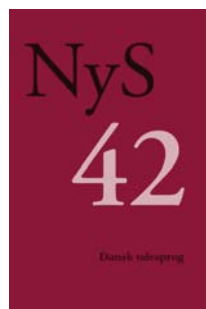


NyS

Titel:	<i>At eller ikke at i tale og skrift</i>
Forfatter:	Kasper Boye, Mads Poulsen, Hannah Bruun Pedersen, Marie Herget Christensen, Line Dalberg og Nicoline Munck Vinther
Kilde:	<i>NyS – Nydanske Sprogstudier 42. Dansk talesprog</i> , 2012, s. 41-61
Udgivet af:	NyS i samarbejde med Dansk Sprognævn
URL:	www.nys.dk



© NyS og artiklens forfatter

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre NyS-numre (NyS 1-36) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

At eller ikke *at* i tale og skrift

KASPER BOYE, MADS POULSEN, HANNAH BRUUN PEDERSEN, MARIE HERGET CHRISTENSEN, LINE DALBERG, NICOLINE MUNCK VINTHER

1. INDLEDNING

Fremkomsten af store talesprogskorpora de seneste årtier er gået hånd i hånd med en stadig stigende interesse for talesprogsgrammatik. Man har søgt at påvise grammatiske fænomener, der er særegne for det talte sprog (danske eksempler er Jensens 2003 “fællesmængdekonstruktion”, “anafor-katafor konstruktion” og “katafor konstruktion”), og det er blevet foreslået, at talesprogsgrammatik er noget helt (f.eks. Brazil 1995) eller delvist (f.eks. McCarthy 1998) andet end skriftsprogsgrammatik. Heroverfor har f.eks. Leech (2000) (med henvisning til Biber et al. 1999) insisteret på, at talesprogs- og skriftsprogsgrammatik grundlæggende er et og samme fænomen (se også Harder & Henriksen 2004).

Mod opfattelsen af tale- og skriftsprogsgrammatik som væsensforskellige kan man i hvert fald indvende følgende:

- 1) Det er uøkonomisk at operere med distinkte tale- og skriftsprogsgrammatikker (såvel deskriptivt som i hovedet – eller hvor grammatik nu sidder).
- 2) Opfattelsen hviler i vid udstrækning på en manglende sondring mellem grammatik som et konventionelt og situationsuafhængigt fænomen og den situationsbundne sprogbrug (jf. Leech 2000: 53-54). Betragt f.eks. et eksempel på Jensens (2003: 297-298) “anaphor-cataphor construction”:

(1) ... *dét* forestiller man sig jo altså ikke at det kan være så svært ...

Helsætningen i (1) indeholder pronomenet *dét*, der ifølge Jensen dels anaforisk repræsenterer et element i den forudgående diskurs, dels kataforisk enten repræsenterer ledsætningen som argument i helsætningen eller repræsenterer et argument i ledsætningen. Jensen mener, dette er en særlig talesprogsgrammatisk kategori (jf. termen “construction”).

Det er imidlertid ikke indlysende, at sætningen i (1) repræsenterer noget som helst konventionelt fænomen. Den lader sig helt ligetil analysere som et sprogbrugsfænomen: en situationsbestemt kontamination af en sætningsknode og en ledsætningskonstruktion (jf. Hansen 1961: 147).

- 3) Langt de fleste grammatiske kategorier (herunder konstruktioner) findes i såvel tale- som skriftsprog; forskelle i hyppighed kan tilskrives de forskellige produktionsbetingelser for talen og skriften (Leech 2000: 61-62, med henvisning til Biber et al. 1999).

I denne artikel vil vi føje endnu et argument til rækken. Vi vil vise, at selv et tilfælde af grammatisk variation, der bedst kan forstås som betinget af forhold, der er karakteristiske for det talte sprog, kan genfindes i det skrevne sprog. Det tilfælde af grammatisk variation, vi vil se på, er variationen i finitte *komplementsætninger* (dvs. nominale ledsætninger, der kræves af et verbal eller et andet *komplementtagende element*) mellem fraværet og tilstedeværelsen af underordningskonjunktionen *at*.

I visse konstruktioner er *at* obligatorisk, f.eks. i konstruktioner, hvor komplementsætningen er foranstillet (står i forfelt) som i (2).

- (2) a. **[Ø Det ikke skulle regne] er usandsynligt.*
b. *[At det ikke skulle regne] er usandsynligt.*

I andre konstruktioner er *at* umuligt, f.eks. i sætningsknuder, hvor knudeelementet er primært argument ("logisk subjekt") i komplementsætningen som i (3).

- (3) a. *Ham tror jeg ikke [Ø klarer det].*
b. **Ham tror jeg ikke [at klarer det].*

Men i de fleste konstruktioner er *at* valgfrit. Det er f.eks. tilfældet i langt de fleste almindelige deklarativer sætninger som i (4), i langt de fleste almindelige interrogative sætninger som i (5) og i sætningsknuder som den i (6), hvor knudeelementet ikke er primært argument i komplementsætningen.

- (4) a. *Jeg tror [Ø det regner].*
b. *Jeg tror [at det regner].*
- (5) a. *Tror du [Ø det regner]?*
b. *Tror du [at det regner]?*
- (6) a. *Ham tror jeg ikke [Ø du kan klare].*
b. *Ham tror jeg ikke [at du kan klare].*

Der kan opstilles flere hypoteser om, hvad der betinger denne variation mellem tilstedeværelse og fravær af *at* (se afsnit 5), og de udelukker ikke hinanden. Vi vil koncentrere os om følgende hypotese:

Hypotese

Proportionen af komplementsætninger med *at* (som i 4b, 5b og 6b) er desto større, jo vanskeligere komplementsætningerne er at producere.

Mere specifikt er hypotesen den, at vi i tilfælde af sprogproduktionsvanskeligheder bruger *at* (som et alternativ til f.eks. *øh*) som udfyldning af en pause for at købe tid til planlægning af den videre sprogproduktion. Når vi skriver, har vi i reglen ikke brug for sproglige pausefisk, men kan løse eventuelle produktionsvanskeligheder i relativ ro og mag. Hypotesen har således klar affinitet med talesprog. I talen er vi underlagt strikte tidsbegrænsninger og dermed forbundne planlægningsbegrænsninger og begrænsninger på muligheden for sproglig “redaktion”. Ikke desto mindre vil vi underbygge hypotesen både i data fra DGCSS’ talesprogs-korpus¹ og i data fra KorpusDK, der langt overvejende er et skriftsprogs-korpus.

Artiklen er disponeret som følger: I afsnit 2 introducerer vi to specifikke og testbare forudsigelser, der følger af den ovenfor nævnte hypotese. I afsnit 3 afrapporterer vi en eftervisning af disse to specifikke forudsigelser i data fra DGCSS’ talesprogs-korpus, og i afsnit 4 afrapporterer vi en eftervisning af den ene af forudsigelserne i data fra KorpusDK. I afsnit 5 beskriver vi kort et par andre hypoteser om, hvad der styrer, om *at* er der eller ej, og vi diskuterer relationen mellem dem og den ovenfor nævnte hypotese. Afsnit 6 er en konklusion og kort diskussion.

2. SPECIFIKKE FORUDSIGELSER

Den hypotese, vi skal beskæftige os med i denne artikel, er, at produktionsforhold har indflydelse på variationen mellem \emptyset og *at*. Her er hypotesen igen:

Hypotese

Proportionen af komplementsætninger med *at* er desto større, jo vanskeligere komplementsætningerne er at producere.

En efterprøvning indebærer en sammenligning af proportionen af *at* for komplementsætninger, der er relativt vanskelige at producere, og proportionen af *at* for komplementsætninger, der er relativt lette at producere. Det forudsætter igen, at der sondres mellem relativt let og relativt svært producerbare sætninger. Disfluenser (f.eks. pauser og ordforlængelser; se afsnit 3) på grænsen til komplementsætningen kan antages langt overvejende at være udtryk for, at taleren har problemer med at få udtrykt det, der skal siges. I overensstemmelse hermed kan det antages, at komplementsætninger, der indledes med disfluens, er vanskeligere at producere end komplementsætninger, der ikke indledes med disfluens. På denne baggrund kan hypotesen ovenfor udmøntes i følgende specifikke forudsigelse:

Specifik forudsigelse 1

Proportionen af komplementsætninger med *at* er større, når der er disfluens på grænsen til komplementsætningen, end når der ikke er det.

En anden sondring mellem relativt let og relativt svært producerbare sætninger hviler på den antagelse, at sætninger med ikke-pronominalt (substantivisk eller komplekst nominalt) subjekt alt andet lige er sværere at producere end sætninger med pronominalt subjekt. Rationalet bag denne antagelse er, at substantiver og komplekse nominalsyntagmer dels repræsenterer sværere tilgængelig information end pronomener (f.eks. Ariel 1990, 2006), dels som oftest er længere og mindre frekvente end pronomener, hvorfor de er sværere at mobilisere. På baggrund af denne antagelse kan den generelle hypotese udmøntes også i følgende specifikke forudsigelse:

Specifik forudsigelse 2

Proportionen af komplementsætninger med *at* er større, når komplementsætningens subjekt er ikke-pronominalt, end når det er pronominalt.

Hypotesen og dens to specifikke udmøntninger understøttes af forskning i engelsk (Ferreira & Dell 2000; Jaeger 2005; Wasow, Jaeger & Orr 2011; Race & MacDonald 2003). F.eks. har Jaeger (2005) i en analyse af et talesprogs-korpus vist, at der var større sandsynlighed for, at en valgfri underordningskonjunktion (relativsætningskonjunktionen *that*) blev ytret, hvis sætningen indeholdt en disfluens. Dette tyder på, at en årsag til, at talere inkluderer en valgfri underordningskonjunktion, er, at konjunktionen giver dem mulighed for at udfylde en pause, der bruges på at planlægge, hvad der så skal siges (jf. også Kaltenböck 2009 om komplementsætningskonjunktionen *that*). Race og MacDonald (2003) har undersøgt, om der er større sandsynlighed for tilstedeværelse af en valgfri underordningskonjunktion i en ledsætning, når sætningens subjekt er ikke-pronominalt, end når det er pronominalt. De fandt tendenser i denne retning både i et amerikansk skriftsprogs-korpus og i et sætningsproduktionsforsøg, hvor forsøgsdeltagere skulle genkalde sig sætninger, hvor det underordnede subjekt enten var pronominalt eller ikke-pronominalt.

Af de to specifikke forudsigelser er det i sagens natur kun den anden, det giver mening at undersøge både i et tale- og skriftsprogs-korpus. Når vi også undersøger den første (i et talesprogs-korpus) er det for at underbygge den overordnede hypotese mest muligt. Som nævnt kan disfluenser antages langt overvejende at være tegn på produktionsvanskeligheder. En eventuel bekræftelse af den første specifikke forudsigelse indebærer således en bekræftelse af den overordnede hypotese, og det styrker den antagelse, at også en eventuel bekræftelse af den anden specifikke forudsigelse kan tilskrives produktionsvanskeligheder. Det er vigtigt for os at bestyrke denne antagelse, eftersom der også kan tænkes andre årsager til en bekræftelse af den anden specifikke forudsigelse. Det vender vi kort tilbage til i afsnit 5. De mellemliggende afsnit er viet til efterprøvningen af de to specifikke hypoteser (foruden en kort perspektivering).

3. TALESPROGSUNDERSØGELSEN

I det følgende afrapporterer vi efterprøvningen af de to specifikke forudsigelser i data fra DGCSS' talesprogskorpus.

3.1 Materiale

DGCSS' talesprogskorpus består af transskriberede sociolingvistiske samtaler (interview og gruppesamtaler) og udgør ca 6½ mio. "ikke-tomme intervaller", hovedsagelig ord, men også pauser udfyldt med tøvelyde og påbegyndte, men ikke fuldendte ord. Til brug for vores undersøgelse uddrog vi data fra to dele af korpusset: 1) det såkaldte "eksplorative delkorpus", som består af transskriptioner af 20 samtaler med personer fra Køge, Nyboder, Næstved, Odder og Vinderup (optaget i perioden 1978-2006), og 2) de såkaldte BySoc1ny- og BySoc2ny-filer, som består af transskriptioner af i alt 44 samtaler med personer fra Nyboder (optaget i perioden 2005-2007) – idet vi slettede den ene kopi af en transskription, der gik igen i begge delkorpora. Vores materiale består således af transskriptioner af i alt 63 samtaler og indeholder ytringer fra 53 forskellige personer (enkelte personer deltager i flere samtaler). I vores undersøgelse har vi kun medtaget data fra egentlige informanter – dvs. vi har set bort fra interviewerens ytringer.

De udvalgte samtaletransskriptioner blev behandlet i Praat. Komplementsætningerne blev lokaliseret manuelt (idet vi dog for det eksplorative delkorpus' vedkommende kunne udnytte, at alle nominale ledsætninger var blevet markeret i forbindelse med en tidligere undersøgelse) og analyseret grammatisk. Efterfølgende blev relevante data uddraget fra korpus og behandlet i Excel.

Til undersøgelsen uddrog vi komplementsætninger, der bl.a. deler følgende træk: 1) De står ikke alene (dvs. de kan analyseres som egentlige komplement, dvs. krævede led; jf. afsnit 1); 2) de har ikke imperativisk struktur; 3) de indledes af *Ø* eller *at*; 4) de har ikke subjektsfordobling; 5) de optræder som eneste komplement eller som det første af to eller flere sideordnede komplement; 6) de er ikke placeret i den overordnede sætnings forfelt; 7) deres overordnede sætning er ikke en knude; 8) komplementet udgør ikke sammen med den overordnede sætning en parentes (som *hun bed* gør det sammen med *tror jeg...* i *det kunne man faktisk godt lave oppe hos hende der fru Madsen tror jeg hun bed*).

Med begrænsningen til komplementsætninger med disse træk sikres det bl.a., at undersøgelsen kun inddrager komplementsætninger med valgfrit *at*.

De uddragne komplementsætninger adskiller sig med hensyn til en række træk, hvoraf følgende er relevante for den talesprogsundersøgelse, der afrapporteres her.

- Komplementsætningsindleder

Herunder skelnes der bl.a. mellem *om*, *Ø* og *at*.

- Disfluens på grænsen til komplementsætningen

Herunder skelnes der mellem fravær af disfluenssymptomer og en række disfluenssymptomer – herunder: 1) ufoneret pause (pause uden lyd), 2) foneret pause (pause med lyd), 3) mere end én pause og 4) forlængelse, gentagelse eller rettelser af komplementsætningsindlederen, sidste ord i matrixsætningen eller første ord i komplementsætningen (uden pause).

- Komplementsætningens subjekt

Herunder skelnes der bl.a. mellem forskellige slags pronominale subjekter, forskellige slags ikke-pronominale subjekter og subjektsfordobling.

- Komplementtagende element

Herunder registreres for hver komplementsætning det element, der tager den som komplement – f.eks. simple verber som *tro* (*jeg tror [(at) hun er der]*), komplekse verber som *tro på* (*jeg tror på [(at) hun er der]*), simple prædikative adjektiver som *sandsynlig* (*det er sandsynligt [(at) hun er der]*) og komplekse prædikativer som *sikker på* (*jeg er sikker på [(at) hun er der]*).

- Informant

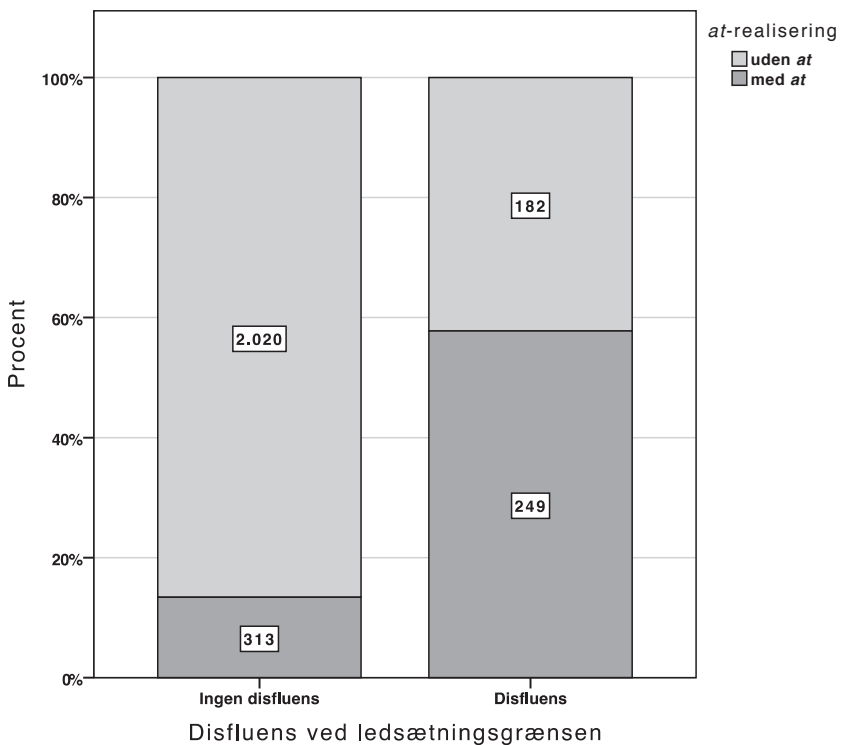
Herunder registreres for hver komplementsætning den af de 53 informanter, der har ytret den.

3.2 Resultater

I alt blev der uddraget 2.764 brugbare komplementsætninger til testning af de to specifikke forudsigelser i talesprog. 20 % af disse komplementsætninger indledes med *at*. 16 % af komplementsætningerne indeholdt en disfluens. I 7 % af komplementsætningerne var subjektet ikke-pronominalt, mens det i de resterende tilfælde var et pronomen.

Figur 1 viser fordelingen af *at*-realisering afhængigt af, om der blev produceret en disfluens eller ej. Fordelingen var i overensstemmelse med den første specifikke forudsigelse: Proportionen af komplementsætninger med *at* var større, når der var disfluens (foneret pause, ufoneret pause, ordforlængelse, ordgentagelse eller rettelse) på grænsen til komplementsætningen, end når der ikke var det. Som det er illustreret i Figur 1, indeholdt 13 % af komplementsætningerne uden disfluens et *at*, mens 58 % af ytringerne med disfluens indeholdt et *at*.

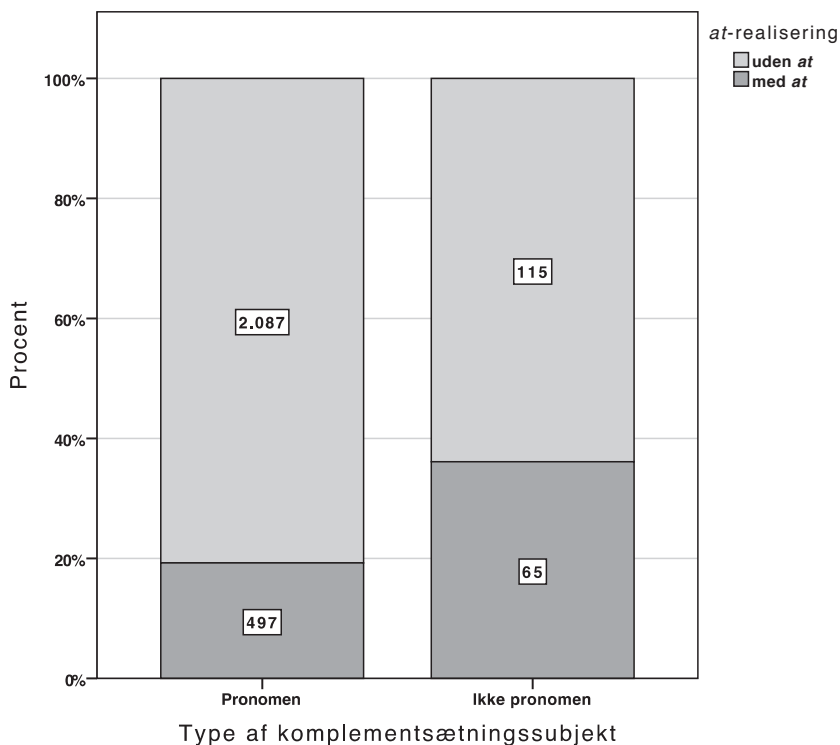
FIGUR 1. TILSTEDEVÆRELSE AF *AT* AFHÆNGIGT AF DISFLUENS PÅ GRÆNSEN TIL KOMPLEMENTSÆTNINGEN I TALESPROG



Figur 2 viser fordelingen af *at*-realiseringer afhængigt af typen af komplementsætningssubjekt. Fordelingen var i overensstemmelse med den anden specifikke forudsigelse: Proportionen af komplementsætninger med *at* var større i komplementsætninger med ikke-pronominalt sub-

jekt end i komplementsætninger med pronominalt subjekt. Som det er illustreret i Figur 2, indeholdt 19 % af komplementsætningerne med pronominalt subjekt et *at*, mens 36 % af sætningerne med ikke-pronominalt subjekt indeholdt et *at*.

FIGUR 2. TILSTEDEVÆRELSE AF *AT* AFHÆNGIGT AF TYPE AF KOMPLEMENTSÆTNINGS-SUBJEKT I TALESPROG



Pålideligheden af det tydelige mønster i figurerne blev bekræftet af en logistisk generaliseret lineær blandet model-analyse (se Baayen 2008; Jaeger 2008) udført ved hjælp af lme4-pakken til statistiksoftwaren R. I analysen behandlede vi tilstedeværelsen af *at* som afhængig variabel og disfluens og type af komplementsætningssubjekt som uafhængige variable. Desuden inkluderede vi informanten og det komplementtagende element som tilfældige variable for at sikre os, at eventuelle effekter ikke var båret af enkelte informanter eller særlige komplementtagende

elementer (jf. Jensens 2003 undersøgelse af variationen mellem *Ø* og *at* for en række komplementtagende prædikater). Både disfluens ($\beta = 2,365, p < ,001$) og type af komplementsætningssubjekt ($\beta = 0,830, p < ,001$) var signifikante prædiktorer i modellen. Det er således usandsynligt, at fordelingen skyldes tilfældige udsving.

4. SKRIFTSPROGSUNDERSØGELSEN

I det følgende afrapporterer vi efterprøvningen af den anden af de to specifikke forudsigelser i data fra KorpusDK.

4.1 Materiale

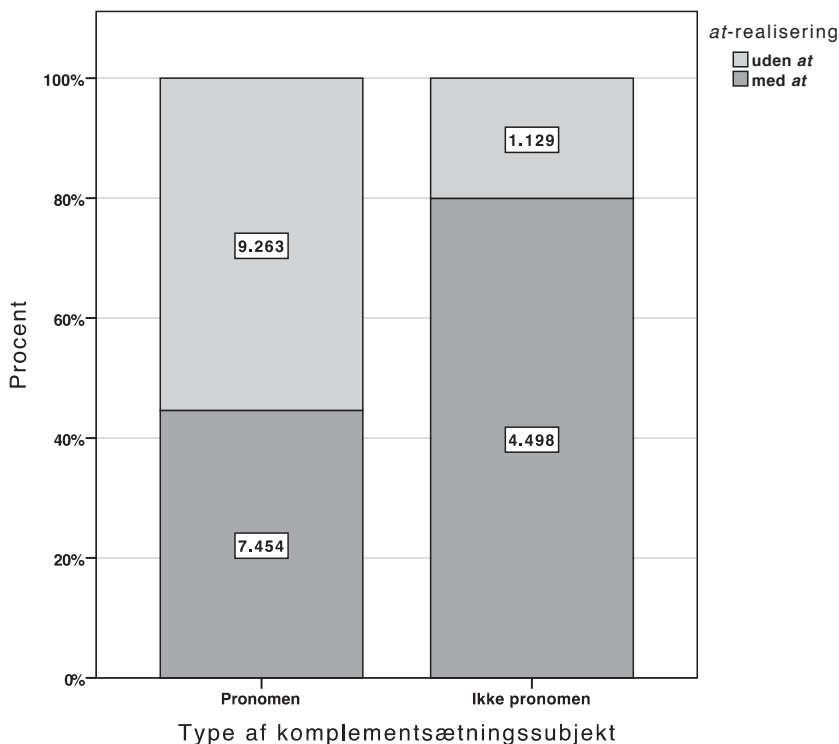
KorpusDK er langt overvejende et skriftsprogskorpus. Det består af to delkorpora, Korpus90 og Korpus2000, som hver især indeholder omkring 28 mio. ord. Vi uddrog data til vores undersøgelse vha. den tilknyttede søgemaskines mulighed for “formel søgning”. For 110 komplementtagende prædikater (verber og verbalforbindelser som *tro*, *bevirke*, *tænke over* og *bedde sig* og adjektiver og adjektivforbindelser som *almindeligt*, *muligt*, *sikker på* og *nervøs for*) søgte vi i hver af de to delkorpora på strenge, hvori prædikaterne efterfølges af først *Ø* eller *at*, derpå et substantiv eller et pronomen og til sidst et finit verbal. Fundene blev efterfølgende tjekket og behandlet manuelt i Excel.

4.2 Resultater

Det samlede materiale bestod af i alt 22.344 komplementsætninger fordelt på de to delkorpora. 54 % af sætningerne indeholder et *at*.

Figur 3 viser fordelingen af *at*-realiseringer afhængigt af typen af komplementsætningssubjekt. Fordelingen var ligesom i talesprogskorpusset i overensstemmelse med den anden specifikke forudsigelse: Proportionen af komplementsætninger med *at* var større i komplementsætninger med ikke-pronominalt subjekt end i komplementsætninger med pronominalt subjekt. Der var *at* i 5 % af komplementsætningerne med pronominalt subjekt, mens 80 % af sætningerne med ikke-pronominalt subjekt indeholdt et *at*.

FIGUR 3: TILSTEDEVÆRELSE AF *AT* AFHÆNGIGT AF TYPE AF KOMPLEMENTÆTNINGS-SUBJEKT I SKRIFTSPROG



Pålideligheden af disse fordelinger blev bekræftet af en logistisk generaliseret lineær blandet model-analyse med tilstedeværelse af *at* som afhængig variabel og type af komplementætningssubjekt som uafhængig variabel. Det komplementtagende element blev inkluderet som tilfældig variabel. Effekten af subjektstype var signifikant ($\beta = 1,310$, $p < ,001$). Det er således usandsynligt, at fordelingen skyldes tilfældige udsving.

5. PERSPEKTIVERING OG DISKUSSION

Resultaterne fra undersøgelserne af både tale- og skriftsprog understøtter hypotesen om, at *at* bruges af sprogbrugere, når de har svært ved at planlægge eller producere en ledsætning. Men variationen mel-

lem fravær og tilstedeværelse af *at* er næppe udelukkende betinget af produktionsvanskeligheder. Den er sandsynligvis styret af en række faktorer. Med skelen til litteraturen om engelsk *that* kan der således fremsættes en række hypoteser om, hvad der influerer på variationen. En gruppe samler sig omkring den generelle hypotese, at der med tiden bliver færre *at*'er. I overensstemmelse hermed er det sandsynligt, at proportionen af *at* er større i konservative dialekter, genrer osv. end i ikke-konservative, og større i skriftsprog end i talesprog. En anden gruppe består af hypoteser om rent grammatiske faktorer. I det følgende diskuterer vi to sådanne rent grammatiske hypoteser og deres relation til hypotesen om produktionsvanskeligheder. Derefter retter vi blikket mod spørgsmålet, hvorfor der overhovedet er variation mellem fravær og tilstedeværelse af *at*.

5.1 Grammatikaliseringshypotesen

Grammatikaliseringshypotesen siger, at proportionen af komplement-sætninger med *at* er desto større, jo mindre grammatikaliseret det komplementtagende element er (jf. Thompson & Mulac 1991 og Cuyckens & Shank 2011 om den engelske underordningskonjunktion *that*).

Denne hypotese tager afsæt i det velkendte forhold, at et komplementtagende element kan undergå grammatikalisering – dvs. udvikle sig til et grammatisk element (se f.eks. Boye & Harder 2007, 2009 og henvisningerne deri). Et muligt dansk eksempel er udviklingen af den komplementtagende verbalforbindelse *må ske* (7a) til en grammatisk størrelse, adverbiet eller partiklen *måske* (7b).

(7) a. *Det må ske* [(*at*) *det regner*].

>

b. *Måske det regner*/ *Det regner måske*.

Grammatikaliseringen indebærer, at det komplementtagende element (*må ske*) mister sin status som komplementtagende element, og – sideløbende hermed – at det, der oprindeligt var en komplementsætning ((*at*) *det regner* i 7a), mister sin status som komplement og bliver til en helsætning (*det regner* i 7b). Det er helt oplagt, at efterhånden som udviklingen skrider frem, bliver komplementsætningskonjunktionen *at*

mere og mere malplaceret: Når der ikke længere er nogen komplementsætning, kan der ikke længere være behov for et *at*.

Grammatikaliseringshypotesen er for nylig blevet bekræftet i en undersøgelse af data fra DGCSS' talesprogs-korpus. Boye og Poulsen (2011) har på baggrund af en ny teori om grammatisk status og grammatikalisering (se Boye & Harder 2012 for seneste version) vist, at hyppigheden af *at* i komplementsætninger er omvendt proportional med antallet af grammatikaliseringssymptomer i de komplementtagende prædikater. I overensstemmelse hermed har Jensen (2011) vist, at *at* er signifikant hyppigere i komplementsætninger med entydig ledsætningsledstilling end i komplementsætninger med entydig helsætningsledstilling.

Grammatikaliseringshypotesen tilbyder en mulig alternativ forklaring på den relation mellem disfluens og *at*-tilstedeværelse, der blev dokumenteret i afsnit 3. Vi har antaget, at disfluens er et udtryk for produktionsvanskeligheder, men man kunne forestille sig, at disfluens i stedet var relateret til grad af grammatikalisering: Man kunne forestille sig, at komplementtagende elementer er desto lettere at producere, jo mere grammatikaliserede de er. Det indebærer muligheden for, at der med begrebet om grammatikaliseringsgrad på en og samme tid kan redegøres udtømmende for såvel variationen mellem fravær og tilstedeværelse af *at* som variationen mellem fravær og tilstedeværelse af disfluenser: Jo mindre grammatikaliseret det komplementtagende element er, desto større er såvel proportionen af komplementsætninger med *at* som proportionen af komplementsætninger, der indledes med disfluens. Ifølge denne mulige alternative forklaring ville korrelationen mellem disfluensvariation og *at*-variation være et tilfælde af samvariation styret af grammatikaliseringsgrad – og ikke, som vi ellers har plæderet for, et udtryk for, at produktionsforhold influerer på *at*-variationen.

Det er imidlertid usandsynligt, at den alternative forklaring kan erstatte en forklaring baseret på produktionsvanskeligheder: Ved et og samme verbum, *tro* (med 1. person singularis-subjekt), fandt vi i en pilotundersøgelse, at proportionen af komplementsætninger med *at* var langt større, når der var disfluens på grænsen til komplementsætningen (30 % komplementsætninger med *at*), end når der ikke var det (3 % komplementsætninger med *at*)².

5.2 Forventelighedshypotesen

Forventelighedshypotesen siger, at proportionen af komplementsætninger med *at* er desto større, jo mindre forventelige komplementsætninger er (jf. Elsnæs 1984 om den engelske komplementsætningskonjunktion *that*).

Denne hypotese er baseret på analysen af *at* som markør af komplementsætninger. Rationalet er, at jo mindre forventelig en komplementsætning er, desto større behov er der for at justere forventningerne ved at introducere komplementsætningen med et *at*. Hypotesen indebærer bl.a., at der er flere *at*'er i forbindelse med komplementtagende elementer, der relativt sjældent tager sætningskomplement, end i forbindelse med komplementtagende elementer, der relativt hyppigt gør det. Den indebærer også, at der er flere *at*'er i kontekster, hvor det første ord i komplementsætningen ikke entydigt kan antages at tilhøre en komplementsætning. Sammenlign f.eks. (8) og (9).

(8) *Jeg er ked af han ...*

(9) *Jeg er ked af det ...*

Når man i (8) har læst *han*, kan man være ret sikker på, at man er i begyndelsen af en komplementsætning. Som nominativform må *han* være subjekt i en sætning, og den sætning må være komplement til *ked af*. I en sætning som (10), der fuldender (8), er der derfor ikke noget særligt behov for at introducere komplementsætningen med *at*.

(10) *Jeg er ked af [(at) han ikke er blevet færdig].*

Det er der derimod i en sætning som (11), som fuldender (9).

(11) *Jeg er ked af [(at) det ikke er blevet færdigt].*

For *det* er ikke en nominativform, og der er mange alternativer til at fuldende (9) med en komplementsætning, f.eks. (12) og (13).

(12) *Jeg er ked af det.*

(13) *Jeg er ked af det m^og.*

Hypotesen indebærer med andre ord bl.a., at der er flere *at*'er i kontekster, hvor der er en midlertidig strukturel flertydighed i begyndelsen af komplementsætningen, end i kontekster, hvor der entydigt kan forventes en komplementsætning. Forklaringen skulle være, at komplementsætningskonjunktionen *at* indsættes for at entydiggøre flertydige kontekster.

Forventelighedshypotesen tilbyder ligesom grammatikaliseringshypotesen en mulig alternativ forklaring på den relation mellem disfluens og *at*-tilstedeværelse, der blev dokumenteret i afsnit 3. Vi har antaget, at disfluens er et udtryk for produktionsvanskeligheder, men man kunne forestille sig, at disfluens i stedet er relateret til forventelighed: Disfluens på grænsen til komplementsætningen indebærer alt andet lige (herunder intonationsforhold) en delvis disintegration af komplementsætningen, og den, der lytter, kan komme i tvivl om, hvorvidt det, der følger efter disfluensen, vitterlig er en komplementsætning (og ikke f.eks. en ny helsætning, der efterfølger en gedigen selvfrydelse). Man kunne således forestille sig, at disfluens på grænsen til komplementsætningen mindsker forventeligheden for komplementsætninger, og man kunne i overensstemmelse hermed forestille sig, at den større proportion af *at* ved disfluens er et udtryk for den mindskede sandsynlighed for komplementsætninger snarere end for produktionsvanskeligheder.

Forventelighedshypotesen tilbyder også en mulig alternativ forklaring på den relation mellem ikke-pronominalt subjekt og *at*-tilstedeværelse, der blev dokumenteret i afsnit 3 og 4. Forklaringen ville være, at der blandt pronomenerne, men ikke blandt substantiverne og de komplekse nominalsyntaxmer, er nominativformer, der entydigt kan antages at være første ord i en komplementsætning og således overflødiggør et entydiggørende *at*. Det kan med andre ord tænkes, at der er mindre tilbøjelighed til at indsætte et *at* før nominativpronomenet i (14), end der f.eks. er til at indsætte et før substantivet i (16), fordi (14) kun kan fuldendes med en komplementsætning som i (15), mens (16) både kan fuldendes som i (17) og som i (18) (jf. diskussionen af (8)-(13) ovenfor).

- (14) *Jeg er ked af hun ...*
 (15) *Jeg er ked af [(at) hun ikke er blevet færdig].*
 (16) *Jeg er ked af huset ...*
 (17) *Jeg er ked af [(at) huset er faldet sammen].*
 (18) *Jeg er ked af huset.*

Der er imidlertid meget, der tyder på, at forventelighedshypotesen ikke holder. Hypotesen kan ganske vist ses som understøttet af det tidligere omtalte forhold, at *at* er obligatorisk i foranstillede komplementsætninger (se afsnit 1): Uden *at* ville man ikke i begyndelsen af en foranstillet komplementsætning vide, at det var en komplementsætning. Men Boye og Poulsen (2011) har i en undersøgelse af data fra DGCSS' talesprogs-korpus ikke kunnet bekræfte, at hypotesen også holder for ikke-foranstillede komplementsætninger. Tilsvarende har Ferreira og Dell (2000) i en række psykolingvistiske forsøg ikke kunnet vise, at talere brugte engelsk *that* til at entydiggøre midlertidigt flertydige sætninger. For så vidt angår talesprog, er dette forventeligt, for i talesproget har vi ofte ikke behov for *at* og *that* for at entydiggøre. Der har vi ofte prosodiske midler til rådighed. Prøv selv at udtale (11) og (12) ovenfor – der er tydelig forskel, allerede når vi når til *af*.

5.3 Årsagen til variationen mellem fravær og tilstedeværelse af *at*

De ovenfor nævnte hypoteser er forsøg på at besvare spørgsmålet, hvad der styrer variationen mellem fravær og tilstedeværelse af *at*. Et mere generelt spørgsmål er dette: Hvorfor er der overhovedet variation? Konjunktionen *om*, den anden kanoniske komplementsætningsindleder i dansk, er til sammenligning altid obligatorisk.

- (19) a. **Hun spurgte [Ø det regner].*
 b. *Hun spurgte [om det regner].*

Svaret på dette spørgsmål har sandsynligvis at gøre med, at *at* er semantisk fattigt. Hvor *om* markerer usikkerhed om den proposition, der udtrykkes af komplementsætningen, yder *at* ikke noget semantisk bidrag til komplementsætningen – *at* er “epistemisk neutralt” (Boye 2008; jf. også Pedersen 2000: 229 om *at*'s betydningsfattighed, og Hansen

& Heltoft 2011: 1659). Det er således påfaldende, at det i germanske sprog som norsk, svensk, færøsk, engelsk og islandsk, der har en kontrast grammatisk og semantisk magen til den mellem *at* og *om*, ligeledes er den betydningsfattige konjunktion (hhv. *at*, *att*, *at*, *that* og *að*), der er valgfri (Nordström & Boye 2011; bemærk, at islandsk *að* i modsætning til sine germanske slægtninge kun er valgfrit, når det følger umiddelbart efter det finitte verbal og selv umiddelbart efterfølges af et pronominalt subjekt; Thráinsson 1994: 186). Betydningsfattigheden kan dog ikke fuldstændigt forklare valgfriheden. Hollandsk og tysk har ligeledes en kontrast, der grammatisk og semantisk er magen til den mellem *at* og *om*, men det hollandske svar på *at*, *dat*, er (med forbehold for dialektale undtagelser) ikke valgfrit, og det tyske svar, *dass*, er obligatorisk til stede i verbalfinale komplementsætninger og obligatorisk fraværende i V2-komplementsætninger.

6. KONKLUSION

I afsnit 3 dokumenterede vi to forhold i dansk talesprog: 1) at proportionen af komplementsætninger med *at* er signifikant større, når der er disfluens (foneret pause, ufoneret pause, ordforlængelse, ordgentagelse eller rettelse) på grænsen til komplementsætningen, end når der ikke er det, og 2) at proportionen af komplementsætninger med *at* er signifikant større i komplementsætninger med ikke-pronominalt subjekt end i komplementsætninger med pronominalt subjekt. Der er flere måder at forstå disse forhold på, men det følger af argumentationen i afsnit 5, hvor et par mulige alternative forklaringer blev overvejet og siden forkastet, at begge forhold bedst kan forstås som manifestationer af det mere generelle forhold, at proportionen af komplementsætninger med *at* i talesprog er desto større, jo vanskeligere komplementsætningerne er at producere. I overensstemmelse hermed kan begge forhold forklares ved en analyse af den valgfrie komplementsætningskonjunktion *at* som et pausemærke: I tilfælde af produktionsvanskeligheder kan man indsætte *at* for at holde gulvet og sikre sig tænketid, som man kan bruge til at komme over vanskelighederne.

Når det gælder det andet forhold, hviler forklaringen på den antagelse, at substantiver og komplekse nominalsyntagmer er mindre

tilgængelige og dermed vanskeligere at producere end pronomener. Forklaringen – og dermed antagelsen – støttes af, at det første forhold helt oplagt kan forklares på samme vis: Disfluenser på grænsen til komplementsætningen kan antages langt overvejende at være udtryk for produktionsvanskeligheder.

Som påpeget i afsnit 1 har forklaringen klar affinitet til talesprog. Når vi skriver, har vi i reglen ikke brug for sproglige pausefisk, men kan løse evt. produktionsvanskeligheder i relativ ro og mag. Ikke desto mindre kunne vi i afsnit 4 dokumentere, at proportionen af komplementsætninger med *at* også i skriftsprog er signifikant større i komplementsætninger med ikke-pronominalt subjekt end i komplementsætninger med pronominalt subjekt.

Vi har således vist, at et tilfælde af grammatisk variation (variationen mellem \emptyset og *at* i komplementsætninger), der bedst kan forstås som betinget af forhold, der er karakteristiske for det talte sprog, kan findes i det skrevne sprog. Dette er et argument mod opfattelsen af tale- og skriftsprogsgrammatik som væsensforskellige. Det er derimod fuldt foreneligt med en opfattelse af talesproget som grundlæggende og skriftsproget som afledt.

Der er mindst to mulige forklaringer på, hvorfor den omtalte variation også findes i skriftsproget. Den ene er, at vi, selv når vi skriver, opfører os, som om vi var i en talesituation, og er tilbøjelige til at indsatte *at*, når vi kommer i produktionsvanskeligheder. Den anden og mere plausible er, at de variationsmønstre, der opstår i talesproget som følge af de for talesproget karakteristiske produktionsbetingelser, bliver kognitivt indgraveret (den engelske fagterm er *entrenched*) og således finder vej til skriftsproget gennem en frekvensstyret opfattelse af, om det i en given sproglig kontekst lyder bedst med eller uden *at* (Race & MacDonald 2003).

Kasper Boye
Københavns Universitet
kabo@hum.ku.dk

Mads Poulsen
Københavns Universitet
m.poulsen@hum.ku.dk

NOTER

- 1 DGCSS er akronym for Danmarks Grundforskningsfonds Center for Sociolinguistiske Sprogforandringsstudier
- 2 I samme pilotundersøgelse blev forskellen på de to komplementtagende verber *tro* og *vide* undersøgt i forhold til disfluensvariation, fordi *vide* udviser færre grammatikaliseringssymptomer end *tro* (Boye & Poulsen 2011; jf. Boye & Harder 2007). Undersøgelsen viste en langt større proportion af disfluensindledte komplementsætninger efter kombinationen af *vide* med 1. person singularis-subjekt (29 % komplementsætninger med disfluens) end efter kombinationen af *tro* med samme subjekt (10 % komplementsætninger med disfluens). Dette kunne tyde på, at grammatikaliseringsgrad influerer på disfluensvariation. Men grammatikaliseringsgrad kan som sagt ikke forklare forskellen på proportionen af *at* med og uden disfluens inden for verbet *tro*. Det tyder på, at der er en klar selvstændig relation mellem disfluensvariation og *at*-variation.

LITTERATUR

- Ariel, M. (1990) *Assessing noun-phrase antecedents*. London: Routledge.
- Ariel, M. (2006) "Accessibility theory". K. Brown (red.) *Encyclopedia of language and linguistic*, vol. I, 2. udg. Oxford: Elsevier. 15-18.
- Biber, D, S. Johansson, G. Leech, S. Conrad & E. Finegan (1999) *Longman grammar of spoken and written English*. London: Longman.
- Baayen, R.H. (2008) *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Boye, K. (2008) "Om om". C. Bache, A. Holsting, H.H. Müller & N. Nørgaard (red.) *Ny forskning i grammatik* 15. Odense: Institut for Sprog og Kommunikation, Syddansk Universitet. 7-29.
- Boye, K. & P. Harder (2007) "Complement-taking predicates: Usage and linguistic structure". *Studies in Language* 31(3). 569-606.
- Boye, K. & P. Harder (2009) "Evidentiality: Linguistic categories and grammaticalization". *Functions of Language* 16(1). 9-43.
- Boye, K. & P. Harder (2012) "A usage-based theory of grammatical status and grammaticalization". *Language* 88(1). 1-44.
- Boye, K. & M. Poulsen (2009) "Om at". K. Faro, A. Holsting, N.-E. Larsen, J.E. Mørgensen & T. Vinther (red.) *Sprogvidenskab i glimt: 70 tekster om sprog i teori og praksis*. Odense: Syddansk Universitetsforlag. 125-128.
- Boye, K. & M. Poulsen (2011) "Complementizer deletion in spoken Danish". Indlæg ved Societas Linguistica Europaea – 44th Annual Meeting, Logroño. 8.9.-11.9.2011.
- Brazil, D. (1995) *A grammar of speech*. Oxford: Oxford University Press.
- Cuyckens, H. & C. Shank (2011) "A diachronic perspective on the grammaticalization of zero-complement clauses in English: The case of *think*, *feel*, and *realize*". Manuskript.
- Elsness, J. (1984) "That or zero? A look at the choice of object clause connective in a corpus of American English". *English Studies* 65. 519-33.
- Ferreira, V. & G.S. Dell. (2000) "Effect of ambiguity and lexical availability on syntactic and lexical production". *Cognitive Psychology* 40(4). 296-340.
- Hansen, Aa. (1961) *Vort vanskelige sprog*. København: Grafisk Forlag.
- Hansen, E. & L. Heltoft (2011) *Grammatik over det danske sprog*. København: Det Danske Sprog- og Litteraturselskab.
- Harder, P. & B. Henriksen (2004) "Mundtligt og skriftligt sprog – en skyttegravskrig?" *Sprogforum* 30. 54-60.

- Jaeger, T.F. (2005) "Optional *that* indicates production difficulty: Evidence from disfluencies". *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop. 10-12 September 2005, Aix-en-Provence, Frankrig.*
- Jaeger, T.F. (2008) "Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models". *Journal of Memory and Language* 59(4). 434-446.
- Jensen, A. (2003) *Clause linkage in spoken Danish*. Ph.d.-afhandling, Københavns Universitet.
- Jensen, T.J. (2011) "Ordstilling i ledsætninger i moderne dansk talesprog". P. Durst-Andersen & H.H. Müller (red.) *Ny forskning i grammatik* 18. Odense: Institut for Sprog og Kommunikation, Syddansk Universitet. 123-150.
- Kaltenböck, G. (2009) "Is *that* a filler? On complementizer use in spoken object clauses". *Vienna English Working Papers* 18(1). 28-63.
- Leech, G. (2000) "Same grammar or different grammar? Contrasting approaches to the grammar of spoken English discourse". S. Sarangi & M. Coulthard (red.) *Discourse and social life*. Harlow, England: Longman. 48-65.
- McCarthy, M. (1998) *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.
- Nordström, J. & K. Boye (2011) "Semantic functions of complementizers in Germanic languages". Indlæg ved LANCHART Workshop on Semantic Functions of Complementizers in European Languages, København. 28.10.-29.10.2011.
- Pedersen, I.L. (2000) "De måtte ud af sengen klokken var fem: Adverbielle ledsætninger med tomt konjunktionalfelt i danske dialekter". L. Elmevik. (red.) *Dialekter och folkminnen. Hyllningsskrift till Maj Reinhammar den 17 maj 2000*. Uppsala. 223-230.
- Race, D.S. & M.C. MacDonald (2003) "The use of 'that' in the production and comprehension of object relative clauses". *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.
- Thompson, S.A. & A. Mulac (1991) "A quantitative perspective on the grammaticalization of epistemic parentheticals in English". E.C. Traugott & B. Heine (red.) *Approaches to grammaticalization*, vol. II. Amsterdam: John Benjamins. 313-329.
- Thráinsson, H. (1994) "Icelandic". E. König & J. van der Auwera (red.) *The Germanic Languages*. London: Routledge. 142-189.
- Wasow, T., T.F. Jaeger & D.M. Orr (2011) "Lexical variation in relativizer frequency". H.J. Simon & H. Wiese (red.) *Expecting the unexpected: Exceptions in grammar*. Berlin: Mouton de Gruyter.