktiv i løbet af det 17. årh.

-(l)ig Formerne på **-(l)ig** (fx *daglig*) kan spores helt tilbage til det ældste nydansk (dvs. omkring 1500), men de begynder først at blive hyppige i løbet af det 17. og 18 årh., i takt med at først formerne på *-(l)ige* og dernæst formerne på *-(l)igen* går af brug.

-(l)igt Formerne på **-(l)igt** (fx *dagligt*) er de yngste af formerne; de optræder fra tidligt i det 19. årh. Efter at formerne på *-(l)ige* og *-(l)igen* gik af brug, opstod muligheden for tilføjelse af *-t* efter *-(l)ig* som markør for adverbial funktion.

4.2. Sprogbrugen i 1800-tallet

I 1800-tallet brugte man overvejende former på *-(l)ig* og *-(l)igt*. Frem til ca. midten af perioden holdt man dog fast ved former på *-(l)igen* i ”gammeldags eller højtidelig stil”, men i ”jævnprosaen” var de forældede (Skautrup 1953:373).

Formerne på *-(li)g* og *-(l)igt* blev brugt i alle adverbial funktioner. Den nyeste af formerne, *-(l)igt*, blev først og fremmest brugt som mådesadverbial (jf. Sørensen 1887:429), men også i denne funktion var former på *-(l)ig* i mange tilfælde de almindeligste. I tabel 1, 2 og 3 anføres en række repræsentative eksempler på brugen af former på *-(li)g* og *-(l)igt* som mådes-, grads- og tidsadverbialer. Optællingen af forekomster er foretaget i et korpus på ca. 2,75 mio. løbende ord. Korpusset indeholder tekster fra hele 1800-tallet, først og fremmest skønlitterære tekster (hentet fra Arkiv for Dansk Litteratur, KB) og breve (hentet fra Danmarks Breve, KB).

Tabel 1		
Mådesadverbialer	1800-talskorpus	Hyppigste form
hastig	186	uden -t signifikant hyppigst
hastigt	51	
Eks. <i>Skyerne jog hastig hen over de gyngende Trætoppe</i> (Gjellerup 1896)		
grundig	15	ikke signifikant forskel
grundigt	7	
Eks. <i>Thi skulle vi filosofere, saa lad os gjøre det grundigt</i> (Bergsøe 1877)		
omhyggelig	54	uden -t signifikant hyppigst
omhyggeligt	33	
Eks. <i>De bleve alle omhyggelig renskrevne og nedlagte i et Skriin</i> (Hauch 1839)		

Tabel 2		
Gradsadverbialer	1800-talskorpus	Hyppigste form
uendelig meget	16	uden -t signifikant hyppigst
uendeligt meget	1	
Eks. <i>Og denne Forvisning er mig uendelig meget mere værd</i> (Kierkegaard 1844)		
rigtig god/godt/gode	23	uden -t signifikant hyppigst
rigtigt god/godt/gode	0	
Eks. <i>"Ikke sandt," sagde den gamle Dame, "rigtig kjønt, rigtig godt"</i> (Topsøe 1875)		
fuldstændig + adj./perf. part.	43	uden -t signifikant hyppigst
fuldstændigt + adj./perf. part.	14	
Eks. <i>det store tempel, hvis søjlegang er næsten fuldstændig bevaret</i> (H.C. Andersen 1837)		

Tabel 3		
Tidsadverbialer	1800-talskorpus	Hyppigste form
daglig	97	uden -t signifikant hyppigst
dagligt	5	
Eks. <i>han betragtede daglig Prindsessens Billede</i> (Ingemann 1826)		
jevnlig/jævnlig	29	uden -t signifikant hyppigst
jevnligt/jævnligt	9	
Eks. <i>Han var jævnlig tvær</i> (Gjellerup 1896)		
øjeblikkelig	22	uden -t signifikant hyppigst
øjeblikkeligt	0	
<i>Smilet forsvandt øjeblikkelig af Karens Ansigt</i> (Wied 1899)		

Som det fremgår af tabel 1, 2 og 3, var de -t-løse former dominerende i 1800-tallets sprogbrug i almindelige adverbialer kontekster. I det følgende afsnit ser vi på nutidens sprogbrug i tilsvarende kontekster.

5. Punktnedslag i nutidens sprogbrug

Som mange sikkert har bemærket, har vi i dag former på *-(l)igt* i mange af de tilfælde hvor former på *-(l)ig* tidligere var de almindelige. Vi har undersøgt den nutidige sprogbrug i et avis-korpus på ca. 57,5 mio. løbende ord. Korpusset indeholder tekster fra 2017 (landsdækkende og regionale aviser).

Brugt som mådesadverbial er former på *-(l)igt* i dag helt dominerende, og sprogbrugen er altså her i overensstemmelse med den norm som fastsættes i de nyere retskrivningsregler og -vejledninger (jf. afsnit 2 og 3), fx

- (1) Lignende tendenser ses **tydeligt** i Europa (Berlingske 2017)
- (2) Nu nærmer forholdet sig **hastigt** frysepunktet (Jyllands-Posten 2017)

Former på *-lig* forekommer så vidt vi kan bedømme, kun pletvis i vores korpus, fx

- (3) Der er ingen tvivl om, at borgerne skal behandles **ordentlig** (JydskeVestkysten 2017)
- (4) På verdensplan går udviklingen også meget **hurtig** (Berlingske 2017)

I funktionen som gradsadverbial er former på *-(l)ig* i visse tilfælde almindelige eller dominerende, men det samlede billede er noget broget, da nogle gradsadverbialer overvejende har former på *-(l)ig*, mens andre overvejende har former på *-(l)igt*. I tabel 4 anføres nogle almindeligt forekommende forbindelser af gradsadverbial + adjektiv/participium.

Tabel 4		
Gradsadverbialer på <i>-(l)ig</i>	Aviser 2017	Hyppigste form
alvorlig syg/sygt/syge	11	
alvorligt syg/sygt/syge	208	med <i>-t</i> signifikant hyppigst
fuldstændig rigtig/rigtigt/rigtige	137	uden <i>-t</i> signifikant hyppigst
fuldstændigt rigtig/rigtigt/rigtige	25	
kraftig faldende/stigende	2	
kraftigt faldende/stigende	42	med <i>-t</i> signifikant hyppigst
rigtig god/godt/gode	3251	uden <i>-t</i> signifikant hyppigst
rigtigt god/godt/gode	475	
særlig stor/stort/store	143	uden <i>-t</i> signifikant hyppigst
særligt stor/stort/store	70	
væsentlig mere	63	
væsentligt mere	106	med <i>-t</i> signifikant hyppigst

Tidsadverbialer har i dag, i modsætning til i 1800-tallet, typisk former på *-(l)igt*, men der er enkelte undtagelser, bl.a. *foreløbig(t)*. I tabel 5 anføres nogle almindeligt forekommende ordforbindelser samt nogle tilfælde hvor vi markerer adverbialets placering i sætningen: forfelt ("F") eller "lille a".

Tabel 5		
Tidsadverbialer på -(l)ig	Aviser 2017	Hyppigste form
er/har jævnlig (lille a)	1	
er/har jævnligt (lille a)	47	med -t signifikant hyppigst
Eks. <i>Problemkomplekset er jævnligt oppe at vende i danske medier</i> (Weekendavisen)		
x gange daglig , ugentlig osv.	5	
x gange dagligt , ugentligt osv.	174	med -t signifikant hyppigst
Eks. <i>Easy Jet flyver mellem København og Bristol to gange ugentligt</i> (Fyens Stiftst.)		
tidlig om morgenen	0	
tidligt om morgenen	96	med -t signifikant hyppigst
Eks. <i>Jeg sender ofte en mail meget tidligt om morgenen</i> (Fyens Stiftst.)		
midlertidig + supinum (lille a)	6	
midlertidigt + supinum (lille a)	55	med -t signifikant hyppigst
Eks. <i>Alicia har midlertidigt forladt scenen</i> (Politiken)		
er/har foreløbig (lille a)	261	uden -t signifikant hyppigst
er/har foreløbigt (lille a)	56	
Eks. <i>I år har Rafael Nadal foreløbig været næsten uovervindelig på grus</i> (Berlingske)		

Holdningsadverbialer skrives altovervejende med -t (tabel 6).

Holdningsadverbialer på -(l)ig	Aviser 2017	Hyppigste form
personlig er/har (F)	2	
personligt er/har (F)	103	med -t signifikant hyppigst
Eks. <i>Personligt er jeg vild med den slags små detaljer</i> (Jyllands-Posten)		
kan rolig (lille a)	2	
kan roligt (lille a)	110	med -t signifikant hyppigst
Eks. <i>Man kan roligt regne med at dommen vil møde modstand i Danmark</i> (Information)		
kan/kunne dårlig (lille a)	1	
kan/kunne dårligt (lille a)	90	med -t signifikant hyppigst
Eks. <i>Københavnerne kan dårligt nok komme frem for vejarbejde</i> (Berlingske)		

Ordet *endelig* brugt som holdningsadverbial skrives dog altovervejende uden -t i forbindelser som fx *Du må **endelig** ikke tage det personligt, når folk er perfide* (Jyllands-Posten 2017). Her

er sprogbrugen i overensstemmelse med Retskrivningsordbogen 1986-2012, som skelner mellem 1. **endelig** *adj.*, -t, -e (*afsluttende, definitiv*) og 2. **endelig** *adv.* (*til slut; omsider; absolut*).

I modsætning til holdningsadverbialerne skrives de epistemiske adverbialer på *-lig* typisk uden *-t* (tabel 7).

Tabel 7		
Epistemiske adverbialer på <i>-(l)ig</i>	Aviser 2017	Hyppigste form
har/er antagelig (lille a)	11	uden <i>-t</i> signifikant hyppigst
har/er antageligt (lille a)	1	
Eks. <i>Personalet gør antagelig alt, hvad de kan</i> (Nordjyske Stiftst.)		
har/er formentlig (lille a)	550	uden <i>-t</i> signifikant hyppigst
har/er formentligt (lille a)	21	
Eks. <i>Dette er formentlig sagt med et glimt i øjet</i> (Jyllands-Posten)		
har/er formodentlig (lille a)	49	uden <i>-t</i> signifikant hyppigst
har/er formodentligt (lille a)	8	
Eks. <i>Ministrene har formodentlig handlet i god tro</i> (Nordjyske Stiftst.)		

Om ordet *antagelig* skal det bemærkes at Retskrivningsordbogen (1986-2012) skelner mellem 1. **antagelig** *adj.*, -t, -e (*tilfredsstillende; ret stor*), og 2. **antagelig** *adv.* (*formodentlig*). I funktionen som epistemisk adverbial skal *antagelig* altså ifølge ordbogen altid skrives uden *-t*. Opslagsordene *formentlig* og *formodentlig* er derimod kun opført som adjektiver (begge med anmærkningen ”*som adv. også uden -t*”). I 1955-ordbogen er *formentlig* opført som ”*bio., to., -t*”, mens *antagelig* og *formodentlig* er opført som ”*to., -t*”. Vi kan ikke se at der er grund til at behandle *antagelig*, *formentlig* og *formodentlig* forskelligt hvad angår muligheden for at tilføje *-t* i funktionen som sætningsadverbial, ligesom ordklassebestemmelsen i ordbøgerne kan forekomme vilkårlig. (Sådanne formers ordklassetilhørsforhold, dvs. om de er bøjningsformer af adjektiver, afledte adverbier eller måske noget helt tredje, diskuteres i Jensen & Schack 2023). En undersøgelse af sprogbrugen i et meget større nutidigt aviskorpus (ca. 2,4 mia. ord) viser at *antagelig* i denne funktion har *-t* i ca. 50 % af forekomsterne, faktisk langt oftere end *formentlig* og *formodentlig*, som i samme korpus har *-t* i henholdsvis ca. 6 % og ca. 10 % af forekomsterne.

6. Afsluttende bemærkninger

Som det er fremgået, har adverbialformer på *-(l)igt* vundet terræn i nyere tid. I 1800-tallet var former på *-(l)ig* dominerende i centrale adverbialer kontekster, inkl. i funktionen som mådesadverbial (jf. afsnit 4.2). I løbet af 1900-tallet er former på *-(l)igt* i visse kontekster blevet

almindelige/dominerende (jf. fx Hansen 1967:209 f.). I nutidigt dansk gælder dette først og fremmest ordformer der bruges som mådesadverbialer, fx *hastig, tydelig*, og som tids- og holdningsadverbialer, fx *daglig, jævnlig; dårlig, rolig*. Ved nogle ordformer der bruges som gradsadverbialer, er former uden *-t* dominerende, fx *fuldstændig, rigtig*; ved andre ordformer er former med *-t* almindeligst, fx *alvorlig, kraftig*. Den lille gruppe af epistemiske adverbialer på *-lig*, *antagelig, formentlig, formodentlig*, er, som helhed betragtet, den mest *-t*-resistente adverbialkategori. Dette kan muligvis hænge sammen med at de tre nævnte ord altovervejende bruges adverbielt og derfor kan opfattes som ”rene adverbier” (Retskrivningsordbogen 1986-2012, § 36), jf. også *-t*-løse epistemiske adverbialer som *givetvis, sandsynligvis* og *tydeligvis*. Men adverbialkategoriernes forskellige præference for former uden og med *-t* fortjener en grundigere undersøgelse. Det samme gælder de forskelle som kan observeres blandt ord som hører til samme adverbialkategori, først og fremmest de forskelle som ses indenfor kategorien gradsadverbialer.

Litteratur

- Falk, Hjalmar & Alf Torp (1900) *Dansk-norskens syntax i historisk fremstilling*. Kristiania: Aschehoug & Co.
- Galberg Jacobsen, Henrik (2019) Yngre nydansk. I *Dansk Sproghistorie. Bind 3. Bøjning og bygning*. København og Århus: Det Danske Sprog- og Litteraturselskab/Aarhus Universitetsforlag.
- Hansen, Aage (1967) *Moderne Dansk. Bind 3*. København: Grafisk Forlag/Det danske Sprog- og Litteraturselskab.
- Jensen, Eva Skafte & Jørgen Schack (2023) Ubøjede adverbialer uden og med *-t* og historien bag. I *Ny forskning i grammatik* 30. 57-73.
- Retskrivningsordbog 1955. Udgivet af Dansk Sprognævn.
- Retskrivningsordbogen 1986-2012. 1. udg., 1986, 2. udg., 1996, 3. udg., 2001, 4. udg. 2012. Udgivet af Dansk Sprognævn.
- Ruus, Hanne (2019) Ældre nydansk. I *Dansk Sproghistorie. Bind 3. Bøjning og bygning*. København og Århus: Det Danske Sprog- og Litteraturselskab/Aarhus Universitetsforlag.
- Saaby, Viggo: Dansk Retskrivningsordbog, 7. udg., 1918.
- Skautrup, Peter (1947) *Det danske sprogs historie*. Bind 2. København: Gyldendal.
- Skautrup, Peter (1953) *Det danske sprogs historie*. Bind 3. København: Gyldendal.
- Sørensen, Axel (1887) En lille bemærkning om de danske biord. *Vor Ungdom*:416-430.

Rækkefølgens betydning i bindestregs-nationalitetsadjektiver som *tysk-dansk* og *dansk-tysk*

Irene Simonsen

Syddansk Universitet

Emnet for denne artikel er kollokationer med bindestregsadjektiver, der betegner nationalitet. Hvorfor hedder det fx *dansk-polsk* i nogen sammenhænge, men *polsk-dansk* i andre?¹ Er rækkefølgen ligegyldig, eller er den udtryk for en bestemt værdisættelse af de to nationaliteter, som adjektiverne betegner, måske en underordning af førsteled under andetled eller eventuelt omvendt? I artiklen sammenlignes brugen af nationale bindestregsadjektiver på tysk og dansk. Det undersøges, om der er semantiske forskelle på de substantiviske højre naboor, der attribueres af bindestregssammensætninger med hhv. *dansk* og *deutsch* som første- og andetled i hhv. udskiftelig og fast rækkefølge. De to Sketch Engine-korpora daTenTen20 og deTenTen18 tjener som empirisk grundlag for sammenligningen.

Som baggrund for undersøgelsen redegøres der i første afsnit for bindestregsbrug, nationalitetsadjektiver og rækkefølgen i adjektivkomposita.

1. Bindestreg, nationalitetsadjektiver og rækkefølgen i adjektivkomposita

1.1 Bindestregen

RO behandler det, der på tysk benævnes *Trennstrich* ('delestreg' (DuG 915)), *Ergänzungsstrich* ('suppleringsstreg' (AR §98), fx *Haupt- und Nebeneingang*) og *Bindestrich* ('bindestreg') under et (RO §57). I de to sprog er der dog enighed om at karakterisere bindestregen som et ordinternt syngrafem, der angiver, at bindestregsforbundne dele skal opfattes som en enhed (DuG 5.1.3, Hansen 1967:368), og som kan have segmenterende eller klassificerende funktion (Gallmann 1996:13 f.). Den segmenterende funktion ses, hvor syngrafemet fx forbinder tal, forkortelser og proprier med appellativer, fx *1900-tallet*, *100-prozentig*; *a-våben*, *UNO-Sicherheitsrat*; *Magritte-udstilling*, *Goethe-Ausgabe* (RO §57; AR §40, §51). Det gælder også

¹ Jf. to vilkårligt valgte korpuseksemppler: "Storstilet *dansk-polsk* gasprojekt rykker tættere på." (<https://ing.dk/artikel/storstilet-dansk-polsk-gasprojekt-rykker-taettere-paa-203721>) vs. "Ministeren fremhæver bl.a. investeringen på et *polsk-dansk* erhvervsmøde [...]" (<http://www.bygtek.dk/artikel/isolering/investerer-over-halv-milliard-i-ny-fabrik-i-polen>).

gruppesammensætninger, som *det-vil-jeg-ikke-finde-mig-i-attitude* (RO§ 57.7), *das Auf-die-lange-Bank-Schieben* (~'denne-putten-i-syltekrukke', AR §43), eller markeringen af ordgrænse i sammensætninger, hvor de enkelte konstituenten ikke umiddelbart kan skelnes fra hinanden, fx *Drucker-Zeugnis/Druck-Erzeugnis* ('trykkercertifikat'/'tryksag', AR §45), *vejret/vej-ret* ofl. (Hansen 1967:369 f.). Klassificerende synes bindestregen at være, hvor den tydeliggør, at en kopulativ sammensætnings dele semantisk skal opfattes som sideordnede (jf. Gallmann 1996:14). Bindestreg sættes således på begge sprog mellem "betydningsmæssigt sidestillede" (RO § 57, 5) eller "sideordnede adjektiver af samme rang" ("gleichrangigen, nebengeordneten"; AR § 44, 2):

"§ 57. Bindestreg	"§ 44
(5) Sammensætninger med sidestillede led	Man sætter en bindestreg i [...]
Bindestreg bruges i sammensætninger (og	2) uoverskuelige sammensætninger af
afledninger med tungere suffikser) hvis dele er	sideordnede adjektiver af samme rang, for
betydningsmæssigt sidestillede:	eksempel:
[...] <i>dansk-norsk-svensk samarbejde</i> [...]"	[...] <i>ein lateinisch-deutsches Wörterbuch;</i>
(RO)	<i>deutsch-österreichische Angelegenheiten</i>
	[...]“ (AR, e.o.)

Sammenfattende kan fastholdes, at bindestregen sættes i kopulative sammensætninger, hvilket dog ikke udelukker, at den også sættes i (uoverskuelige) determinative sammensætninger.

1.2 Nationalitetsadjektiver

Hvad jeg i denne artikel har kaldt 'nationalitetsadjektiver', kan ses som en delgruppe af 'etniske adjektiver' (Alexiadou & Stavrou 2011) eller 'gruppeadjektiver' (Ørsnes & Markantonatou 2002), som igen er en underkategori til 'relationelle adjektiver' (jf. Horn ea. 2014)². Semantisk dækker de relationelle adjektiver en række forskellige domæner, herunder kategorier, som "materiale, formål, funktion, status og rang, oprindelse og modus operandi" (Horn ea. 2014:5 f.); gruppeadjektiver inkluderer heroverfor kun afledninger af stednavne og professionsbetegnelser (Ørsnes & Markantonatou 2002:139), etnicitetsadjektiver kun adjektiver med reference til oprindelse og herkomst ("origin or provenance", Alexiadou & Stavrou, 2011:118), og denne undersøgelse altså kun adjektiver med reference til nationalitet.

² Tak til Jan Rijkhoff for henvisning.

Fælles træk for de nævnte adjektiver er, at de er nominalt afledte, og at de i modsætning til øvrige adjektiver kan fungere klassificerende, idet de kan specificere ”en bestemt underkategori af den enhed, som det attribuerede substantiv/hovedsubstantivet betegner” (Horn ea. 2014:3), altså fx *en dansk dronning* (klassificerende) vs. *dronningen er dansk* (kvalificerende). I de nævnte kilder er negative fællestræk ved adjektiverne i deres klassificerende funktion (ift. kvalificerende adjektiver), 1) at de ikke kan optræde prædikativt, 2) at de ikke kan modificeres, herunder kompareres eller negeres på samme måde som kvalificerende adjektiver (*ikkedansk* (klassificerende) vs. *udansk* (kvalificerende)), 3) at de godt kan attribueres, men ikke kan koordineres med kvalificerende adjektiver (*elsket dansk dronning* vs. *elsket og dansk dronning*), og 4) at de ikke kan adskilles fra deres hoved (**en dansk elsket dronning*) (jf. Horn et al. 2014:13, Ørsnes & Markantonatou 2002:140; Alexiadou & Stavrou 2011:121). Herudover er det som nævnt karakteristisk for adjektiverne, at de er flertydige: Dels udviser de, hvad Horn et al. (2014:19) beskriver som ”intersektiv gradiens”, nemlig ”overlap i forhold til ordklassemedlemskab” (substantiv/adjektiv), dels er de flertydige, i og med at de altså kan fungere som og læses som klassificerende eller kvalificerende (ibid.; jf. Ørsnes & Markantonatou 2002:140). Specifikt ift. det tredje træk peger Ørsnes & Markantonatou (2002:175) på, at et syntagma med bindestregs-kompositum (*den dansk-tyske beslutning*) i modsætning til et syntagma, hvor nationalitetsbeskriverne kobles med ’og’ (*den danske og tyske beslutning*), ikke har distributiv, men kollektiv reference.

At de relationelle adjektiver, og herunder altså nationalitetsadjektiverne, udgør en særskilt gruppe adjektiver, der tilsyneladende kun har det positive definitionskriterie til fælles med de fleste øvrige adjektiver, at de kan anvendes som attributiv til nominal (jf. Horn et al. 2014:12), synes at spille en rolle, når disse adjektiver optræder i komposita, hvilket er emnet for det følgende afsnit.

1.3 rækkefølgen i adjektivkomposita

I referencegrammatikkerne på både dansk og tysk skelnes der uafhængigt af ordklasse grundlæggende mellem to slags komposita, nemlig dels ”determinative” eller ”attributive sammensætninger”, dels ”kopulative sammensætninger” eller ”dvandvasammensætninger” (DuG 1074; DGS §30.1, §30.3c, Anm.). I udgangspunktet afgrænses de to typer ud fra følgende forskelle: Mens determinativkomposita har en binær struktur, er det kendetegnende for kopulativkomposita, at de kan være flerleddede. Determinativkompositumets konstituenten kan tilhøre forskellige ordklasser, hvorimod kopulativkompositumets konstituenten tilhører samme ordklasse, og ifølge DuG 1372, hvis der er tale om et adjektivkompositum, også samme semantiske

kategori. I determinativkompositummet determinerer førsteleddet andetleddet, mens det er definerende for kopulativkompositummet, at leddene er sideordnede og som sådan ligeværdige eller ”symmetriske” (DuG 1372). Determinativkompositumets andetled dominerer semantisk og repræsenterer hele kompositummet, mens det kopulative kompositums betydning fremgår additivt som en sum af de enkelte, i princippet vilkårligt ordnede konstituenters betydning (GDS §30; DuG 1074, 1075, 1258). Specifikt om adjektivkomposita – og indskrænkende ift. denne modstilling – fremhæves det i DuG 1372 om kopulative adjektivkomposita, at sidsteleddet også i kopulativkomposita betragtes som hoved, fordi det bestemmer kompositumets grammatiske egenskaber, hvilket må ses som en indskrænkning af ligeværdigheden mellem leddene. Samtidig understreges det ift. rækkefølgen af det kopulative kompositums konstituenters, at selvom den ganske vist anses for vilkårlig, er der oftest en bestemt rækkefølge, der er leksikaliseret. Ydre forhold kan desuden være afgørende for rækkefølgen, herunder fx rækkefølgen af et flags farver eller en ordbogs sprogpar (*russisch-deutsch* vs. *deutsch-russisch*).

Den grundlæggende todeling i determinative og kopulative komposita har bl.a. Donalies (1996) kritiseret (jf. også Breindl & Thurmair 1992:46). Donalies betvivler flere af de opridsede kriterier, ja forkaster ift. kopulative komposita kriteriet om konstituenternes ”vilkårlige rækkefølge uden væsentlig betydningsændring”³ som generelt gyldigt (Donalies 1996:284), og konkluderer om kopulativkomposita: ”minimalkriteriet ’samme ordklasse’ gælder” (ibid.:285 e.o.). ”Udslagsgivende er kriteriet ’sideordning af konstituenterne’, som viser sig i bestemte parafraser” (Donalies 1996:285, e.o.). Ift. adjektivkomposita foreslår Donalies (1996, afsnit 3) en inddeling, som søges illustreret i figur 1.

Donalies skelner mellem 1) farveadjektiver, 2) adjektiver, der betegner tilhørsforhold til person(grupp)er og 3) øvrige adjektiver. Interessant ift. nationalitetsadjektiver er undergrupperne 2.1 og 2.2, hvor Donalies parafraserer de kopulative komposita til hhv. ’*både* egenskab et og to’ og ’*mellem* de personer/institutioner, som de nominaler, som adjektiv 1 og 2 er afledt af, betegner’ (Donalies 1996:279, e.o.; jf. hertil også Gudiksen 2019:302; Pümpel-Mader et al. 1992:36). *Ottotisk* og *salisk* i 2.3 er afledninger af hhv. *proprierne* Otto I/II/III (10. årh.) og det østfrankiske kejserslægtsnavn *Salierne* (11.-12. årh.). Mens konstituenternes rækkefølge i denne type kompositum til betegnelse af bestemt kronologi er fast, gælder for type 2.1 og 2.2 ifølge Donalies (1996:279), at rækkefølgen principielt er arbitrær med den indskrænkning, at semantisk vægtning og bestemt navngivning (ibid.) kan have medført leksikaliseringen af en bestemt rækkefølge.

³ Egen oversættelse, markeres i teksten i øvrigt med ’e.o.’.

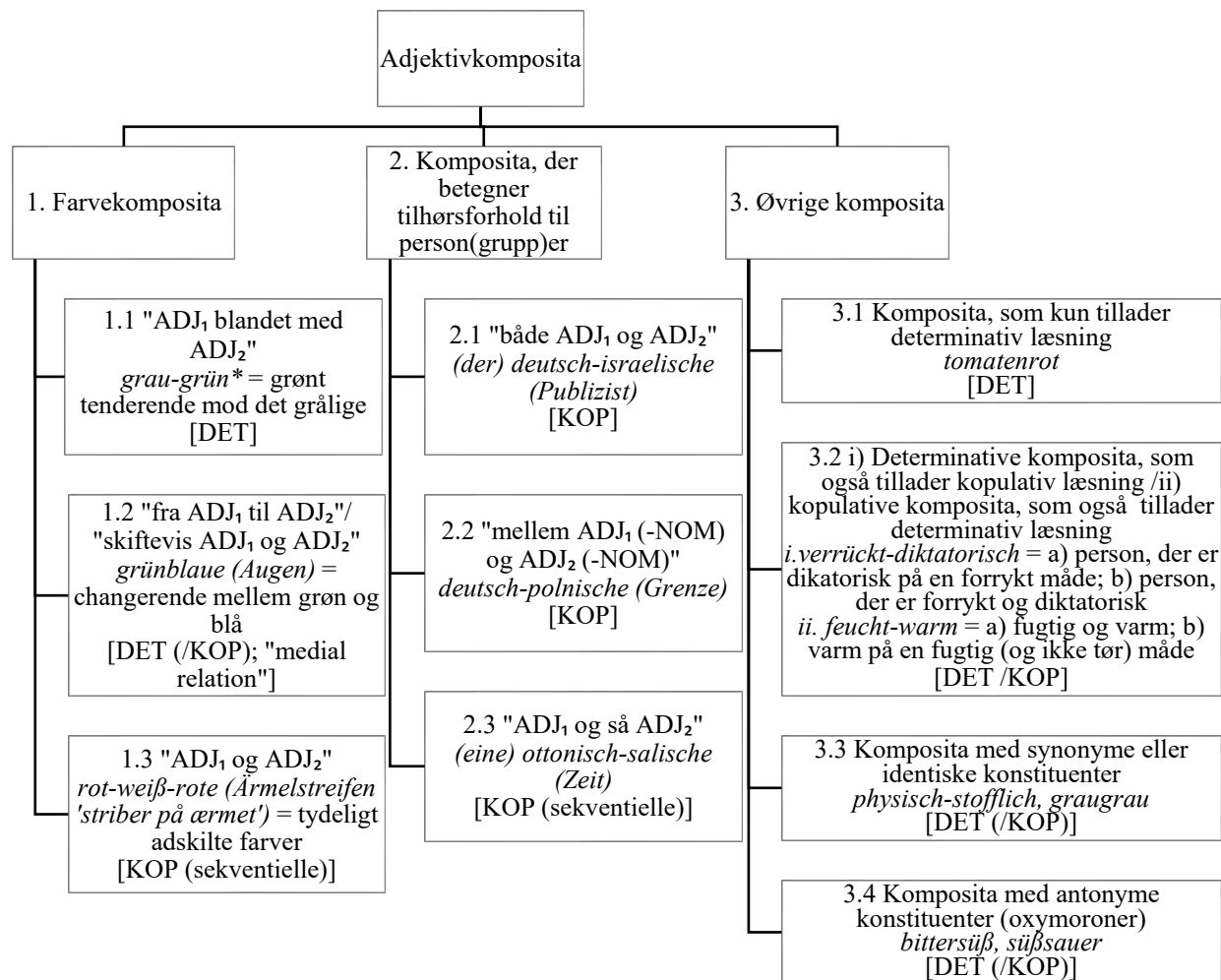


Fig. 1. Oversigt over inddelingen af adjektivkomposita hos Donalies (1996: afsnit 3).

(DET = entydigt determinative komposita; KOP: entydigt kopulative komposita; DET/KOP: både kopulativ og determinativ læsning er mulig. DET(/KOP): i udgangspunktet determinativ, men tilladende kopulativ læsning).⁴

Undersøgelsen begrænser sig til en undersøgelse af attributive kollokationer, dvs. den toleddede syntagmatiske kombination af [NADJ₁-NADJ₂] + [substantiv (appellativ)], og udelukkende umiddelbare højre naboorde til bindestregskomposita undersøges.

At der fokuseres på kollokationer, begrundes i forventningen om, at deres leksikaliserede form er udtryk for stereotypi og kan blotlægge ”kollektiv, socialt generel viden eller viden, der er udbredt i bestemte grupper” (jf. Stein og Stumpf 2019:180, e.o.). Den leksikaliserede form,

⁴ *grau-grün* skriver Donalies (1996:277) med bindestreg, selvom det betragtes som determinativt kompositum og fx står sammen med andre bindestregsløse blandingsfarve-komposita, som *rotbraun*. Jf. også DGS 33b Anm. om det ”hypotaktiske kompositum” *rødbrun*.

som kollokationen har, ses altså som udtryk for faste tankemønstre ("Denkmuster", Stein & Stumpf 2019:193). I denne forstand er undersøgelsens nationaladjektiv-kollokationer samtidig nært beslægtede med "associative semantiske stereotyper" ("assoziative semantische Stereotype", som de beskrives af Quasthoff og Hallsteinsdóttir:

Associative semantiske stereotyper er potentielle betydninger af sproglige tegn [...] generaliserende forestillinger, som i tidens løb er blevet kollektivt associeret med ordet af sprogbrugere eller allerede er blevet leksikaliseret som en del af ordets betydning. (Hallsteinsdóttir 2016:37-38 e.o.; jf. også Quasthoff&Hallsteinsdóttir 2016:350).

Quasthoff og Hallsteinsdóttir (2016) undersøger signifikant sætningskontekst og signifikante naboord til ord med hhv. *dän* og *tysk* som morfologiske komponenter og konkluderer i deres undersøgelse, at signifikante kookkurens partnere kan fungere som indikator for stereotyper (Quasthoff og Hallsteinsdóttir 2016:373).

2. Data og fremgangsmåde

I undersøgelsen anvendes de to Sketch Engine-korpora, deTenTen18 (5.346.041.196 løbende ord) og daTenTen20 (3.480.275.804 løbende ord). Det er korpora, som er automatisk genererede med materialer fra nettet, og deres ensartede opbygning gør, at de udgør et godt grundlag for sammenligning, samtidig med at de på grund af deres størrelse sikrer kvantitativt pålidelige resultater (Jakubíček et al. 2013).⁵

Der er knap 2 mia. ords forskel på de to korpora, hvilket afspejler sig i det samlede antal komposita med *dansk* ift. *deutsch*, ligesom det også afspejler sig i antallet af de adjektiv-substantiv-forbindelser, som kollokationerne ekstraheres fra.

Nationaladjektiverne *dansk* og *deutsch* undersøges sammen med de 10 nationaladjektiver for Danmarks og Tysklands nabolande, nemlig *norsk/norwegisch*, *svensk/schwedisch*, *nederlandsk* og *hollandsk/niederländisch*, *belgisk/belgisch*, *luxemb(o)urgsk/luxemburgisch*, *fransk/französisch*, *schweizisk/schweizerisch*, *østrigsk/österreichisch*, *tjekkisk/tschechisch* og *polsk/polnisch*.

⁵ Det nyere tyske korpus deTenTen20 (17.512.733.172 lb. ord), som ganske vist er genereret samtidig med det danske daTenTen20 i sommeren 2020, anvendes ikke, fordi det tyske deTenTen18 harmonerer bedre med det danske i størrelse.

I et første trin undersøges, hvordan det kvantitative forhold mellem nationaladjektivkomposita med hhv. *dansk* og *deutsch* som første- eller andetled er. Det undersøges med en simpel konkordanssøgning på kompositaene og en talmæssig sammenligning.

I et næste trin trækkes de basisord ud, som kollokerer med hhv. de 11 danske og de 11 tyske komposita. I denne proces anvendes værktøjet *Word Sketch Difference* i *Sketch Engine*. Værktøjet sorterer to søgeords kollokatorer efter, om kollokatorerne optræder sammen med begge eller kun et af søgeordene. De enkelte kollokationers typikalitet udtrykkes samtidig vha. en *logdice score*, som er et standardiseret mål på en skala med en fast maksimal værdi på 14. Det vil sige, jo højere score, jo højere sandsynlighed er der for, at en kollokator optræder sammen med søgeordet som kollokation, og jo lavere score, jo højere sandsynlighed er der for, at en kollokator optræder lige så hyppigt sammen med andre søgeord. Som standard viser *Word Sketch Difference* maksimalt 100 kollokatorer eller her basisord, og den grænse er for sammenlignelighedens skyld blevet bibeholdt i undersøgelsen.

I et tredje trin sorteres basisordene efter tematik og valorisering i en datadrevet proces, hvor Quasthoff og Hallsteinsdóttirs (2016) undersøgelse tjener som forbillede (jf. også Dalmas et al. 2015:21).

3. Resultater

1) Hvordan er den kvantitative fordeling mellem nationaladjektivkomposita med *dansk* eller *deutsch* som henholdsvis første- og andetled?

Det er et generelt mønster på begge sprog, at adjektivet for den nationalitet, der svarer til sproget, hyppigere står som førsteled end som andetled (jf. tabel 1 i appendiks). Den eneste undtagelse til denne regel er kompositummet *deutsch-schwedisch*, som forekommer sjældnere end *schwedisch-deutsch* (42 : 58). Fordelingen mellem komposita med *dansk* eller *deutsch* som førsteled ift. komposita med *dansk* eller *deutsch* som andetled er meget ens. Den gennemsnitlige danske fordeling er 85 : 15, og den tyske 84 : 16. Størst forskel er der, hvor der er hhv. flest og færrest hits, nemlig ved *dansk-luxemb(o)urgsk* vs. *luxemb(o)urgsk-dansk* (100 : 0) samt *deutsch-luxemburgisch* vs. *luxemburgisch-deutsch* og *deutsch-französisch* vs. *französisch-deutsch*, hvor forholdet i begge tilfælde er 95 : 5. Den mest ligelige fordeling ses ved *dansk-polsk* vs. *polsk-dansk*, hvor forholdet er 54 : 46.

Det optællingen også viser, er, at komposita med *-tysk* som andetled på dansk udgør den største gruppe med 9306 forekomster, mens sammensætninger med *-norsk* og *-svensk* som andetled kommer på anden- og tredjepladsen med hhv. 2831 og 2264 forekomster. På tysk er komposita med *-französisch* som andetled på samme måde entydigt mest frekvente (21.882

hits), idet disse forekommer ca. 2½ gange hyppigere end *deutsch-polnisch* på andenpladsen (8.836 hits).

2) Er der forskel på sprogene med hensyn til de substantiver, der kollokerer udskifteligt med hhv. NADJ₁-NADJ₂ og NADJ₂-NADJ₁?

Samlet ser talgrundlaget for ekstraktion af kollokationer i *Word Sketch Difference* sådan ud: *dansk*:- 12.637, *-dansk*: 1597, *deutsch*:- 28.800, *-deutsch*: 2.531. Der er sammentalt 159 forskellige basisord efter komposita med adjektiver i udskiftelig rækkefølge, 83 på dansk og 76 på tysk (jf. tabel 2 i appendiks). Af de 159 har sprogene 52 forskellige basisord til fælles, 51 basisord forekommer kun på tysk, 56 kun på dansk.

Det leksem, der optræder efter flest komposita med udskiftelige adjektiver, er *grænse* på begge sprog. Fælles er basisord som *grænseområde*, *grænseland* og *grænseregion* (jf. hertil Quasthoff & Hallsteinsdóttir 2016: 365).

Det temaområde, der er repræsenteret med flest fælles kollokationer, er 'Kultur, uddannelse og forskning'. Der kan især identificeres tre underområder her, nemlig film og filmproduktion, repræsenteret ved ord som *co-produktion* og *spillefilm*; forskning er repræsenteret ved sammensætninger med *-team*, *-projekt* og *-samarbejde*, og endelig er sprog også et i forhold til frekvens centralt underemne, repræsenteret fx ved *ordbog*.

Hvad angår basisordenes semantiske værdi, kan *krig* identificeres som det eneste fælles entydigt negative basisord, mens der er flere positivt værdiladede basisord, som *alliance*, *overenskomst*, *samarbejde* og *udveksling*.

I en sammenligning af de ikke-fælles kollokationer, dvs. kollokationer med basisord, der kun forekommer i det ene sprog efter komposita med adjektiver i udskiftelig rækkefølge, er der to resultater, som kan fremhæves: Dels er basisord, som de netop nævnte, altså positivt værdiladede betegnelser, som anvendes på tværs af temaområderne, hyppigere på tysk, herunder fx *Abkommen*, *Aussöhnung*, *Freundschaft*, *Initiative*, *Kooperation*, *Nachbarschaft*, *Partnerschaft*, *Versöhnung*, *Verständigung*, *Vorschlag* og *Zusammenleben*. Dels er 28 % af de danske basisord efter komposita med adjektiver i udskiftelig rækkefølge personbetegnelser, som fx *arkitekt*, *billedhugger*, *billedkunstner*, *ekspressionist*, *keramiker*, *kok*, *krigshelt*, *køkkenchef*, *kunstner*, *kobberstikker*, *missionær*, *præst*, *sanger* og *søhelt*. På tysk er der kun tre basisord i denne gruppe (6 %), nemlig *Doppelbürger*, *Komponist* og *Politiker*.

3) Er der forskel på sprogene med hensyn til substantiver, der kun kollokerer med enten NADJ₁-NADJ₂ eller NADJ₂-NADJ₁?

Samlet set er fordelingen af denne type kollokationer mere ulige på tysk end på dansk, idet adjektivkomposita med *-deutsch* som andetled kun optræder foran 55 forskellige basisord, mens

adjektivkomposita med *deutsch-* som førsteled optræder foran 242 forskellige basisord. På dansk er forholdet 233 (*-dansk*) : 333 (*dansk-*).

En intersproglig sammenligning viser klare ligheder i fordelingen på emner. 'Kultur, uddannelse og forskning' er det største emneområde på begge sprog uanset adjektivernes rækkefølge. Det fylder mere på tysk, idet mere end 40 hhv. 50 % af basisordene kan tilordnes dette emneområde, mens det er en tredjedel af de danske basisord.

Intrasprogligt er der på tysk 10 procentpoint flere substantiver i gruppen 'Kultur, uddannelse og forskning', som står efter kompositum med *-deutsch* som andetled end omvendt.

Trods disse talmæssige forskelle er der ikke grundlæggende semantiske forskelle at iagttage, hverken inter- eller intrasprogligt⁶. De leksemer, der dominerer på begge sprog inden for emnerne 'Kultur, uddannelse og forskning', 'Politik' og 'Økonomi', er disse:

- 'Kultur, uddannelse og forskning': **kultur**, **forschung*/forskning**, **austausch**, **zusammenarbeit*/samarbejde**, **projekt**
- 'Politik': **vertrag*/aftale**, **abkommen*/overenskomst**
- 'Økonomi': **handel**, **netzwerk*/netværk**

Det hyppigste leksem på tysk er også i denne gruppe **grenz**, som optræder i 29 forskellige ordformer.⁷ Også på dansk er **grænse** et af de hyppigste leksemer, i ikke mindre end 14 forskellige basisord, hvoraf flere dog ikke falder i temaområdet 'Geografi', men fx i temaområderne 'Politik' (*grænsedragning, grænsegængerregler grænsehandel, grænsehandelsarrangement*).

Der er størst intersproglig og samtidig også størst intrasproglig forskel på dansk på basisord, som betegner personer. 30 % af de danske basisord efter et kompositum med *-dansk* som andetled betegner personer. Det er også hyppigere med et kompositum med *-deutsch* som andetled på tysk ved personbetegnelser end den omvendte form, men ikke i samme omfang som på dansk (jf. tabel 3 i appendiks).

⁶ Der kan dog iagttages enkelte interessante detaljer, som fx at substantiverne *integration* og *migration* kun forekommer efter komposita med *-dansk* og *-deutsch* som andetled, fx i *svensk-danske integration eller polnisch-deutsche Migration*.

⁷ Fx *Grenzbahnhof, -bereich, -berg, -frage, -gänger, -gebirge, -gemeinde, -handel, -kampf, -landgeschichte, -linie, -örtchen, -ort, -stadt, -strecke, -verkehr, -vertrag, -ziehung, Landesgrenze, Sprachgrenze*.

4. Konklusion

Der kan ikke generelt konstateres semantiske forskelle mellem basisord efter komposita med adjektiver i udskiftelig rækkefølge og efter komposita med adjektiver i fast rækkefølge, dvs. rækkefølgen synes ud fra denne analyse ikke at spille nogen rolle.

Der er dog rent kvantitativt en klar præference for *deutsch*- eller *dansk*- som førsteled, hvilket tyder på, at en sprogbrug med omvendt rækkefølge kan være synonym med en pragmatisk eller stilistisk markerethed. Nogen forklaring på den klare præference kan ligge i vægtning. Fx er alle de adjektiver, *deutsch* kobles med, to- eller flerstavelsesadjektiver, hvilket ville kunne forklare rækkefølgen på tysk. Den ene undtagelse i forhold til rækkefølge på tysk, *schwedisch-deutsch*, som forekommer hyppigere end *deutsch-schwedisch*, kan samtidig begrundes i det lydlige sammenstød af *-sch-sch-* i sidstnævnte form. På dansk synes vægtning omvendt ikke at udgøre nogen fyldestgørende forklaring, idet fem af de 11 nationaladjektiver er enstavelsesord, nemlig *norsk, svensk, tysk, fransk og polsk*.

Tematisk kan der konstateres en meget ens fordeling af basisordene på de to sprog: Leksemet **grenz*/*grænse** er absolut hyppigste, og temaområdet 'Kultur, uddannelse og forskning' dominerer. Især begreber, som har med underemnerne film og filmproduktion at gøre, kollokerer med komposita med adjektiver i udskiftelig rækkefølge. Denne tematiske fordeling kan sikkert til dels forklares med korporaenes sammensætning⁸.

Hvad angår basisordenes konnotative værdi, har de nationale bindestregskomposita – uanset rækkefølgen af adjektiverne – præference for basisord, der betegner positiv udveksling i videste forstand, som fx *aftale, forbindelse, forsoning, forståelse, Kooperation, møde, nabo-skab, partnerskab, samarbejde, samliv, venskab*. Kun 17 ud af de 1022 basisord, som er registreret til sammen i de forskellige grupper, og det vil sige under 2 %, kan indordnes som negativt værdiladede. Hertil hører ord som *grænseproblem, modsætning, krig, krise og stereotyp*.

Vigtigste undtagelse til den generelle iagttagelse, at rækkefølgen tilsyneladende ikke er begrundet i semantiske forhold, udgør de danske personbetegnelser, der altså forekommer hyppigere efter komposita med *-dansk* som andetled. Her vil en detaljeret undersøgelse af både de referenter, som kompositaene attribuerer, og af kontekst, evt. kunne give et mere nuanceret svar på undersøgelsens spørgsmål.

⁸ En tematisk beskrivelse af de to anvendte korpora foreligger ikke i Sketch Engine, men for korpus deTenTen20 er der foretaget en tematisk opdeling af materialet på 11 områder, og her udgør "Arts" og "Recreation" knap en tredjedel af materialet (hhv. 17,7 og 11,9 %), jf. <https://www.sketchengine.eu/detenten-german-corpus>.

Litteratur

- Alexiadou, Artemis & Stavrou, Melita (2011) Ethnic Adjectives as Pseudo-Adjectives: A case study on syntax–morphology interaction and the structure of DP. *Studia Linguistica* 65 (2):117-146.
- [Amtliche Regelung] (2018): *Deutsche Rechtschreibung. Regeln und Wörterverzeichnis. Aktualisierte Fassung des amtlichen Regelwerks entsprechend den Empfehlungen des Rats für deutsche Rechtschreibung 2016*. Mannheim. (§ 44). [AR]
- Breindl, Eva & Thurmaier, Maria (1992) Der Fürstbischof im Hosenrock. Eine Studie zu den nominalen Kopulativkomposita des Deutschen. I: *Deutsche Sprache* 20:32-61.
- Dalmas, Martine., Dobrovol'skij, Dmitrij., Goldhahn, Dirk & Quasthoff, Uwe (2015) Bewertung durch Adjektive Ansätze einer korpusgestützten Untersuchung zur Synonymie. *Zeitschrift für Literaturwissenschaft und Linguistik* 45. Jg., 177:12-28.
- Dansk Sprognævn (2012a): *Retskrivningsordbogen*. København: Alinea. (§ 57(5)). [RO]
- Donalies, Elke (1996) Da keuchgrinste sie süßsauerlich. Über kopulative Verb- und Adjektivkomposita. *Zeitschrift für Germanistische Linguistik*, Berlin, vol. 24, iss. 3::73-286.
- Gallmann, Peter (1996) Interpunktion (Syngrapheme). I Hartmut, Günther & Ludwig, Otto (red.): *Schrift und Schriflichkeit. Writing and Its Use. Ein interdisziplinäres Handbuch internationaler Forschung*. Berlin: de Gruyter.
- Gudiksen, Asgerd (2019) Døvtum og sortstjernet – en diskussion af termen og begrebet dvandvasammensætning. *MUDS – Møderne om Udforskningen af Dansk Sprog*, Bind 17:93-305.
- Hallsteinsdóttir, Erla (2016) Deutsch-dänische Nationalstereotype in der Sprache. I Erla Hallsteinsdóttir/Klaus Geyer/Katja Gorbahn/Jörg Kilian (Hrsg.): *Perspektiven der Stereotypenforschung*. Frankfurt a.M.: Peter Lang:35-57.
- Hansen, Erik & Heltoft, Lars (2011) *Grammatik over det danske sprog bd. I-III*. København: Det Danske Sprog- og Litteraturselskab. [GDS]
- Hansen, Aage (1967) *Moderne dansk. Bd. I. Analyse*. København: Det Danske Sprog- og Litteraturselskab.
- Horn, Nynne T., McGhee, Heidi A.; Rijkhoff, Jan (2014) Relationelle adjektiver på dansk. *Occasional Papers in Linguistics from Aarhus* 1485-621, nr. 1. 1-28.
- Jakubiček, Miloš, Kilgarriff, Adam, Kovář, Vojtěch, Rychlý, Pavel & Suchomel, Vít (2013) The TenTen corpus family. *7th International Corpus Linguistics Conference CL*. 125-127.
- Pümpel-Mader, Maria; Gassner-Koch, Elsbeth; Wellmann, Hans; Ortner, Lorelies (1992) *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache*. Berlin: de Gruyter.
- Quasthoff, Uwe und Hallsteinsdóttir, Erla (2016) Stereotype in Webkorpora: Strategien zur Suche in sehr großen Datenmengen. *Linguistik Online* 79, 5. 347-379.
- Stein, Stephan, Stumpf, Sören (2019) *Muster in Sprache und Kommunikation*. Berlin: Erich Schmidt Verlag.
- Wöllstein, Angelika; Kunkel-Razum, Kathrin; Münzberg, Franziska; Neuhaus, Laura (red.) (2022) *Duden. Die Grammatik*. Berlin: Dudenverlag. [DuG]
- Ørsnes, Bjarne; Markantonatou, Stella (2002) Group Adjectives. *Journal of Greek Linguistics* 3. 139-178.

Appendiks

Tabel 1. Den kvantitative fordeling af komposita med hhv. <i>dansk</i> og <i>deutsch</i> som første- og andetled							
daTenTen20				deTenTen18			
	Hits	NADJ ₁ - NADJ ₂ :NADJ ₂ - NADJ ₁	/mio. lb. ord	/mio. lb. ord	NADJ ₁ - NADJ ₂ :NADJ ₂ - NADJ ₁	Hits	
<i>dansk-tysk</i>	9306	93:7	2,25	0,3	90:10	1902	<i>deutsch-dänisch</i>
<i>tysk-dansk</i>	681		0,16	0,03		209	<i>dänisch-deutsch</i>
<i>dansk-norsk</i>	2831	90:10	0,69	0,03	73:27	221	<i>deutsch-norwegisch</i>
<i>norsk-dansk</i>	307		0,07	0,01		81	<i>norwegisch-deutsch</i>
<i>dansk-svensk</i>	2264	78:22	0,55	0,18	42:58	1130	<i>deutsch-schwedisch</i>
<i>svensk-dansk</i>	623		0,15	0,24		1554	<i>schwedisch-deutsch</i>
<i>dansk-hollandsk</i> (/nederlandsk)	181 (26)	83:17	0,04; 0,01	0,33	88:12	2094	<i>deutsch-niederländisch</i>
<i>hollandsk</i> (/nederlandsk)-dansk	31 (11)		0,01; <0,01	0,05		299	<i>niederländisch-deutsch</i>
<i>dansk-belgisk</i>	36	92:8	0,01	0,09	87:13	563	<i>deutsch-belgisch</i>
<i>belgisk-dansk</i>	2		<0,01	0,01		81	<i>belgisch-deutsch</i>
<i>dansk-luxemb(o)urgsk</i>	28	100:0	0,01	0,04	95:5	277	<i>deutsch-luxemburgisch</i>
<i>luxemb(o)urgsk-dansk</i>	0		-	< 0,01		15	<i>luxemburgisch-deutsch</i>
<i>dansk-fransk</i>	837	76:24	0,2	3,43	95:5	21882	<i>deutsch-französisch</i>
<i>fransk-dansk</i>	271		0,07	0,19		1214	<i>französisch-deutsch</i>
<i>dansk-schweizisk</i>	96	92:8	0,02	0,16	81:19	1004	<i>deutsch-schweizerisch</i>
<i>schweizisk-dansk</i>	8		<0,01	0,04		234	<i>schweizerisch-deutsch</i>
<i>dansk-østrigsk</i>	53	93:7	0,01	0,43	85:15	2762	<i>deutsch-österreichisch</i>
<i>østrigsk-dansk</i>	4		<0,01	0,08		479	<i>österreichisch-deutsch</i>
<i>dansk-tjekkisk</i>	72	86:14	0,02	0,51	92:8	3282	<i>deutsch-tschechisch</i>
<i>tjekkisk-dansk</i>	12		<0,01	0,05		303	<i>tschechisch-deutsch</i>
<i>dansk-polsk</i>	249	54:46	0,06	1,38	92:8	8836	<i>deutsch-polnisch</i>
<i>polsek-dansk</i>	211		0,05	0,12		771	<i>polnisch-deutsch</i>
		85:15			84:16		

Tabel 2. Basisord efter komposita med udskiftelige nationaladjektiver (NADJ ₁ -NADJ ₂ + NADJ ₂ -NADJ ₁)			
	Kun i danske kollokationer (56)	Fælles (e.o. af tyske basisord) (52)	Kun i tyske kollokationer (e.o. af tyske basisord) (51)

Rækkefølgens betydning i bindestregs-nationalitetsadjektiver som tysk-dansk og dansk-tysk

Geografi	<i>bassin, landsdel, landegrænse</i>	<i>grænse (de8, da1), grænseområde (de8, da1), grænseland, grænseregion (de2, da2)</i>	<i>grænseflod (2), grænseovergang, grænserum, integrationsområde (Verflechtungsraum), statsgrænse (3), vadehav</i>
Kultur, uddannelse, forskning	<i>fælleslitteratur, gårdhund, krimiserie, kunstforbindelser, missionsorganisation, ordliste (3), [x]-projekt (2), skriftsprog, studentereksamen, thrillerserie, tv-serie, venskabsår, vidensbase</i>	<i>(kunstner)duo (de3, da2) forskerteam (de3, da1) forskningsprojekt, forskningsamarbejde, koproduktion (de12, da3), spillefilm (de2, da1), (online-)ordbog (de4, da8)</i>	<i>duel, ekspedition, eventyrfilm, fjernsynskanal, forfatterpar, fredsgran, jury, kulturfilm, kulturudveksling, nationalpark, naturpark, [x]-stiftelse, tandem, transmissionssystemoperatør trio, videostudie</i>
Politik	<i>barsebäckkomité, gasledning, informationstjeneste. regeringsaftale, regeringsudvalg, retsforhold, udenrigshandel</i>	<i>dobbeltbeskatningsaftale/dba (de1, da8)</i>	<i>akse, krise, modsætning, motor, parlamentariker-gruppe, specialforbund, statsaftale, topmøde</i>
Økonomi	<i>biotekselskab, fyrtårnsprojekt, handelskammer (4) onlinetøjbutik, postselskab, rederi</i>	<i>[x]-koncern (de1, da2)</i>	<i>clearing, turismevirksomhed, økonomisk relation (Wirtschaftsbeziehung) (3)</i>
Historie og militær	<i>flåde (2), flådesamarbejde, fælleskommando, grænseproblem, grænsespørgsmål, samarbejdspolitik, skæbneår</i>	<i>krig (de2, da1)</i>	-
Positiv relation (uspec., tematisk forbindelse til flere områder)	-	<i>alliance, (erfarings-)ud- veksling, konsortium (de1, da2), overenskomst, samarbejde (de 2, da1), forhold (de4, da1), relation</i>	<i>aftale, forslag, forsoning, forståelse, initiativ, Kooperation (2), naboskab, partnerskab (2), samliv, udsoning (2), venskab (4)</i>
Person- betegnelse	<i>arkitekt, billedhugger (6), billedkunstner, ekspressionist, keramiker, kobberstikker (2), kok, krigshelt, kunstner, køkkenchef, landskabsarkitekt, missionær, præst, sangerinde, søhelt</i>	<i>forfatter (de5, da1)</i>	<i>dobbelt statsborger (Doppelbürger), komponist, politiker</i>

Familiær relation	<i>familie</i>	<i>ægtepar (de2, da1)</i>	<i>forældre</i>
--------------------------	----------------	---------------------------	-----------------

Tabel 3. Person- og stillingsbetegnelser, som kun kollokerer med enten NADJ₁-NADJ₂ eller NADJ₂-NADJ₁

<i>dansk-NADJ</i>	<i>anatom, animator, baron, børneteatermand, chefkok, dagplejemor, dejlighed, drabsefterforsker, dukkemager, dirigent, eksdiplomat, erhvervsleder, filmlektor, filmskaber, finansmand, fodboldtræner, forhandler, forpagter, fredsforsker, freelanceoversætter, gallerist, gæstesolist, hitmager, hvalfanger, københavner-professor, litograf, lydkunstner, mademoiselle, malerinde, mimekunstner, musikstjerne, mæcen, nynazist, overretsråd, plantageejer, portrætmaler, rigmand, salgsdirektør, skipper, sopransolist, sømand, tennisspiller, verdenshypnotisør, vinbonde</i>	<i>Aktionskünstler, Aphoristiker, Bassist, Brillenmacher, Buchmacher, Chemiker, Designerin, Doppelbürgerin, Gesellschaftstheoretiker, Gitarristin, Grenzgänger, Kommissarin, Powermetall-Kollege, Psychiater, Publizist, Schuhmacher, Schriftstellerin, Sopranistin, Staatsmann</i>	<i>deutsch-NADJ</i>
<i>NADJ-dansk</i>	<i>adelskvinde, arbejderforfatter, astronom, auteur, auschwitz-overlever, avantgarde-guitarist, balletdanser, barokkomponist, billedhugger, billedskærer, bogtrykker, cellist, dessertkok, domkirkearkitekt, ekspressionist, falkoner, filmfotograf, finansmand, fløjtevirtuos, forfatterinde, galanterihandler, gehejmeråd, gravør, handelsdirektør, handelsmand, harmonikaspiller, havearkitekt, heraldiker, hofembedsmand, hofkunstner, hofmand, hofmesterinde, illustrator, internetjurist, kammerherre, kancelliråd, kapelmester, kemiker, kirkemand, klaverpædagog, klokkenist, komedieforfatter, konceptkunstner, landskabsgartner, lensgreve, maler, matematiker, musikforsker, musikkritiker, nationaløkonom, oboist, orgelbygger, ornitolog, performer, robotudvikler, sceneinstruktør, skønånd, socialpsykolog, sociolog, stenhugger, storkøbmand, søofficer, vingøglar, violinist, virksomhedsdirektør, virksomhedsleder, visesangerinde, xylograf</i>	<i>Choreografin, Großmeister, Organist, Pianistin, Publizistin, Schlagersängerin</i>	<i>NADJ-deutsch</i>

Resumé:

Nominalfrasens kompleksitet i dansk som L1 og L2

– en undersøgelse af polske lørneres og danske gymnasieelevers skriftlige produktion

Mikołaj Sobkowiak

Universitetet i Poznań

Selv om polakker fortsat er på førstepladsen blandt Danmarks største indvandrergupper (IID 2021:13), er der forholdsvis lidt forskning vedrørende polske lørneres tilegnelse af dansk. Det overordnede formål med min undersøgelse er at bidrage til forståelsen af, hvordan voksne polakkers lørnerdansk udvikler sig. I mit bidrag fokuserer jeg på syntaktisk kompleksitet og diversitet i nominalfraser. Syntaktisk kompleksitet har i de sidste årtier været anset for en god indikator af andetsproglørneres udvikling (Bulté/Housen 2014), og der er en lang række mere og mindre sofistikerede kompleksitetsindekser, man har anvendt som målestok for at beskrive lørnersyntaksen, også i forbindelse med dansk som fremmedsprog (Sobkowiak 2019). Nogle af de hyppigst anvendte kompleksitetsindekser inden for L2-skrivning er generelle længdemål på sætnings-, clause- og fraseniveau (Bulté/Housen 2018), men der er også forskere, som arbejder med meget detaljerede indekser inden for fraser (f.eks. Kyle 2016). I en del undersøgelser (f.eks. Biber et al. 2011, Parkinson & Musgrave 2014, Berggren/Sørland 2016, Nilsen 2022) fokuserer man til gengæld på nominalfrasens udvikling, der anses for at være en god indikator af lørnerens progression, bl.a. i forbindelse med akademisk skrivning.

I mit oplæg vil jeg gerne præsentere resultaterne af en længdeundersøgelse, hvor jeg analyserer eksamensopgaver skrevet af polske danskstuderende (n=25) efter bachelorstudiets første, andet og tredje år (n=3 opgaver per studerende). Jeg vil fokusere på nominalfrasens udvikling, og for at følge progressionen vil jeg anvende en række kompleksitetsmål, f.eks. nominalfrasens gennemsnitlige og maksimale længde, gennemsnitligt antal af adled per frase.

Derudover vil jeg undersøge nominalfrasernes diversitet, og det vil jeg gøre ved at se nærmere på bestemmelsernes typer og frekvens. Her vil jeg bruge Christensen og Christensens klassifikation af nominalfrasens bestemmelser som udgangspunkt (2019:182 ff.). De nævnte

lørnerdata vil jeg sammenligne med tilsvarende L1-tekster, skrevet af danske gymnasieelever ud fra de samme opgaveformuleringer som L2-teksterne.

Litteratur

- Berggreen, H. & Sørland, K. (2016). Syntaktisk kompleksitet i et skriftlig innlærerspråkmateriale. *NOA norsk som andrespråk*, Årgang 32, 1-2/2016, s. 31–75.
- Biber, D., Gray, B. & Poonpon, K. (2011). Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development? *TESOL Quarterly*, 45, s. 5–35.
- Bulté, B., & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics* 28:1, s. 147–164.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, s. 42–65.
- Christensen, R.Z. & Christensen, L (2019). *Dansk grammatik*. Odense: Syddansk Universitetsforlag.
- IID = Indvandrere i Danmark 2021. Udgivet af Danmarks Statistik. <https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=34714&sid=indv2021>.
- Kyle, K. (2016). *Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication*. Dissertation, Georgia State University. http://scholarworks.gsu.edu/alesl_diss/35.
- Nilsen, G. K. (2022). *Substantivfrasens kompleksitetsutvikling i innlærerspråk. En konstruksjonsgrammatisk analyse av skriftlig produksjon fra A1- til B2-nivå*. Ph.d.-afhandling, Universitetet I Stavanger.
- Parkinson, J. & Musgrave, J. (2014). Development of Noun Phrase Complexity in the Writing of English for Academic Purposes Students. *Journal of English for Academic Purposes*, 14, s. 48–59.
- Sobkowiak, M. (2019). Hvordan skriver polske førsteårsstuderende? En korpusundersøgelse af syntaktisk kompleksitet i dansk som fremmedsprog. *Skandinavskaja Filologija* 17(1), s. 36–5.

Konstruktioner af typen *en glædens dag* – en korpusanalyse

Andrzej Szubert

Adam Mickiewicz universitetet, Poznań

1. Indledning¹

Blandt mange konstruktioner i dansk sprog er der i hvert fald én, der ser ugrammatisk ud, nemlig den af typen *et åbenbaringens lys*. Selve strukturen i konstruktionen er iøjnefaldende ved, at bestemmeren (i dette tilfælde en ubestemt artikel) ikke kongruerer med det efterfølgende substantiv, dvs. *åbenbaringens* (der er i bestemt form genitiv), men med det sidste substantiv (*lys*). Dette ser i en udlændings øjne usædvanligt ud. Det 'normale' ville være den lineære kongruens, dvs. *en åbenbarings lys* eller *åbenbaringens lys*. Der findes arkaiske og leksikaliserede udtryk som *et lettelsens suk* eller *i dette herrens år*, men det viser sig, at konstruktionen er i live, dvs. den er produktiv. Til at begynde med troede jeg, at det var en trykfejl eller en redaktionel fejl, da jeg så de første eksempler på konstruktionen. Antallet af eksempler, der forekom ikke kun i historiske romaner, men også i helt moderne kriminalromaner, fik mig til at undersøge konstruktionens grammatiske og semantiske forhold. I de gængse grammatikker, dvs. i Mikkelsen (1911), Diderichsen (1947), Hansen (1967) og Hansen/Heltoft (2011), findes eksempler på konstruktionen, men det er hovedsageligt genitiven af substantivet, der beskrives. Ingen af dem nævner sådanne konstruktioner som en selvstændig grammatisk konstruktion med sine egne regler. Af pladshensyn bliver der i det følgende ikke henvist til de nævnte grammatikker (bort set fra Mikkelsen 1911, se nedenfor), men artiklen er et forsøg på at beskrive og fortolke konstruktionen; en henvisning til de nævnte beskrivelser findes i Szubert (2021).

¹ Jeg vil gerne takke Nina Thirup, der var oversætter af Per Olov Enquists roman *Livlægens besøg*, hvis bemærkninger angående konstruktionens semantiske og stilistiske lag var utroligt værdifulde for mig. En stor tak gælder også Ole Tøgeby for hans meget relevante bemærkning om, at mange af eksemplerne kunne fortolkes som leksikaliserede ordgrupper (set bort fra faste udtryk) bestående af et substantiv i bestemt form genitiv og et efterfølgende substantiv i nøgen form, hvilket gør, at et bestemt antal eksempler ville falde fra, da de kan fortolkes som helheder og ikke dele af konstruktionen. Dette var i denne fase af undersøgelsen ikke det vigtigste, for hensigten var at se, om konstruktionen er produktiv. Ikke desto mindre er det en vigtig faktor, som der under alle omstændigheder vil blive taget hensyn til i fremtiden. En ikke mindre tak skylder jeg Martin Palludan-Müller for uvurderlige diskussioner og rettelser.

2. Konstruktionens grammatiske og semantiske forhold

Selve konstruktionen er interessant af to grunde. Sædvanligvis stemmer de grammatiske og semantiske relationer i en konstruktion overens med hinanden. I dette tilfælde ser det imidlertid aldeles anderledes ud.

2.1. Grammatiske forhold

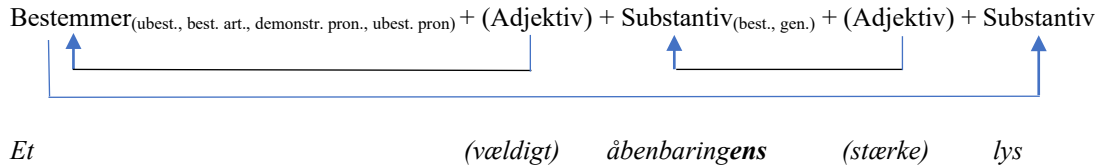
De grammatiske forhold udviser ret enkle relationer mellem de enkelte elementer i konstruktionen, der i sin enkleste udformning består af en bestemmer og to substantiver. Der kan derudover indsættes adjektiver mellem bestemmeren og det første substantiv og mellem substantiverne. Som helhed forekommer de enkelte elementer i følgende rækkefølge: Som første element i konstruktionen optræder bestemmeren (B), der fx kan være en ubestemt eller bestemt artikel, et demonstrativt pronomen m.fl. Som et næste fast element i konstruktionen er et substantiv (S1), der er i bestemt form genitiv singularis eller pluralis, og som bruges som attribut (attributsubstantivet) og til sidst et substantiv (S2) i nøgen form, der kan kaldes for kernesubstantivet. Efter bestemmeren kan der, men behøver ikke, komme et adjektiv (A1), og der kan også, men behøver ikke, komme endnu et adjektiv (A2) mellem begge substantiverne. Strukturen i konstruktionen ser således ud: B + (A1) + S1 + (A2) + S2.

Grammatisk kongruerer bestemmeren (B) med kernesubstantivet (S2) i genus og numerus, mens substantivet i bestemt form genitiv (S1) fungerer som attribut til kernesubstantivet (S2) og er grammatisk uafhængigt (dvs. altid i bestemt form genitiv), bortset fra den faste position i konstruktionen. Hvis der forekommer et adjektiv (A1), der placeres mellem bestemmeren (B) og attributsubstantivet (S1), så kongruerer adjektivet med bestemmeren og kernesubstantivet i genus og numerus (fx *et godt troens ord*). Et adjektiv (A2), der står mellem attributsubstantivet og kernesubstantivet, er grammatisk afhængigt af attributsubstantivet og udviser den svage bøjning (dvs. bestemt form med *e*-endelsen). Mikkelsen (1911:194) betragter den svage bøjning således:

[...] undertiden er dog ved lighedannelse med de tilfælde, hvor tillægsfald er bestemmesord, tillægsordenes bestemte form trængt ind, f. eks. *efter ti timers uafbrudte vandring* (ved siden af det regelmæssige *uafbrudt vandring*).

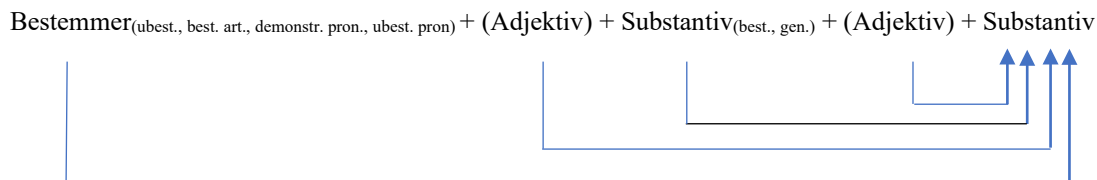
Dette betyder, at adjektivet i den dengang regelmæssige bøjning kongruerede med bestemmeren og kernesubstantivet. Med tiden er det, der dengang 'trængte ind', blevet til det

almindelige. De grammatiske forhold kan fremstilles grafisk i form af følgende skema (jf. Szubert 2021:135):



2.2. Semantiske forhold

Semantisk er forholdene enklere, men de afviger fra den almindelige brug af substantiver i genitiv og attributter. Den almindelige brug ses i sætningen *En gammel bils ejer*, hvor adjektivet henviser til *bil* og ikke til *ejer*. I dette tilfælde kongruerer adjektivet med artiklen og henviser til *bil* (jf. Szubert 2021:135). I sætningen *Den gamle bils nye ejer* henviser begge adjektiverne til det substantiv, der direkte følger efter. I den i nærværende artikel beskrevne konstruktion er situationen dog anderledes. De semantiske forhold stemmer ikke overens med de grammatiske, da både bestemmeren, begge adjektiverne og attributsubstantivet henviser semantisk til kerne-substantivet, som kan vises af følgende skema (jf. Szubert 2021:135):



2.3. Kombinatoriske muligheder

De kombinatoriske muligheder er begrænsede, fordi der potentielt er mulighed for at besætte fem positioner, der kan have et begrænset antal elementer i paradigmet. Jeg har beregnet i forhold til de grammatiske regler, hvad der kan stå som bestemmer foran et substantiv. Dette betyder, at der i position 1 kan placeres bl.a. to ubestemte artikler (*en* og *et*), ubestemte pronomener (*nogen*, *noget*, *nogle*), tre former for bestemt artikel (*den*, *det*, *de*), tre former for demonstrativt pronomener (*denne*, *dette*, *disse*), og de possessive pronomener (*min*, *mit* osv.). Dette gør, at der potentielt kan forekomme 22 forskellige ord i position 1. Om alle kombinationer opfylder konstruktionens krav, bør undersøges nærmere ved en senere lejlighed. I positionerne 2 og 4 er der kun tre muligheder at vælge imellem, da der kun findes tre adjektiviske bøjningsformer. I

position 3 er der tre muligheder (bestemt form singularis genus commune og neutrum, og én pluralisform), mens der i position 5 findes to muligheder (ubestemt form singularis og pluralis).

Rent kombinatorisk er der ret mange sammenstillingsmuligheder, og hvis alle positionerne er besat, inklusive de 'ugrammatiske' former, så er der potentielt 1188 forskellige kombinationer:

1 2 3 4 5
bestemmer + adjektiv + substantiv + adjektiv + substantiv
fx **et stort virksomhedens interne arbejdsredskab*.

Hvis der til gengæld kun er besat tre positioner (uden adjektiver), hvilket er det mest almindelige, så er der 132 kombinationer:

1 3 5
bestemmer + substantiv + substantiv
fx *en fredens mand*.

Den tredje kombination er med kun ét af adjektiverne i enten position 2 eller 4, hvilket giver 396 muligheder:

1 (2) 3 (4) 5
bestemmer + (adjektiv) + substantiv + (adjektiv) + substantiv
et (stort) virksomhedens (interne) arbejdsredskab

Der er alt i alt rigtig mange kombinationsmuligheder, men, som det fremgår af korpusanalysen, er det ikke alle kombinationerne, der finder anvendelse.

3. Korpusanalysen

De kombinatoriske muligheder giver et stort spillerum for konstruktionens potentielle brug, og den bruges hyppigere, end man skulle tro. For at undersøge, hvor hyppig den er, blev der foretaget en korpusanalyse i internetkorpusset daTenTen17², ved hjælp af søgemaskinen Sketch-Engine. Korpusset daTenTen2017 består af meget forskellige typer tekster, som fx blogs, ved

² Korpusset indeholder forskellige typer af tekster, der findes danske hjemmesider. Det er artikler, bøger, blogs m.m.

siden af avisartikler i bl.a. Information, så det er ikke kun de skønlitterære tekster, konstruktionen forekommer i; det omfatter 2.170.690.492 tokens³. Til søgningsformål har jeg formuleret søgningsalgoritmer iht. SketchEngines CQL (*Corpus Query Language*). Der blev også defineret 'forkerte', dvs. ugrammatiske algoritmer, for at se, om de forekommer; disse er også medtaget i undersøgelsen. I denne del af korpusundersøgelsen blev der fokuseret på formerne med ubestemt artikel. Nedenfor præsenteres nogle eksempler på fraser med forudbestemt mønster, fx som i a) ubestemt artikel + adjektiv i nøgen form + substantiv i fælleskøn genitiv + substantiv i fælleskøn singularis, og deres omskrivning til tagging i en code i *query language*, der bruges ved søgning i SketchEngine, hvilket giver følgende sekvenser, der svarer til de enkelte strukturer og deres tagging codes, fx

a) *en kirkens mand*

(Art_{ubst, com} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, com})⁴

[word="en"][tag="NC:sdgc:.*"][tag="NC:siuc:.*"]

b) *et kærlighedens smil*

(Art_{ubst, neut} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, neut})

[word="et"][tag="NC:sdgc:.*"][tag="NC:siun:.*"]

c) *en særlig juryens pris*

(Art_{ubst, com} + Adj_{sing, ubst, com} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, com})

[word="en"][tag="AC:siuc:--:p---.*"][tag="NC:sdgc:.*"][tag="NC:siuc:.*"]

e) *et kærlighedens taknemmelige smil*

(Art_{ubst, neut} + Sb_{sing, bst, gen, com} + Adj_{sing, best} + Sb_{sing, ubst, neut})

[word="et"][tag="NC:sdgc:.*"][tag="AC:sdu#:--:p---.*"][tag="NC:siun:.*"]

Korpusset blev indtil videre gennemført mht. mulige kombinationer, og der var et betydeligt antal mønstre, der ingen resultater gav, som fx

³ https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fdatenten17_tt2_2&corp_info=1.

⁴ Forkortelser: bst = bestemt, com = (genus) commune, gen = genitiv, neut = neutrum, pl = pluralis, sing = singularis, ubst = ubestemt.

**et kærlighedens taknemlig smil*

(Art_{ubst, neut} + Sb_{sing, bst, gen, com} + Adj_{sing, ubest, com} + Sb_{sing, ubst, neut})

[word="et"][tag="NC:sdgc:.*"][tag="AC:siu#:--:p---.*"][tag="NC:siun:.*"]

hvor adjektivet *taknemlig* er i ubestemt form, men burde være i bestemt. I dette tilfælde kongruerer det heller ikke i genus med kernesubstantivet. Ifølge konstruktionens grammatiske regler burde det hedde *et kærlighedens taknemlige smil*. I det følgende eksempel er der to problematiske steder, fordi den ubestemte artikel kongruerer med det efterfølgende substantiv, der er i bestemt form, hvilket grammatisk er umuligt, og adjektivet kongruerer med kernesubstantivet, hvilket også i moderne dansk er ualmindeligt (jf. Mikkelsen 1911:612):

**en fremtidens ungt potentiale*

(Art_{ubst, com} + Sb_{sing, bst, gen, com} + Adj_{sing, ubst, neut} + Sb_{sing, ubst, neut})

[word="en"][tag="NC:sdgc:.*"][tag="AC:siun:--:p---.*"][tag="NC:siun:.*"]

Det blev undersøgt, hvilke kombinationer, der er de hyppigste. Hvis man tager i betragtning, at ca. 75 % af alle substantiver er genus commune (Hansen / Lund 1994 :113), så vil det vel nok også påvirke resultaterne af de enkelte kombinationer.

I det følgende præsenteres kun få eksempler på strukturer indenfor konstruktionen samt udvalgte korpusbelæg med ubestemte artikler *en* eller *et*, der svarer til den pågældende struktur. Derudover angives antallet af samtlige og af de individuelle eksempler, dvs. hvis et eksempel forekommer flere gange, bliver det kun medregnet én gang. Rækkefølgen er ikke alfabetisk, men efter antallet af forekomster. Hvis ikke der er angivet noget tal, udgør de nævnte eksempler samtlige korpusbelæg.

3.1. Kongruerende strukturer

Blandt de kongruerende strukturer er der nogle, i hvilke kun tre af de fem positioner er besat, og nogle hvor enten den ene eller den anden adjektiviske position er besat.

1) *en formens mester*

(Art_{ubst, com} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, com})

Antal belæg i alt: 5775, individuelle eksempler: 3409.

Andre eksempler: *en kirkens mand, en ondskabens akse, en brudens bror, en djævelens opfindelse, en lykkens musik, en fredens mand, en skæbnens ironi, en fremtidens folkeskole.*

2) *et samtidens ord*

(Art_{ubst, neut} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, neut})

Antal belæg i alt: 2913, individuelle eksempler: 1489.

Andre eksempler: *et tavshedens råd, et oplysningens kulturmiljø, et smertens rum, et industriens selvstyre, et krigens lyn, et industrikulturens folkelivsmuseum.*

3) *en folkets krig*

(Art_{ubst, com} + Sb_{sing, bst, gen, neut} + Sb_{sing, ubst, com})

Antal belæg i alt: 1389, individuelle eksempler: 783.

Andre eksempler: *en fremskridtets mand, en mørkets tid, en opbruddets tid, en øjets fest*

4) *en damernes mand.*

(Art_{ubst, com} + Sb_{pl, bst, gen} + Sb_{sing, ubst, com})

Antal belæg i alt: 660, individuelle eksempler: 413.

Andre eksempler: *en kulturernes mangfoldighed, en modsætningernes dag, en avisernes dækning, en mulighedernes ramme, en detaljernes mester.*

5) *et mørkets spil*

(Art_{ubst, neut} + Sb_{sing, bst, gen, neut} + Sb_{sing, ubst, neut})

Antal belæg i alt: 374, individuelle eksempler: 239.

Andre eksempler: *et menneskets synspunkt, et menneskets blod, et håbets vindue, et lysets sted, et selskabets fradrag.*

6) *et mulighedernes spil*

(Art_{ubst, neut} + Sb_{pl, bst, gen} + Sb_{sing, ubst, neut})

Antal belæg i alt: 341, individuelle eksempler: 184.

Andre eksempler: *et malernes andelsselskab, et nationernes hus, et kontrasternes sted, et videnskabernes tempel, et soldaternes bordel, et borgernes udvalg, et følelsernes sprog.*

7) *en særlig juryens pris*

(Art_{ubst, com} + Adj_{sing, ubest, com} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, com})

Antal belæg i alt: 338, individuelle eksempler: 314.

Andre eksempler: *en kedelig klassens time, en ordentlig hudens fugtighed, en ny glædens dag, en retfærdig fredens mand.*

8) *et godt troens ord*

(Art_{ubst, neut} + Adj_{sing, ubest, neut} + Sb_{sing, bst, gen, com} + Sb_{sing, ubst, neut})

Antal belæg i alt: 151, individuelle eksempler: 129.

Andre eksempler: *et uformelt videnskabens imperium, et sådant frihedens opgør, et stabilt fredens land, et hemmeligt gadens sprog.*

9) *en fornuftens stenrige stemme*

(Art_{ubst, com} + Sb_{sing, bst, gen, com} + Adj_{sing, best} + Sb_{sing, ubst, com})

Antal belæg i alt: 84, individuelle eksempler: 81.

Andre eksempler: *en dekorationens visuelle effekt, en bestyrelsens konkrete vurdering, en globaliseringens lingvistiske pointe, en hævnens voldelige vej.*

10) *et klarhedens store øjeblik*

(Art_{ubst, neut} + Sb_{sing, bst, gen, com} + Adj_{sing, best} + Sb_{sing, ubst, neut})

Antal belæg i alt: 26, individuelle eksempler: 26.

Andre eksempler: *et overvejelsens intellektuelle blik, et identitetens uudtømmelige område, et frygtens mørke imperium, et enhedens barmhjertige lys, et virksomhedens interne arbejdsredskab.*

3.2. Opsummering af korpusundersøgelsen

Ifølge optællingen består de hyppigste konstruktioner af ubestemt artikel fælleskøn, attributsubstantivet i fælleskøn og kernesubstantivet i fælleskøn (5775 tokens med 3409 individuelle eksempler). Dette er måske ikke så overraskende, hvis man tager i betragtning, at ca. 80 % af alle substantiver i dansk er fælleskøn. På andenpladsen ligger den konstruktion, hvor strukturen har bestemmeren i intetkøn, attributsubstantivet i fælleskøn og kernesubstantivet i intetkøn (2913 tokens med 1489 individuelle eksempler), hvilket udgør lidt mere end halvdelen af den hyppigste struktur i konstruktionen. Tredjepladsen hører til en struktur, hvor bestemmeren er i fælleskøn, attributsubstantivet i intetkøn og kernesubstantivet i fælleskøn (1389 tokens og 783 individuelle eksempler). Derefter er der et betydeligt fald i antallet af forekomster. Resultaterne viser, at eksempler på konstruktionen, hvor strukturen har et adjektiv, er meget mindre hyppig, mens den hyppigste var en struktur med et adjektiv i nøgen form af typen *en særlig juryens pris*

med 338 tokens i alt og 314 individuelle eksempler. Eksempler med adjektivet i den anden position (typen *en fornøftens stenrige stemme*) er meget mindre hyppige (84 tokens og 81 individuelle eksempler). En tydelig tendens har også vist sig mht. andelen af individuelle eksempler. Deres antal rækker helt op til 100 % (jf. typen *et klarhedens store øjeblik* med 26 tokens i alt og 26 individuelle eksempler). Der blev kun fundet enkelte eksempler på strukturer, hvor adjektiver i begge positioner forekommer på én gang (jf. *en eventuel arbejdsgiverens lovpligtige arbejdsskadeforsikring*). De hyppigste strukturer i konstruktionen er uden adjektiver

4. Konklusioner

Man skulle tro, at konstruktionen er arkaisk og nærmest ikke bruges, men undersøgelsen har vist, at dette ikke er tilfældet, fordi tallene på eksemplerne er meget større jeg havde forestillet mig. Den bruges ikke kun i høj stil, og eksemplerne stammer ikke kun fra ældre men også nutidige kilder, som fx kriminalromaner eller økonomisprog (fx *et selskabets fradrag*). Rent kombinatorisk er der mange mulige kombinationer indenfor selve konstruktionen. Som det fremgår af undersøgelsen er der mange belæg på de mest typiske strukturer i konstruktionen, der er i overensstemmelse med 'reglerne', men der er også forholdsvis mange eksempler, der viser ikke-kongruerende strukturer. Det er ikke alle potentielle kombinationer, der er blevet undersøgt, men det var heller ikke hensigten med undersøgelsen.

Hensigten var at se, om konstruktionen er 'levende', og resultaterne har uden tvivl påvist, at den er produktiv og ikke en relik, der kun bruges i højere stil eller i tekster stiliserede til arkaisk sprog. Der kan i mange tilfælde også være tale om leksikaliserede helheder (fx *min dagens fidus*, *min trøstens kilde*), og de er ikke blevet trukket fra i optællingen, men det var heller ikke relevant at undersøge, hvor mange eksempler der findes, men om konstruktionen er produktiv.

Når man taler om konstruktionens tekstuelle funktion, kan man imidlertid sige, at genitiven i konstruktionen er fokusskabende. Udtryk som *et samtidens ord* er mere emphatisk og magtfuld i sammenligning med tilsvarende kompositum, jf. *et samtidsord*, hvis en sådan overhovedet kan dannes. I faste udtryk dominerer den emphatiske funktion, og nogle gange ligger der et vist element af højtidelighed, som fx i *nogle trøstens ord*, *denne nødens stund* og *et herrens vejr*. Opsummerende kan det siges, at der er tale om en regelmæssig konstruktion med interne grammatiske regler, og at den har sine semantiske funktioner, og at den er ikke arkaisk og stadig i brug.

Litteratur

Diderichsen, Paul (1957) *Elementær Dansk Grammatik*. 2. udg. København: Gyldendal.

Hansen, Erik & Heltoft, Lars (2011) *Grammatik over det Danske Sprog*. Bind II. Odense: Syddansk Universitetsforlag.

Hansen, Erik & Lund, Jørn (1994) *Kulturens Gesandter. Fremmedordene i dansk*. København: Munksgaards Sprogserie.

Hansen, Aage. (1967) *Moderne Dansk*. Bind II. København: Grafisk Forlag.

Mikkelsen Kristian (1911) *Dansk Ordføjningslære med sproghistoriske Tillæg*. København: Hans Reitzels Forlag.

Szubert, Andrzej (2021) *I dette åbenbarings lys – et eksempel på orddannelse, inkongruens eller måske noget helt tredje?* I *Skandinavskaja filologija* 19/1:126-138.

Resumé:

Klausale eksklamativer med *hvor*, *hvilken* og *sikke* – og deres ordstilling

Søren Sandager Sørensen

Aarhus Universitet

Der findes en række forskellige typer eksklamative sætninger i dansk, fx ”hvor det dog er en smuk julebuk” og ”sikke vi kan”. Både eksklamative sætninger med *at* (D’Hertefeldt & Verstraete 2014) og *sikke* (Heltoft 2007) er kendt for at være eksempler på insubordination (som kan oversættes til afunderordning) i den forstand at de har ledsætningstræk (fx konjunkional og neutral ordstilling), men kan bruges som selvstændige ytringer.

På baggrund af eksempler fra KorpusDK kan man se at eksklamativer med *hvor*, *hvilken* og *sikke* som subjektive partikler udviser variation i ordstilling. Det gælder både variation i om de kan have deklarativ eller neutral ordstilling, og om fx et adjektiv optræder ved siden af substantivet det modificerer, eller eksklamativytringens indledende *hvor*, *hvilken* eller *sikke*, og altså om frasen er diskontinuert (Delsing 2010).

Eksklamativerne fordeler sig efter typer som dem beskrevet i Zevakhina (2016). Det viser sig at *hvor* kan bruges i de alle typerne og med begge ordstillinger. *Hvilken* (og former deraf) findes også med begge ordstillinger, men er begrænset til nominale typer. *Sikke* kan derimod bruges i alle typer, men ser ud til at være begrænset til neutral ordstilling. Frasal diskontinuitet er valgfri ved *hvor*, men *hvilken* eller *sikke* står altid sammen med det led de modificerer.

På baggrund af dette kan man diskutere eksklamativernes træk hvad angår afunderordning og den diskontinuerte frases status, hvilket kan pege mod ideer til fremtidige studier og pointer om fx frasale eksklamativers rolle og evt. interaktionelle perspektiver.

Artiklen udgives i Skrifter i Samtalegrammatik som: Sørensen, Søren Sandager (2023). Exclamative clauses in Danish with *hvor*, *hvilken* and *sikke* – insubordination and discontinuity. Skrifter om Samtalegrammatik 9(1).

Litteratur

- D’Hertefelt, Sarah, & Verstraete, Jean-Christoph (2014). Independent complement constructions in Swedish and Danish: Insubordination or dependency shift? *Journal of Pragmatics*, 60:89–102.
<https://doi.org/10.1016/j.pragma.2013.11.002>
- Delsing, Lars-Olof (2010). Exclamatives in Scandinavian. *Studia Linguistica*, 64(1):16–36.
<https://doi.org/10.1111/j.1467-9582.2010.01168.x>
- Heltoft, Lars (2007). Subjektive partiklers grammatikalisering i nyere dansk – Eksemplet sikke. *Danske Studier*, 102:17–42.
- Zevakhina, Natalia (2016). The hypothesis of insubordination and three types of wh-exclamatives. *Studies in Language*, 40(4):765–814. <https://doi.org/10.1075/sl.40.4.02zev>

Forfeltsvariation i elevtekster fra 5. og 9. klassetrin

Morten Tannert, VIA University College

Ulf Dalvad Berthelsen, Aarhus Universitet

Indledning

Diderichsen (1966, s. 24) skriver om udfyldning af forfeltet: ”Behandlingen af dette Indlederled, der paa een Gang er bestemmende for hele Sætningens indre Struktur og det Element der fortrinsvis betegner Sætningens Forhold til det foregaaende, er al Prosas Nerve”. Tydeligere kan det vist ikke siges. Beherskelsen af forfeltets muligheder er et helt afgørende element i beherskelsen af skriftsproget og et vigtigt grundelement i dansk syntaks, idet det fungerer som et vigtigt tekstgrammatisk fundament for en teksts generelle sagsfremstilling, fx som grundlag for tema-remastrukturer (Jørgensen, 2000; Lindgren, 2021). Selvom der findes få begrænsninger for, hvad man kan placere i forfeltet, er det samtidig klart, at der er visse sætningsled, som forekommer i forfeltet hyppigere end andre, ligesom der er forskelle imellem tekster, hvis man ser på forfeltslængde og graden af variation. Der er derfor også god grund til at interessere sig for, hvad der kendetegner skoleelevers anvendelse af forfeltet på tværs af klassetrin. Forfeltet har af denne grund været et fokuspunkt i den hidtidige internationale skriveforskningslitteratur, som har undersøgt syntaktisk udvikling i tekster skrevet af skoleelever. Her har det bl.a. vist sig, at forfeltslængde er et af de sproglige træk, der bedst kan forudsige kvaliteten af elevernes tekster (typisk forstået som tekstens tildelte karakter) (Crossley, 2020), samt at anvendelsen af bl.a. adverbelle led eller ledsætninger i forfeltet er et tegn på, at eleverne har mere avancerede skrivefærdigheder (Myhill, 2008). Der er imidlertid ikke foretaget lignende systematiske undersøgelser af forfeltsvariation i elevtekster i en dansk kontekst.

I denne artikel præsenterer vi derfor en kvantitativ, eksplorativ analyse af forskelle i grundskoleelevers realisering af forfeltet i et tekstkorpus bestående af elevtekster fra 5. (n = 86) og 9. klassetrin (n = 88). Mere specifikt undersøger vi, om der er statistisk signifikante stigninger mellem de to klassetrin i det gennemsnitlige antal ord i forfeltet samt i antallet af lange forfelter relativt til det totale antal forfelter i teksten. På baggrund af disse kvantitative analyser undersøger vi dernæst kvalitativt seks elevtekster, som har fået forskellige værdier i de kvantitative analyser, med henblik på at beskrive, hvilke syntaktiske træk der kendetegner tekster med

forskellige numeriske værdier i de kvantitative analyser. Analyserne er således designet som mixed methods-analyser og giver både et billede af generelle tendenser i tekstkorpuset samt et kvalitativt og grammatisk indblik i de bagvedliggende syntaktiske strukturer.

Tidligere forskning

At undersøge sproglige strukturer i elevtekster fra et kvantitativt perspektiv er på én gang både en ny og gammel forskningstradition. Det er en gammel tradition i den forstand, at korpuslingvistikken har dybe rødder i 1940'ernes datalingvistik og 1960'ernes og 1970'ernes korpuslingvistiske landvindinger (Hirst 2013; McEnery & Hardie 2013). I en skandinavisk kontekst har dansk og norsk skriveforskning primært fokuseret på skrivedidaktik og læse- og skriveprocesser forstået som sociokulturelle praksisformer, mens man i Sverige har haft en større tradition for korpuslingvistisk og tekstorienteret skriveforskning (Bremholm et al. 2021; Tannert 2022). I et historisk perspektiv var det særligt Hultman & Westmans *Gymnasistsvenska* fra 1977, som markerede et afgørende nybrud, idet projektet med stor gennemslagskraft introducerede kvantitative operationaliseringer af syntaktisk kompleksitet og leksikal variation som nye metodologiske trædesten for den lingvistiske skriveforskning i Sverige. Dette blev op igennem 80'erne og 90'erne fulgt op af en lang række kvantitative studier af skriveudvikling i elevtekster fra grundskolen og forskellige ungdomsuddannelser (fx Chrystal & Ekvall 1999; Garne 1988; Josephson et al. 1998; Larsson 1984; Pettersson 1980, 1989). Kendetegnende for mange af disse studier er imidlertid, at de er baseret på manuelt indtastede og/eller annoterede tekstkorporer. Det vil sige, at både analysearbejdet og den første bearbejdning af teksterne er gjort i hånden.

Den kvantitative skriveforskning, som den ser ud i dag, er imidlertid også etableret i en nyere forskningstradition i den forstand, at nye teknologiske muligheder i form af maskinlæring har gjort værktøjer tilgængelige, der automatisk kan processere og analysere store mængder af tekst. Det er disse værktøjer, som kendes under samlebetegnelsen Natural Language Processing (NLP) (se fx Eisenstein 2019). Sammenholdt med det forhold, at elevtekster i dag som følge af digitaliseringen af uddannelsessystemet ofte eksisterer i digitalt format, er der således kommet gode muligheder for at undersøge skrivning og skriveudvikling i grundskolen fra et digitalt og korpuslingvistisk perspektiv. I den internationale forskningslitteratur inden for området er der to parallelle tendenser. I den ene undersøges det typisk, hvordan forskellige lingvistiske træk i teksterne forandrer sig over tid (et diakront perspektiv), og i den anden undersøges det typisk, hvordan de lingvistiske træk korrelerer med eller er i stand til at forudsige en række ekstralingvistiske forhold, fx en teksts tildelte karakter eller skribentens køn (et synkront perspektiv). Af diakrone studier kan bl.a. nævnes Beard & Burell (2010) og Beard et al. (2016), som undersøger

udviklingen af sætningstyper og -længde i argumenterende og berettende tekster skrevet af 9-11-årige elever, samt Troia et al. (2019), der undersøger klassetrinsrelaterede forskelle i op mod 100 forskellige lingvistiske variable i tekster fra 4.-6. klassetrin. Kendetegnen for den diakrone tilgang er således, at skriveudvikling primært anses som en udviklingsproces, der påvirkes over tid af elevernes alder/klassetrin. Den synkrone tilgang ses derimod hos bl.a. McNamara et al. (2010) og Crossley & Kim (2022), som undersøger, hvilke lingvistiske variable i elevernes tekster der bedst forudsiger teksternes tildelte karakter, samt hos Weber (2022) og Reilly et al. (2019), der undersøger sammenhængen mellem lingvistiske variable i elevernes tekster og henholdsvis elevernes sociale baggrund og køn. Den synkrone tilgang er således kendetegnet ved at interessere sig for årsagerne til den variation, der findes i tekster skrevet af elever med samme alder. I denne artikel er det imidlertid ikke denne synkrone variation, vi interesser os for, men derimod de diakrone forskelle i anvendelsen af forfелtet, der kan identificeres mellem tekster skrevet af elever med fem klassetrins forskel.

Der findes et forholdsvis begrænset antal studier med rødder i den lingvistiske og kvantitative skriveforskning, som specifikt har haft forskellige former for kompleksitet og variation i forfелtet som analysegenstand. Dette hænger sandsynligvis sammen med, at der dels er en række tekniske udfordringer forbundet med at isolere forfелtet computationelt, dels at forfелtet ikke er et lige så velfunderet teoretisk fænomen, som det i Danmark er blevet i kraft af Diderichsens (1946) indflydelsesrige introduktion af helsætningsskemaet i dansk strukturel lingvistik. Det amerikansk-udviklede Coh-Metrix, som er et af de største automatiserede, kvantitative værktøjer til tekstanalyse (McNamara et al. 2014), har forfелtet (defineret som 'number of words left of the main verb in the main clause') som et af over 100 lingvistiske variable, og i kraft af dette er analyser af forfелtet inkorporeret i en række amerikanske studier af skrivning og skriveudvikling (se fx Crossley & McNamara 2014; Graesser et al. 2011; Perin & Lauterbach 2018), dog uden at være genstand for særskilt fokus. I flere studier viser forfелtet sig imidlertid at være et af de syntaktiske træk, der har den stærkeste korrelation med tekstkvalitet (Crossley et al. 2011; McNamara et al. 2010, 2013), hvilket altså peger på, at forfелtet er et vigtigt fokuspunkt for at forstå børns skriveudvikling. I et større dansk skriveforskningsprojekt, ATEL, er forfелtet ligeledes inkorporeret som én blandt flere syntaktiske variable i en analyse af et tekstkorpus af elevtekster fra indskolingen (Bremholm et al. 2022; Kabel et al. 2020). Studiet peger bl.a. på, at mere avancerede elevskrivere i indskolingen kan variere forfелtet ved bl.a. at placere direkte objekt, interrogative pronominer eller tidsadverbialer i forfелtet, hvilket står i modsætning til de mest simple tekster, som udelukkende har korte subjektsudfyldninger i forfелtet. Projektet peger således på, at evnen til at variere forfелtet kan ses som et tegn på skriveudvikling på grund-

skolens begyndertrin. I forlængelse heraf undersøger vi i denne artikel, hvordan denne variation ser ud, når man sammenligner tekster skrevet på 5. klassetrin (grundskolens mellemtrin) og 9. klassetrin (grundskolens afsluttende trin) – både fra et kvantitativt, strukturelt perspektiv og fra et kvalitativt, grammatisk perspektiv.

Data og metode

Teksterne, der udgør datamaterialet for denne undersøgelse, stammer fra et ph.d.-projekt om skriveudvikling i grundskolen (Tannert 2022). Tekstkorpusset består af to teksttyper, henholdsvis berettende og beskrivende. Teksterne er skrevet af elever fra 5. og 9. klassetrin, og korpusset indeholder i alt 174 tekster, heraf 87 beskrivende tekster (5. kl. = 43, 9. kl. = 44) og 87 berettende tekster (5. kl. = 43, 9. kl. = 44). Vi vil i det følgende henvise til de fire grupper af tekster på følgende måde: 5D = 5. klassetrins deskriptive tekster, 5N = 5. klassetrins narrative tekster, 9D = 9. klassetrins deskriptive tekster, 9N = 9. klassetrins narrative tekster. Eleverne har skrevet teksterne på baggrund af en forskerinitieret skriveprompt, som består af to opgaver. Skrivepromptens overskrift er ”Sundhedskrise i verden”. I den første opgave skal eleverne vælge mellem to fotografier og lave en beskrivelse af situationen på fotoet. Begge fotos relaterer sig til coronapandemien, som var højaktuel på dataindsamlingsstidspunktet. I den anden opgave skal eleverne fortælle om en typisk dag for dem derhjemme under den første nedlukning af skolerne i foråret 2020.

Som forberedelse til de kvantitative analyser blev teksterne konverteret til .txt-format. De gennemgik desuden en ortografisk standardisering, og syntaktiske perioder (fra punktum til punktum) blev splittet i helsætninger. Dette kan umiddelbart virke som en drastisk transformation af det rå datamateriale, men da formålet med undersøgelsen ikke er at belyse elevernes retskrivningsfærdigheder, vurderede vi, at det var acceptabelt indgreb, fordi det til gengæld forbedrede NLP-modellens performance, der ved de indledende afprøvninger fungerede dårligt på elevtekster præget af ortografiske afvigelser og manglende tegnsætning. Herefter foregik processeringen af teksterne i to skridt. Først blev de rensede og standardiserede tekster annoteret vha. SpaCy v. 3.5.0 (<https://spacy.io>), et NLP-modul til Python, der blev anvendt til at parse teksterne samt tilføje part-of-speech-tags og dependency-tags. Herefter blev forfelterne inklusive POS-tags og DEP-tags udtrukket i listeform for hver tekst vha. af en regelbaseret algoritme baseret på den automatiske annotering af teksterne. At bruge automatiserede metoder til præprocessing, annotering og dataudtræk er forbundet med en vis usikkerhed, idet der er tale om statistisk baserede, probabilistiske sprogmodeller, der kort sagt genererer kvalificerede gæt.

SpaCy-modulet er imidlertid et relativt velfungerende modul med en nøjagtighed for dansk på mere end 95% (for en sammenligning med andre alternativer se Enevoldsen et al. 2021).

På baggrund af dataudtrækket gennemførte vi dels en deskriptiv statistisk analyse med henblik på at sammenligne den gennemsnitlige forfælslængde samt andelen af lange forfæls for de to klassetrin, dels en række permutationstest med henblik på at fastslå om de gennemsnitlige forskelle mellem grupperne var signifikante. En permutationstest indikerer, hvor sandsynligt det er, at et resultatet ikke er et udtryk for tilfældighed. Vi valgte permutationstesten fordi den i modsætning til andre typer af hypotesetest, fx t-test, ikke forudsætter en antagelse om, at datamaterialet er normalfordelt. Signifikansniveauet er fastsat til $p < 0.05$, hvilket vil sige, at der er mindre end 5% sandsynlighed for, at den observerede forskel skyldes tilfældig variation.

Resultater fra de kvantitative analyser

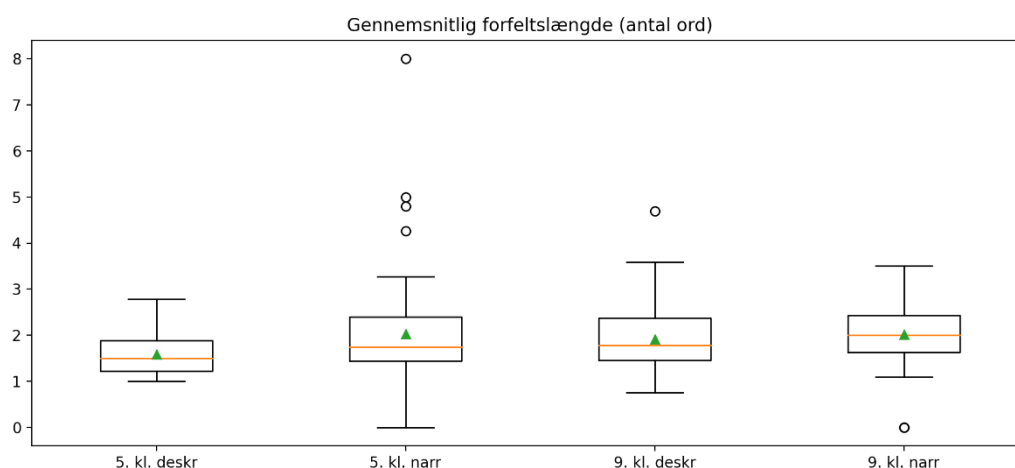
I den deskriptive statistiske analyse har vi fokuseret dels på teksternes gennemsnitlige forfælslængde målt på antal ord, dels på teksternes gennemsnitlige andel af lange forfæls målt som andelen af forfæls, der består af mere end to ord, relativt til det totale antal forfæls i teksten. Grænsen på to ord er valgt, fordi mere end to ord indikerer brugen af andet end pronomener og simple præpositionsforbindelser, der ellers præger mange af elevteksterne. Resultaterne af den deskriptive statistiske analyse er opsummeret i tabel 1.

Tabel 1

Oversigt over resultaterne af den deskriptive statistiske analyse

	Gns. 5. kl. deskriptiv		Gns. 5. kl. narrativ		Gns. 9. kl. deskriptiv		Gns. 9. kl. narrativ		Ratio 5. kl. deskriptiv		Ratio 5. kl. narrativ		Ratio 9. kl. deskriptiv		Ratio 9. kl. narrativ	
Antal	43	43	44	44	43	43	44	44	43	43	44	44	43	43	44	44
Gns.	1,58	2,02	1,91	2,02	0,16	0,22	0,21	0,21	0,16	0,22	0,21	0,21	0,16	0,22	0,21	0,21
SD	0,45	1,43	0,77	0,74	0,11	0,23	0,14	0,13	0,11	0,23	0,14	0,13	0,11	0,23	0,14	0,13
Min.	1,00	0,00	0,75	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
25%	1,23	1,44	1,46	1,62	0,09	0,11	0,11	0,12	0,09	0,11	0,11	0,12	0,09	0,11	0,11	0,12
50%	1,50	1,75	1,78	2,00	0,14	0,17	0,21	0,20	0,14	0,17	0,21	0,20	0,14	0,17	0,21	0,20
75%	1,88	2,40	2,37	2,44	0,25	0,27	0,33	0,29	0,25	0,27	0,33	0,29	0,25	0,27	0,33	0,29
Max.	2,79	8,00	4,69	3,50	0,38	1,00	0,50	0,50	0,38	1,00	0,50	0,50	0,38	1,00	0,50	0,50
IQR	0,65	0,96	0,91	0,81	0,16	0,16	0,23	0,16	0,16	0,16	0,23	0,16	0,16	0,23	0,16	0,16

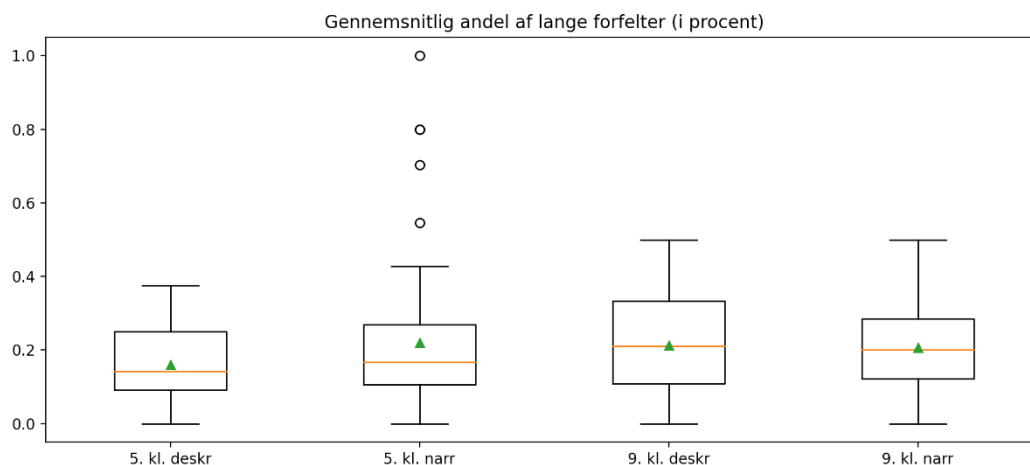
Den gennemsnitlige forfeltslængde (se figur 1) er 1,58 for 5D, 2,02 for 5N, 1,91 for 9D og 2,02 for 9D og svinger således mellem 1,58 som det laveste og 2,02 som det højeste gennemsnitlige antal ord per forfelt. Der er en relativt stor forskel mellem 5D og 5N på 0,44, mens der er en noget mindre forskel mellem 9D og 9N på 0,11 ord. Sammenlignes klassetrinnene, ser vi, at der er en forskel på 5D og 9D på 0,33, mens der ingen forskel er på 5N og 9N. I alle fire grupper er der stort set overensstemmelse mellem middelværdi og median. Dette indikerer sammen med gruppernes lave interkvartilbredde (se IQR i tabel 1), at middelværdierne i rimelig grad karakteriserer grupperne, fordi fordelingen ikke er skævvredet, og spredningen er relativt lille. Den eneste af grupperne, der afviger en lille smule fra dette mønster, er 5N, der, som det fremgår af figur 1, er præget af outlierere, hvilket resulterer i en forskel mellem middelværdi og median på 0,27.



Figur 1: Boxplot over fordelingen af gruppernes gennemsnitlige forfeltslængde.

Ser vi herefter på andelen af lange forfelter, tegner der sig et lignende billede. Den gennemsnitlige andel af lange forfelter (se tabel 1) er 16% for 5D, 22% for 5N, 21% for 9D og 21% for 9D og svinger således mellem 16% som den laveste og 22% som den højeste gennemsnitlige andel af lange forfelter. Som det var tilfældet med den gennemsnitlige forfeltslængde, er der en vis forskel på 5D og 5N og mellem 5D og 9D. Til gengæld er der stort set ingen forskel mellem de øvrige grupper. Som det også var tilfældet med den gennemsnitlige forfeltslængde, gælder det for gruppernes gennemsnitlige andel af lange forfelter, at der, bort set fra i 5N der er præget af outlierere, stort set er overensstemmelse mellem median og middelværdi samt en relativt lille

spredning (se figur 2).



Figur 2: Boxplot over fordelingen af gruppernes gennemsnitlige andel af lange forfелter.

På baggrund af den deskriptive statistiske analyse kan det være vanskeligt at vurdere, i hvilket omfang de observerede forskelle er udtryk for reelle forskelle eller blot for tilfældig variation. Vi gennemførte derfor en række permutationstest med henblik på at afgøre, om der er tale om signifikante forskelle. Som det fremgår af tabel 2 (signifikante resultater er markeret med *), er der signifikante forskelle mellem den gennemsnitlige forfелtslængde mellem 5D og 5N ($p=0.02$) og mellem 5D og 9D ($p=0.01$). Sammenligner vi gruppernes gennemsnitlige andel af lange forfелter, viser det sig, at der er en signifikant forskel mellem 5D og 9D ($p=0.03$) samt en forskel mellem 5D og 5N (markeret med +), der ikke er signifikant, men dog ligger meget tæt på det fastsatte signifikansniveau ($p=0.07$). De øvrige test viste ingen signifikante forskelle.

Tabel 2

Oversigt over resultaterne af de gennemførte permutationstest

Klasse/teksttype	Forskel i gennemsnitlig forfелtslængde (antal ord)				Forskel i gennemsnitlig andel af lange forfелter (procent)			
	5D/5N	9D/9N	5D/9D	5N/9N	5D/5N	9D/9N	5D/9D	5N/9N
Forskel	0,43	0,11	0,32	0,00	0,06	0,01	0,05	0,01
p-værdi	0,02*	0,24	0,01*	0,50	0,07+	0,58	0,03*	0,6

(* = $p < 0.05$ / + = $p < 0.1$)

De kvantitative analyser giver ikke i sig selv svar på, hvordan de observerede forskelle skal fortolkes. Selvom visse af resultaterne er signifikante, er der ikke tale om store forskelle i absolute tal. Det er heller ikke klart på baggrund af de statistiske beskrivelser af datamaterialet, om

det er klassetrin eller genre, der er årsag til de observerede forskelle. For 5D gælder, at den både er forskellig fra 5N, hvilket kunne indikere en genrebetinget forskel, og forskellig fra 9D, hvilket kunne indikere en klassetrinsrelateret forskel. På den anden side er der ingen forskel på hverken 9D og 9N eller på 5N og 9N, hvilket ikke undersøger disse fortolkninger.

For at komme nærmere på, hvad forskellene dækker over samt på at forstå den interne variation i grupperne, kaster vi i det følgende et kvalitativt blik på seks tekster, der er udvalgt med henblik på at belyse, hvilke tekstlige strukturer der ligger bag resultaterne af den kvantitative beskrivelse af korpusset.

Et kvalitativt blik på seks elevtekster

Tekstudvalget fordeler sig sådan, at der er tre tekster fra hver af de to teksttyper, hvoraf én tekst har korte forfelter, én har en gennemsnitlig forfeltslængde omkring medianværdien for årgangen, og én tekst har en forfeltslængde i tredje/fjerde kvartil for årgangen. Af pladsmæssige hensyn er ikke alle elevtekster gengivet i deres fulde længde. Forfelterne i hver tekst er markeret ved, at der efter hvert forfelt er en kantet parentes, hvori det er angivet, hvilket sætningsled der er i forfeltet. Der er kodet for subjekt [S], direkte objekt [DO], indirekte objekt [IO], adverbial [A], formelt/foreløbigt subjekt [FS], interrogativt pronomen [IP] og for tomme forfelter [T].

Tre berettende elevtekster

Vi starter med at se nærmere på tre berettende tekster. Den første tekst er skrevet af en elev fra 5. årgang. Teksten har en gennemsnitlig forfeltslængde på 0,9. Det vil sige, at teksten er placeret i første kvartil for årgangen. Teksten er gengivet i sin helhed i tabel 3.

Tabel 3		
<i>Skematisk oversigt over berettende elevtekst (5. årgang)</i>		
Interjektion/ konjunktion	Forfelt	Centralfelt/slutfelt
	Nu [A]	vil jeg fortælle om en dag ligesom alle dage
men	[T]	ikke helt fordi der var nedlukning i Danmark
og	det [S]	var kedeligt fordi man kunne ikke lege sammen med nogle af sine venner.
Nå men	[T]	lad os starte
	jeg [S]	stod op klokken 7.

	Min mor [S]	sagde der var morgenmad.
	Jeg [S]	spiste cornflæks ligsom alle dage.
	Efter [A]	havde vi "Skole" hjemme.
	Vi [S]	skulle starte med at lave matematik.
	Efter der [A]	lavet vi dansk.
	Til Sidst [A]	var det madkunstskab.
	Der [A]	skulle vi lave banan kage.
	Efter [A]	skulle vi hag aftensmad.
	Derefter [A]	skulle i seng.
Godnat	[T]	
og	[T]	sov godt.

Teksten består af i alt seksten forfælder. Syv af disse består af forskellige typer af adverbialer, hvoraf de fleste er enkeltstående tidsadverbialer som 'nu', 'efter' og 'derefter'. Eleven anvender således disse adverbialer til at sekvensere tekstens handlingssekvenser kronologisk. Der er fem forfælder med subjekt, hvoraf fire af dem er enkeltstående pronominer, og blot et enkelt forfælt har et nominalsyntaxme med en foranstillet bestemmer ('min mor'). Der er desuden fire tomme forfælder ved imperativformer. At teksten er endt med en gennemsnitlig forfæltslængde på blot 0,9, skyldes således, at der er ni forfælder med ét ord, tre med to ord og fire tomme forfælder. Teksten indeholder således en smule variation i sætningsled i forfæltet, idet der veksles mellem subjekt og adverbialer realiseringer, mens der til gengæld ikke er megen variation i længden og den grammatiske kompleksitet i forfælterne.

Den næste elevtekst, vi skal se på, er skrevet af en elev på 9. årgang. Teksten har en gennemsnitlig forfæltslængde på 2,1, hvilket placerer teksten lige lidt over medianen for årgangen. Teksten er forholdsvis lang (391 ord), så i tabel 4 nedenfor er der blot gengivet et uddrag af den første halvdel af teksten.

Tabel 4		
<i>Skematisk oversigt over berettende elevtekst 2 (9. årgang)</i>		
Interjektion/ konjunktion	Forfælt	Centralfælt/slutfælt
	Da vi skulle starte med at have skole hjemmefra [A],	stod jeg op meget sent. Omkring kl 10 eller 11.

Men	jeg [S]	fandt hurtigt ud af, at jeg skulle stå tidligere op.
	Ellers [A]	ville jeg ikke have en chance for at nå at lave mine lektier, før det blev sent.
	Derfor [A]	begyndte jeg at stå op omkring 8-10 tiden.
	Jeg [S]	ville spise morgenmad, glemme at børste tænder og tjekke hvad dagens lektier byder på.
	I nogle tilfælde [A]	tog jeg ud og løb som det første nogle morgener før morgenmad,
og	[T]	gik derefter i bad efterfølgende.
	Lektierne [S]	tog det meste af dagen,
men	jeg [S]	snakkede nogle gange med mine venner, som hurtigt kunne hjælpe mig.
	Vi [S]	havde også møder nogle gange, som gjorde det lidt mindre kedeligt, og sjovere fordi jeg fik lov at se mine klassekammerater igen.
	Som sagt [A]	tog jeg ud at løbe nogle gange,
men	jeg [S]	dansede også Just Dance
eller	[T]	gik ture med min mor [...]

Den samlede tekst består af niogtyve forfelter. Ni af disse er adverbialer enten i form af enkeltstående adverbialer eller adverbielle ledsætninger. Der er tre tomme forfelter ved ellipsekonstruktioner, mens de resterende sytten forfelter består af subjekt. Den væsentligste årsag til, at vi i denne tekst ser mere end en fordobling af den gennemsnitlige forfeltslængde sammenlignet med den forrige tekst, er, at der er flere eksempler på lange adverbialer i forfeltet, fx ledsætninger som "Da vi skulle starte med at have skole hjemmefra" eller tidsadverbielle udtryk som "i nogle tilfælde". Det er altså i høj grad i længden og variationen af de adverbielle udtryk, at man ser de største forskelle på de to berettende elevtekster. Vi ser ligeledes enkelte eksempler på nogle andre typer af adverbialer i forfeltet i denne tekst, som i højere grad er knyttet til betingelser og årsagsforhold i teksten (fx 'ellers' og 'derfor'), hvorimod de i forrige tekst forholdsvis entydigt var knyttet til kohæsion gennem temporale sætningskoblinger. Der er til gengæld ikke forskel på de sætningsled, som fremgår af forfeltet i de to tekster. Begge steder varieres der mellem forskellige adverbielle udtryk og subjektsudfyldninger med enkeltstående personlige pronominer.

Den sidste berettende elevtekst, vi skal se på, er ligeledes skrevet af en elev på 9. årgang. Teksten har en gennemsnitlig forfelseslængde på ca. 3,1, hvilket placerer teksten i fjerde og øverste kvartil på årgangen. Som ved den forrige tekst er der af pladsmæssige hensyn kun gengivet et uddrag af teksten (se tabel 5).

Tabel 5		
<i>Skematisk oversigt over berettende elevtekst 3 (9. årgang)</i>		
Interjektion/ konjunktion	Forfelt	Centralfelt/slutfelt
	En typisk hverdag hjemme hos mig under coronanedlukningen [S]	ville cirka foregå følgende.
	Jeg [S]	står op klokken 6:30 (nogle gange også klokken 7:00), som jeg normalt også ville have gjort, hvis vi havde været i skole som normalt.
	Så [A]	går jeg i bad,
	[T]	tager tøj på
	[T]	spiser morgenmad.
	Så [A]	er den normale morgenrutine egentligt overstået.
	Nu [A]	kommer vi til alt det, som har ændret sig betydeligt meget ved min hverdag i forhold til før.
	I stedet for at tage i skole på det her tidspunkt [A],	går jeg ovenpå, op på mit værelse med en kop te og
	[T]	starter computeren.
	Først [A]	kigger jeg lektierne igennem
og	[T]	begynder måske på en enkelt individuel opgave vi har fået for.
	Når klokken bliver cirka 8:30 [A]	kontakter jeg min studiegruppe [...]

Teksten består af i alt sytten forfelter, hvoraf hele ti af dem består af forskellige adverbielle led. Tre forfelter består af subjekt og fire forfelter er tomme, idet de forekommer ved ellipsekonstruktioner. Der er altså en større andel af adverbielle led i forfeltet i denne elevtekst end i den forrige, og ser man nærmere på disse adverbielle led, er det tydeligt, at det især er disse led, der er med til at øge forfeltslængden. Det kommer bl.a. til udtryk ved tidsadverbielle ledsætninger som ”Når klokken bliver cirka 8:30” og længere fraser som ”I stedet for at tage i skole på det her tidspunkt” og ”Så hver morgen”. I tekstens første sætning er der ligeledes et meget langt nominalsyntaxme på otte ord (”En typisk dag hjemme...”) i forfeltet, som er med til at øge den gennemsnitlige forfeltslængde i teksten. Det, der adskiller denne elevtekst, fra de to forrige er således, at der er flere eksempler på meget lange forfelter i teksten blandet med korte forfelter, som typisk består af enkeltstående pronominer eller tidsadverbialer.

Tre beskrivende elevtekster

Vi vender nu blikket mod tre beskrivende elevtekster. Den første elevtekst har en gennemsnitlig forfeltslængde på 1,0 og er skrevet af en elev på 5. klassetrin. Teksten er gengivet i sin helhed i tabel 6.

Tabel 6		
<i>Skematisk oversigt over beskrivende elevtekst 1 (5. årgang)</i>		
Interjektion/ konjunktion	Forfelt	Centralfelt/slutfelt
	Jeg [S]	ser nogle piger der godt ville lege også selvom der er corona.
	De [S]	er i byen næsten helt alene.
	Der [FS]	er ingen mennesker på restauranterne i baggrunden.
	De [S]	er på vej ud at lege i springvandet.
	Jeg [S]	tror det er ved at være forår fordi pigerne ikke har jakker på.
	Jeg [S]	tror ikke det er så varmt
	det [FS]	er heller ikke isnende koldt.
	Det [FS]	er nok sådan lidt lundt.
	De [S]	er nok glade men kede af at de skal holde afstand og bære maske.

Af de tekster, vi har set på indtil videre, er denne tekst den med mindst variation både i forhold til længde og sætningsled. Alle ni forfelter består af præcis ét ord, og alle forfelter har enten subjekt eller formelt/foreløbigt subjekt i forfeltet. Flere af sætningerne i teksten indeholder imidlertid andre former for syntaktisk kompleksitet såsom efterstillede ledsætninger eller indskudte sætninger, så et kort forfelt er i denne tekst ikke det samme som, at teksten har en simpel syntaks eller korte syntaktiske perioder. Eleven anvender forfelterne til dels at positionere sig selv som tekstens fortæller, der kommenterer og vurderer forhold på billedet, dels til deiktisk at referere til pigerne på billedet. Den tilsvarende berettende elevtekst, som havde en gennemsnitlig forfeltslængde på 1,0, havde en større variation i sætningsled, idet den narrative form i den berettende teksttype både kræver en sammenhænge temporal struktur og en aktivt agerende og vurderende fortæller. I denne beskrivende tekst kan eleven til gengæld uproblematisk udføre skriveordren uden variation i, hvilket sætningsled vedkommende placerer i forfeltet.

Den næste beskrivende elevtekst, vi skal se på, har en gennemsnitlig forfeltslængde på 1,3, som altså er en anelse højere end den forrige tekst og samtidig er placeret i andet kvartil

lige under medianværdien. Teksten er ligeledes skrevet af en elev på 5. klassetrin. Teksten er gengivet i sin helhed i tabel 7.

Tabel 7		
<i>Skematisk oversigt over beskrivende elevtekst 2 (5. årgang)</i>		
Interjektion/ konjunktion	Forfelt	Centralfelt/slutfelt
	Jeg [S]	har valgt billede nummer 1.
	Her [A]	ser man to piger på cirka 5 og 6 år stå midt ude på en hel tom gågade.
	Der [FS]	er ikke nogle mennesker fordi at corona eller covid 19 hersker verden over.
	Det [FS]	er en ny helt ukendt virus
og	forsker hele verden over [S]	kæmper på livet løs for at få en vaccine.
	Der [FS]	er normalt liv i gaderne
men	nu [A]	er virkeligheden sat på pause.
	De [S]	har visir på for at beskytte sig selv og andre.
	De [S]	står inde på torvet uden for en restaurant.
	Det [FS]	ligner at der står at restauranten hedder Janne.
	Den ene af pigerne [S]	sidder på en sten.
	De [S]	har nok vejret hjemsendt i et lille stykke tid.
	Nu [A]	er de blevet træt af at være indenfor
og	[T]	våger sig ud på gaden.
	Jeg [S]	håber at det snart er over.
	Det [S]	bliver nok først når de har fundet en vaccine.

Teksten indeholder i alt seksten forfelter, hvoraf tolv består af subjekt eller foreløbigt/formelt subjekt, tre består af et adverbialt led og et enkelt forfelt er tomt ved en ellipse. Anvendelsen af adverbialle led giver således en anelse større variation i udfyldningen af forfeltet i denne tekst sammenlignet med den forrige. Det er dog stadig subjekt og formelt/foreløbigt subjekt, der i overvejende grad er de dominerende sætningsled i forfelterne. Ser man på længden af forfelterne, består tretten af dem kun af ét ord. Det, der øger tekstens gennemsnitlige forfeltslængde

sammenlignet med forrige tekst, er således to nominalsyntaxmer på fire ord ('forsker hele verden over' og 'Den ene af pigerne'). Teksten har desuden generelt en større syntaktisk kompleksitet forstået på den måde, at vi ser større grad af variation mellem parataktiske og hypotaktiske sætningskonstruktioner, hvilket altså også viser sig ved, at eleven formår at anvende mere komplekse nominalsyntaxmer i forfeltet.

Den tredje og sidste beskrivende elevtekst har en gennemsnitlig forfelseslængde på hele 3,0 og er skrevet af en elev på 9. klassetrin. Teksten er forholdsvis lang, så i tabel 8 er blot gengivet et uddrag af første halvdel af teksten.

Tabel 8		
<i>Skematisk oversigt over beskrivende elevtekst 3 (9. årgang)</i>		
Interjektion/ konjunktion	Forfelt	Centralfelt/slutfelt
	Jeg [S]	har valgt billede nr. 2.
	Grunden til at jeg har valgt billede nr. 2 [S]	er fordi, at jeg føler at der er mere betydning i det billede end det billede nr. 1.
	Selvfølgelig [A]	kan jeg ikke lade vær med at referere til corona krisen, som er det der hovedsageligt er fokus på i medierne for tiden.
	Først og fremmest [A]	vil jeg beskrive miljøet og hvordan stedet ser ud.
	Der hvor billedet er taget, [S]	kunne godt ligne en opgang fra en togstation eller et sted inde i en by, hvor der er trafikeret.
	Der [FS]	er mange mennesker - især mange UDEN mundbind i det her offentlige område.
	Helt bagerst i baggrunden [A]	er der nogle træer samt en flok af mennesker.
	På bebyggelserne [A]	er der skilte, som kan tyde på at der er et offentligt sted inde i byen.
	På billedet [A]	ser man en mand i højre side, der ryger en cigaret.
	Det [FS]	ligner at manden er fra en middelklasse, da hans tøj ligner noget, man kunne se fra en normal hverdag på et kontorarbejde.
	Men selvom at det tyder på at hans fornuft er i top og at	kan man stadig ikke lade vær med at se cigaretten han ryger[...]

	han har et ordentligt arbejde [A],	
--	------------------------------------	--

Teksten består af i alt enogtyve forfelter, hvoraf tretten indeholder subjekt eller formelt/foreløbigt subjekt, syv indeholder et adverbialt led og kun et enkelt forfelt er tomt ved en ellipse. Elleve af forfelterne består blot af et enkelt ord (typisk et pronomen), så den øgede gennemsnitlige forfeltslængde i denne tekst sammenlignet med de forrige skyldes især, at der er enkeltteksempler på meget lange forfelter. Disse forfelter indeholder typisk enten nominalsyntagmer med indbyggede relativsætninger (fx ”Der hvor billedet er taget”) eller længere adverbialle fraser med deiktisk funktion (”Helt bagerst i baggrunden” eller ”Ud af alle dem omkring hende”), der peger ud på selve billedet på skriveordren. Tekstens længste forfelt består af atten ord og består af hele tre ledsætninger (”... selvom at det tyder på at hans fornuft...”), hvilket viser, at teksten således har en meget stor spredning og variation i længden af forfelterne. Når den gennemsnitlige forfeltslængde øges, tyder det altså på, at det skyldes enkelte meget lange og komplekse forfelter snarere end en generel øgning i længden af forfelterne generelt i teksten. De korte forfelter med enkeltstående pronomener eller adverbier er således dominerende på tværs af alle de elevtekster, vi har set på.

Forfeltet og elevers skriveudvikling

At det netop er forfeltet, som har været genstand for denne undersøgelse, er som anført i indledningen ikke et tilfælde. Forfeltet kan indeholde stort set alle sætningsled og har som følge heraf en vigtig funktionel rolle i de syntaktiske valg, som eleverne foretager i skriveprocessen. De kvantitative og kvalitative analyser peger samlet set på, at der findes to typer af forfeltersvariation, og at disse to typer af variation kan fungere uafhængigt af hinanden og på forskellig vis gøre os klogere på elevers skriveudvikling. Den første type af variation handler om variation i det sætningsled, som placeres i forfeltet. Her indikerede de kvalitative analyser, at det i høj grad er teksttypen og skriveordrens udformning, der afgør denne type af variation. I de berettede tekster så vi fx, at den narrative fortællestruktur i teksterne gjorde, at der ofte var tidsadverbialer i forfeltet, mens der omvendt ofte var subjektudfyldninger i de beskrivende tekster, idet eleverne enten anvendte forfeltet til at positionere sig selv som fortællere (fx ”Jeg ser nogle piger...”) eller til deiktisk at pege ud på forskellige elementer af det billede, de skulle beskrive (”De står inde på torvet...”). Der var dog også eksempler på adverbialle led i forfeltet i de beskrivende tekster, men også her havde de i høj grad en deiktisk funktion, idet de pegede ud og identificerede de områder af billedet på skriveordren, som den pågældende sætning beskrev

(fx ”Helt bagerst i baggrunden er der nogle træer...”). Det er altså tydeligt ved de beskrivende tekster, at skriveordren sætter eleverne i en bestemt kommunikationssituation, hvor de eksplicit skal forholde sig til et billede, og her er det i høj grad forfелtet, som eleverne anvender til at løse den specifikke kommunikative udfordring, det er at henviser til et billede på en eksternt givet skriveordre.

Den anden type af forfелtsvariation er knyttet til kompleksiteten og længden af forfелterne. Her viste de kvantitative analyser, at der var statistisk signifikante forskelle mellem 5. og 9. klassetrin på både forfелtslængde og ratioen af lange forfелter i de beskrivende tekster, mens der omvendt ikke var signifikant forskel mellem de to klassetrin, når man sammenlignede de berettende tekster. Fra et rent kvantitativt og strukturelt perspektiv ligner de to grupper af berettende tekster altså hinanden. At der til gengæld er signifikante forskelle ved de beskrivende tekster, kan skyldes flere ting. Der er eksempelvis tegn i de kvalitative analyser, der indikerer, at nogle elever omfortolker den beskrivende teksttype og gør den til en mere analyserende og vurderende tekst, hvor eleven fx kommer med begrundelser for valg af billede (fx ”Grunden til at jeg har valgt billede nr. 2 er fordi...”) eller kommer med mere vurderende udsagn om forhold på billedet (fx ”Men selvom at det tyder på at hans fornuft er i top og at han har et ordentligt arbejde, kan man...”). Denne fortolkning af skriveordren gør altså, at nogle elever kan bruge forfелtet til mere komplekse syntaktiske forhold som fx at opstille betingelsesledsætninger eller anvende mere vurderende adverbialer. Hvor variation i sætningsleddet i høj grad er noget, der afgøres af teksttypen og skriveordren, er der i de kvalitative analyser noget, der tyder på, at variation i kompleksitet og længde af forfелtet er noget, der kendetegner mere erfarne skrivere. I hvert fald er det primært i tekster fra 9. årgang, at vi ser eksempler på meget lange forfелter. De kvalitative analyser viser således, at det i høj grad er enkeltteksempler på meget lange forfелter, der får den gennemsnitlige forfелtslængde til at stige, snarere end det er en generel kvantitetsforøgelse af alle forfелterne i teksten. Det prototypiske forfелt på tværs af alle de kvalitativt analyserede tekster består således af ét enkelt ord (oftest et pronomen eller et tidsadverbium), og det, der adskiller tekster med forskellig gennemsnitlig forfелtslængde, er således i hvilken grad disse prototypiske forfелter varierer i sammenhæng med enkeltteksempler på lange forfелter med længere nominalsyntagmer, tidsadverbialer eller ledsætninger.

Fra et forfелtsperspektiv kommer skriveudvikling således ikke entydigt til udtryk som gennemsnitligt set mere kvantitet over tid, men som en funktionel evne til at variere og tilpasse forfелtet til teksttype og skriveordre. Med andre ord er variabilitet, dvs. evnen til at tilpasse skriftsproget til situationen, et tegn på skriveudvikling, idet denne type af adaptive adfærd er funktionelt og kommunikativt betinget (Fogal 2020). Der er således ikke noget i dette studie,

der indikerer, at bestemte længder eller typer af forfelter er udtryk for bedre skrivefærdigheder eller mere komplekse tekster end andre, ligesom det heller ikke kan udledes af analyserne, at lange forfelter i sig selv er bedre end korte. Analyserne bekræfter således fundene fra andre studier af syntaktisk udvikling i elevtekster, der viser, at skriveudvikling ikke kommer til udtryk som en lineær og kontinuerligt øget længde af en bestemt syntaktisk struktur, men som evnen til at variere mellem forskellige former for grammatisk kompleksitet i teksterne som svar på funktionelt betingede forhold i skrivesituationen eller skriveordren (se fx Crossley et al. 2014; Hardy & Römer 2013; Tannert 2022), og her er forfaltet et af flere steder, hvor denne adaptive evne kommer til udtryk i teksterne.

Afsluttende betragtninger

Skriveudvikling, som det kommer til udtryk i grundskolen, er et komplekst fænomen, og det kan som følge heraf studeres på forskellige niveauer og fra mange forskellige metodiske perspektiver. I denne artikel har vi undersøgt en lille afgrænset del af dette felt ved at sætte særskilt fokus på, hvordan elever anvender og varierer forfaltet på tværs af to klassetrin og to teksttyper. Ved blandt andet at fokusere på betydningen af forfelterslængde har vi forsøgt at bidrage til den voksende gruppe af forskningslitteratur, der undersøger tekster ved hjælp af automatiserede og digitale analyseværktøjer. En væsentlig metodologisk pointe i forskningslitteraturen er, at selv om længdebaserede mål af forskellige sproglige træk i elevtekster godt kan identificere udviklingstræk i elevernes tekster og ofte kan være med til at forudsige, hvilken karakter en tekst vil få (dvs. har en høj prædiktiv validitet), så kan de være vanskelige at fortolke lingvistisk (se fx diskussion heraf hos Norris & Ortega 2009). I denne artikel har vi derfor suppleret de kvantitative analyser med kvalitative analytiske nedslag i seks elevtekster for at få en dybere og grammatisk detaljeret analyse og forståelse for, hvordan eleverne anvender forfaltet. Et væsentligt fokuspunkt fremadrettet er imidlertid at forskningsfeltet for automatiseret tekstanalyse udvikler nye og mere detaljerede kvantitative mål for forskellige dele af elevernes tekster. Det svarer til det, som Biber et al. (2020) omtaler som en bevægelse fra prædiktive mål til mål af grammatisk kompleksitet. For eksempel er det i denne artikels analyser blevet klart, at teksttype og skriveordre har stor betydning for, hvilket sætningsled eleverne placerer i forfaltet, idet skriveordren tvinger eleverne til at anvende forfaltet deiktisk, mens den berettende teksttype stiller krav til, at forfaltet særligt anvendes til at skabe kohæsion i teksten gennem temporale sætningskoblinger. Tilsvarende kan man i forlængelse af Togeby (2014:146-150) forestille sig, at forskellige former for grammatisk kompleksitet i forfaltet varierer systematisk relativt til kompositions-mønstrene i en tekst – altså at bestemte sætningsled og kompleksitetsformer er knyttet til be-

stemte dele af en teksts generelle sagsfremstilling. Disse forhold kan således være fokuspunkter i fremtidige studier af forfeltsvariation i elevtekster fra grundskolen.

Litteratur

- Beard, R., & Burrell, A. (2010) Investigating narrative writing by 9–11-year-olds. *Journal of Research in Reading*, 33(1):77–93. <https://doi.org/10.1111/j.1467-9817.2009.01433.x>.
- Beard, R., Burrell, A., & Homer, M. (2016) Investigating persuasive writing by 9–11 year olds. *Language and Education*, 30(5), 417–437. <https://doi.org/10.1080/09500782.2015.1133638>.
- Biber, D., Gray, B., Staples, S., & Egbert, J. (2020) Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes*, 46, 1–15. <https://doi.org/10.1016/j.jeap.2020.100869>.
- Bremholm, J., Bundsgaard, J., & Kabel, K. (2022) Proficiency scales for early writing development. *Writing and Pedagogy*, 13(1–3), 121–154. <https://doi.org/10.1558/wap.21490>.
- Bremholm, J., Kabel, K., Liberg, C., & Skar, G. B. (2021) A review of Scandinavian writing research between 2010 and 2020. *Writing and Pedagogy*, 13(1–3), 7–49. <https://doi.org/10.1558/wap.21637>.
- Chrystal, J.-A., & Ekvall, U. (1999) Planering och revidering i skolskrivande. I L.-G. Andersson (Red.), *Svenskans beskrivning 23* (s. 57–76). Lund University Press.
- Crossley, S. A. (2020) Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3), 415–443. <https://doi.org/10.17239/jowr-2020.11.03.01>.
- Crossley, S. A., & Kim, M. (2022) Linguistic Features of Writing Quality and Development: A Longitudinal Approach. *Journal of Writing Analytics*, 6, 59–93. <https://doi.org/10.37514/JWA-J.2022.6.1.04>.
- Crossley, S. A., & McNamara, D. S. (2014) Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79. <https://doi.org/10.1016/j.jslw.2014.09.006>.
- Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014) What Is Successful Writing? An Investigation Into the Multiple Ways Writers Can Write Successful Essays. *Written Communication*, 31(2), 184–214. <https://doi.org/10.1177/0741088314526354>.
- Crossley, S. A., Weston, J. L., Sullivan, S. T. M., & McNamara, D. S. (2011) The Development of Writing Proficiency as a Function of Grade Level: A Linguistic Analysis. *Written Communication*, 28(3), 282–311. <https://doi.org/10.1177/0741088311410188>.
- Diderichsen, P. (1946) *Elementær dansk grammatik*. Gyldendal.
- Diderichsen, P. (1966) Prolegomena til en metodisk dansk Syntax. I *Helhed og struktur: Udvalgte sprogvidenskabelige afhandlinger*, s. 21–24. Gads Forlag.
- Eisenstein, J. (2019) *Introduction to natural language processing*. The MIT Press.
- Enevoldsen, K., Hansen, L., & Nielbo, K. (2021) DaCy: A unified framework for Danish NLP. *Ceur Workshop Proceedings*, 2989, 206–216.
- Fogal, G. G. (2020) Investigating Variability in L2 Development: Extending a Complexity Theory Perspective on L2 Writing Studies and Authorial Voice. *Applied Linguistics*, 41(4), 575–600. <https://doi.org/10.1093/applin/amz005>.

- Garne, B. (1988) *Text och tanke: Om skrivstrategier i elevuppsatser*. Liber.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011) Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, 40(5), 223–234.
<https://doi.org/10.3102/0013189X11413260>.
- Hardy, J. A., & Römer, U. (2013) Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, 8(2), 183–207.
<https://doi.org/10.3366/cor.2013.0040>.
- Hirst, G. (2013) Computational Linguistics. I K. Allan (Red.), *The Oxford Handbook of The History of Linguistics* (s. 707–726). Oxford University Press.
- Hultman, T. G., & Westman, M. (1977) *Gymnasistsvenska*. Liber Läromedel.
- Josephson, O., Melin, L., & Oliv, T. (1998) *Elevtext: Analyser av skoluppsatser från åk 1 till åk 9*. Studentlitteratur.
- Jørgensen, H. (2000) *Indføring i Dansk Syntaks* (3. udg.). Århus Universitet.
- Kabel, K., Bremholm, J., Bundsgaard, J., & Berthelsen, U. D. (2020) *Elevers tidlige skriveudvikling: Identificerede trin* (Working paper i ATEL). DPU, Aarhus Universitet.
- Larsson, K. (1984) *Skrivförmåga: Studier i svenskt elevspråk*. Liber Förlag.
- Lindgren, J. (2021) Fundamentet: Syntaktisk struktur möter textgrammatisk funktion. I J. Brandtler & M. Kalm (Red.), *Nyanser av grammatik: Gränser, mångfald, fördjupning* (s. 179–192). Studentlitteratur AB.
- McEnery, T., & Hardie, A. (2013) The History of Corpus Linguistics. I K. Allan (Red.), *The Oxford Handbook of The History of Linguistics* (s. 727–745). Oxford University Press. <https://doi.org/10.1093/oxford-hb/9780199585847.013.0034>.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010) Linguistic Features of Writing Quality. *Written Communication*, 27(1), 57–86. <https://doi.org/10.1177/0741088309351547>.
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013) Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45(2), 499–515. <https://doi.org/10.3758/s13428-012-0258-1>.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014) *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>
- Myhill, D. (2008) Towards a Linguistic Model of Sentence Development in Writing. *Language and Education*, 22(5), 271–288. <https://doi.org/10.1080/09500780802152655>.
- Norris, J. M., & Ortega, L. (2009) Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>.
- Perin, D., & Lauterbach, M. (2018) Assessing Text-Based Writing of Low-Skilled College Students. *International Journal of Artificial Intelligence in Education*, 28(1), 56–78. <https://doi.org/10.1007/s40593-016-0122-z>.
- Pettersson, Å. (1980) *Nittonårssvenska: Språket i 1977 års centrala prov i svenska för gymnasiets årskurs 3*. Lärarhögskolan.
- Pettersson, Å. (1989) Utvecklingslinjer och utvecklingskrafter i elevernas uppsatser. I C. Sandqvist & U. Teleman (Red.), *Språkutveckling under skoltiden* (s. 159–184). Studentlitteratur.
- Reilly, D., Neumann, D. L., & Andrews, G. (2019) Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445–458. <http://dx.doi.org/10.1037/amp0000356>.

Forfælsvariation i elevtekster fra 5. og 9. klassetrin

- Tannert, M. (2022) *Skriftsproglig udvikling i grundskolens danskfag: Et korpusbaseret studie af elevtekster fra 5.-9. klassetrin*. Ph.d.-afhandling. DPU, Aarhus Universitet. <https://doi.org/10.7146/aulsps-e.466>.
- Togeby, O. (2014) *Bland blot genrene—Ikke teksterne! Om sprog, tekster og samfund*. Samfundslitteratur.
- Troia, G. A., Shen, M., & Brandon, D. L. (2019) Multidimensional Levels of Language Writing Measures in Grades Four to Six. *Written Communication*, 36(2), 231–266. <https://doi.org/10.1177/0741088318819473>.
- Weber, M. (2022) Language skills in student essays: Social disparities and later educational attainment. *Longitudinal and Life Course Studies*, 13(2), 239–261. <https://doi.org/10.1332/175795921X16244506861047>

Danskerne og det engelske sprog i 1990'erne og 2020'erne¹

Jacob Thøgersen

Københavns Universitet

Indledning

MUDS er som vi alle ved Møderne om Udforskningen af Dansk Sprog, så hvordan passer en artikel om danskernes forhold til det *engelske* sprog ind i denne samling? Ja, det gør den måske heller ikke, men det er min påstand at engelsk i dag fylder så meget i alle danskeres sproglige hverdag at en fyldestgørende sociolingvistisk beskrivelse af dansk nødvendigvis også må forholde sig til danskernes brug af engelsk. Jeg håber at jeg på de næste sider kan overbevise om rimeligheden i sådan en påstand.

For godt 20 år siden publicerede Bent Preisler (1999) den til dato måske mest omfattende undersøgelse af danskeres holdninger til og brug af engelsk. Han viste at engelsk blev brugt af et flertal af danskere, men også at omkring 20 % af især de ældre var ude af stand til at bruge engelsk.

I årene efter Preislers undersøgelse gennemførtes også det fællesnordiske projekt *Moderne Importord i Sprogene i Norden* (2002 –) (Kristiansen & Sandøy 2010)². Projektet viste at engelsk spillede en vis rolle i alle nordiske sprogområder; og at danskere, når de blev spurgt eksplicit, var de mest positive over for engelsk, men når de blev 'spurgt' implicit ved at reagere på oplæsninger med flere eller færre engelske ord, var de blandt de mest negative (Kristiansen & Vikør 2006, Kristiansen 2006). Jarvad (2001) undersøgte status for engelsk i Danmark med særlig fokus på om der var anledning til bekymring for 'domænetab' især inden for forskning og videnskab.

Interessen (og bekymringen) for engelsk i Danmark gav også anledning til adskillige regeringsudspil (Kulturministeriet 2003, 2008) med målet at 'styrke' og 'bevare' dansk og at bekæmpe domænetab. Alt i alt var der altså en stor interesse for engelsk i Danmark i årene omkring årtusindskiftet.

¹ Denne artikel er en kortere oversigtsversion af undersøgelsen. For en mere udfoldet diskussion, se Thøgersen & Preisler (under udgivelse).

² Se også <https://folk.uib.no/hnohs/Bokserien.htm> og <https://novus.no/collections/moderne-importord>.

Det sproglige klima i Danmark er ændret markant siden da. Dansk er fortsat det eneste officielle sprog i Danmark (på trods af politisk tale om at give engelsk en mere formel status), men vi har set en øget sproglig mangfoldighed, bl.a. pga. migration; og engelsk er blevet et endnu mere udbredt og nærmest usynligt 'andet' sprog i Danmark (Lønsmann, Mortensen & Thøgersen 2022). Heidemann Andersen & Jarvad (2018) kan ses som en opfølgning på Moderne Importords-projektets spørgsmål om låneord og låneordstilpasning, altså spørgsmålet om 'sprogforandring'. Med projektet *English and Globalisation in Denmark* (eller: Engelsk i Danmark, herefter Enida)³ forsøger vi at beskrive hvilke 'sociolingvistiske forandringer' (Coupland 2009, 2014) der er sket mht. engelsk i Danmark. Vi ser altså ikke på hvor mange engelske ord der er i dansk, men på hvad brugen af engelsk og andre sprog betyder for danskere og det danske sprogsamfund i 2020'erne.

Historiske udviklinger

Vi forventer at især tre historiske udviklinger over det sidste kvarte århundrede har sat deres spor i danskeres forhold til engelsk: 1) den øgede europæiske integration, 2) liberaliseringer og internationalisering af medielandskabet, 3) internettets nærmest eksplosive udvikling. De gennemgås i hovedtræk nedenfor.

Efter Berlinmurens fald og sovjetblokkens opløsning i starten af 1990'erne iværksattes et europæisk (re)integrationsprojekt. Efter at Europa i en vis forstand havde været delt i to i halvtreds år, skulle de to halvdele (og ikke mindst de to halvdele af Tyskland) (gen)forenes. De objektive indikationer på dette projekt er fx EU's/EF's udvidelse fra 11 medlemsstater i 1992 til 27 i 2022, og diverse EU-støttede programmer for at øge udveksling og integration mellem staterne, herunder princippet om 'arbejdskraftens frie bevægelighed' og Bologna-processen der har til mål at gøre det lettere for europæere at tage (dele af) deres uddannelse på højere læreanstalter i andre af EU's medlemslande. Begge dele har medført en øget mobilitet, og en øget tilstedeværelse af 'internationale' sprog – i praksis engelsk – på europæiske arbejdspladser og uddannelsesinstitutioner.

Indtil 1988 havde Danmarks Radio monopol på at udsende landsdækkende radio og tv i Danmark. Nogle kunne høre og se nabolandskanaler, men for det store flertal var radio og tv dansksproget eller teksten på dansk. Op igennem 90'erne fik vi adgang til satellit-tv, men det var intet imod det udbud af udenlandsk tv som vi i dag har adgang til med bare den mindste kabel-tv-pakke. Samtidig er tv blevet frikoblet fra det klassiske flow-tv sådan at mere og mere

³ <https://engerom.ku.dk/forskning/centre/english-and-globalisation-in-denmark/>.

tv-kigning foregår via streamingtjenester eller via gør-det-selv-videotjenester som YouTube, TikTok, Twitch osv. Det er jo ikke sådan at al international tv og video er på engelsk, men engelsk fylder meget i dette internationale medielandskab og fungerer for mange danskere som adgang til globalt tv.

Det bringer os videre til den måske mest slående samfundsudvikling, nemlig internettets udbredelse. I 2001 da Moderne Importord-projektet gennemførte sin spørgeskemaundersøgelse (Kristiansen & Vikør 2006) gav det mening at spørge respondenterne om de havde en computer hjemme, og i givet fald om det var tilsluttet internettet. Et tilsvarende spørgsmål ville være meningsløst i 2022 hvor alle 'har internettet i lommen'. Internettet har for mange betydet en langt højere grad af kontakt til engelsk, fx i form af underholdning, debatter og nyheder formidlet via nettet, handel på nettet osv. Igen er engelsk absolut ikke eneste sprog der bruges, men landegrænser betyder mindre på nettet, og engelsk fylder meget hvis man går ud over landets grænser.

Formålet med Enida-undersøgelsen er at finde ud af hvilke sociolingvistiske forandringsprocesser disse samfundsforandringer har givet anledning til. Preislers (1999) undersøgelse havde to hoveddele: Dels en 'overfladekortlægning', med et engelsk lån 'survey', af danskernes brug af og holdninger til engelsk udført som en landsdækkende spørgeskemaundersøgelse; dels mere dybdegående kvalitative undersøgelser af forskellige ungdomskulturer. Her er det survey-undersøgelsen som vi fokuserer på. Preislers undersøgelse blev udført med ca. 850 repræsentativt udvalgte respondenter og indsamlet i 1995. I foråret 2022 gennemførte vi en diakron gentagelse af (dele af) denne survey med et ligeledes repræsentativt udvalg af danskere af samme størrelsesorden.

Metode

Metoden er altså et spørgeskemabaseret survey. Spørgeskemaspørgsmålene var i al væsentlighed ordret de samme i 1999-undersøgelsen og 2022-undersøgelsen for at sikre maksimal sammenlignelighed. Dog blev "De" ændret til "du" konsekvent, og nogle spørgsmål blev omformuleret for at passe bedre til moderne kontekster. Fx spurgtes der i 1999 om respondenter ønskede en funktion så de kunne slå undertekster fra på deres tv. Sådan en funktion findes nu, så spørgsmålet er forældet og udgik derfor. Omvendt er der kommet en masse nye (it-)teknologier til, ikke mindst forskellige former for sociale medier. Dem er der spurgt til i den nye undersøgelse. Tilsvarende bruger mange danskere engelsk forbindelse med deres arbejde. Det er også noget vi spørger til i den nye undersøgelse. En udtømmende undersøgelse af hvor meget og hvordan danskere bruger engelsk i dag, må selvfølgelig også inkludere disse spørgsmål. Her fokuserer

vi dog på det som kan sammenlignes over tid. Spørgsmål om fx sociale medier og engelsk på arbejde er udeladt af sammenligningen her fordi de ikke blev stillet i 1999-undersøgelsen.

I denne gennemgang fokuserer vi kun på spørgsmål som er de samme (med undtagelse altså af ”De”/”du”-ændringen). I tilfælde hvor andet er omformuleret nævner vi det. ’De samme’ skal selvfølgelig tages med et stort gran salt. Man kan præsentere respondenterne for samme ordlyd, men det betyder absolut ikke at spørgsmålet betyder det samme når det er stillet to gange med ca. 25 års mellemrum. Vi vender tilbage til problemet i diskussionen, men en helt oplagt udfordring er at den større tilstedeværelse af engelsk i dag, kan betyde en del for respondenternes opmærksomhed om engelsk. På kontraintuitiv vis kan det således vise sig at man er *mindre* opmærksom på engelsk fordi det ikke længere stikker ud som et fremmedelement.

Respondenter blev i begge undersøgelser udtrukket og kontaktet af et (semi)offentligt statistikinstitut. I Preislers tilfælde Socialforskningsinstituttet, i 2022-gentagelsen af Danmarks Statistik. Det betyder at stikprøven er sandsynlighedsudtrukket på baggrund af CPR-registreret. I 1999-undersøgelsen blev respondenterne opsøgt og interviewet hjemme. I 2022 blev de kontaktet via deres eBoks og blev bedt om at udfylde spørgeskemaet online.

Delvis pga. respondenternes selvsektion (altså om de valgte at svare eller ej), og delvis pga. samfundsforandringer er stikprøverne noget forskellige. I 1999-undersøgelsen var der relativt flere unge mens der i 2022 var relativt flere ældre. I forlængelse af det var der i 1999-undersøgelsen flere arbejdsaktive (og færre pensionister) end i 2022-undersøgelsen – 62 % mod 51 % erhvervsaktive. Uddannelsesniveaue er også ret forskelligt i de to stikprøver. I 1999 var der en klar overvægt af respondenter med kortere uddannelser, i 2022 nærmest lige mange med kortere og med videregående uddannelser.

I gennemgangen her kontrollerer vi for den skæve alderssammensætning ved at vise resultater nedbrudt på aldersgrupper (i 10-års-intervaller). Vi sammenligner altså ikke stikgrupperne samlet, men aldersgruppe per aldersgruppe.

Vi analyserer den sociolingvistiske forandring i danskeres forhold til engelsk ved at slå ned på tre forhold: (Selvoplevet) KOMPETENCE, (selvoplevet) KONTAKT og HOLDNINGER. Man kan med god ret hævde at både kompetence- og kontaktspørgsmål er udtryk for holdninger til engelsk når man *spørger* i stedet for at *teste* eller *observere* faktisk kompetence og kontakt. Når man vurderer hvor god man er til engelsk, siger man sikkert også noget om sin holdning til engelsk; og det samme når man siger hvor ofte man bruger engelsk.

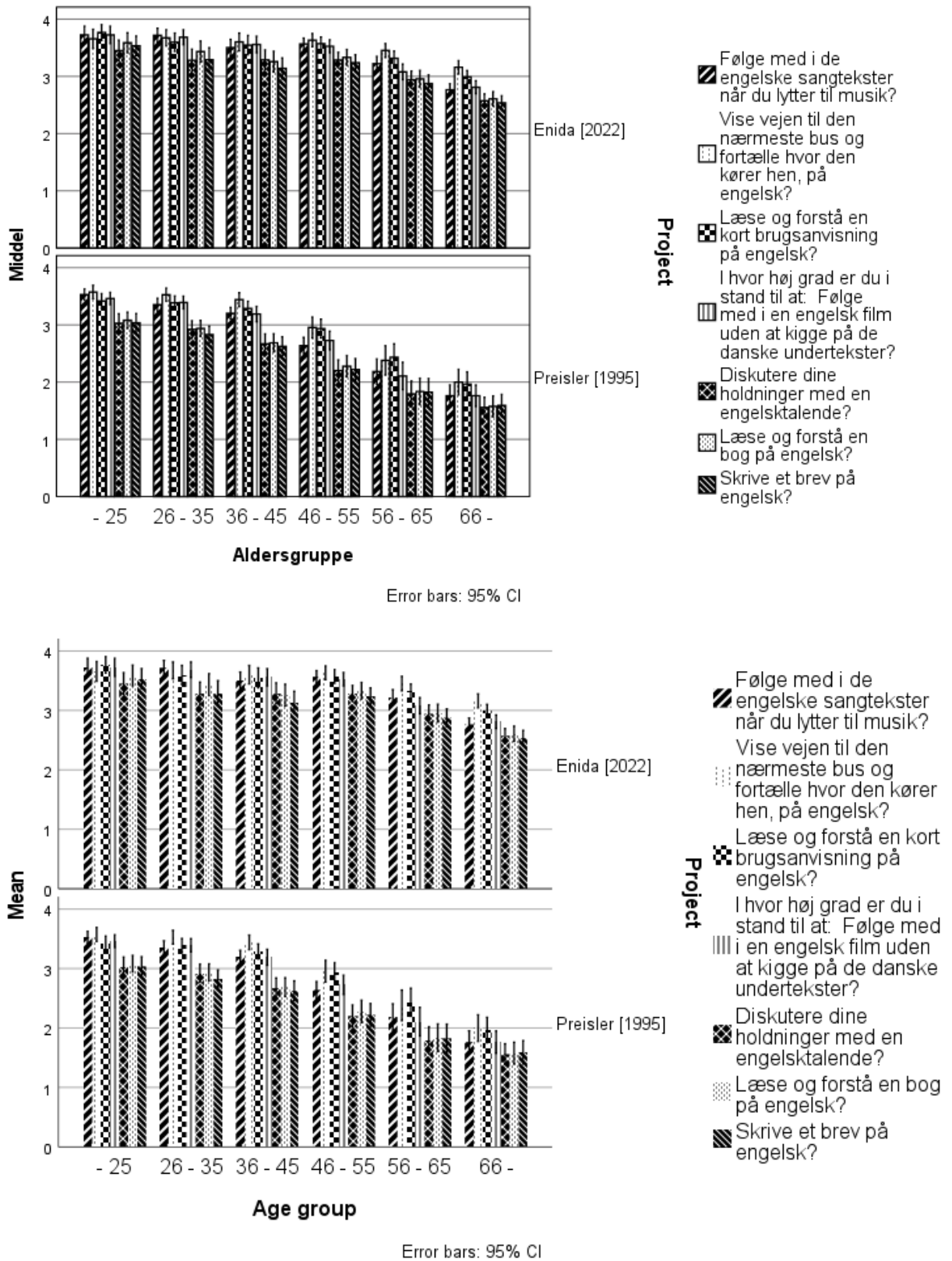
KOMPETENCE

Respondenternes (selvrapporterede) KOMPETENCE blev målt med syv spørgsmål af formen: ”i hvor høj grad er du i stand til at...”:

- ...følge med i de engelske sangtekster når du lytter til musik
- ...vise vejen til den nærmeste bus og fortælle hvor den kører hen, på engelsk
- ...læse og forstå en kort brugsanvisning på engelsk
- ...følge med i en engelsk film uden at kigge på de danske undertekster
- ...diskutere dine holdninger med engelsktalende
- ...læse og forstå en bog på engelsk
- ...skrive et brev på engelsk

Bortset fra ændringen af ”De”/”Deres” til ”du”/”dine” var formuleringerne ordret de samme i de to undersøgelser. Svarmulighederne var givet på en 4-trins ordinalskala med punkterne ”i høj grad”, ”i nogen grad”, ”i ringe grad” og ”slet ikke”.

Figur 1 viser svar på de 7 spørgsmål opgjort som gennemsnitsværdi per aldersgruppe med Enida (2022-projektet) i øverste halvdel af figuren og Preisler (1999-projektet) i nederste halvdel af figuren.



Figur 1: KOMPETENCE opdelt i aldersgrupper og datasæt.

Vi bemærker især tre ting: For det første at engelskkompetencen for den enkelte er relativt konstant over den 25-årige periode. Hvis man tænker på at de respondenter der var under 25 i 1999, svarer til dem der er 46–55 i 2022, 25 år senere osv., kan man mentalt 'parallelforskyde' de to sæt af søjler med tre trin. Hvis man gør det, vil man se at gruppen af under 25-årige i 1999 ligner 46–55-årige i 2022 til forveksling, at 26–35-årige i 1999 ligner 56–65-årige i 2022 osv. Det tyder altså som sagt på relativ konstans i den enkeltes selvopfattede engelskkompetence. Alderseffekterne som man tidligere har fundet (altså at unge vurderer deres kompetencer højest) er altså ikke en effekt af 'age-grading', at unge altid vurderer sig selv højere. Der er tale om en reel forandring.

For det andet er den relative kompetence væsentligt højere i 2022 end i 1999. Blandt de alleryngste er forskellen ikke meget stor (måske pga. såkaldt loftseffekt), men blandt de andre aldersgrupper er der klare forskelle. Den gennemsnitlige selvrapporterede engelskkompetence er altså steget.

For det tredje ser vi at alderseffekten, eller gradienten for alder, er meget mindre i 2022 end den var i 1999. Det er i begge datasæt sådan at yngre vurderer deres kompetencer højere end de ældre; men effekten var væsentligt større i 1999 end den er i 2022.

Man må have in mente at metoden er lidt forskellig i 1999 og 2022: I 1999 sad respondenterne ansigt-til-ansigt med en interviewer, i 2022 sad de alene foran deres computer og svarede. Der var altså ingen at 'stå til ansvar for' i besvarelsen. Det er muligt at denne ændrede situation er medvirkende til de højere selvrapporterede kompetencesvar. Omvendt kan man også sige at det potentielle ansigtstab ved at erkende at man ikke kan engelsk, er tilsidesat, og det kunne potentielt bevirke at flere ville rapportere lavere kompetencer når de svarer på computer. Det er ikke til at sige med sikkerhed. Disse evt. effekter vil dog kun påvirke den overordnede sammenligning af respondenterne i 1999 og 2022. Der er ingen grund til at tro at den forskellige grad af 'socialitet' skulle påvirke de forskellige aldersgrupper forskelligt.

English have-nots

En anden måde at undersøge engelskkompetence over tid er ved at se på antallet af, som Preisler kaldte dem, 'English have-nots', altså danskere som svarer at de "slet ikke" kan udføre de forskellige opgaver.

I tabel 1 opgør vi andelen af respondenter som svarer "slet ikke" på hvert af de 7 kompetencespørgsmål. I sidste række opgør vi andelen for hvem "slet ikke" er det eneste svar de giver, det er deres 'maksimale engelskkompetence'.

KOMPETENCE:	Andel 'Slet ikke'-svar	
	Preisler	Enida
"I hvor høj grad er du i stand til at..."	(1999)	(2022)
Følge med i de engelske sangtekster når du lytter til musik?	15,3%	5,4%
Vise vejen til den nærmeste bus og fortælle hvor den kører hen, på engelsk?	17,1%	5,5%
Læse og forstå en kort brugsanvisning på engelsk?	14,9%	6,1%
Følge med i en engelsk film uden at kigge på de danske undertekster?	17,9%	6,6%
Diskutere dine holdninger med en engelsktalende?	32,2%	11,0%
Læse og forstå en bog på engelsk?	29,6%	12,0%
Skrive et brev på engelsk?	32,7%	13,0%
Engelskkompetence _{Max}	9,7%	2,8%

Tabel 1: 'English have-nots' i 1995 og 2022.

Tabellen viser generelt et fald til i størrelsesordenen en tredjedel af antallet af 'English have-nots'. Rækkefølgen af opgaverne, altså hvad der opfattes som lette og svære opgaver, er stort set den samme, men væsentligt færre mener at de "slet ikke" kan udføre opgaverne. Det forklares bedst med at de ældste og engelsksvageste respondenter i 1995-undersøgelsen ikke længere er til stede i 2022-undersøgelsen.

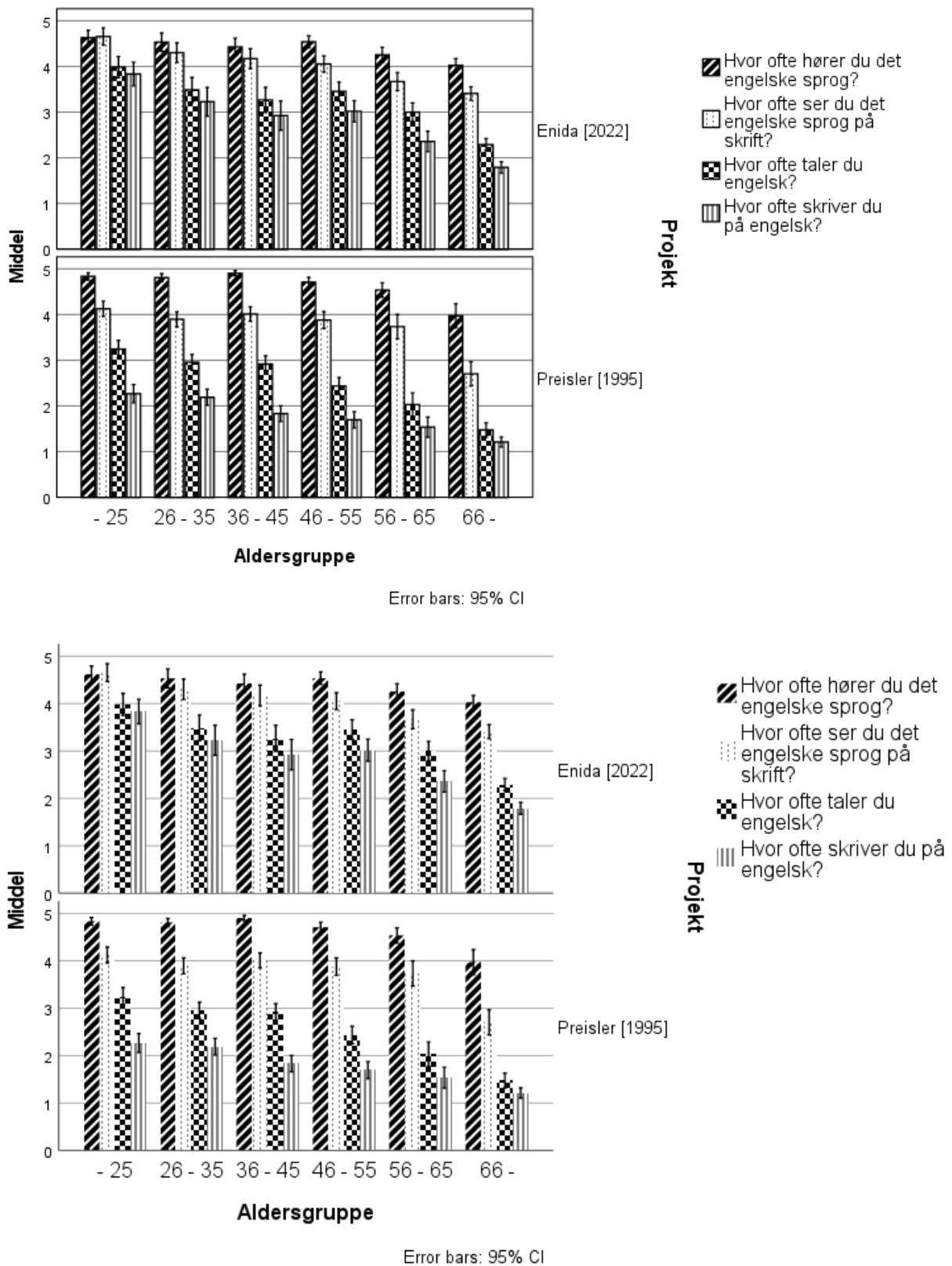
KONTAKT

KONTAKT med engelsk blev afdækket med fire spørgsmål der spørger respondenter hvor tit de henholdsvis "taler", "skriver" og "hører" engelsk og hvor tit de "ser engelsk på skrift". Svarmulighederne var givet på en 5-trins ordinalskala med punkterne "én til flere gange om dagen", "én til flere gange om ugen", "én til flere gange om måneden", "én til flere gange om året" og "sjældnere end én gang om året/aldrig".

Figur 2 viser igen svar på spørgsmålene opgjort som gennemsnitsværdi per aldersgruppe med Enida (2022-projektet) i øverste halvdel af figuren og Preisler (1999-projektet) i nederste halvdel af figuren.

Igen hæfter vi os ved tre ting: For det første: vi bemærkede med hensyn til KOMPETENCE at respondenterne var relativt konstante over deres liv. For KONTAKT gælder der ikke det samme. Hvis man igen parallelforskyder søjlerne tre aldersgrupper til siden, vil man se at de ikke ligner hinanden så meget for KONTAKT som de gjorde for KOMPETENCE. Respondenter af samme generation har alt i alt oplevet en stigende kontakt med engelsk over deres liv – der er altså tale om en sociolingvistisk forandring i hele samfundet.

For det andet: Da vi så på alderseffekten mht. KOMPETENCE så vi at den var mindre i 2022 end i 1999. Det er i mindre grad tilfældet for KONTAKT, og i det omfang det er gældende, så gælder det kun for det man kunne kalde 'produktiv' kontakt. Mao.: de ældre respondenter i 1999 oplevede stort set aldrig at tale eller skrive engelsk, mens de yngre gjorde det i et vist omfang. Også i 2022 oplever de yngre oftere at skrive og tale engelsk, men forskellen på de unge og de ældre er blevet mindre, ligesom det gjaldt for KOMPETENCE. Men når det kommer til 'receptiv' KONTAKT, altså at "høre engelsk" og at "se engelsk på skrift", er forskellene mindre markante. Der er i begge datasæt en effekt af alder, men den er mindre, og den ser ikke ud til at være forandret nævneværdigt.



Figur 2: KONTAKT opdelt i aldersgrupper og datasæt.

Det peger på den tredje ting, nemlig at *typen* af engelskkontakt er forandret. Hvor der i 1999 for alle aldersgrupper var en markant forskel på graden af 'receptiv' og 'produktiv' kontakt, så er denne forskel stort set udjævnet for de yngste i 2022 – de taler og skriver næsten lige så meget engelsk som de hører og ser.

Igen kan man supplere opgørelsen af forandringer ved at se på hvor mange der "aldrig" oplever engelsk – 'English do-nots' kunne man kalde dem. Det har vi gjort i tabel 2. Vi opdeler spørgsmålene i 'receptiv' og 'produktiv' kontakt, og ser igen også på dem der svarer "aldrig" til alle fire kontakttyper, dvs. dem for hvem "aldrig" er deres 'maksimale engelskkontakt'. I nederste række opgiver vi dem der som deres 'maksimale engelskkontakt' har "én til flere gange hver dag". Vi skelner i denne række ikke mellem om de "hører", "taler", "skriver" eller "ser" engelsk hver dag, men registrerer bare det som de gør oftest.

Do nots

KONTAKT: "Hvor ofte ..."	Andel 'Sjældnere end en gang om året' / 'Aldrig'-svar	
	Preisler (1999)	Enida (2022)
Hører	2,3%	4,4%
Ser på skrift	8,1%	7,8%
Middel _{Receptiv}	5,2%	6,1%
Taler	23,9%	16,1%
Skriver	54,3%	35,9%
Middel _{Produktiv}	39,1%	26,0%
Engelskkontakt _{Max}	2,5%	2,2%
Møder engelsk dagligt	82,4%	64,8%

Tabel 2: 'English do-nots' i 1999 og 2022.

Resultaterne er meget forskellige for den receptive og den produktive kontakt. Mht. den produktive kontakt, så er andelen der aldrig taler eller skriver engelsk, faldet med omkring en tredjedel, fra 39 % til 26 %. Mht. den receptive kontakt, så er andelen der aldrig hører eller ser engelsk, *steget*. Der er altså i 2022 lidt flere som aldrig oplever at høre engelsk eller se engelsk på skrift. Bemærk dog de relativt små forskelle, henholdsvis 0,3 %-point (en 4 % stigning) for 'engelsk på skrift', og 2,1 %-point (en 91 % stigning) for 'høre engelsk'. Vores bedste forklaring på dette lidt kontraintuitive resultat er at engelsk er blevet mindre *markeret* eller mindre synligt. Man er måske mindre opmærksom på at man faktisk støder på engelsk. Hvis det er rigtigt, rejser det

nogle interessante spørgsmål ved metoden med at stille 'de samme spørgsmål to gange med 25 års mellemrum'. For er spørgsmålene 'de samme', når konteksten er så forandret?

Andelen der aldrig (oplever at de) møder engelsk overhovedet, er stort set konstant, 2,5 % mod 2,2 %. Og andelen der møder engelsk hver dag, er *faldet* fra ca. 82 % til ca. 65 %. Som antydet gætter vi på at det skyldes en mindre grad af opmærksomhed på den 'receptive' engelskkontakt fx i form af engelsksproget musik og film eller i form af engelsksprogede skilte og busreklamer. Vi vil undersøge det videre i kommende undersøgelser (se Thøgersen & Preisler, under udgivelse).

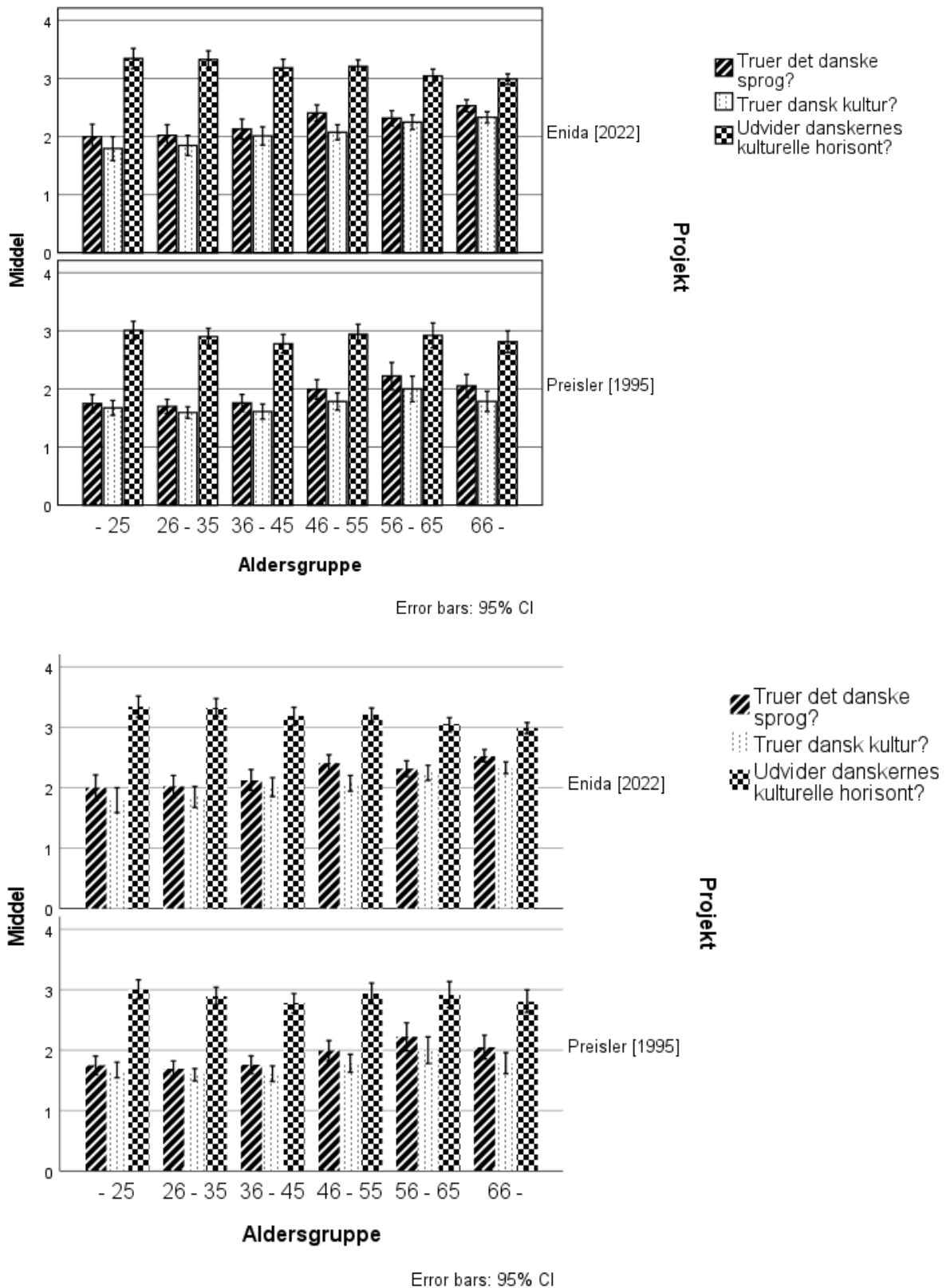
HOLDNINGER

HOLDNINGER blev undersøgt med tre ret generelle eller abstrakte spørgsmål om konsekvenserne af engelsk i Danmark, nemlig "Det engelske sprogs tilstedeværelse i hverdagen..."

- ...truer dansk sprog?
- ...truer dansk kultur?
- ...udvider danskernes kulturelle horisont?

Som man kan se, er de formuleret som to 'negative' holdningsspørgsmål og et 'positivt'. Resultaterne i figur 3 viser en højere grad af enighed i *begge* holdningsdimensioner. Respondenterne er altså mere enige i både de positive og de negative aspekter af engelsks tilstedeværelse.

Danskerne og det engelske sprog i 1990'erne og 1920'erne



Figur 3: HOLDNINGER opdelt i aldersgrupper og datasæt.

Både i 1999 og i 2022 er der en tendens til at de yngre grupper er mere enige i det positive udsagn og de ældre er mere enige i de negative. Det er bemærkelsesværdigt at aldersforskellene

er *større* i 2022 end de var i 1999. Hvor vi for KOMPETENCE- og KONTAKT-spørgsmålene overordnet set så en mindre effekt af alder i 2022, så er alderseffekten på holdnings-spørgsmålene blevet *større* over tid.

Samlet set er det mest de 'negative' holdninger der er øget. Den øgede KOMPETENCE og KONTAKT har altså ikke generelt set ført til mere positive holdninger, men i højere grad til mere negative holdninger. En mulig tolkning er at det skal ses i sammenhæng med den stigende nationalisering i Europa. Hvor man i årtusindets første årti talte meget om internationalisering og nedbrydning af nationale grænser, så har de politiske vinde i Europa i de sidste år blæst i retning af øget *nationalisering*. I den danske universitetsverden har vi meget tydeligt oplevet denne ny-nationalisering eller "internationalisering 2.0" som Janus Mortensen har kaldt det: I start 2000'erne begyndte man gennemførslen af 'Bolognaprocessen' som skulle øge uniformiteten og dermed lette mobiliteten for uddannelse inden for EU, og universiteterne konkurrerede om at have flest udenlandske studerende og flest engelsksprogede kurser. I 2021 fik vi med "Aftale om reduktion af engelsksprogede videregående uddannelser" et regeringsindgreb *imod* engelsksprogede uddannelser for at: "nedbringe SU-udgifterne til vandrende arbejdstagere ved at reducere tilgangen til engelsksprogede uddannelser" (ibid.).

Uden for universitetsverden kan man fx pege på Storbritanniens udtræden af EU som ses som en reaktion på hvad der opleves som overdreven europæisk indvandring til Storbritannien. Og i Danmark kan man pege på den kraftige fremgang for partier med en kraftig national orientering gennem det sidste årti, fx Dansk Folkeparti, Nye Borgerlige og Danmarksdemokraterne.

En anden tolkning er at de sprogpoltiske debatter fra starten af årtusindet måske har sat sine spor. Dengang blev det nærmest etableret som en politisk sandhed at dansk var truet af 'domænetab' fra engelsk. Både Jarvad (2001) og de sprogpoltiske redegørelser *Sprog på spil* (Kulturministeriet 2003) og *Sprog til tiden* (Kulturministeriet 2008) fastslog at dansk var under pres fra engelsk. I Jarvad (2001:110) hedder det fx i forbindelse med videnskabelige publikationer at: "... domænet videnskab inden for de naturvidenskabelige-tekniske fag er tabt, og aktiviteten på dansk er så lille at det danske videnskabssprog stagnerer og dør ud", og mht. dansk kultur siger *Sprog på spil* (Kulturministeriet 2003:11) "den dansksprogede kunst befinder sig i skærpet konkurrence med nye medieprodukter og stærke internationale producenter både inden for litteratur, film og scenekunst". Når vi i spørgeskemaet spørger om engelsk "truer dansk sprog" og "truer dansk kultur", kan det altså opfattes ikke så meget som et holdningsspørgsmål, men mere som et spørgsmål om 'viden', om man har fulgt med i hvad der er sagt og skrevet i diverse sprogpoltiske debatindlæg.

Konklusion

Vores resultater viser at danskere i 2022 opfatter sig som mere kompetente end i 90'erne, og at de oplever større kontakt med engelsk. Den stigende kontakt er især den 'produktive' kontakt, mens 'receptiv' kontakt er mere konstant eller måske endda (oplevet) faldende. Resultaterne viser også at den øgede KOMPETENCE og KONTAKT har ført til at nogle kompetenceforskelle mellem aldersgrupperne er blevet mindre, og det samme med forskellen i 'produktiv' kontakt mellem aldersgrupperne.

Det er sådanne tendenser der gør at jeg indledningsvist påstod at denne undersøgelse har en plads i MUDS-regi. Hvis man vil give en komplet sociolingvistisk beskrivelse af (danskernes brug af) dansk, bliver man næsten nødt til at forholde sig til engelsk også. Der var engang hvor engelsk havde samme status som andre fremmedsprog: Det kom op når man rejste eller hvis man opsøgte fremmede kulturer – det var for de særligt interesserede. Det er ikke længere tilfældet. Snart sagt alle danskere uanset alder oplever engelsk i deres i øvrigt dansksprogede hverdag. Det gjorde mange også i 1990'erne, men den store forskel er hvor mange der aktivt *bruger* engelsk i forskellige hverdagsaktiviteter, fx læser og skriver i forbindelse med deres arbejde eller opsøger engelsksproget information og underholdning på nettet. Hvis man ikke tænker danskere i dag som tosprogede sprogbrugere og engelsk som et de facto andetsprog i Danmark, tror jeg man misser en vigtig pointe i udforskningen af dansk sprog.

Mht. HOLDNINGER: I 1990'erne fandt Preisler at engelsk både opfattedes som "en trussel" og "en mulighed". Vores resultater antyder at det samme er tilfældet nu, og dermed også at vi ikke entydigt er vidne til en udvikling hvor engelsk opfattes mere og mere positivt. Tværtimod ser vi mere 'ekstreme' holdningsudsagn, både positive og negative. De ældre er generelt mere negative, og de yngre generelt mere positive, og det ser ud som om den holdningsmæssige opposition mellem yngre og ældre er ved at blive større.

Et forbehold for resultaterne er dog problemet om man kan stille 'det samme spørgsmål' to gange med 25 års mellemrum, eller om konteksten (både den samfundsmæssige kontekst og den personlige kontekst som respondenter sidder i mens hun svarer) gør en sådan sammenligning meningsløs. Man kan som bekendt ikke træde ned i den samme flod to gange... En sammenligning der bruger 'det samme' måleinstrument to gange, er kun ét nedslag i sammenligningen af engelsks ændrede status og de sociolingvistiske forandringsprocesser vi har været vidne til over det sidste kvarte århundrede. Med det samlede Enida-projekt håber vi at kunne bringe flere brikker til det store puslespil. Artiklen her giver én puslespilsbrik, Lønsmann, Mortensen & Thøgersen (2022) gav en anden, og i de næste år vil vi forsøge at give flere.

Tak

Artiklen er et produkt af projektet *English and Globalisation in Denmark: A Changing Sociolinguistic Landscape*, et 4-årigt projekt finansieret af Danmarks Frie Forskningsfond | Kultur og Kommunikation, <https://engerom.ku.dk/forskning/centre/english-and-globalisation-in-denmark/>.

Jeg vil gerne takke projektets øvrige deltagere, Dorte Lønsmann, Janus Mortensen, Kamilla Kraft og Marianne Haugaard Skov samt projektets advisory board, Nikolas Coupland, Sari Pietikäinen, Jette G. Hansen, Tore Kristiansen, Barbara Soukup, for deres værdifulde input og frugtbare kommentarer undervejs. Og tak til Andrzej Szubert fra Adam Mickiewicz Universitetet i Poznan, Bror Julebæk Hansen fra Københavns Universitet og Anne Mette Nyvad fra Aarhus Universitet for konstruktive kritiske kommentarer til en tidligere version af denne artikel.

En særlig tak skylder jeg til Bent Preisler for at have lagt grunden som overhovedet muliggjorde en diakron undersøgelse og for yderst frugtbare diskussioner om analyser og fortolkninger i forbindelse med artiklens tilblivelse. Alle fejl og uklarheder er naturligvis mit eget ansvar. Bent nåede desværre ikke selv at opleve udgivelsen af artiklen, før han gik bort i sommeren 2023.

Litteratur

- Coupland, Nikolas (2009) Dialects, standards and social change. Marie Maegaard et al. (red.). *Language attitudes, standardization, and language change*. Oslo: Novus: 27–49.
- Coupland, Nikolas (2014) Sociolinguistic change, vernacularization and broadcast British media. Jannis Androutsopoulos (red.). *Mediatization and sociolinguistic change*. Berlin: De Gruyter: 67–96.
- Heidemann Andersen, Margrethe & Pia Jarvad (red.) (2018) *Yes, det er coolt*. Sprognævnets konferenceseerie bind 6. Bogense: Dansk Sprognævn.
- Jarvad, Pia (2001) *Det danske sprogs status i 1990'erne med særligt henblik på domænetab*, Dansk Sprognævnsskrifter 32. København: Dansk Sprognævn. <https://dsn.dk/wp-content/uploads/2021/01/det-danske-sprogs-status.pdf>.
- Kristiansen, Tore (red.) (2006) *Nordiske sprogholdninger: en masketest*. Moderne Importord i Språka i Norden, vol. 5, Oslo: Novus forlag.
- Kristiansen, Tore & Helge Sandøy (2010) Introduction. The linguistic consequences of globalization: the Nordic laboratory. *International Journal of the Sociology of Language* 204: 1-7.
<https://doi.org/10.1515/ijsl.2010.027>
- Kristiansen, Tore & Lars S. Vikør (2006) *Nordiske språkhaldningar. Ei meningsmåling*. Moderne Importord i Språka i Norden, vol. 6 Oslo: Novus forlag.

Danskerne og det engelske sprog i 1990'erne og 1920'erne

Kulturministeriet (2003) *Sprog på spil – Et udspil til en dansk sprogpolitik*. https://kum.dk/fileadmin/_kum/5_Publikationer/2003/Sprog_paa_spil.pdf.

Kulturministeriet (2008) *Sprog til tiden – rapport fra sprogudvalget*. https://kum.dk/fileadmin/_kum/5_Publikationer/2008/sprog_til_tiden_netpub.pdf.

Lønsmann, Dorte, Janus Mortensen & Jacob Thøgersen (2022) Er engelsk stadig et fremmedsprog i Danmark? Et spørgsmål om kollektiv sproglig identitet. *NyS, Nydanske Sprogstudier* 61: 126–179. <https://doi.org/10.7146/nys.v1i61.132242>.

Preisler, Bent (1999) *Danskerne og det engelsk sprog*. Frederiksberg: Roskilde Universitetsforlag.

Regeringen (2021) *L 59 Forslag til lov om ændring af lov om erhvervsakademiuddannelser og professionsbacheloruddannelser*. <https://www.ft.dk/samling/20211/lovforslag/159/index.htm>.

Thøgersen, Jacob & Bent Preisler (under udgivelse). *Danskerne og det engelske sprog. Dengang og nu. NyS – Nydanske Sprogstudier* 64.

Lad os bare håbe!

Adverbialer, sætningsskema og tryk

Ole Togeby

Aarhus Universitet

Diderichsens sætningsskema (Diderichsen 1957) er ikke blot en beskrivelse af sætningen og dens led, men semiotisk set et diagram som er andet og mere end en sproglig beskrivelse. Mens en sproglig sætning alene giver information på grundlag af konvention, og billeder alene angiver information på grundlag af analogi, er sætningsskemaet som tegn et diagram som både angiver ved analogi og konvention, som har regler for transformation mellem analoge og konventionelle informationer (fx en målestok eller signaturforklaring), som ikke har noget fokus, og som derfor kan give modtagere der interagerer med diagrammet mere information end afsenderen har lagt ind i det.

Forskellen på effekten af typen af det beskrivende metasprog kan illustreres af følgende eksempel. Kristian Mikkelsen beskriver i sin *Dansk Ordføjningslære* (1911:629) en ledstillingsregel således:

Nogle ord og ordforbindelser har altid eller ofte en mere fremskudt stilling, end vedkommende sætningsled plejer at have, idet de i hovedudsagnssætninger med almindelig ordstilling, der ikke begynder med *gid, bare, blot, måske, kanske*, samt i spørgehovedsætninger eller udråbssætninger, der begynder med *så*, sættes lige efter omsagnet eller hjælpeudsagnsordet, i andre sætninger efter grundledet. Dette er tilfældet med: A) (...). B) 1) svage biord uden hensyn til betydningen, f. Eks. (...) *Det kan da være dig det samme.*(...) *Det kommer der nu ingenting ud af.*

Det er lidt svært helt at forstå hvad denne sprogligt formulerede regel egentlig er en regel om, og hvor generel den er. Når man bruger Diderichsens sætningsskema, kan reglen formuleres således:

Tryksvage dialogpartikler som *da* og *nu* kan kun stå på **a**-pladsen.

Nu er det en grammatisk regel som afspejler realiteter, som man kan antage at sprogbrugerne må kende (uden at kunne formulere dem) og bruge heuristisk til at forstå meningen med en sætning.

I denne artikel skal det undersøges om sætningsskemaerne som diagrammer kan forbedres så de beskriver ledstillingen i dansk videnskabeligt bedre end Diderichsens to skemaer, dvs. mere modsigelsesfrit, udtømmende og simpelt end tidligere udgaver af skemaet – efter Hjelmsevs empiriprincip (Hjelmselev 1966).

Et skema for både helsætninger og ledsætninger

Diderichsen foreslog to skemaer, et for helsætninger med ledstillingen

[F [v n a" [V" N" A"]]]

og et for ledsætninger med ledstillingen

[k [n a" v [V" N" A"]]]

Signaturforklaringen: **n** = plads til nominal (i nominativ hvis det kan ses); **N** = plads til nominal i oblik kasus; **v** = finit verbum; **V** = infinit verbum; **a** = neksusadverbial; **A** = indholdsadverbial; **F**- pladsen (før det finitte verbum i helsætninger) kan være tomt, eller der stå et eksemplar af alle typer af led – undtaget **v**. Det finitte verbum skal altid stå på **v**, som således er skemaets forankringspunkt. **k** = underordningskonjunktional; **X**" = led af typen X kan forekomme flere gange på samme plads.

[F	[v	n	a"	[V"	N"	A"]]]
<i>Hun</i>	<i>ville</i>	<i>(-)</i>	<i>ikke</i>	<i>give</i>	<i>Ib bogen</i>	<i>i går.</i>
<i>Hun</i>	<i>gav</i>	<i>(-)</i>	<i>jo</i>	<i>(-)</i>	<i>Ib bogen</i>	<i>i går.</i>
<i>I går</i>	<i>gav</i>	<i>hun</i>	<i>ikke</i>	<i>(-)</i>	<i>Ib bogen</i>	<i>(-)</i>
<i>(-)</i>	<i>Kom</i>	<i>hun</i>	<i>ikke</i>	<i>(-)</i>	<i>(-)</i> <i>(-)</i>	<i>i går?</i>

[k	[n	a"	v	[V"	N"	A"]]]
<i>Hvis</i>	<i>hun</i>	<i>ikke</i>	<i>ville</i>	<i>give</i>	<i>Ib bogen</i>	<i>i går, ...</i>
<i>Fordi</i>	<i>hun</i>	<i>ikke</i>	<i>gav</i>	<i>(-)</i>	<i>Ib bogen</i>	<i>i går, ...</i>

I skemaerne er der tre typer af tomme pladser, dels en tom plads fordi den aktuelle sætning tilfældigvis ikke har et led af denne type, fx tomt **V''** når det er hovedverbet der er i finit form på **v**, fx *Hun gav ham – bogen*, dels fordi det led der skulle stå der, er placeret i **F**, så det får en særlig mening, fx *Ib gav hun ikke – bogen*, dels når tomt **F** angiver at det er en spørgesætning, fx *– Gav hun ham bogen?* Man må antage at sprogbrugerne bruger kendskab til sådanne regler når de fx forstår at *hun* er subjekt både på **F** og **n**, og *Ib* indirekte objekt, og at en sætning med tomt **F** er et spørgsmål.

Disse to skemaer er (i denne generaliserede og forenkede form), som det kan ses, nok modsigelsesfrie og udtømmende, men de er faktisk ikke den simplest mulige beskrivelse. I stedet for to skemaer kan man nøjes med ét skema (Togebj 2023).

To skemaer:

F	v	n	a''		V''	N''	A''
k		n	a''	v	V''	N''	A''

Et fælles skema:

F	[[v ¹	n]	a''	[v ²	[V''	N''	A'']]]
<i>Hun</i>	<i>ville</i>		<i>ikke</i>		<i>give</i>	<i>Ib bogen</i>	<i>i går.</i>
<i>Hun</i>	<i>gav</i>		<i>jo</i>			<i>Ib bogen</i>	<i>i går.</i>
<i>I går</i>	<i>gav</i>	<i>hun</i>	<i>ikke</i>			<i>Ib bogen</i>	<i>(-)</i>
<i>Kom</i>		<i>hun</i>	<i>ikke</i>				<i>i går?</i>
<i>Hvis</i>		<i>hun</i>	<i>ikke</i>	<i>ville</i>	<i>give</i>	<i>Ib bogen</i>	<i>i går, ...</i>
<i>Fordi</i>		<i>hun</i>	<i>ikke</i>	<i>gav</i>		<i>Ib bogen</i>	<i>i går, ...</i>
<i>Gid</i>		<i>vi</i>	<i>snart</i>	<i>var</i>		<i>hjemme.</i>	
<i>Kommer</i>		<i>hun</i>	<i>ikke?</i>				

Signaturforklaringen for ét diagram bliver så lidt anderledes end for to. Forankringspunktet kan ikke være **v**-pladsen (for det finite verbum kan stå tre forskellige steder); forankringspunktet bliver således **F** (som står for *første felt* og ikke for *forfelt* eller *fundamentfelt*), og denne første plads kan så ikke være tom. Der kan ikke komme noget før **F**, for **F** er defineret som den første plads, og der er ikke nogen **k**-plads (konjunktioner står på **F**).

Dette ene skema giver til gengæld flere oplysninger end de to skemaer, hvor man jo på forhånd skulle vide om den sætning man skal til at analysere, er en helsætning eller en ledsætning. Med det ene fælles skema vil udfyldningen give den oplysning at hvis finittet står på v^1 , er sætningens sandhedsværdi fremsat (som i en helsætning) og hævdet eller benægtet, mens hvis finittet ikke står på v^1 , er sætningen sandhedsværdi ikke fremsat. Står finittet på v^2 , angiver det at sandhedsværdien er neutral (som i en ledsætning), og står finittet på **F**, angiver det en helsætning der er et spørgsmål. Diagrammet giver således som output, når man bruger det, om sætningen har fremsat sandhedsværdi eller ej, uafhængigt af om det er en sætning der er led i en anden sætning eller ej.

Fordelen ved det ene fælles skema er dels at analysen bliver simplere (man skal ikke på forhånd kende sætningens type), og det vil bedre afspejle den viden om ledstilling som sprogbrugerne har og bruger når de forstår en sætning mens den løber, fx at sætningen *Gid vi snart var hjemme* ikke har fremsat sandhedsværdi fordi det finite verbum står på v^2 , selv om den ikke er led i en anden sætning, og at sætningen *Kommer hun ikke?* er et spørgsmål.

Letledspladser

I den forenkede form er sætningsskemaerne ikke en udtømmende beskrivelse; der findes ledstillinger som skemaerne ikke kan tillægge betydning (Diderichsen 1964, Hansen og Heltoft 2001, Heltoft 2011, Togeby 2003). Man kan fx ikke sætte sætningen *Hun gav ham den ikke* i skema, og ikke se at *ham* er indirekte objekt og *den* direkte objekt når de ikke står på **N''**. Derfor må man i skemaet indsætte to **L**-pladser før **a** til at beskrive disse leds rækkefølge:

F	[[v^1 n]	[(L''	a'']	[v^2	[V''	N''	A'']]]
<i>Hun</i>	<i>gav</i>	<i>ham den</i>	<i>ikke</i>				
<i>Hun</i>	<i>gav</i>		<i>ikke</i>			<i>Ib bogen</i>	

Meningen med dem er at lette (tryksvage) pronominer der er indirekte objekt og direkte objekt, står før negationen for ikke at falde i dens virkefelt. Trykstærke objekter er i indholdsfeltet benægtet: *Hun gav ikke Ib bogen, men billedet. *Hun gav ham den ikke, men billedet. Hun gav ham ikke bogen, men billedet. Hun så ikke \ham, men \han så \hende.*

Det er dog også således at letledspladserne ikke findes når der står et infinit verbum i **V''**, eller et finit verbum i v^2 (som i en ledsætning): *Hun ville ikke give ham den. *Hun ville ham den ikke give. Hvis hun ikke gav ham den, så *Hvis hun ham den ikke gav, så ...*

Så det at trykstærke led i indholdsfeltet står i nægtelsens virkefelt, er ikke hele forklaringen. For at forklare letledspladserne må man derudover antage at tryktab altid sker klitisk, dvs. således at tryktabsordet læner sig op ad et trykstærkt ord som både overtager dets tryk og inkorporerer dets mening, enten (kontinuert) enklitisk, *Hun |slog_osis |ikke*, hvor *sig* læner sig op ad *slog* som får betydningen 'føle smerte', eller (eventuelt diskontinuert) proklitisk: *Hun |stod |ikke |op*, hvor *stod* taber sit tryk til *op* så helheden får betydningen 'rejse sig', som er den betydning som benægtes (i modsætning til *Hun |stod |ikke |op*).

Grunden til at objekter der er tryksvage pronominer, står før negationen, er at de ellers ville læne sig enklitisk op ad negationen og således være benægtet. Det sker ikke når **v2** eller **V"** er udfyldt, for så læner pronominerne op ad disse to ord: *Hvis hun ikke gav_ham_den, så ...*

Udfyldte letledspladser angiver altså at objekterne læner sig op ad hovedverbet og kun kan benægtes som en inkorporeret del af dette. Dette betyder at letledspladserne kun er der når negationspladsen er udfyldt, og ellers står lette pronominer, som de skal, på **N"**. Det noteres i skemaet således at tilstedeværelsen af **L"**-pladserne er betinget af udfyldningen af **a**-pladsen, og det noteres med runde parenteser: (**L"**)**a**. Læg mærke til at Diderichsens to skemaer skulle notere at **L"**-pladserne ikke findes i ledsætninger og ikke i helsætninger med sammensat tid. Dette behøver man ikke når **L**-pladserne er betinget af udfyldt **a**.

Lad os bare håbe

Piet Hein har i Politiken 2/4 1950 formuleret følgende GRUK:

Lad os bare håbe, men ikke bare håbe.

De to forekomster af *bare* står begge på **a**-pladsen i det forenklede skema, men de betyder tydeligvis to forskellige ting, nemlig 'lad os uden forbehold håbe, men ikke kun håbe'. Det kan forklares ved at de to ord står på to forskellige pladser i skemaet. Ifølge DDO er kernebetýdningen af *bare* 'ikke andet/andre end X'. Men det får også mening af sin placering, og de to forekomster står på en plads hhv. før og efter negationen, og det er det der giver forskel i mening:

Hun spiste bare ikke maden.

= 'det var ikke andet end tilfældet at hun ikke spiste maden'.

= 'hun spiste simpelthen ikke maden'.

X = deklARATIONEN af udsagnets sandhed, og *bare* ≈ 'simpelthen'.

Hun spiste ikke bare maden.

= 'det var ikke tilfældet at hun ikke spiste andet end maden'.

= 'hun spiste også andet end maden'.

X = det der kommer i det efterfølgende fokus, og *bare* ≈ 'kun', og *ikke bare X* ≈ 'også andet end X'.

I et udvidet skema må man altså dele **a**-pladsen op i tre pladser: **a¹ – neg – a²**, således at skemaet kommer til at se således ud:

F	[[v ¹ n] [(L'') [a ¹ neg a ²]] [v ² [V'' N'' A'']]]
<i>Hun spiste</i>	<i>bare ikke maden</i>
<i>Hun spiste</i>	<i>ikke bare maden</i>

På **a²** har *ikke bare* også et virkefelt (markeret ved understregning) som ordene taber deres tryk til:

Han tog ikke obare |én kage (– men to).

Han tog ikke obare oen |kage (– men også brød og frugt).

Han |tog ikke obare oen |kage (– men spurgte først).

Betydningen af netop *bare* afhænger dog også af andre formelle træk i sætningen end placeringen på **a¹** eller **a²**, nemlig af om det finitte verbum er imperativ eller indikativ, om finittet står på **v¹** eller **v²**, og om *bare* står på **F**. Fx betyder ordet i *Han ventede bare* noget andet end i *Bare vent!* I *Bare jeg går forbi en bager, tager jeg to kilo på* betyder det 'hvis blot'+, og i *Bare han ikke kører galt* 'gid'.

De fem kontekstbetydninger fremgår af følgende skema:

FORM OG PLACERING AF ORDET <i>bare</i>		TYPE	BETYDNING	EKSEMPEL
v i imperativ og <i>bare</i> på F eller a¹		forstærkende sætningsadverbial	'uden forbehold'	<i>Kom bare indenfor!</i> <i>Bare vent!</i>
v i indi- kativ	<i>bare</i> på F med finit på v²	i ledsætning	logisk operator	'hvis blot' ... <i>Bare De ser en pige, bliver De forelsket.</i>
		i helsætning	ønskemarkør	'gid'... <i>Bare han ikke kører galt.</i>
	<i>bare</i> på a¹ , før neg		sætningsadverbial	'simpelthen' <i>Hun spiste bare ikke maden.</i>
	<i>bare</i> på a² , efter neg		logisk operator	'kun' <i>Hun spiste ikke bare maden.</i>

Og da *Lad os* er imperativ, går *bare* i *Lad os bare håbe* betydningen 'uden forbehold', 'under alle omstændigheder' og i *ikke bare håbe* betydningen 'også andet end'.

Lad os ^{a1}*bare* ('uden forbehold') *håbe*, men *ikke* ^{a2}*bare* ('kun') *håbe*.

Meningen med ledstilling

Med denne beskrivelse kan man formulere den generalisering at en bestemt type af adverbier har en kernebetydning (fx *bare*: 'ikke andet/andre end X') som suppleres med kontekstbetydningen, dvs. meningen med placeringen i konteksten (og X = fokus), til den komplette aktuelle mening (fx *ikke bare* = 'også andet end'). På **a¹** modificerer de den sproglige handling og evt. deklarerer af sandheden af udsagnet, mens de på **a²** modificerer det der fokuseres på i sagforholdet:

Han kommer bare ('simpelthen') *ikke* — *sådan er det*.

Han kommer ikke ¹*bare* ('kun') — *han bliver en uge*.

Bare ('hvis blot') *han kommer, går hun*.

Hun var 30 år; han var *oda* ('så vidt jeg ved'+) *ikke 20 år*.

Hun var 30 år; han var ikke ¹*da* ('på den tid') *20 år*.

Da (på den tid hvor=) han var 20 år, var hun 30 år.

Siger du omu (= 'egentlig') alligevel at han er gået.

Siger du alligevel |nu (= 'i dette øjeblik') at han er gået.

Nu ('når') du siger det, kan jeg godt huske det.

Han kom altså (som det fremgår) ikke.

**? Han kom ikke altså (er det det du mener?).*

Altså ('ergo') kom han ikke.

Litteratur

Diderichsen, Paul (1946, 1957) *Elementær Dansk Grammatik*. København: Gyldendal.

Diderichsen, Paul (1964) Sætningsleddene og deres stilling tredive år efter i Paul Diderichsen (1966) *Helhed og struktur*. København: G:E:C: Gads Forlag.

Hansen, Erik og Heltoft, Lars (2011) *Grammatik over det Danske Sprog*. København: Det Danske Sprog- og litteraturselskab.

Heltoft, Lars (2011) Lette pronomeners placering: klise og topologisk integritet i *Ny Forskning i Grammatik 18*, 2011. Odense: Institut for Sprog og Kommunikation, Syddansk Universitet:61-80.

<http://ojss.statsbiblioteket.dk/index.php/nfg/issue/archive>.

Hjelmslev, Louis (1943, 1966) *Omkring Sprogteoriens grundlæggelse*. København: Akademisk Forlag.

Mikkelsen, Kristian. (1911, 1975) *Dansk Ordføjningslære med sproghistoriske tillæg*. København: Hans Reitzels Forlag.

Togeby, Ole (2003) *Fungerer denne sætning?* København: Gads Forlag.

Togeby, Ole (2023) *Sætningsleddene og deres stilling 77 år efter*. København: Universitets-Jubilæets danske Samfund.

Resumé:

***Altså* udtaler og funktioner i DanPASS-korpusset**

Christophe Zerakitsky Vies

Københavns Universitet

Undersøgelser som Heegård & Mortensen (2014) og Heegård (2013, 2015) viser, at et ords udtale til tider korrelerer systematisk med dets funktion i ytring. Hvis man kigger i en dansk ordbog (fx Den Danske Ordbog), præsenteres altså som et enkelt leksem med et særdeles bredt betydningsspænd, og dets brug spænder fra et konkluderende/sammenfattende adverbial til en dialogisk partikel til et turindledende eller præciserende interjektional (Hansen & Heltoft 2019, Heinemann & Steensig 2018). Derudover kan man i DanPASS-korpusset observere, at *altså* udtale varierer fra meget distinkt [ˈalʔsʌ] til stærkt reduceret [a].

Oplægget er baseret på Vies & Heegård Petersen (indsendt). Målet med undersøgelsen er at vurdere, hvorvidt altså stadig kan beskrives som ét polysemt leksem, eller om det er i færd med at splitte sig op i flere tegn. Det gør vi ved at sammenholde udtalen af *altså* med dets funktion i ytringen vha. et reduktionshierarki, der kvantificerer grader af udtalereduktion. I alt blev 428 forekomster af *altså* fra DanPASS analyseret topologisk og funktionelt ud fra den ortografiske transskription, hvorefter vi koblede analysen til den fonetiske realisering fra lydskriftangivelsen i korpusset. Vi sammenlignede dernæst reduktionsgraden for de forskellige funktioner ved hjælp af en ordinal logistisk regression.

Resultaterne indikerer, at udtalen af *altså* systematisk korrelerer med dets funktion i ytringen på en sådan måde, at man kan opstille to overordnede klynger for dets brug: en sætningsekstern og en sætningsintern. Førstnævnte består af interjektionaler i ytringsmarginen eller som parentetiske indskud. Denne gruppe udtales komprimeret, dvs. med færre segmenter. Sidstnævnte består af sætningsadverbier og dialogiske partikler, der udtales relativt ekspanderet, dvs. med flere segmenter.

Litteratur

- DanPASS (u.å). *Udtaleordbog*: www.danpass.hum.ku.dk/links_2014/dictionaries/pronouncing_dictionary_with_references.pdf. (Sidst tilgået 10. december 2021).
- Den Danske Ordbog (DDO). <https://ordnet.dk/ddo>.
- Hansen, Erik & Lars Heltoft (2019) *Grammatik over det danske sprog* (2. udg.). Odense: Syddansk Universitetsforlag.
- Heegård, Jan (2013a) Functional phonetic reduction? The case of the Danish adverb faktisk ('actually'). I Jan Heegård & Peter Juel. Henrichsen (red.) *New Perspectives on Speech in Action. Proceedings from the 2nd SJUSK Conference on Contemporary Speech Habits* (Copenhagen Studies in Language, 43). Frederiksberg: Samfundslitteratur Press: 59–78.
- Heegård Jan (2015) Ordvarighed og grammatisk og pragmatisk funktion. Tilfældet egentlig. *NyS – Nydanske Sprogstudier*, 49, 61–97. DOI: 10.7146/nys.v1i49.22904.
- Heegård, Jan & Janus Mortensen (2014) Fonetisk reduktion og kommunikative kontraster: Tilfældet faktisk. *Ny Forskning i Grammatik*, 21, 87–102. DOI: 10.7146/nfg.v22i21.23564.
- Heinemann, Trine & Jakob Steensig (2018) Justifying departures from progressivity: the Danish turn-initial particle altså. I John Heritage & Marja-Leena Sorjonen (red.) *Between turn and sequence: Turn-initial particles across languages*, 435–465. Amsterdam: John Benjamins. DOI: 10.1075/slsi.31.15hei.
- Zerskitsky Vies, Christophe & Jan Heegård Petersen (2022). "Åha, vi er altså på den, for tegnet har delt sig i to." Om altså's udtaler og funktioner. *NyS, Nydanske Sprogstudier*, 62(1), 46-75. DOI: 10.7146/nys.v1i62.134594

Hvor kommer subjektet fra?

– VP-intern oprindelse som grund til skopus-tvetydighed

Sten Vikner

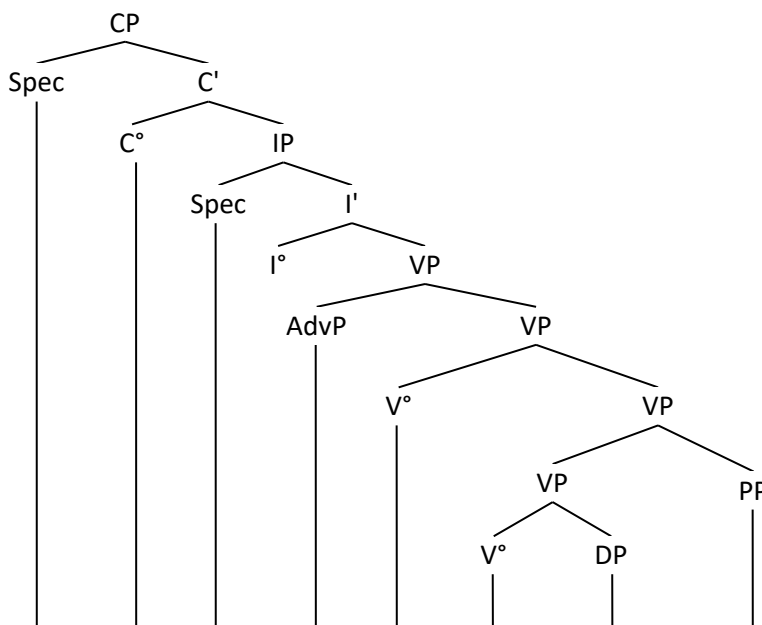
Aarhus Universitet

1. Indledning

Mange analyser af dansk sætningsstruktur går ud fra at et subjekt 'fødes'/basisgenereres i Diderichsens (1946) lille **n** (som vi i den generative lingvistik betegner som specifikator-positionen i IP), hvorefter det så enten forbliver i denne position, (1)c, eller flytter til fundamentfeltet **F** (= specifikator-positionen i CP), (1)d.

(1)c,d er selve de to eksempler, (1)b viser hvordan (1)c,d analyseres ifølge Diderichsen (1946), og (1)a viser en generativ analyse af (1)c,d. For detaljer af disse analyser, se fx Vikner & Jørgensen (2017).

(1) a.



b.	Fund.	Neksusfelt			Indholdsfelt		
	F	v	n	a	V	N	A
c.	<i>Nu</i>	<i>har</i>	<i>Peter</i>	<i>igen</i>	<i>poleret</i>	<i>bilen</i>	<i>med ståluld.</i>
d.	<i>Peter</i>	<i>har</i>		<i>igen</i>	<i>poleret</i>	<i>bilen</i>	<i>med ståluld.</i>

De to analyser i (1) er enige om at enten bliver subjektet stående i subjektpositionen (lille **n** = specifikator-positionen i IP), som i (1)c, eller også flytter det op i topikpositionen (fundamentfeltet **F** = specifikator-positionen i CP), som i (1)d, også selvom Diderichsen (1946) her ikke direkte taler om ”flytning”.

Denne artikel vil imidlertid foreslå en yderligere komplikation, nemlig at selv når subjektet befinder sig i subjektpositionen (lille **n** = specifikator-positionen i IP) i fx (1)c, så er subjektet ikke basisgenereret i denne position. Jeg vil i stedet gøre mig til fortaler for at subjektet altid begynder sin løbebane længere nede i sætningen.

Jeg vil begynde min argumentation med at diskutere en type sætninger hvor der er meget klare tegn på at subjektet kommer fra en position længere nede i sætningen. Passive sætninger viser nemlig at det er muligt for et subjekt at blive basisgenereret i fx en objektsposition.

Derefter vil jeg generalisere denne analyse ved at foreslå at også i aktive sætninger er der fordele ved at antage at et subjekt basisgenereres længere nede i sætningen end i subjektpositionen lille **n** (= specifikator-positionen i IP). Jeg vil argumentere for at i de fleste aktive sætninger basisgenereres subjektet i en position til højre for Diderichsens lille **a** og til venstre for Diderichsens store **V** (altså i den position som i den generative analyse kaldes specifikator-positionen i VP). Fra denne position flytter subjektet så til Diderichsens (1946) lille **n** (= specifikator-positionen i IP).

I afsnit 3-4 vil jeg vise forskellige fordele ved ovenstående analyse, hvoraf kan nævnes


- (afsnit 3) en forklaring på hvorfor subjektet kan befinde sig så langt væk fra det verbum som det får sin tematiske rolle fra - der kan fx stå adskillige hjælpeverber imellem dem (... *da **hun** kunne have **stjålet** min bil ved andre lejligheder*): Subjektet er nødt til at flytte til IP-spec for at få tilskrevet kasus. Dette gælder ikke længere kun subjekter i sætninger med passiv men nu altså alle subjekter.
- (afsnit 4) en forklaring på de tilfælde hvor skopus mellem fx en negation og det kvantificerede subjekt kan forstås på to måder, hvilket inkluderer en forklaring på hvorfor der er så mange eksempler af typen *alle danskere har ikke et gyldigt pas*, (16)a, der ikke har den betydning man skulle forvente. (Man ville forvente at eksemplet betød *Ingen danskere har et gyldigt pas*, men den intenderede betydning er derimod *Det er ikke alle danskere der har et gyldigt pas*.)

Hensigten med artiklen er altså at vise hvorfor det i den generative tilgang antages at subjektet typisk basisgenereres i specifikator-positionen i VP. Hovedvægten ligger på de fordele denne antagelse har for analysen af dansk.

Artiklen henvender sig også til læsere der ikke er så vant til at beskæftige sig med generativ lingvistik, med det formål at vise i detaljer hvad en generativ analyse har at tilbyde specifikt for dansk, og dermed også at vise analyser af danske data som andre tilgange til dansk grammatik også kunne tænkes at byde ind på.

2. Ikke alle subjekter er født i subjektpositionen: Passiv

Lad os begynde med passiv. Her er en aktiv og en passiv sætning:

- (2) a. ... da hun stjal **min bil**. AKTIV
- b. ... da **min bil**₁ blev stjålet *t*₁ (af hende). PASSIV
- 

Den **aktive** sætning, (2)a, har et subjekt, *hun*, og et objekt, *min bil*. Der er også to tematiske roller, AGENS, som *stjæle* tilskrives til sit subjekt, og TEMA, som *stjæle* tilskrives til sit objekt.

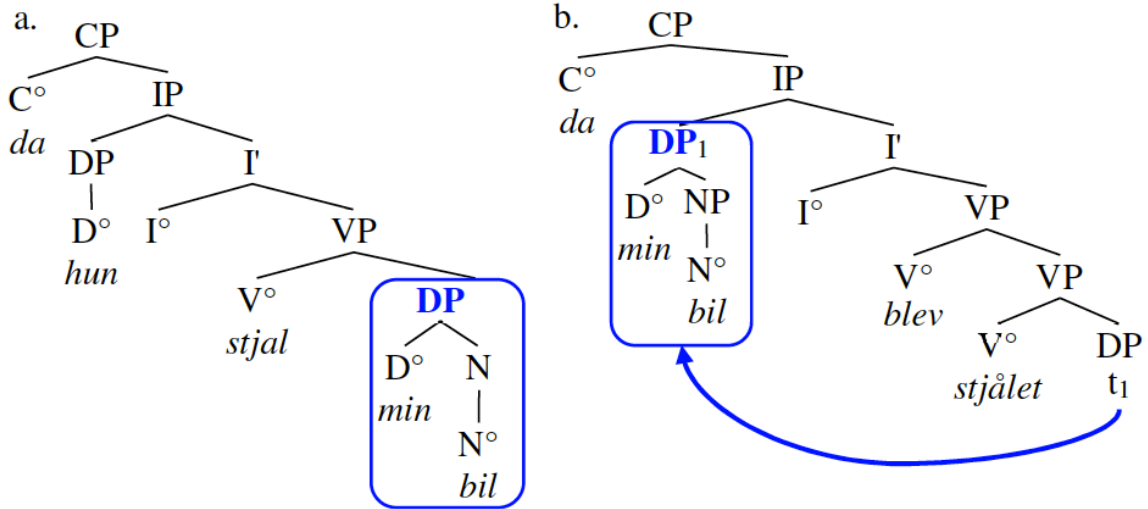
I den **passive** sætning, derimod, (2)b, er *min bil* ikke objekt, men subjekt, også selvom *min bil* stadig har den tematiske rolle TEMA.

2.1 Analysen af passiv-sætninger

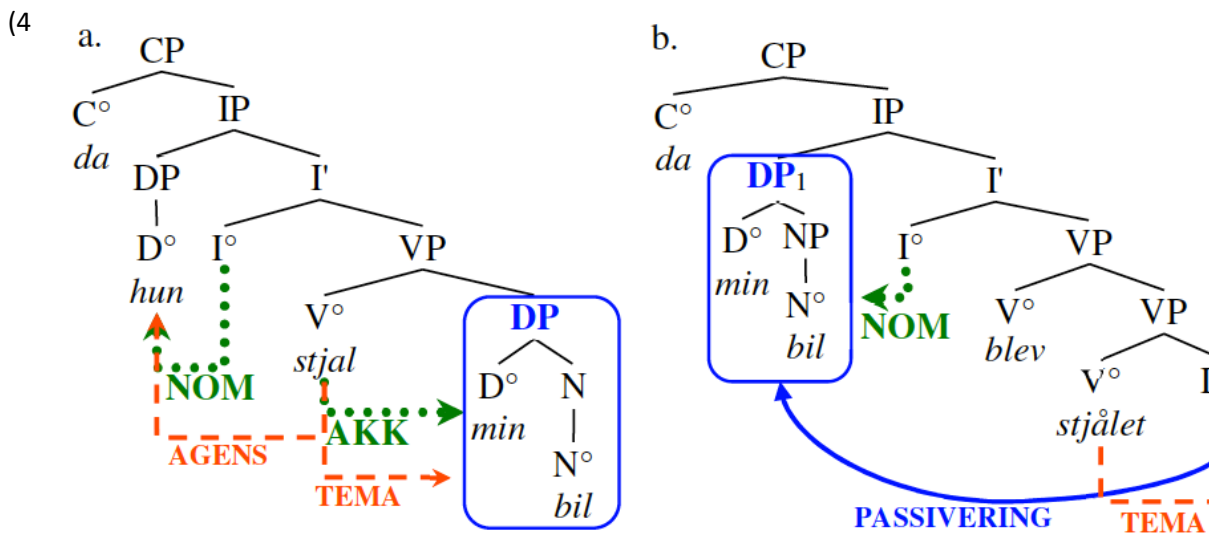
Vi antager derfor at *min bil* er flyttet til subjektpositionen, og at den efterlader et spor i objektpositionen (*t* for "trace"). I (2)b/(3)b er **denne flytning angivet med en blå ubrudt pil**. I passiv kan AGENS udelades eller realiseres i en PP med præpositionen *af*.

Med andre ord: Det der adskiller en **aktiv** fra en **passiv** sætning, er de **grammatiske funktioner** og ikke de tematiske roller. *Min bil* er det der bliver stjålet i både (2)a,b/(3)a,b, og *hun* er den der stjæler i både (2)a,b/(3)a,b - også selvom vi kan vælge at udelade denne omstændighed i passiv, (2)b/(3)b.

(3)



Hvis vi også angiver **tematiske roller (med røde stiplede pile)** og **kasus (med grønne ”prikkede” pile)**, ser det således ud:



- I **aktiv**, (4)a, tilskriver det aktive hovedverbum *stjal* de to tematiske roller AGENS og TEMA, og *stjal* tilskriver også akkusativ kasus til objektet. Nominativ tilskrives til subjektet fra I° (men kun i finitte sætninger).
- I **passiv**, (4)b, tilskriver det passive hovedverbum *stjålet* kun den tematiske rolle TEMA. Den passive form af verbet har mistet både evnen til at tilskrive den tematiske rolle AGENS og

evnen til at tilskrive akkusativ kasus (og de to ting hænger formentlig sammen)¹. Nominativ bliver også her tilskrevet af I^o, det påvirkes altså ikke af passiveringen.

Min bil bliver således tilskrevet akkusativ i **aktiv**, (4)a, jf. *hun stjal dem*, og nominativ i **passiv**, (4)b, jf. *de blev stjålet*. (jeg bruger flertal *de* vs. *dem* her for at få en synlig kasusforskel). Vi antager derfor at flytningen i en passiv sætning er motiveret af kasus, dvs. at DP'en må flytte for at få tilskrevet kasus, idet den ikke får tilskrevet kasus i sin basisposition. I både aktiv og passiv bliver DP'en *min bil* basisgenereret i objektspositionen, og den får tilskrevet sin tematiske rolle (TEMA) sammesteds (DP = determiner phrase ≈ nominalgruppe).

I passiv er det dog lige ved at gå galt: Idet et passivt verbum ikke tilskriver kasus til sit objekt, kan *min bil* ikke overleve i denne position, en DP skal nemlig have tilskrevet kasus. Heldigvis er der i den passive sætning en anden position som både får tilskrevet kasus, og som ikke er besat, nemlig subjektspositionen. Således kan det der i aktiv ville have været et objekt, nu flytte op i subjektspositionen og få tilskrevet kasus der.

2.2 Passivering af perceptionsverber

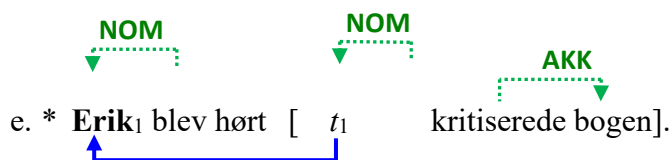
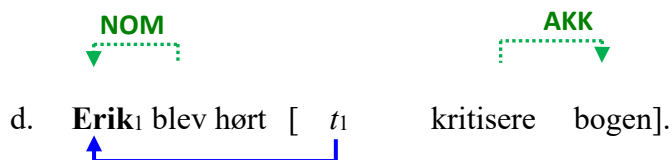
Med denne analyse af passiv kan man også redegøre for hvad der sker når man passiverer perceptions-verber som *høre*, der kan tage en infinitivsætning som komplement (se fx Vikner 2014a):

(5) a. Vi hørte [**Erik** kritisere bogen].

b. * Det blev hørt [**Erik** kritisere bogen].

c. Det blev hørt [at **Erik** kritiserede bogen].

¹ I præpositionspassiver er der en yderligere komplikation, idet det er en præposition der taber evnen til at tilskrive kasus når sætningen bliver passiveret, fx *af* i *Han blev grinet af*, jf. diskussionen i Vikner (2017: 381-383). Upersonlige passiver, fx *Der blev danset hele aftenen*, udgør derimod ikke nogen komplikation: Der er bare ikke nogen DP der i aktiv ville have fået akkusativ tilskrevet fra *danse* og dermed 'mister' kasus i passiv, jf. diskussionen i Vikner (1995: 209-213).



- I aktiv, (5)a, får *Erik*, subjektet i den indlejrede infinitivsætning, akkusativ fra perceptionsverbet *høre*, som er hovedverbet i matrixsætningen. Det at kasus kommer fra et hovedverbum, ikke fra I° , er grunden til at *Erik* får tilskrevet akkusativ (*Erik* kan erstattes med *ham* og ikke med *han*). Det er i sig selv ikke så tit et subjekt får tilskrevet akkusativ (jf. **Dem er gode*), og hertil kommer at *Eriks* kasustilskriver (*høre*) står uden for den sætning der indeholder *Erik*, hvilket også er usædvanligt, deraf navnet ECM ("exceptional case marking"). Det kaldes også ACI ("akkusativ med infinitiv").
- I passiv, (5)b-e, kan det passive verbum *hørt* ikke tilskrive akkusativ, og derfor har *Erik* brug for at få kasus et andet sted fra, hvilket dog ikke sker i (5)b, så derfor er (5)b ugrammatisk.
- I (5)c er den indlejrede sætning finit, så subjektet *Erik* får nominativ fra den indlejrede sætning I° .
- I (5)d er *Erik* flyttet til subjektspositionen i matrixsætningen. Det er kun i en finit sætning at nominativ tilskrives til DP'en i subjektspositionen, så hvis *Erik* ville blive stående i den indlejrede infinitive sætning i (5)d, ville den ikke få kasus. Til gengæld er matrixsætningen finit, så *Erik* kan flytte til subjektspositionen i matrixsætningen og få tilskrevet nominativ fra matrixsætningens I° .
- I (5)e får *Erik* kasus hele to gange, både fra I° i matrixsætningen og fra I° i den indlejrede sætning, idet begge sætninger er finitte. Som man ser, er det at få tilskrevet kasus to gange lige så problematisk som det er ikke at få kasus, og derfor er (5)e ugrammatisk.

2.3 Passivering i idiomatiske udtryk

Et andet fænomen som man kan gøre rede for ved hjælp af denne analyse af passiv, er hvad der sker når man passiverer idiomatiske udtryk.

Den figurative (idiomatiske) læsning af subjektet i passiv er kun mulig i præcis de tilfælde, dvs. (7)a, hvor sådan en læsning af objektet er mulig i aktiv, dvs. (6)a. Det er netop det vi ville vente hvis passiv involverer en flytning til subjektspositionen af den DP der ville have været objekt i aktiv.

- | | | | | |
|-----|----|---|------------------------------------|-------------------------------|
| (6) | a. | ... at Bo tog | skeen i den anden hånd. | IDIOMATISK ELLER BOGSTAVELIGT |
| | b. | ... at Bo tog | gaflen i den anden hånd. | KUN BOGSTAVELIGT |
| | c. | ... at Bo flyttede | skeen til den anden hånd. | KUN BOGSTAVELIGT |
| | | | | |
| (7) | a. | ... at skeen ₁ blev taget | t ₁ i den anden hånd. | IDIOMATISK ELLER BOGSTAVELIGT |
| | b. | ... at gaflen ₁ blev taget | t ₁ i den anden hånd. | KUN BOGSTAVELIGT |
| | c. | ... at skeen ₁ blev flyttet | t ₁ til den anden hånd. | KUN BOGSTAVELIGT |

Vi er under alle omstændigheder nødt til at sige hvorfor (6)a har en idiomatisk læsning når (6)b,c ikke har det. Vi kunne fx sig at en idiomatisk læsning bl.a. kræver både en form af *tage* som verbum og en form af *ske* som objekt. Efterfølgende er spørgsmålet så om vi derudover er nødt til at sige noget mere for også at redegøre for at (7)a har en idiomatisk læsning, og at (7)b,c ikke har det.


Hvis vi antager at passiv involverer en flytning fra objektspositionen til subjektspositionen, så slipper vi for at skulle sige noget særligt om (7)a vs. (7)b,c. Med en flytningsanalyse af passiv vil forskellen (7)a vs. (7)b,c nemlig følge af det man i forvejen siger om den tilsvarende forskel i aktiv, altså (6)a vs. (6)b,c, fordi (7)a er afledt fra (6)a og (7)b,c fra (6)b,c.

2.4 Konklusion på afsnit 2

Det jeg har prøvet at vise i afsnit 2, er altså at der findes i hvert fald nogle subjekter, nemlig subjekter i passiv-eksempler som (8)b, der ikke bliver basisgenereret i subjektspositionen (IP-spec), men som først flytter til subjektspositionen efterfølgende.

- | | | | | |
|-----|----|---|----------------------------|--------|
| (8) | a. | ... da hun | stjal min bil. | AKTIV |
| | b. | ... da min bil ₁ blev stjålet | t ₁ (af hende). | PASSIV |

Det samme kan man vise for raising, som i (9)b (se også Vikner 2023). Subjektet i et raising-eksempel bliver heller ikke basisgenereret i den subjektspostion hvor det forekommer, men flytter først dertil i løbet af derivationen:

- (9) a. ... da det lod til at [**ingen** forstod spørgsmålet]. UDEN RAISING
- b. ... da **ingen**₁ lod til [*t*₁ at forstå spørgsmålet]. MED RAISING
- 

3. Hypotesen om en VP-intern subjektspostition

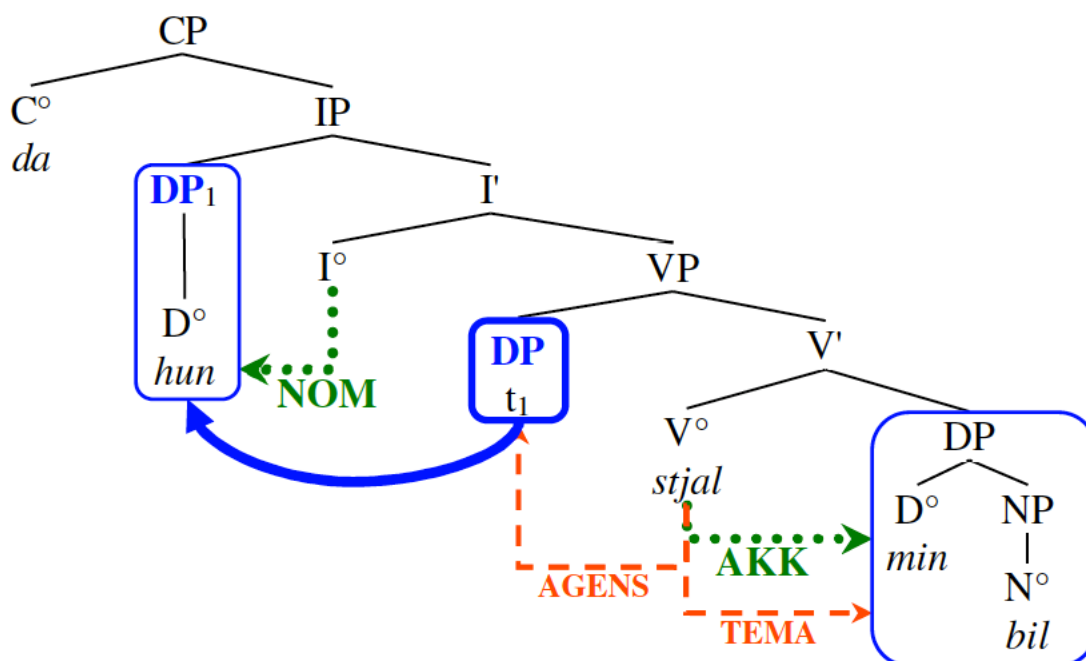
Ovenfor så vi at kasus kan være anledningen til en flytning ind i subjektspostitionen af en DP som har modtaget en tematisk rolle i en anden position – dvs. en DP som er basisgenereret i en anden position. Nu er spørgsmålet så om vi kan være sikre på at dette ikke er tilfældet for alle subjekter, altså om også fx subjekter i simple aktive transitive (og ditransitive) sætninger kan være genereret (og dermed have modtaget deres tematiske rolle) et andet sted før de flyttede til subjektspostitionen.

Denne såkaldte VP-internal subject hypothesis går helt tilbage til Fillmore (1968) og McCawley (1970), men det er først hos Fukui (1986: 55), Sportiche (1988), Koopman and Sportiche (1991), McCloskey (1997) og mange senere analyser at den tager den form som jeg her vil præsentere, og som nu er ret udbredt inden for den generative lingvistik.

Indtil nu er jeg gået ud fra at et hovedverbum tilskriver sine tematiske roller inden for VP'en, med undtagelse af én bestemt rolle, den rolle som subjektet får, og som bliver tilskrevet fra V° helt op til IP-spec. Fra nu af kan vi derimod sige at også fx i en aktiv transitiv sætning som (8)a bliver alle tematiske roller tilskrevet af et hovedverbum inden for dettes egen VP. Tilskrivningen af tematiske roller vises som **røde stiplede pile** i (10): TEMA til objektet i komplementspositionen til V°, AGENS til VP-spec.

Subjektet ses dermed ikke længere som basisgenereret i IP-spec, men derimod i VP-spec (og enkelte gange endda endnu længere nede, som ved passivering og raising). I VP-spec er der imidlertid ingen kasus, og derfor er subjektet nødt til at flytte fra VP-spec (se den **blå ubrudte pil** i (10)) til IP-spec (lidt som i passiv eller raising). I IP-spec er der kasus, idet dette er en subjektspostition i en finit sætning, og nominativ tilskrives fra I° i finitte sætninger. De **grønne ”prikkede” pile** i (10) viser kasus.

(10)



Hvis vi er parate til at betale en vis (beskeden?) pris, nemlig den at alle subjekter (også fx subjekter i aktive transitiver sætninger) basisgenereres et andet sted end subjektspositionen og først efterfølgende flytter til denne subjektsposition (\approx lille **n** = IP-spec), så opnår vi:

- en forklaring på hvorfor subjektet kan befinde sig så langt væk fra det verbum som det får sin tematiske rolle fra – der kan fx stå adskillige hjælpeverber imellem dem (... *da hun kunne have stjålet min bil ved andre lejligheder*). Subjektet er nemlig nødt til at flytte til IP-spec for at få tilskrevet kasus. Dette gælder ikke længere kun passive subjekter som i (8)b, men også aktive subjekter som i (10).
- større ensartethed mellem de forskellige tematiske roller, idet alle tematiske roller nu tilskrives af et hovedverbum inden for dettes egen VP.
- større ensartethed mellem de forskellige fraser (som fx CP, IP, DP, se (1)a), idet vi ikke længere behøver at sige at VP mangler en specifikatorposition i modsætning til CP, IP og DP, idet alle disse fraser nu kan siges at have en specifikatorposition: CP-spec er Diderichsens **F** (1946) fundamentfelt, **F**; IP-spec er Diderichsens lille **n** (1946), se fx Vikner & Jørgensen (2017), og DP-spec er possessorpositionen, se fx Vikner (2014b). Dette er selvfølgelig kun en fordel for så vidt som sådanne fraser overhovedet antages at eksistere, som i Vikner & Jørgensen (2017), Vikner (2014a,b), Platzack (2010), Åfarli & Eide (2003), m.fl.

4. Skopus-tvetydighed mellem negation og kvantificeret subjekt

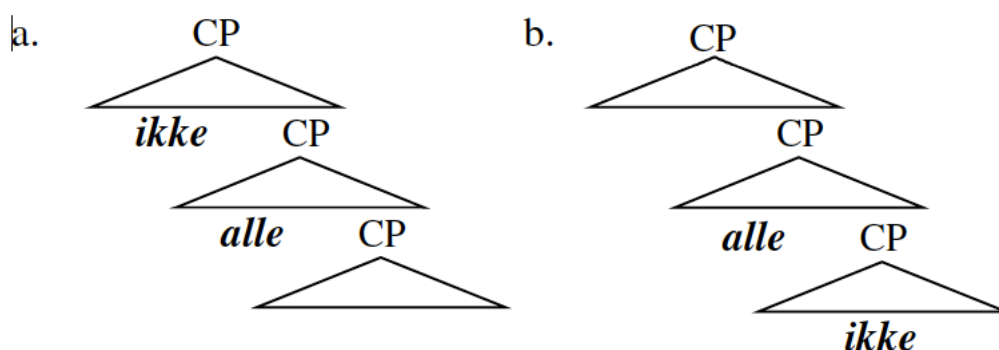
Nu vil jeg så vise endnu et område hvor det kan være fordelagtigt at antage at subjektet basis-genereres i VP-spec. Her skal vi imidlertid begynde et helt andet sted.

Ingen af de to følgende eksempler er tvetydige:

- (11) a. ... at det faktisk **ikke** var tilfældet [_{CP} at det var **alle** branddøre [_{CP} der lukkede]].
 b. ... at det faktisk var tilfældet [_{CP} at det var **alle** branddøre [_{CP} der **ikke** lukkede]].

- I (11)a står sætningsnegationen *ikke* i en sætning der også indeholder sætningen med det kvantificerende udtryk *alle*, og derfor kan betydningen kun være at i hvert fald nogle af branddørene ikke lukkede (*ikke* har skopus over *alle*).²
- I (11)b står det kvantificerende udtryk *alle* i en sætning der også indeholder sætningen med sætningsnegationen *ikke*, og derfor kan betydningen kun være at ingen af branddørene lukkede (*alle* har skopus over *ikke*).

(12)



Disse eksempler viser at *ikke*'s og *alle*'s skopus her kan aflæses direkte af sætningsstrukturen:

- I (11)a/(12)a, hvor kun nogle af branddørene ikke lukkede, har *ikke* skopus over *alle* (*ikke* er i sætningen ovenover sætningen med *alle*).
- I (11)b/(12)b derimod, hvor ingen af branddørene lukkede, har *alle* skopus over *ikke* (*alle* er i sætningen ovenover sætningen med *ikke*).

² I det førstnævnte tilfælde, (11)a, er det selvfølgelig i hvert fald teoretisk muligt at ingen af branddørene lukkede, men det vil i så fald udgøre en overtrædelse af et pragmatisk maksime, for hvorfor sagde taleren så ikke i stedet fx (11)b eller *ingen af branddørene*? Forskellen mellem (11)a og (11)b er at jeg som taler i (11)a kun står inde for at nogle af branddørene ikke lukkede, mens jeg i (11)b står inde for at ingen af dem lukkede.

Derfor er følgende observation angående en formulering i TV-avisen i 1991 ret interessant:

- (13) ”Det blev fastslået at **alle branddøre ikke** lukkede” (TVA 8.7.91).
Guvé hvad det betyder: at ikke alle lukkede, eller at ingen lukkede?
(fra Hansen & Lund 1991)

Den kursiverede sætning i (13) har med andre ord to læsninger, både den der svarer til (11)a (”ikke alle lukkede”), og den der svarer til (11)b (”ingen lukkede”).

Egentlig kunne man måske vente at sådanne sætninger var relativt sjældne, i begge læsninger.

Læsningen hvor *alle* har skopus over *ikke*, (11)b, er faktisk ikke så hyppig, måske fordi man – som Hansen & Lund (1991) gør opmærksom på – kan anvende det utvetydige *ingen* i stedet. Jeg har dog fundet et par stykker:

- (14) a. *Én af udmeldingerne var blandt andet, at frisører landet over blev beordret til at lukke grundet deres tætte kontakt med kunderne. Det betyder, at **alle danskere ikke** har mulighed for at få klippet, plejet og passet håret frem til den 30. marts.*
[https://newsbreak.dk/a-historier/ingen-frisor-sadan-plejer-du-haret-hjemmefra/119939!/,](https://newsbreak.dk/a-historier/ingen-frisor-sadan-plejer-du-haret-hjemmefra/119939!/) 19.03.2020
- b. *Det offentlige behøver ikke at spørge, for **alle instanser** er **ikke** i tvivl om, at samfundets medlemmer vil byde den nye teknik velkommen. Man vil ikke kunne forstå, at man har klaret sig uden.*
KorpusDK, TV-Aktuelt, 1983
- c. *I Saxoforskningen har **alle disse kvindeskikkelser** **ikke** vakt større interesse før i nyeste tid.*
KorpusDK, fagbog, 1984

Intentionerne her er at det gælder for **alle** danskere at de **ikke** kan gå til frisør (under den første coronanedlukning i marts 2020), at det gælder for **alle instanser** at de **ikke** er i tvivl, og at det gælder for **alle disse kvindeskikkelser** at de **ikke** har vakt interesse, alle tre parallelle med (11)b.

Den modsatte læsning af denne type sætninger, altså den hvor *ikke* har skopus over *alle*, som i (11)a, er derimod slet ikke så sjældnen, selvom man måske kunne have forventet det, idet skopus og rækkefølge her ikke passer sammen. Med andre ord, selvom *alle og enhver* i (15) går forud for *ikke*, er læsningen helt klart at *ikke* har skopus over *alle og enhver*:

- (15) **Alle og enhver** har **ikke** adgang til ejendomsregistreringsportalen – det har bl.a. advokater, landinspektører og banker, men ikke privatpersoner.

<https://landmaalercompagniet.dk/landinspektoer-hvad-maa-jeg-selv-goere/>, tilgået 09.10.2022

Intentionen her er ikke at det gælder for **alle og enhver** at de **ikke** har adgang, (11)b, men derimod at det **ikke** gælder for **alle og enhver** at de har adgang, (11)a.

Denne (lidt forvirrende) type eksempler, altså hvor *ikke* har skopus over *alle*, selvom *alle* kommer før *ikke*, findes både som indlejrede sætninger, (13), (16)a-c, og som hovedsætninger, (16)d-j:

- (16) a. *Og det lader til, at det er kommet bag på den styrelse [Digitaliseringsstyrelsen], at **alle danskere ikke** har et gyldigt pas.*

02.12.2020, Ugeavisen Grindsted, Sektion 1, Side 20

- b. *Fx er der indenfor de seneste 4 år solgt 6,2 mio. smartphones. Da **alle danskere ikke** har smartphones, er der altså mange, som har købt flere smartphones i løbet af de seneste fire år.*

<https://soendagaften.dk/2017/10/udlicitering-af-kulturpolitik/>, 02.10.2017

- c. *Generaldirektøren tilføjer imidlertid, at det er en vanskelig opgave at ændre hele organisationens måde at tænke på. ”Jeg er enig i, at **alt ikke** er perfekt, og vi har stadig meget at forbedre”.*

KorpusDK, Politiken, 1999

- d. *En institution som denne kan **alle forældre** selvfølgelig **ikke** nøjes med, men vi er en del der kan.*

KorpusDK, 1989

- e. *I et hierarkisk system som officerernes kan alle jo ikke blive generaler og admiraler.*
KorpusDK, Jyllands-Posten, 1998
- f. *Men selvfølgelig er alle tyske arbejdsgivere ikke ude på at snyde. I de andre job hun har haft i Tyskland har lønnen har været højere og kontrakterne i orden.*
KorpusDK, Sprog & Erhverv, 1999
- g. *Vi mener, at indlæring fungerer bedst, hvis den bærer lønnen i sig selv, d.v.s. imødekommer et behov hos eleven, der drives af sin interesse. Selvfølgelig kan al indlæring ikke baseres på dette princip.*
KorpusDK, 1991
- h. *Per Dalbjerg, formand for Studenterrådet, anbefaler udlandsophold til de, der har lyst, men han advarer også mod at overvurdere betydningen: ”Alle studerende har ikke brug for at tage til udlandet, og et godt studiejob kan også give noget til cv’et.”*
JP Aarhus, 10.10.2013, Sektion 4, Side 8
- i. *Vi skal have fundet en måde at få dem, der er udvist, ud på. Kriminelle gæster i landet skal ud. Enhver har ikke en ret til at komme hertil. Det har de egentligt aldrig haft, selv om man fra politisk hold har syntes at mene det.*
Camilla-Dorthea Bundgaard, Jyllands-Posten, 07.12.2018, Sektion 1, Side 24 (Debat)
- j. *Hvad har Postdanmark tænkt sig med den service, som ikke eksisterer på nordlandet, ja på hele øen? Alle har ikke bil, jeg har kun rollator, og en gang imellem har man da brug for et posthus.*
Bornholms Tidende, 02.09.2014, Sektion 1, Side 6, læserbrev

Det viser sig desuden at ovenstående tvetydighed ikke er specifik for *ikke* når det følger efter *alle/enhver*, idet vi finder en helt tilsvarende situation når *aldrig* følger efter *alle/enhver*. Også her er læsningen hvor *alle* har skopus over *aldrig*, ikke særlig hyppig, men jeg har dog fundet (17). Den modsatte læsning, altså den hvor *aldrig* har skopus over *alle* er derimod mindre sjælden, (18)a-c, selvom det er den hvor skopus og rækkefølge ikke passer sammen. Med andre

ord: Hvor (17) betyder at det er **alle** der **aldrig** ville betale så meget, så betyder fx (18)a at det **aldrig** vil være tilfældet at **alle** bliver mønsterbrydere (og (18)a betyder IKKE at det er **alle** der **aldrig** bliver mønsterbrydere):

(17) *Hvordan reagerede jeres omgangskreds at I ville ud og have sådan en race kat ?? - Hvordan reagerede de, da de fandt ud af hvad de kostede ? ... **Alle** ville **aldrig** give så meget for en kat ...*

<https://kattegade.dk/t/hvordan-reagerede-familien/20052>, sept. 2010

(18) a. ***Alle** vil **aldrig** blive mønsterbrydere under noget som helst system (så ville de jo ikke være mønsterbrydere...), men forudsætningerne for at de opstår er minimaliserede i DK.*

<https://ing.dk/blog/space-x-i-ballerup-120997>, 07.08.2011

b. *Men de seneste år, og især efter Marlon kom, har jeg forsøgt at tvinge mig selv til at være mere ligeglad. **Alle** vil **aldrig** kunne lide mig alligevel, ...*

<https://christinasander.bloggersdelight.dk/2020/01/19/ska-do-nu-til-at-vaere-blaaaarger/>, 19.01.2020

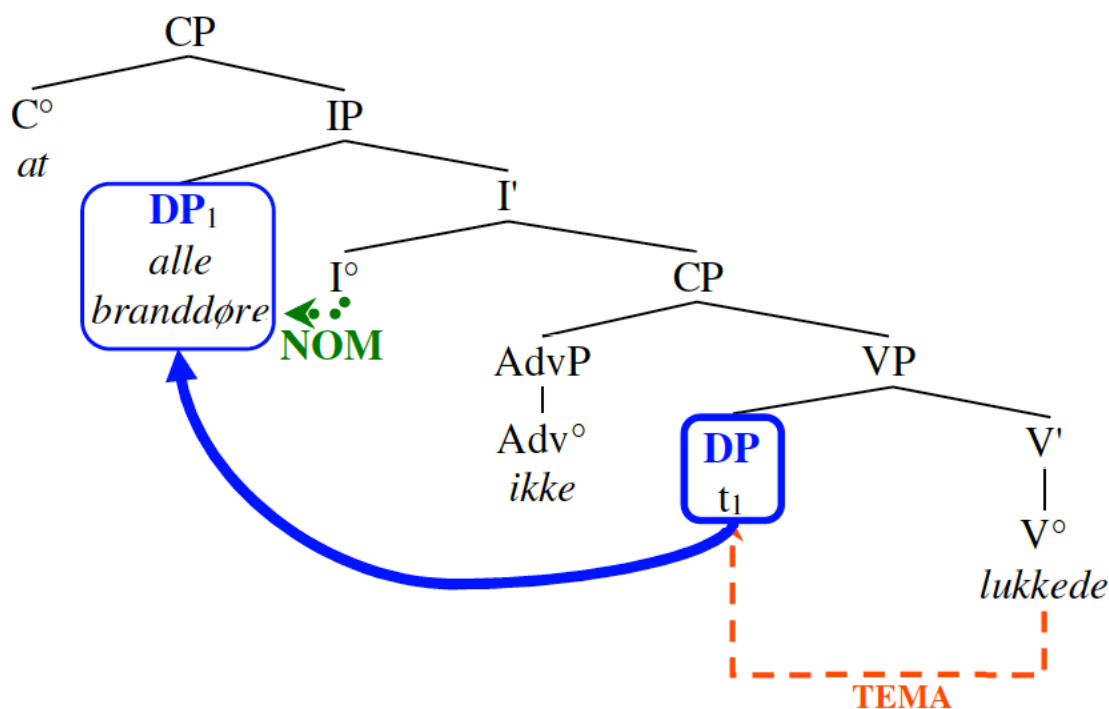
c. *Utrecht er ualmindeligt utilfredse med, at man ikke er blevet tildelt en plads i Europa League efter den hollandske Æresdivision i dag blev aflyst. ... **Alle** vil **aldrig** blive glade når tingene afgøres på skrivebordet.*

https://bold.dk/snak/nyheder/utrecht-protesterer-over-misset-el-plads?room_id=14&thread_id=476435#, 25.04.2020

Vi har altså her set en del eksempler der viser at selve ordstillingen hvor *alle/enhver* efterfølges af *ikke/altid*, er principielt tvetydig, således at man må støtte sig til konteksten for at nå frem til den rette læsning. Hvordan kan vi nu redegøre for denne principielle tvetydighed?

For det første vil jeg foreslå en analyse af de tvetydige eksempler ved hjælp af hypotesen om en VP-intern subjektspostition, som i afsnit 3 ovenfor, således at fx (13) analyseres som følger:

(19)



For det andet vil jeg følge bl.a. Beck & Gergel (2014: 139-141, 191-199) og Radford (2009: 241) i at grunden til at (19) både kan forstås som om *alle* har skopus over *ikke*, og som om *ikke* har skopus over *alle*, er at man **både** kan fortolke (19) som om *alle* står i IP-spec (hvilket ville give *alle* skopus over *ikke*), **og** som om *alle* står i VP-spec (hvilket ville give *ikke* skopus over *alle*).

I eksemplerne i (11)a,b hvor *alle* og *ikke* står i to forskellige sætninger (jf. strukturerne i (12)a,b), forventer vi stadig ikke nogen tvetydigheder, idet *alle* ikke her antages at blive basisgenereret i en position neden for *ikke* og så senere flytte til en anden position oven for *ikke*.

5. Konklusion

Jeg har i denne artikel³ forsøgt at vise at det godt kan betale sig at antage at subjektet ikke er basisgenereret i subjektspositionen (dvs. IP-spec = Diderichsens lille **n** (1946)), men derimod længere nede i sætningen, dvs. for subjektet i de fleste passive sætninger: i objektpositionen, og for subjektet i de fleste aktive sætninger: i en position til højre for Diderichsens lille **a** og til venstre for Diderichsens store **V**, nærmere bestemt i specifikator-positionen i hovedverbets VP.

³ Tak til Ken Ramshøj Christensen, Anna Sofie Hartling, Peter Juel Henriksen, Anne Mette Nyvad, Kathrine Thisted Petersen, Ole Togeby og Carl Vikner, til publikum ved MUDS 19 i oktober 2022 og ikke mindst til de studerende i mine kurser om komparativ 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*.

Derved opnår vi

- (afsnit 3) en forklaring på hvorfor subjektet kan befinde sig så langt væk fra det verbum som det får sin tematiske rolle fra – der kan fx stå adskillige hjælpeverber imellem dem (... *da hun kunne have stjålet min bil ved andre lejligheder*): Subjektet er nødt til at flytte til IP-spec for at få tilskrevet kasus. Dette gælder ikke længere kun subjekter i passive sætninger, men nu altså alle subjekter.
- (afsnit 3) større ensartethed mellem de forskellige tematiske roller, idet alle tematiske roller nu tilskrives af et hovedverbum inden for dettes egen VP.
- (afsnit 3) større ensartethed mellem de forskellige fraser (syntagmer), idet vi ikke længere behøver at sige at VP mangler en specifikatorposition i modsætning til CP, IP og DP, idet alle disse fraser nu kan siges at have en specifikatorposition: CP-spec er Diderichsens **F** (1946) fundamentfelt, **F**; IP-spec er Diderichsens lille **n** (1946), se fx Vikner & Jørgensen (2017), og DP-spec er possessorpositionen, se fx Vikner (2014b).
- (afsnit 4) en forklaring på de tilfælde hvor skopus mellem fx en negation og det kvantificerede subjekt kan forstås på to måder, hvilket inkluderer en forklaring på hvorfor der er så mange eksempler af typen *alle danskere har ikke et gyldigt pas*, (16)a, der ikke har den betydning man skulle forvente. (Man ville forvente at eksemplet betød *Ingen danskere har et gyldigt pas*, men den intenderede betydning er derimod *Det er ikke alle danskere der har et gyldigt pas*.)

Yderligere fordele ved en sådan analyse forsøger jeg at påvise i Vikner (2023).

Litteratur

- Beck, Sigrid & Remus Gergel (2014) *Contrasting English and German Grammar: An Introduction to Syntax and Semantics*. Berlin: De Gruyter Mouton.
- Diderichsen, Paul (1946) *Elementær dansk Grammatik*. Copenhagen: Gyldendal.
- Fillmore, Charles (1968) The Case for Case. I Emmon W. Bach & Robert T. Harms (red.). *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston: 1-88.
- Fukui, Naoki (1986) A Theory of Category Projection and Its Applications. Ph.d.-afhandling, Massachusetts Institute of Technology. <https://dspace.mit.edu/handle/1721.1/15105>.
- Hansen, Erik & Jørn Lund (1991) Sprogbrevet nr. 63. København: Danmarks Radio. <https://sproget.dk/raad-og-regler/artikler-mv/sprogbrevet-dr/sprogbrevetdr-nr-63>.
- Koopman, Hilda & Dominique Sportiche (1991) The position of subjects. *Lingua* 85 (2): 211-58. [https://doi.org/10.1016/0024-3841\(91\)90022-W](https://doi.org/10.1016/0024-3841(91)90022-W).

Hvor kommer subjektet fra?
– VP-intern oprindelse som grund til skopus-tvetydighed

- McCawley, James D (1970) English as a VSO Language. *Language* 46 (2): 286-99.
<https://doi.org/10.2307/412279>.
- McCloskey, Jim (1997) Subjecthood and Subject Positions. I Liliane Haegeman (red.). *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Springer: 197-235. https://doi.org/10.1007/978-94-011-5420-8_5.
- Platzack, Christer (2010) *Den fantastiska grammatiken: en minimalistisk beskrivning av svenskan*. Stockholm: Norstedt.
- Radford, Andrew (2009) *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511801617>.
- Sportiche, Dominique (1988) A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19 (3): 425-49. <https://www.jstor.org/stable/25164903>.
- Vikner, Sten (1995) *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.
- Vikner, Sten (2014a) Kan en konstituent være både subjekt og objekt på samme tid? – om indlejrede infinitiv-sætninger på dansk. I Ole Togeby, Sten Vikner & Henrik Jørgensen (red.). *Problemer og perspektiver i dansk syntaks – med Kristian Mikkelsen som anledning*. Odense: Syddansk Universitetsforlag: 171-191. <https://tildeweb.au.dk/au132769/papers/vikn14a.pdf>.
- Vikner, Sten (2014b) Possessorens status inden for nominalgruppen. *Ny Forskning i Grammatik* 21: 193-211. <https://doi.org/10.7146/nfg.v22i21.23570>.
- Vikner, Sten (2017) Germanic Verb Particle Variation. I Enoch Oladé Aboh, Eric Haeberli, Genoveva Puskás & Manuela Schönenberger (red.). *Elements of Comparative Syntax - Theory and Description*. Berlin: Mouton de Gruyter: 371-398. <https://doi.org/10.1515/9781501504037-014>.
- Vikner, Sten (2023) Hvor kommer subjektet fra? VP-intern oprindelse kan belyse både idiomatiske udtryk og flytninger ud af koordinationer. Submitted. <https://tildeweb.au.dk/au132769/papers/vikn23b.pdf>.
- Vikner, Sten & Henrik Jørgensen (2017) En formel vs. en funktionel tilgang til dansk sætningsstruktur. *Nydanske Sprogstudier - NyS* 52–53: 135–68. <https://doi.org/10.7146/nys.v1i52-53.24954>.
- Åfarli, Tor Anders & Kristin Melum Eide (2003) *Norsk generativ syntaks*. Oslo: Novus.

Jo som dialogisk partikel i tidlig ældre nydansk

Lennart Westergaard

Lunds Universitet

1. Indledning¹

Moderne dansk har en særskilt gruppe ord, de såkaldte dialogiske partikler. Den inkluderer blandt andre *jo*, *sgu*, *skam*, *nu*, *da*, *nok*, *vist* og *vel*. Disse er kendetegnet ved kun at kunne stå i centralfeltet og her at stå før alle øvrige sætningsadverbialer (sml. Hansen & Heltoft 2011: 1037-9). Dette eksemplificeres ved *jo* i følgende konstruerede eksempel:

- (1) Måske er partikelgrammatik *jo* faktisk sjovere, end de fleste tror.

De fleste dialogiske partikler udtrykker betydninger, der vedrører dialogens struktur. I moderne dansk udtrykker *jo* en forventning om, at modtageren “vil være enig (...) og ikke modsige” afsenderen (Hansen & Heltoft 2011:1050; sml. Boeg Thomsen 2015:142-143; Davidsen-Nielsen 1996:285; 290; Engberg-Pedersen 2009:226; Engberg-Pedersen et al. 2019:95; delvist også Harder 1975:107).

Man har også analyseret *jo* som udtryk for, at sætningens proposition er kendt eller i hvert fald indlysende (fx *jo* i *Ordbog over det Danske Sprog* (ODS) og *Den Danske Ordbog* (DDO)), men eksempler som i (2) fra talesprogs-korpusset BySoc illustrerer tydeligt fordelene ved at analysere *jo* som udtryk for, at modtageren ikke forventer at blive modsagt:

- (2) men vi har **jo** også oplevet alt det der med at natmanden er kommen for vi havde **jo** en tante og onkel der havde have på A% % % % og der var vi **jo** tit ude da vi så kom hjem om aftenen så skulle vi **jo** spise lidt
(BySoc, citeret i Christensen 2007:147, tilpasset).

¹ Artiklen har nydt godt af mange kommentarer fra Kaja Borthen, Kasper Boye og Eva Skafte Jensen. Det skylder jeg dem en stor tak for.

Her er der tale om informationer, der kun dårligt kan regnes som kendte eller indlysende for den person, der laver det sociolingvistiske interview. Eksemplerne er derimod helt uproblematisk for en analyse, der argumenterer for, at man med *jo* markerer, at man forventer enighed og ikke forventer at blive modsagt: Kun afsender kan afgøre propositionernes sandhedsværdi og kan dermed med rette forvente ikke at blive modsagt.

2. Modalpartikler historisk

Der er ikke skrevet synderligt meget om, hvornår eller hvordan *jo* har fået denne betydning. I det hele taget er der ikke skrevet meget om de danske dialogiske partiklers historie (men se fx Jensen 2000 og Heltoft 2019a; 2019b om *nok*, *vel* og *vist* og Westergaard 2021 om udviklingen af *nu* og Westergaard 2022 om de dialogiske partiklers rolle for udviklingen af imperativ med subjekt-konstruktionen). Heltoft (2019a; 2019b) daterer opståen af de dialogiske partikler til 1800-tallet. I Westergaard (2021) har jeg argumenteret for, at i hvert fald den dialogiske partikel *nu* er betydeligt ældre. ODS antager, at *jo* med betydningen som dialogiske partikel er et middelnedertysk lån, dvs. det må senest være blevet indlånt i løbet af 1500-tallet. I *Svenska Akademiens Ordbok* mener man også, at *jo* i betydninger, der kunne minde om den moderne dialogiske partikel, er belagt tilbage til fornsvensk.

I min ph.d.-afhandling undersøger jeg udviklingen af en række dialogiske partikler. I middeldansk² finder jeg ingen entydige eksempler på, at *jo* udtrykker en forventning om enighed. I nærværende artikel vil jeg argumentere for, at *jo* i ældre nydansk udtrykker en forventning om enighed, dvs. den moderne partikelfunktion. I afsnit 3 vil jeg meget kort redegøre for mit metodiske afsæt. Dernæst vil jeg i afsnit 4 komme ind på, hvilket materiale jeg har brugt. I afsnit 5 præsenterer jeg *jo*'s betydningspotentiale i middeldansk og ældre nydansk. I afsnit 6 viser jeg, at *jo* havde en betydningsvariant, der svarer til den betydning, den moderne dialogiske partikel udtrykker. Afsnit 7 opsummerer artiklen.

3. Kodet vs. pragmatisk betydning

Når man forsøger at rekonstruere betydninger på et ældre sprogtrin, løber man hurtigt ind i den udfordring, at det ikke altid er helt let at afgøre, om en given funktion er kodet indhold af et tegn eller kan afledes pragmatisk, dvs. som en kontekstuel berigelse af en anden betydning. Et hævdevundet kriterium for kodet betydning er, at betydningen skal optræde i kontekster, hvor alternative betydninger ikke er mulige (sml. Paul [1880]1995:77 og Heines 2002 kontekst-

² Jeg følger Skautrup's (1947:1; 120) periodeinddeling af yngre middeldansk (1350-1500) og ældre nydansk (1500-1700).

typologi). Ideen er, at den betydning, man antager er kodet, i sådanne kontekster ikke kan være afledt af en anden betydning.

For at illustrere denne metode kan man se på konjunktionen *siden* i moderne dansk. Konjunktionen *siden* har en tidslig (3) og en kausal betydning (4). Denne polysemi skyldes, at tidslig rækkefølge ofte implicerer kausalitet (som en *post hoc ergo propter hoc*-fejlslutning) som i (5):

- (3) Jeg har spillet tennis lige siden jeg var ni år. (eksempel fra DDO)
- (4) Derfor er det ikke en bog for børn, siden det vil være svært for dem både at forstå og fordybe sig i teksten. (eksempel fra internettet)
- (5) Han er fri af hoste og har fået en meget bedre vejrtrækning, siden han stoppede med at ryge i sommer. (eksempel fra internettet)

Eksempel (5) alene ville ikke være tilstrækkeligt til at argumentere for, at *siden* faktisk har en kodet kausal betydning: Om end eksemplet kunne tolkes kausalt, kan vi ikke udelukke en tidslig tolkning, og det er nok også den intenderede betydning i eksemplet. I (4) kan man derimod ikke tolke *siden* tidsligt: Årsagen er jo blot tænkt. Den kausale betydning kan derfor ikke være pragmatisk afledt af den tidslige betydning, da selve den pragmatiske afledning jo antages at tage afsæt i en anden kodet betydning. På baggrund af en sætning som (4) kan man derfor konkludere, at *siden* har en kodet kausal betydning.

For at kunne afgøre, om en given betydning er kodet eller ej, kræver det derfor et overblik over, hvilke øvrige betydninger et sprogligt udtryk har eller havde. Dette overblik gives i afsnit 5.

4. Materialet

For at afdække *jo*'s brug i middel- og tidlig ældre nydansk har jeg sammensat et korpus på ca. 3.000.000 løbende ord, som dækker perioden fra ca. 1300 til slutningen af 1500-tallet. Jeg har undgået kilder på vers.

Korpusset³ inkluderer tekster fra:

- Digitale Undersøgelser af Dansk Sprog (DUDS)
- tekstnet.dk (Studér Middelalder på Nettet)

³ Jeg skylder en stor tak til Ida Larsson og *Det Danske Sprog- og Litteraturselskab*, der har gjort de digitale filer bag *Arkiv for Dansk Litteratur* og *Studér Middelalder på Nettet* tilgængelige for mig.

- ADL (især en lang række tekster fra Poul Helgesen og Peter Palladius)
- en række yderligere tekster såsom *Lucidarius*, tingsvidner og brevkorrespondencer blandt adelige i 1500-tallet.

Teksterne er blevet exciperet maskinelt ved hjælp af *AntConc* (Anthony 2022). Ved hjælp af søgemaskinen har jeg søgt på strengene *jo, ju, jw, joo, jor, io, iu, iw, ioo, ior*⁴ (stikprøvevis på *ja* og *ia*). Baggrunden for analysen udgøres af 1240 eksempler (670 eksempler fra regnet former af *ja*).

5. *Jo* i middeldansk og ældre nydansk

I afsnit 5.1. præsenteres først *jo*'s tidlige betydning, dernæst præsenteres de ikke-tidlige betydninger. I min ph.d.-afhandling diskuteres denne del af analysen mere udførligt, men idet hensigten med artiklen er at illustrere, at *jo* har en betydning, der svarer til den betydning, den dialogiske partikel *jo* udtrykker i moderne dansk, vil jeg ikke gå i dybden med denne del af analysen her. Jeg vil heller ikke diskutere kronologien bag disse betydninger.

5.1. Tidlig betydning: 'tidlig uafgrænsethed'

Jo's tidlige betydning kan bruges på en del forskellige måder. Fælles for disse er betydningen 'tidlig uafgrænsethed'.

Den tidlige betydning bruges typisk for at udtrykke 'altid':

- (6) [forud går en beskrivelse af, hvordan en kvinde bør forholde sig, hvis hendes mand har uretmæssigt erhvervet gods]

J al thesse fal estu **jw** pliktogh at vara lydhogh thinum husbonda i edhro rætte hionalagh. 'i alle disse tilfælde er du **altid** pligtig til at adlyde din mand i jeres retfærdige/gode ægteskab'

(SjT, 1425:99).

I forbindelse med udtryk, der betegner højere intensitet, dvs. adjektiver i komparativ eller udtrykt med *mere*, angives det, at denne intensitet får tidlig kontur og bliver ved med at stige svarende til moderne dansk *stadig*:

- (7) Oc HERREN gjorde Salomon **io** mere oc mere ypperlig faar al Jsrael 'og herren gjorde Salomon **stadigt** mere fornem i al Israel'
(Bibel, 1550, 1 Krøn. XXX 25).

⁴ Herfra skal *jo* forstås som en stedfortræder for denne ortografiske variation.

Samme betydning ser man i forbindelsen *jo ... jo*:

- (8) Men **io** mere hand forbød/ **io** mere kundgiorde de det
'Men **desto** mere han det forbød, **desto** mere kundgjorde de det.'
(Alterbogen, 1556:443).

5.2. Ikke-tidslige betydninger

Jo lader i hvert fald til at have de følgende 5 ikke-tidslige varianter i ældre nydansk ud over den moderne partikelfunktion, som behandles i afsnit 6:

1. 'i hvert fald'
2. 'præcis', 'lige'
3. modal nødvendighed
4. emfasemarkør i forbindelse med ønsker og assertioner
5. markør for positiv polaritet (i ledsætninger med negativ polaritet)

En betydningsvariant, der indtager en mellemposition mellem tidslige og ikke-tidslige betydninger, er betydningen 'i hvert fald':

- (9) Oc giff ved din hellig Aand/ at wi kunde foruare oss fra Synd oc bliffue i den wskyldighed/ til huilcken wi komme ved Daaben/ Eller **io**/ om wi falde/ oc kunde icke bliffue bestandige for Menniskens skrøbelighed/ at wi icke bliffue liggende i Synden/ men omuende oss ved ret anger
'og gør det ved din hellige ånd sådan, at vi kunne bevare os fra synd og blive i den uskyldighed, til hvilken vi er kommet gennem dåben, eller **i hvert fald**, hvis vi ikke skulle kunne stå imod på grund af menneskelig skrøbelighed, sådan, at vi ikke forbliver liggende i synden, men omvender os ved ret anger.'
(Alterbogen, 1556:349).

Det er svært at afgøre, om man skal regne med en separat kodet betydning 'i alle fald', eller om denne betydning kan betragtes som værende del af betydningen 'tidslig uafgrænsethed', men i hvert fald i (8) er det svært at holde fast i en tidslig tolkning af *jo*. Dette taler for at betragte betydning 'i hvert fald' som en selvstændig betydning.

I (10) modificerer *jo* frasen *saa slemme som synden er selff* med betydningen 'præcist' eller 'lige':

- (10) Sliige mange seeder oc logher findis i mange land, ther ere **iw** saa slemme som synden er selff
'Så mange sæder og love findes i mange lande, der er **præcist/lige** så slemme som synden selv.'
(En kristen Fyrstes Lære, 1534:222)

Jo kunne i ældre nydansk også udtrykke en **modal 'nødvendighed'**. Hvilken type nødvendighed, der er tale om, afgøres af konteksten. I (11)-(14) er det modalverbets modalitetstype, der er med til at entydiggøre *jo*'s modalitet. I (11) er modalitetstypen dynamisk, dvs. det er et kausalt forhold i verden (spedalskheden), der nødvendiggør sagforholdets realisering. I (12) er modalitetstypen deontisk, dvs. at nødvendigheden stammer fra nogen andres vilje end subjektets. I (13) er modalitetstypen volitiv, dvs. at nødvendigheden stammer fra subjektets egen vilje. I (14) er modalitetstypen epistemisk, dvs. *jo* markerer, at afsenderen er sikker på, at propositionen er sand:

- (11) Tha swarathe thæn spitalske brodherin at han vilde hældir døø. keysaren saghdhe at han skulde thæt **jo** ændeliga giøra.
'Da svarede den spedalske broder, at han hellere ville dø. Kejseren sagde, at han **nødvendigvis** skulle gøre det.'
(SjT 1425:72).
- (12) Der som ieg **io** skal rose mig / da vil ieg rose mig aff min skrøbelighed.
'Da/hvis jeg **nødvendigvis** skal rose mig, da vil jeg rose mig for min skrøbelighed.'
(Bibelbog, 1550, 2 Kor XI 30).
- (13) Men dersom du vilt **io** gaa fra huße till helligdom, da gach her heden til Haggsted der skalt du finde den fattige gammel sogneprest her lens paa sin sotteseng.
'Da/hvis du **absolut** vil gå fra hus til helligdommen, da gå her hen til Haggsted. Der vil du finde den fattige gamle sogneprest, herre Jens, på sit sygeleje.'
(Visitatsbogen, 1543:132).
- (14) Och thror jeg well, att y **jo** well kiennder hanns forellder
'Og jeg tror godt, at I **sikkert** godt kender hans forældre'
(BreveKGBB, 1578:292).

Måske med undtagelse af den epistemiske betydning lader nødvendighedsbetydningen i særlig høj grad til at optræde med modalverber.

Jo kunne også bruges som emfase­markør i forbindelse med ønsker som i imperativen i (15), i ønske­ledsætninger som i (16) og i assertioner som i (17):

(15) Den gamle Mand sagde / Fred vere met dig / alt huad som dig fattis det skalt du finde hoss mig / bliff **io** icke i nat paa gaden.

Den gamle mand sagde: Fred være med dig. Alt det, du mangler, skal du finde hos mig. Men bliv **bare/endelig** ikke på gaden i nat.

(Chr. III.-bibel, 1550, 2. Makk VII 30).

(16) Therfor, kierre her Monnss, er mynn flitig bønn tiil etther, y wilie wel giøre och fli thet saa, at thenne betaling matte **jo** wiis komme tiil stede for winterenn. 'Derfor, kære herre Mogens, er min ivrige bøn til jer, at I vil gøre det rigtige og sørge for, at denne betaling **endelig** med sikkerhed ankommer før vinteren.'

(BreveMGAS, 1565:248).

(17) O gode Gud, ieg bekiender **io** at mine Gierninger eller fortieniste, icke kand vdslette mine Synder.

'Åh gode Gud, jeg bekender **bestemt**, at mine gerninger eller fortjenster ikke kan udslætte mine synder.'

(Bønner, 1577:36).

I disse ytringer understreger taleren aspekter ved illokutionen, enten at afsenderen ønsker sag­forholdets realisering⁵, eller at afsenderen står inde for den proposition, der fremsættes.

Endelig kan *jo* annullere negativ polaritet i ledsætninger. Hermed mener jeg ledsætninger, der er negeret i kraft af et negativt element i oversættelsen som *aldrig* i (18):

(18) dagen ehr aldrig saa lang, at afftenen **io** kommer paa

'Dagen er aldrig så lang, at aftenen ikke kommer.'

(Visitatsbogen, 1543:39).

Ledsætningen fremsættes som ikke-negeret på trods af den negative kontekst: 'aftenen kommer'. I moderne dansk skulle man have en dobbelt-negation for at få samme semantiske effekt som illustreret i oversættelsen.

⁵ Ligesom de fleste andre betydninger, *jo* kunne udtrykke i middeldansk og ældre nydansk, er denne betydning udgået i moderne (rigs-)dansk. Betydningen findes dog stadig i en række dialekter som fx sønderjysk. Jysk Ordbog illustrerer betydningen med følgende eksempler:

(a) kom 'juø i jæn (Bjerrum 1931-1935, som citeret i Jysk Ordbog)

'kom **endelig** igen (en anden gang)'

(b) Håws **jo**', å sæt e Pråp o æ Fla-sk, æjsøn fådoftør æ Sprit (Ottosen 1961, som citeret i Jysk Ordbog)

'husk **endelig** at sætte proppen i flasken, ellers fordufter spritten.'

6. Belæg for kodet status af betydningen 'forventet enighed'

Som påpeget i afsnit 3 kræver et godt belæg, at man kan udelukke andre konkurrerende betydninger – altså at belægget er entydigt. Med afsæt i den korte gennemgang i det foregående afsnit er det nu muligt at karakterisere de træk, en kontekst for sådan et belæg skal have. På baggrund af disse træk vil jeg argumentere for, at *jo* i ældre nydansk havde en betydning svarende til den moderne partikelbetydning.

6.1. En idealkontekst for betydningen 'forventet enighed'

Betydningen 'altid' kan udelukkes i kontekster, hvor afsenderen refererer til enkeltstående begivenhed. Dette kan illustreres med følgende eksempel:

(19) Gud skabte altid verden.

En ytring som (19) er ikke nødvendigvis ugrammatisk: Betydningen 'altid' kan i en sådan kontekst fremtvinge en iterativ eller habituel læsning af *skabe*. Ytringen er dog påfaldende, da den begivenhed, der refereres til, er en enkeltstående begivenhed.

Tolkninger, der kan udløses af betydningen 'i hvert fald', kræver en proposition, som den proposition, der fremsættes, står i opposition til. En ytring som (20) ud af det blå vil derfor være påfaldende, da der netop ikke er nogen kontrasterende proposition:

(20) Gud findes i alle fald.

Kan man ikke finde en kontrasterende proposition i sætningens kontekst, kan man altså også udelukke denne betydning.

En betydning, man let kunne forveksle *jo*'s betydning med, er epistemisk nødvendighed. Paradoksalt nok lader tegn, der betegner epistemisk nødvendighed, altid til at implicere en anelse tvivl (sml. Brandt 1999:28). Dette er grunden til, at man ville studse, hvis man hørte paven ytre (21):

(21) Gud findes sikkert.

I en sådan ytring viser afsenderen en større grad af usikkerhed end i en ytring uden modalitetsmarkør som (22):

(22) Gud findes.

Om usikkerheden, som den kommer til udtryk i sætninger som (21), er kodet eller en pragmatiske slutning, er sekundært. Det væsentlige er, at markører for epistemisk nødvendighed forudsætter en grad af usikkerhed. I en kontekst, hvor man ikke kan antage, at afsenderen tillader det mindste tvivl, kan man derfor udelukke betydningen 'epistemisk nødvendighed'.

Andre former for modal nødvendighed lader som påpeget i høj grad til at korrelere med forekomsten af modalverber. Indeholder sætningen ikke noget modalverb, er en sådan tolkning som minimum usandsynlig. Den kan usandsynliggøres yderligere i kontekster, hvor man ikke med rimelighed kan antage, at der er noget, der nødvendiggør sagforholdets realisering. Ved et sagforhold som det, der udtrykkes i (22), kan man fx ikke umiddelbart finde en vilje eller et kausalt forhold, der nødvendiggør sagforholdets realisering. Det kan dog være svært at udelukke den generelle modale nødvendighed, da man meget ofte kan tænke sig til, at der er nogen eller noget, der nødvendiggør sagforholdets realisering. Hvis et sådant forhold ikke er fremtrædende i informationsstrukturen, vil jeg derfor antage, at den modale nødvendighedsbetydning er en usandsynlig læsning.

Jo's moderne partikelfunktion kan også forveksles med emfasebetydningen. Den første variant af emfasebetydningen, nemlig den, der intensiverer et ønske, optræder kun i sætninger, der udtrykker ønskede sagforhold, som er en kontekst, hvor den moderne partikelbetydning er udelukket. Der vil derfor aldrig være fare for forveksling. Hvad angår emfasebetydningen i forbindelse med assertioner, er der to konteksttræk, der kræves for at usandsynliggøre sådan en læsning: På den ene side kræver emfatiske ytringer som (23), at der er en grund til, at man understreger, at man står inde for den proposition, man fremsætter, fx at modtageren ikke tror på propositionen, men har et modsatrettet synspunkt eller lignende:

(23) Gud findes sgu.

Derfor kan denne brug af emfase-betydningen udelukkes i kontekster, hvor modtageren ikke kan antages at være i tvivl om sandheden af propositionen.

På den anden side kan udtryk for emfase også bruges for at udtrykke overraskelse. Hvis jeg fx fik en åbenbaring, ville jeg også kunne ytre (23). Afsenderen må derfor heller ikke antages at kunne være overrasket over propositionen, hvis emfasebetydningen skal kunne udelukkes.

Der er ingen forvekslingsfare med betydningen 'præcis', da den kun optræder i umiddelbar nærhed af det led, der modificeres (sml. (10)). Den positive polaritetsmarkør optræder udelukkende i ledsætninger med negativ polaritet. Dette er heller ikke en kontekst, hvor vi ville forvente det moderne *jo*, og der er derfor heller ingen forvekslingsfare her.

Idealkonteksten for et godt belæg på den moderne partikelfunktion har altså følgende træk:

- der refereres til enkeltstående begivenheder;
- der er ingen proposition, ytringen står i opposition til;

- der er ingen tvivl om propositionens sandhed hos afsender eller modtager;
- der er ikke noget modalverbum eller entydigt forhold, der nødvendiggør sagforholdets realisering;
- afsenderen er ikke overrasket over propositionen;
- der refereres til en real proposition; og
- *jo* optræder ikke i en ledsætning med negativ polaritet.

Konteksten skal selvfølgelig også tillade en tolkning af *jo* som udtryk for forventet enighed.

6.2. Eksemplificering

Med de nævnte konteksttræk ved hånden er det nu muligt at gennemgå nogle eksempler på *jo* med en funktion svarende til den moderne dialogiske partikel. Eksempler som (24) begynder at optræde omkring 1500-tallet:

(24) Du maat iche thie met nogen troldqvinde, di faae nu deris rette løn, de kunde nu iche lenger bare dem udi denne klare evangelij liuþ oc dag, di faa nu en verdens skam aff verden med dem, det er och dieris fortiente løn.

De brende **jo** en hob tilforne aff dem udj Malmøe, udj Kiøge, och anderstedze, och hører vi thill at der sidder atther en hob greben i Malmøe och schulle brendes.

'Du må ikke skjule troldkvinder. De får nu deres rette løn. De kan nu ikke længere opholde sig i dette klare evangeliske lys og denne dag. De får nu verdens skam. Af verden med dem! Det er deres fortjente løn.

De brændte **jo** en flok af dem for nylig i Malmø, i Køge og andre steder. Og vi hører, at der atter sidder en flok fanget i Malmø, som skal brændes'

(Visitatsbogen, 1543:110).

I (24) er der reference til en enkelte fortidige begivenhed (vs. 'altid'), en række specifikke heksebrændinger, som vi med en vis sikkerhed kan antage, hverken afsender eller modtager var i tvivl om (vs. epistemisk nødvendighed og emfase). Begivenheden er real (vs. 'emfase af ønsker'). Hekseafbrændingerne er ikke sket på trods af noget, der er givet i teksten (vs. 'i alle fald').

Sagforholdet er i høj grad et sagforhold, der kan være blevet nødvendiggjort af nogens vilje, men dette passer ikke umiddelbart i tekstens informationsstruktur (sml. en parafrase med *de skulle brænde en hob tidligere* i samme passage), og derudover er der ikke noget modalverbum, som denne betydning ellers plejer at optræde sammen med.

(25) er fra biblen, hvor det bruges til at oversætte Luthers *ja*. Jeg mener, at begge *jo*'er er eksempler på betydningen 'forventet enighed', men kun den anden forekomst lever op til ideal-kontekst-kravene:

(25) Men ieg er HERREN din Gud aff Egypti land / Oc du skulde io ingen anden Gud kende vden mig / oc ingen Frelsere / vden mig alene. Jeg tog mig dig **io** til i Ørcken vdi det tørre Land.

'Men jeg er herren, din Gud fra Ægypten. Og du skulle jo ikke have andre Guder end mig og heller ikke andre frelsere end mig alene. Jeg tog mig **jo** af dig/indgik en pagt med dig i ørkenen i det tørre land.'

(Bibel, 1550, Hos. 13).

Igen er der reference til en enkeltstående og en i tekstuniverset real begivenhed. Endvidere kan Gud næppe antages at være i tvivl om eller overrasket over, at han tog sig af eller indgik en pagt med israelitterne, og det samme gælder israelitterne. Der er ligeledes ingen modalfaktor, der nødvendiggør Guds handling. Derudover er der ingen åbenlys proposition, som *jo*-sætningens propositioner står i modsætning til.

Eksempel (26) er fra en bøn, hvor Gud mindes om løftet om at tilgive synderne, hvis synderen angreder:

(26) Du Gud est allene Gud/ alene sandru/ men alle Menniske ere Løgnere. Du haffuer **io** sagt/ Paa huilken dag/ en Syndere sucker oc giffuer sig/ oc det angreder hannem/ at hand haffuer syndet/ da vilt du icke mere komme hans synd ihu/

'Du, Gud, er den eneste Gud, den eneste sandhed, mens alle mennesker er løgnere. Du har **jo** sagt: På den dag, hvor en synder sukker, erkender og angreder, at han har syndet, vil du glemme hans synd.'

(SUM, 1551:167).

Også her er der reference til en enkel, real begivenhed, hverken afsender eller modtager (Gud) kan antages at være i tvivl om sagforholdet, det er ikke nærliggende at betragte sagforholdet som nødvendiggjort, og der er ingen proposition, som står i opposition til påmindelsen.

Jeg vil afrunde denne kvalitative argumentation med et kvantitativt argument, der taler for den præsenterede analyse. I middeldansk finder jeg som påpeget ikke gode eksempler på *jo* som udtryk for den moderne partikelfunktion. Sammenligner man den tekst fra middeldansk, hvor *jo* optræder hyppigst, nemlig *Sjælens Trøst* (1425), med en tekst fra ældre nydansk, nemlig *Visitatsbogen* (1543), kan man se, at *jo* (uanset betydningsvariant) har en næsten ti gange så høj

frekvens i sidstnævnte: 16 ud af 44.227 tokens (0,036 %) i *Sjælens Trøst* over for 97 ud af 41.129 tokens (0,236 %) i *Visitatsbogen*. Dette er i overensstemmelse med, at den betydning, den dialogiske partikel *jo* udtrykker, har det med at være meget frekvent (sml. fx Christensen 2012:7 f. for *jo*'s frekvens i moderne dansk).

7. Opsummering

Målet med denne artikel var at argumentere for, at *jo* i ældre nydansk havde betydninger, der svarer til *jo*'s moderne partikelfunktion. På baggrund af en gennemgang af de betydningsvarianter, *jo* har i ældre nydansk, har jeg udarbejdet, hvilke træk en idealkontekst for den moderne partikelfunktion har. Ved at finde ytringer, der svarer til en sådan idealkontekst, har jeg argumenteret for, at man med en vis sikkerhed kan antage, at *jo* havde en betydningsvariant i ældre nydansk, der svarer til den moderne danske partikelfunktion.

Kilder

- Alterbogen = *Alterbogen*. 1556 I: Palladius, Peder: *Peder Palladius' Danske Skrifter*. Udg. af Lis Jacobsen, H.H. Thiele, 1911-1926.
- BreveMGAS = Marquard, E. (udg.) *Breve til og fra Mogens Gyldenstjerne og Anne Sparre*. København: Gad. 1936.
- BreveKGBB = Bang, Gustav (udg.) 1898-1899: *Breve til og fra Kristoffer Gøje og Birgitte Bølle*.
- Bibel = *Biblia, Det er den gantske Hellige Scrifft, vdsæt paa Danske*. Udg. i faksimile 1928. Udg. 2018 af Marita Akhøj Nielsen, netpub. på *dsl.dk*.
- Bønner = Hans Christensen Sthen *Christelige oc vdkaarne Bøner, for alle Stater, oc for allehaande nød oc trang i den gantske Christenhed, vddelte at bede, paa huer besynderlig dag vdi Vgen, met Tacksigelser, oc Guds paakaldelse, om Velsignelse oc Beuaring Morgen oc Afften. Tilhobe sæt, aff den hellige Scrifft, ved D. Iohannem Auenarium. Oc nu nylige fordanskede, oc igien paa ny corrigerede, aff Hans Christensøn Roschildense, Predickere oc Guds ords Tienere vdi Helsingør*. 1577. I: Hans Christensen Sthens Skrifter. Jens Lyster (udg., under medvirken af Jens Højgård). København: Det Danske Sprog- og Litteraturselskab, 2003, Bd. 2, s. 14-149
- En kristen fyrstes lære = Helgesen, Poul *Skrifter af Paulus Helie* 5. Bind, ved Marius Kristensen. 1935. Udgivet af Det Danske Sprog- og Litteraturselskab.
- Gavebreve = 3 *Gavebreve fra Dronning Margrethe fra året 1411*. I: Kirkehistoriske Samlinger. *Tredie række, Udgivne af Selskabet for danmarks kirkehistorie*. Udg. af Holger Rørdam. 3. Bd. København: Gad (Thieles bogtrykkeri). 1881-82.
- SjT = *Sjælens Trøst*, Udg. for Universitets-Jubilæets Danske Samfund v. Niels Nielsen, J. H. Schultz Forlag, Kbh. 1937
- SUM = Oversættelse af Savonarola's Udlæggelse af Miserere 1551. I: Palladius, Peder: *Peder Palladius' Danske Skrifter*, udg. af Lis Jacobsen, H.H. Thiele, 1911-1926

Visitatsbog = Peder Palladius' Visitatsbog. Danmarks folkeminder 30. Udg. af Lis Jacobsen. København: H. H. Thieles Bogtrykkeri. 1911-1926.

Litteratur

- Anthony, Laurence (2022) *AntConc* (Version 4.1.2) [Computer Software]. Tokyo, Japan: Waseda University. tilgængelig på <https://www.laurenceanthony.net/software>
- Boeg, Thomsen, Ditte (2015) Én stavelse, to synsvinkler. Børns tilegnelse af dialogiske partikler. I *Nydanske Sprogstudier* 48:138-163.
- Bjerrum, Anders (1931-1935) Optegnelser. Seddelsamling i Afdeling for Dialektforskning, Nordisk Forskningsinstitut. Københavns Universitet.
- Brandt, Søren (1999) *Modal Verbs in Danish*. København: C A Reitzel.
- Christensen, Tanya Karoli (2007) *Hyperparadigmer – en undersøgelse af paradigmatiske samspil i danske modus-systemer*. Ph.d.-afhandling. Roskilde: Roskilde Universitetscenter.
- Christensen, Tanya Karoli (2012) Epistemiske sætningsadverbialer i dansk talesprog. I *Ny forskning i grammatik* 19:5-28.
- Daidsen-Nielsen, Niels (1996) Discourse particles in Danish. I Elisabeth Engberg-Pedersen, Michael Fortescue, Peter Harder, Lisbeth Falster Jakobsen, Lars Heltoft (red.) *Content, Expression and Structure. Studies in Danish Functional Grammar*. Amsterdam/Philadelphia: John Benjamins: 283 ff.
- Engberg-Pedersen, Elisabeth (2009) Det er *jo* Lisbeths fødselsdag! - om nytten af instrukssemantik. I Ken Farø, Alexandra Holsting, Niels-Erik Larsen, Jens Erik Mogensen og Thora Vinther (red.), *Sprogvidenskab i glimt: 70 tekster om sprog i teori og praksis*. Odense: Syddansk Universitetsforlag: 225 ff.
- Engberg-Pedersen, Elisabeth, Kasper Boye & Peter Harder (2019) *Semantik*. Samfundslitteratur.
- Hansen, Erik, & Heltoft, Lars (2011). *Grammatik over det danske sprog*. Syddansk Universitetsforlag.
- Harder, Peter (1975) Prædikatsstruktur og kommunikativ funktion. *Nydanske Sprogstudier* 8, 102-112.
- Heine, B. (2002): On the role of context in grammaticalization. I: New Reflections on Grammaticalization. I. Wischer & G. Diewald (red.). Amsterdam: John Benjamins, 83-101.
- Heltoft, Lars (2019a) ”Syntaks” i Hjorth, E., Jacobsen, H. G., Jørgensen, B., Jacobsen, B., Jørgensen, M. K., & Fahl, L. K. (2019). *Dansk Sproghistorie: Bøjning og bygning*. Aarhus Universitetsforlag.
- Heltoft, Lars (2019b) ”Dialogiske partikler” supplement til Hjorth, E., Jacobsen, H. G., Jørgensen, B., Jacobsen, B., Jørgensen, M. K., & Fahl, L. K. (2019). *Dansk Sproghistorie: Bøjning og bygning*. Aarhus Universitetsforlag.
- Jensen, Eva Skafte (2000a) *Danske sætningsadverbialer og topologi i diakron belysning*. Ph.d.-afhandling ved Københavns Universitet.
- Ottsen, Mathias Bendix (1961) *Hostrup-Dansk. Lydlære*. (Udg. Ella Jensen). Udvalg for Folkemaals Publikationer. Serie A nr. 20. København: Københavns Universitet.
- Paul, H. ([1880]1995) *Prinzipien der Sprachgeschichte*. Tübingen: Max Niemeyer Verlag.
- Skautrup, P. (1947): Det danske sprogs historie II. København: Gyldendal.
- Westergaard, Lennart (2021) Grammatikalisering af *nu*. Fra tidlig kontrast til polyfoni, *Ny forskning i grammatik* 28, 175-194.

Westergaard, Lennart (2022) Saa tag du bort den daarlighed, huilcken der staar i de vnge Menniskis Hierter:
Udviklingen af imperativ med subjekt-konstruktionen. *Ny Forskning I Grammatik* 29, 174–194.

Det Centrale Ordregister (COR): Hvad, hvor, hvorfor og hvordan?

Thomas Widmann
Dansk Sprognævn

Hvad er COR?

Det Centrale Ordregister (COR) er en standardiseret og centraliseret database over alle lemmer og ordformer på dansk, med tilhørende unikke og permanente ID-numre. COR er et samarbejde mellem Dansk Sprognævn, Det Danske Sprog- og Litteraturselskab og Center for Sprogteknologi (Københavns Universitet), og projektet er støttet af Digitaliseringsstyrelsen. Opdelingen i lemmer i COR's grundregister er baseret på Retskrivningsordbogen fra Dansk Sprognævn, og den første version af denne grundlæggende del af COR blev offentliggjort den 30. september 2022.

Hvordan er COR designet?

Fælles for alle COR-resurser er et lemmanummer. De enkelte resurser kan have deres egen struktur, altså egne underopdelinger og datafelter.

Som et eksempel kan vi se på grundregisteret. Her består en indgang af præfiksset/domænet 'COR' efterfulgt af et 5-cifret indeksnummer der angiver det specifikke lemma.

Fx COR.29645 samsisk/samsk

Derefter kommer der et 3-cifret nummer som angiver ordklasse og bøjning for ordformen:

Fx COR.29645.302 samsiske/samske (302 er adj.sg.best)

Slutteligt har et COR-indeks i grundregisteret altid et 2-cifret løbenummer som angiver ortografisk variation.

Fx COR.29645.302.02 samske (samsiske er COR.29645.302.01)

Den 'rå' databaseindgang for *COR.29645.302* ser således ud:

```
COR.29645.302.01    samsisk adj.sg.best samsiske    1
COR.29645.302.02    samsisk adj.sg.best samske      1
```

Man ser her at grundregisteret tillige rummer et normeringsfelt ("1" sidst på linjen). Det er 0 for enkelte former (fx visse komparativer og superlativer) som ikke er blevet tjekket manuelt endnu og derfor ikke er omfattet af Retskrivningsordbogens status som normativ for det danske skriftsprog.

Hvad er et lemma?

Hvad angår definitionen af et lemma, følger vi Retskrivningsordbogens definition af opslagsord:

Opdelingen i opslagsord er principielt uafhængig af ordenes betydning. Det bevirker at ord med forskellig betydning er slået sammen i ét opslagsord hvis de i øvrigt har samme stavemåde, udtale, ordklasse og bøjning, og hvis de indgår i sammensætninger på samme måde.

Det betyder fx at fisken *skade* og fuglen *skade* har samme COR-id da forskellen mellem dem udelukkende er semantisk.

Tre niveauer

COR består af tre niveauer, hvilket giver mulighed for at inkludere data fra forskelligartede kilder.

- Niveau 1 svarer til Retskrivningsordbogen (grundregisteret, også kaldet COR1).
- Niveau 2 rummer andre resurser fra de professionelle sprogmiljøer i Danmark (specifikt dem som er medlem af Dansk Sprognævns repræsentantskab). Allerede nu ligger her bl.a. en del ekstra lemmaer fra DDO: **COR.EXT**.
- Niveau 3 vil rumme alle andre resurser, og her er der ingen begrænsninger – alle og enhver kan få allokeret et præfiks/domæne og et ID-interval ved at henvende sig til Dansk Sprognævn.

Relationer

I Det Centrale Ordregister (COR) er relationer en måde at forbinde lemmaer og ordformer på, så man kan indikere hvordan de er forbundet til andre. Relationer kan kun “pege indad”, således at Niveau 3-resurser kan pege på alle resurser, Niveau 2-resurser kun på sig selv, hinanden og grundregisteret, og grundregisteret på Niveau 1 kun på sig selv; det hviler således i sig selv og refererer aldrig til andre resurser.

Disse relationer gør det nemmere at organisere og søge efter data i COR da man kan følge relationerne mellem lemmaer og ordformer for at finde relaterede sproglige elementer.

Derudover er relationerne også nyttige for udviklere af sprogteknologiske værktøjer; de kan fx bruge relationerne til at forstå hvordan lemmaer og ordformer er forbundne, og de kan dermed forbedre deres værktøjers funktionalitet.

Der er mange slags mulige relationer, fx:

forkortelse	betydning
fus	fusion af to eller flere COR-indekser
rep	erstatning for et eller flere COR-indekser
spl	opsplitning i to eller flere COR-indekser
sms	sammensætning af to COR-indekser (til sammensatte ord)
hyr	hyperonym (overbegreb) til to eller flere COR-indekser
hyp	hyponym (underbegreb) til et andet COR-indeks
rim	rim som involverer et eller flere COR-indekser (til rimordbøger)

Ovenstående er blot eksempler. Hver resurse kan definere sine egne relationer, og grundregisteret bruger p.t. ingen relationer.

Hvordan kan man få adgang til COR?

COR's grundregister og enkelte andre resurser kan tilgås på to måder:

1. Det hele kan downloades som en CSV-fil fra *ordregister.dk*. Dette giver mulighed for at arbejde med bl.a. grundregisteret offline og integrere det i egne systemer.
2. Der er også en onlinegrænseflade der kan bruges til at søge i grundregisterets data og tilgå information om lemmaer og ordformer med. Disse informationer kan enten vises i HTML eller tilgås fra et program i CSV- eller JSON-format. Denne grænseflade findes på samme adresse: *ordregister.dk*.

Andre resurser vil kun have lister over definerede lemmaer og former liggende på *ordregister.dk*. Der vil så være et link til den URL som ressourcen kan hentes fra (hvis den er offentligt tilgængelig).

COR og Retskrivningsordbogen

Hvordan afviger Det Centrale Ordregister fra Retskrivningsordbogen?

Det Centrale Ordregisters grundregister er baseret på Retskrivningsordbogen, men der er nogle væsentlige forskelle mellem de to resurser:

- COR er lavet så et computerprogram nemt kan bruge det; RO er designet til mennesker.
- COR rummer de almindeligt brugte fuldformer for alle lemmaer; RO har en del huller i dækningen (og de vises ikke).
- COR kan bruges uden nogen restriktioner; RO må ikke bruges til at lave ordbøger med.
- COR har ingen eksempler; det har RO.
- COR har ingen regelhenvisninger; det har RO.
- COR har færre og kortere glosser end RO.

Opdatering til nyeste retskrivning

Det Centrale Ordregister (COR) er baseret på Retskrivningsordbogen (RO) og afspejler derfor den nyeste retskrivning. Når der sker ændringer i retskrivningen, vil COR også blive opdateret for at reflektere disse ændringer.

Helt konkret vil der ske flg. når den næste udgave af Retskrivningsordbogen offentliggøres:

1. Den hidtidige udgave omdøbes (sandsynligvis til *COR.RO2012*) og vil fortsat kunne tilgås på *ordregister.dk*.
2. Ord som udgår af Retskrivningsordbogen, vil fortsat beholde deres unikke ID. Hvis vi fx hypotetisk set forestiller os at *COR.62542 prokrustesseng* udgår i den næste udgave, vil *COR.62542* stadigvæk være tilknyttet lemmaet *prokrustesseng*, men lemmaet vil ikke have nogen former i den gældende retskrivning, og der vil være en note om at lemmaet er udgået. I *COR.RO2012* vil *prokrustesseng* stadig være defineret (altså som *COR.RO2012.62542*).
3. Ord som ændrer stavemåde i den nye retskrivning, vil ikke få noget nyt ID-nummer. Hvis man igen rent hypotetisk forestiller sig at *Odense* mister sit *d* (som jo er stumt på fynske dialekter), vil *COR.05162* stadig være det relevante COR-ID – det vil blot rumme stave-

måden *Oense*, og *Odense* vil nu være forvist til *COR.RO2012.05162*. Det samme gør sig gældende hvis der indføres ortografiske alternativer uden betydningsforskel (altså hvis både *Odense* og *Oense* tillades).

4. Nye ord vil få ID-numre i intervallet fra 0 til 99999, og disse vil naturligvis ikke defineres i *COR.RO2012*.
5. Hvis et lemma splittes op, vil de nye lemmaer få nye ID-numre, og vi laver relationer som viser sammenhængen mellem det gamle og de nye numre. Se eksemplet med *fjeder/fjer* i afsnittet om *RO^{hist}* nedenfor.
6. Hvis to lemmaer slås sammen, vil det fusionerede lemma også få et nyt ID-nummer, og igen vil vi lave en relation som beskriver sammenhængen.

Deling af egne resurser

Oprettelse af resursen

For at oprette sin egen COR-resurse skal man kontakte Dansk Sprognævn, som forvalter uddeelingen af nummerserier, og oplyse følgende:

- Et navneforslag (fx *Ordbog over dansk udtale i 1922*).
- Et forslag til et COR-domænenavn (fx *COR.UDT1922*).
- En oversigt over hvilke oplysninger man vil angive om hvert enkelt COR-indeks (fx glosse, udtale i Dania, udtale i IPA).
- Antal indeksnumre man regner med at få brug for (fx 10.000 – man skal ikke medregne de id-numre fra andre resurser som man regner med at genbruge).
- Licensbetingelser.
- Ejerforhold.
- Kontaktperson.

Når resursen er klar, bedes man sende en liste over de lemmaer som er defineret i den (men ikke indholdet af datafelterne), til Dansk Sprognævn. Det vil tillade os for et givet COR-ID at oplyse hvilke resurser der har data for dette lemma.

Tildeling af COR-ID-numre til resursen

For hvert lemma man har, er der basalt set fire scenarier:

1. Perfekt match – brug det eksisterende ID: *COR.OPEN.MUDS.0003919 skål!*

2. Blot ortografisk variation – brug eksisterende ID (evt. med underinddeling):
COR.OPEN.MUDS.0003919.02 sjkål!
3. Intet match – brug nyt ID fra egen nummerserie: *COR.OPEN.MUDS.2220002 cheers!*
4. Dårligt match – brug nyt ID med relationsopmærkning:
COR.OPEN.MUDS.2220022 whiskey (irsk); i grundregisteret er der kun *COR.94863 whisky*, så lav en relation mellem dem i *COR.OPEN.MUDS*-resursen.

Hvis man er i tvivl, er man naturligvis velkommen til at kontakte Dansk Sprognævn.

Eksempler på anvendelse af COR

Opmærkning af tekst

I COR's grundregister skelnes der som sagt mellem enslydende ordformer med forskellig grammatisk funktion. Fx kan formen *skål* svare til fire forskellige COR-ID-numre:

- ”Her er en *skål* med øl i”: Det drejer sig om *COR.82511.110.01*, sb.fk.sg.ubest af **skål** (*beholder*).
- ”Lad os udbringe en *skål*”: Her er det *COR.55947.110.01*, sb.fk.sg.ubest af **skål** (*det at skåle*).
- ”Kom nu, *skål* med mig!”: I denne sætning er det *COR.32888.209.01*, vb.imp af **skåle**.
- ”*Skål!*”: Sidst, men ikke mindst, har vi her *COR.03919.870.01*, udråbsordet **skål**.

Hvis man – manuelt eller automatisk – tilføjer links til de korrekte COR-numre til hvert ord i en løbende tekst, kaldes det COR-linking. Se Henrichsen (dette bind) for mere information.

Opmærkning af ordbog

Man kan opmærke en ordbog med COR-ID-numre på mange måder. Et typisk eksempel kan fx se således ud:

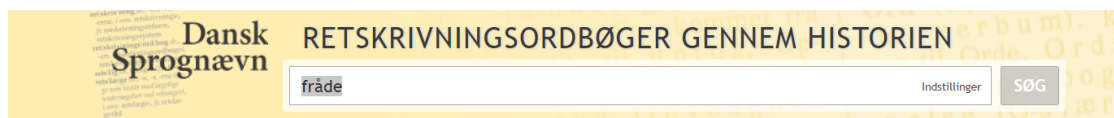

```
<entry>
  <orth COR="55947">skål</orth>
  <pos>sb</pos>
  <tran>toast</tran>
  <example>De udbragte en <hw/> for MUDS.</example>
  <tran>They toasted the MUDS conference.</tran>
</entry>
```

Det vil gøre det muligt at tjekke at man staver opslagsordene i overensstemmelse med nyeste retskrivning, og man kan tilføje bøjningsoplysninger automatisk. Man vil også kunne opdatere ordbogen automatisk når den næste udgave af Retskrivningsordbogen udkommer.

Hvis nogen på et tidspunkt publicerer en COR-opmærket udtaleordbog, vil udtale-angivelser herfra også automatisk kunne tilføjes til ens ordbog (hvis man i øvrigt har rettighederne hertil).

RO^{hist}

Dansk Sprognævns RO^{hist}-projekt (*rohist.dk*) er en søgemaskine (baseret på det svenske SAOLhist-system) hvormed man kan sammenligne retskrivningerne fra 1872 til 2012:



Tidslinje

Seneste opslagsform	DHO 1872	SRO 1892	SRO 1918	DRO 1923	DRO 1946	RO 1955	RO 1986	RO 1996	RO 2001	RO 2012
fråde, sb.	Fraade		Fraade			fråde ¹	1. fråde	1. fråde	1. fråde	1. fråde
fråde, vb.	fraade		fraade			fråde ²	2. fråde	2. fråde	2. fråde	2. fråde
Antal forekomster	DHO 1872: 2	SRO 1892: 0	SRO 1918: 2	DRO 1923: 0	DRO 1946: 0	RO 1955: 2	RO 1986: 2	RO 1996: 2	RO 2001: 2	RO 2012: 2

Et eksempel på brug af RO^{hist}.

Den langsigtede plan er at udvide RO^{hist} med alle historiske retskrivningsordbøger og andre ortografiske ”rettesnore”. Fx er vi for tiden ved at omdanne *Ove Mallings Store og gode Handlinger af Danske, Norske og Holstenere* fra 1777 til en ordbog som kan tilføjes til RO^{hist} (se Hartling og Widmann 2020).

Vi vil også bruge den information som ligger i RO^{hist}, til at tildele COR-ID-numre til de historiske ordbøger. Dette vil tillade os at simplificere og forbedre implementeringen, og det vil gøre det muligt at lægge de historiske ordbøger ind som COR-resurser på *ordregister.dk*.

Det vil foregå således at det samme ID-nummer bruges hvis lemmaet er det samme – også hvis stavemåden har ændret sig. Et ord som *fråse* (RO2012) vil altså få samme COR-nummer som *frådse* i RO1996 og som *fraadse* i Dansk Haandordbog (1872):

ID-nummer	lemma
<i>COR.RO2001.59309</i>	fråse
<i>COR.RO1996.59309</i>	frådse
<i>COR.DHO1872.59309</i>	fraadse

Hvis man søger efter *fråse* i den fremtidige RO^{hist}, vil man basalt set blot skulle finde COR-nummeret (her 59309) og derefter slå dette op i de andre ordbøger.

Det er ikke nogen forhindring hvis et ord ikke er med i RO2012 da hver ordbog vil få tildelt sit eget nummerområde til ekstra lemmaer. Ordet *backfisch* var fx sidste gang med i 2001; hvis vi forestiller os at RO2001's ekstra nummerområde er 4002000-4004999, kan man så ved COR'ificeringen af denne ordbog tildele dette lemma ID-nummeret *COR.RO2001.4002123*.

Andre ordbøger kan så genbruge dette – *Backfisch* i DHO1872 vil altså kunne hedde *COR.DHO1872.4002123*. Nummeret viser at det oprindeligt blev defineret af RO2001 ved at ligge i intervallet 4002000–4004999.

Det vil være muligt at indikere sammenfald eller opdeling af stavemåder ved brug af relationer. Fx var der i DRO1923 fri variation mellem formerne *Fjeder* og *Fjer* uafhængigt af betydningen, hvorimod vi i dag har to separate lemmaer:

ID-nummer	lemma	glosse	relation
<i>COR.RO2012.79703</i>	fjeder		
<i>COR.RO2012.83551</i>	fjer		
<i>COR.DRO1923.4008020</i>	Fjeder/Fjer		fus:79703+83551

Det samme gør sig gældende hvis to lemmaer i historiske ordbøger svar til ét i dag. Fx skelnedes der i RO1955 mellem **skade** (*en fisk*) og **skade** (*en fugl*), mens de i dag som sagt er slået sammen i ét opslagsord da de kun adskiller sig semantisk:

ID-nummer	lemma	glosse	relation
<i>COR.RO2012.03094.110.01</i>	skade	en fugl; en fisk	
<i>COR.RO1955.4011080.110.01</i>	skade	en fugl	spl:03094
<i>COR.RO1955.4011081.110.01</i>	skade	en fisk	spl:03094

RO^{hist} vil altså også skulle tjekke relationsfeltet for at kunne præsentere fyldestgørende resultater.

Vi kan således opmærke alle de historiske ordbøger i rækkefølge, og vores liste over kendte lemmaer vil gradvist vokse i processen. Vi véd endnu ikke hvor meget af COR-opmærkingen af de historiske ordbøger vi kan foretage automatisk, men vi håber at de eksisterende links mellem ordbøgerne i den eksisterende RO^{hist}-database vil hjælpe os i processen.

Sådan laver man et lille sprogpil

COR er nemt at bruge fra et program da man ikke behøver at downloade noget som helst på forhånd. Se fx flg. Python-program:

```
from urllib.request import urlopen # hent data fra en URL
from urllib.parse import quote     # formater URL-parametre
import json                        # håndtér JSON-data
from random import randint         # generér tilfældige tal

def beregnsvar(ord, gaet):
    """
    Funktionen tager to parametre: ordet der skal gættes, og det
    gættede ord. Den returnerer et svar hvor bogstaverne i det
    gættede ord er enten store (hvis de er korrekte på korrekt plads),
    små (hvis de er korrekte på forkert plads) eller bindestreger
    (hvis de er forkerte).
    """
    svar = ""
    for pos in range(len(gaet)):
        if len(ord) > pos and gaet[pos] == ord[pos]:
            svar = svar + gaet[pos].upper() # korrekt bogstav på korrekt plads
        elif ord.find(gaet[pos]) != -1:
            svar = svar + gaet[pos].lower() # korrekt bogstav på forkert plads
        else:
            svar = svar + "-" # forkert bogstav
    return svar

def tjeksvar(ord, gaet):
    """
    Funktionen tager to parametre: ordet der skal gættes, og brugerens gæt.
    Den tjekker om det gættede ord findes i ordregisteret, og hvis det gør
    det, returnerer den svaret fra beregnsvar-funktionen.
    """
    url = "https://ordregister.dk/fuldform/COR/json/" + quote(gaet) # lav URL
    data = json.loads(urlopen(url).read())
    status = data['status']
```

```
if status == 'ok':
    return beregnsvar(ord, gaet)
else:
    return "Det ord findes ikke!"

# Vælg et tilfældigt lemma fra COR med en længde på 5 bogstaver
found = False
while not found:
    try: # hvis noget går galt, havner vi i except og begynder forfra
        r = randint(1, 99999)
        url = "https://ordregister.dk/id/COR." + str(r) + ".json"
        data = json.loads(urlopen(url).read()) # går galt hvis COR.r ikke
findes
        ord = data['lemma']
        if len(ord) == 5:
            found = True # COR-ID findes, og lemmaet er på fem bogstaver -
hurra!
    except:
        continue

# Beder brugeren om at gætte på ordet og tjekker om det er korrekt.
# Prøver igen hvis det er forkert.
while True:
    gaet = input("Jeg tænker på et ord på fem bogstaver ... Hvad er det? ")

    if gaet == ord:
        break
    else:
        print(tjeksvar(ord, gaet))

input("Tillykke, det var rigtigt!")
```

Programmet laver simple JSON-kald til *ordregister.dk* for at tilgå lemmaer og fuldformer. Man kan sagtens forbedre programmet – fx fungerer *beregnsvar*-funktionen ikke på samme måde som i Wordle, og der tjekkes ikke om det fundne lemma indholder ikke-alfabetiske tegn (såsom bindestreger) – men det overlades her som en øvelse til læseren.

Konklusion

Vi har i denne artikel vist hvad Det Centrale Ordregister er, hvor man kan tilgå det henne, hvorfor man bør begynde at anvende det, og hvordan man rent praktisk gør det. Vi har også set eksempler på konkrete anvendelser af COR, inklusive et spil skrevet i Python. Forhåbningen er at læseren nu føler sig inspireret til at begynde at bruge COR – og gerne på måder som vi slet ikke har forestillet os!

Litteratur

- DHO1872 = Grundtvig, Sven (1872) *Dansk Haandordbog med den af Kultusministeriet anbefalede Retskrivning*. 2. udgave. København: C. A. Reitzel.
- DRO1923 = Glahder, Jørgen (1923) *Dansk Retskrivningsordbog*. Udgivet af Undervisningsministeriets Retskrivningsudvalg. København: Gyldendal.

Det Centrale Ordregister (COR): Hvad, hvor, hvorfor og hvordan?

- Dideriksen, Christina, Peter Juel Henriksen og Thomas Widmann (2022) Det Centrale Ordregister. I Nyt fra Sprognævnet 2022/3.
- Hartling, Anna Sofie & Thomas Widmann (2020) Den første ortografiske rettesnor for dansk – fra læsebog til ordbog: Malling (1777) på <http://rohist.dsn.dk>. I Goldshtein, Yonatan, Inger Schoonderbeek Hansen & Tina Thode Hougaard (red.). 18. Møde om Udforskningen af Dansk Sprog. Aarhus: NORDISK, Institut for Kommunikation og Kultur, Aarhus Universitet:213-230.
- Henriksen, Peter Juel (2023a) Det Centrale Ordregister. Et indeks for det danske ordforråd – en gave til dansk sprogteknologi. Nordisk Forening for Leksikografi. (Endnu ikke udkommet).
- Henriksen, Peter Juel (2023b) Diktatoriske befølelser. I denne konferencerapport.
- RO1955 = *Retskrivningsordbog*. Dansk Sprognævn (1955) København: Gyldendalske Boghandel.
- RO1996 = *Retskrivningsordbogen. 2. udgave*. Dansk Sprognævn (1996) København: Aschehoug.
- RO2001 = *Retskrivningsordbogen. 3. udgave*. Dansk Sprognævn (2001) København: Alinea – Aschehoug Dansk Forlag.
- RO2012 = *Retskrivningsordbogen. 4. udgave*. Dansk Sprognævn (2012) København: Alinea.
- Widmann, Thomas (2023) Det Centrale Ordregister og dets leksikografiske anvendelser. Nordisk Forening for Leksikografi. *Endnu ikke udkommet*.