

# Træstruktur i en skov af fortolkninger

KEN RAMSHØJ CHRISTENSEN

## ABSTRACT

Denne artikel fremfører først et logisk argument for at det ikke er muligt at redegøre for sprogtilegnelsen og sprogbrugen ved hjælp af funktion og hyppighed alene. Der er også behov for at tage form-delen – dvs. syntaksen – alvorligt, for den er forudsætningen for funktionen. Dernæst gennemgår artiklen resultaterne fra en række danske neuro- og psykolingvistiske undersøgelser der understøtter den formelle lingvistik grundtese, nemlig at det er muligt at adskille sproglig form og sproglig funktion. Forsøgene viser at vi laver usandsynlige fortolkninger undervejs i sprogforståelsen, og at sandsynlighed og hyppighed ikke forklarer hvorfor eller hvordan vi misforstår visse sætningstyper, eller hvorfor vi tilskriver mening til sætninger der ikke har en fast betydning. Selvom visse semantiske og syntaktiske kriterier normalt skal være opfyldt, er det muligt så at sige at kortslutte det sproglige fortolkningssystem og dermed adskille form og funktion. Der argumenteres for at hjernen er en forudsigelsesmaskine der konstant konstruerer og opdaterer mentale modeller af verden omkring os, og at én af disse modeller er sprogets grammatik. Ligesom enhver anden form for søge- eller læringsalgoritme må sprogforståelsen nødvendigvis være styret af visse formelle principper.

**EMNEORD:** parsing, psykolingvistik, neurolingvistik, sprogtilegelse, hyppighed

## 1 INTRODUKTION

Sproglige tegn har som bekendt både form, dvs. grammatisk kategori og struktur, og funktion i form af betydning og anvendelse, herunder anvendelsesfrekvens. Jeg vil i det følgende fremføre et logisk argument for at det ikke er muligt at redegøre for sprogtilegnelsen og sprogbrugen ved hjælp af funktion alene. Der er også behov for at tage form-delen alvorligt, for den er forudsætningen for funktionen. Der-

næst vil jeg gennemgå en række danske publicerede psykolingvistiske forsøg der viser at form (her: syntaks) og funktion (her: semantik og pragmatik) kan skilles ad. Eller rettere, funktionsdelen (betydningen) kan nærmest kortsluttes uden at det bliver bemærket, men dog kun hvis formdelen (syntaksen) er i orden. Desuden viser forsøgene også at hyppighed forudsætter form. Disse resultater understøtter den formelle lingvistik grundtese, nemlig at det er muligt at adskille sproglig form og sproglig betydning/funktion, som også argumenteret indgående for af fx Newmeyer (2003).

Hjernen kan betragtes som en slags forudsigelsesmaskine (Bubic et al. 2010, Clark 2013, Fuster 2015: 398-399). Den forsøger hele tiden at forudsige perception og stimuli fra omverdenen ud fra hvad den ved i forvejen, og derfra træffe beslutning om passende reaktion (*decision making*). Den mentale model af omverdenen opdateres og justeres alt efter om forudsigelserne (forventningerne) passer eller ikke passer (*error detection*). Det forudsætter en vis mængde medfødte kategorier og adfærdsmønstre, men fører på den anden side også til tillærte eller ændrede kategorier og adfærdsmønstre. Interaktionen mellem disse medfødte og tillærte mekanismer danner grundlag for videre forudsigelse og adaption. Én af de modeller som hjernen kontinuerligt danner forudsigelser og opbygger forventninger ud fra, er sprogets grammatik. Det skal for en god ordens skyld allerede fra starten understreges at der naturligvis er behov for såvel en formel beskrivelse som en funktionel beskrivelse af sproget, og at de to tilgange komplementerer hinanden (Newmeyer 2010), et faktum der ofte glemmes i lingvistiske debatter (men se Newmeyer 2010, Togeby 2017).

## 2 SPROGTILEGNELSEN OG DEN SORTE BOKS

Formelle og funktionelle lingvistiske tilgange er (blandt andet) uenige om hvad der ligger til grund for sprogtilegnelsen, hvad der er medfødt, og hvad der er tillært. Det er i bund og grund forskellen mellem nativism (kort sagt: medfødt sprogspecifik viden) og empirisme (kort sagt: ingen medfødt sprogspecifik viden). For empirismen spiller hyppighed en central rolle idet det antages at jo hyppigere man udsættes for noget (sprogligt), jo mere 'indlejret' bliver det. Problemet er bare at en tilgang

til sprogtilegnelsen baseret primært på hyppighed er umulig, men det vender jeg tilbage til i næste afsnit.

Den tidlige behaviorisme (Skinner 1957) var en version af rendyrket empirisme. Der var kun stimulus og respons. Det mellemliggende, det ”mellem ørerne”, var et *black box*-fænomen man ikke kunne udtale sig om. Læring blev anset som dannelse af associationer mellem stimulus, respons og belønning (eller straf) (*reinforcement*). Jo hyppigere en stimulus (og dermed respons og belønning) var, jo stærkere blev associationen – der var altså tale om frekvens- eller hyppighedsbaseret læring. Selvom behaviorismen for længst er gået af mode (efter Chomskys sønderlemmende kritik, Chomsky 1959), er ideen om at mennesket fødes uden specifikke medfødte evner ikke. Empirismen lever stadig.

Michael Tomasello er udtalt modstander af ideen om medfødte sproglige mekanismer og kategorier. Ifølge ham er småbørn gode til at finde mønstre i auditive sekvenser, og denne evne forbereder dem til at tilegne sig grammatiske konstruktioner (Tomasello 2003: 19). Sprogtilegnelsen er desuden baseret på fælles opmærksomhed mellem barn og voksen (*joint attention*), evnen til at læse andres hensigt (*attention reading*) og kulturel læring (*cultural learning*). Ifølge Tomasello er sproget intet andet end en (speciel) slags evne til at bruge fælles opmærksomhed (2003: 21), og børn lærer således sprog, herunder ord og sætningskonstruktioner, på samme måde som de lærer alle andre slags kulturelle konventioner (2003: 90).

The most important point is that constructions are nothing more or less than patterns of usage, which may therefore become relatively abstract if these patterns include many different kinds of specific linguistic symbols. But never are they empty rules devoid of semantic content or communicative function (Tomasello 2003: 100).

De tidligste ”konstruktioner” som børn ifølge Tomasello tilegner sig, er ”verbum-øer” (*Verb Islands*): ”each verb seemed like its own island of organization in an otherwise unorganized language system” (2003: 117). Sprogtilegnelsen, eller rettere sproglæringen, sker ved hjælp af analogidannelse: ”the pattern-finding skill of analogy, which basically

categorizes together complex wholes on the basis of commonalities in their relational structures” (2003: 193).

When people repeatedly use the same particular and concrete linguistic symbols to make utterances to one another in “similar” situations, what may emerge over time is a pattern of language use, schematized in the minds of users as one or another kind of linguistic category or construction (Tomasello 2003: 99).

Bemærk de mange meget vage termer, der giver anledning til flere spørgsmål end de besvarer: ”ligheder” (*commonalities*), ”gentagne gange” (*repeatedly*), ”samme” (*same*), ”lignende situationer” (*similar situations*), ”kan opstå” (*may emerge*), ”en eller anden slags sproglig kategori eller konstruktion” (*one or another kind of linguistic category or construction*), og sidst men ikke mindst, ”analogidannelse” (*pattern-finding skill of analogy*). Men som Steven Pinker (2004: 951) også har påpeget, er henvisning til analogidannelse tom snak med mindre man specificerer hvordan. Selvom Tomasello ikke mener at læring via ren association kan forklare sprogtilegnelse (fx Tomasello 2003: 20), er det svært at se at hans model er meget forskellig fra en model der er baseret på hyppighed (”samme”, ”gentagne gange”) og en slags ”analogi” (se fx Tomasello 2003: 48, 56, 106-7, 173-175). Det er stort set den samme samling af vage ideer der fremføres af Evans (2014), og hans fremstilling er således også temmelig problematisk, som også påpeget af blandt andre Allott og Rey (2017) og Adger (2015a, 2015b). Denne neo-empiristiske tilgang gør os således ikke klogere på hvad der sker i ’den sorte boks’.

Et forsøg på at beskrive og forklare hvordan læringsprocessen foreløber, altså hvad der foregår inde i ’den sorte boks’, er *connectionismen* (også kaldet parallelt distribueret processering), der baserer sig på computerbaserede såkaldte kunstige neurale netværker (Elman et al. 1996, Reilly og Sharkey 1992). Netværker lærer at associere stimuli som det udsættes for hundredvis af gange, ved hjælp af feedback der som sådan svarer til behaviorismens *reinforcement*, dvs. belønning eller straf. Der er ingen medfødte symbolske repræsentationer (sproglige kategorier, såsom verbum, substantiv, etc.) eller sprogspecifikke strukturelle principper. Netværket lærer ved hjælp af hyppighed (frekvens), generel læring (via association) og feedback. En sprogbrugers viden om

sit sprog forstås her som en statistisk samling af sproglige erfaringer ("a statistical ensemble of language experiences") der konstant ændrer sig ved brug (Ellis 2002: 64) – og det til trods for at connectionismen ser sig selv om et modsvar til behaviorismen, eftersom det jo netop handler om netværker inde i 'den sorte boks' (Elman et al. 1996: 103-104). Problemet er bare at for at disse generiske, domæne-generelle netværker overhovedet skal kunne gøre noget som helst, er man nødt til at specificere hvad netværket skal lære, dvs. give det medfødte kategorier, regler eller konstruktioner, se fx Pinker (1997: 112-31, 2002) og Newmeyer (2003, 2005). Neoempiristiske henvisninger til association, hyppighed og analogi er utilstrækkelige, og der er således brug for medfødte læringsmekanismer dedikeret til sprogtilegnelsen. Hertil kommer at antallet af ord og deres potentielle kombinationsmuligheder, altså det som der skal indlæres, er ubegrænset.

### *2.1 Et væld af sætninger*

For at illustrere hvilken enorm udfordring sprogtilegnelsen er, og hvorfor jeg også mener at en hyppighedsbaseret tilgang er utilstrækkelig, vil jeg først vise hvor stort omfanget af mulige sætninger er.<sup>1</sup> Dernæst vil jeg beskrive den ligeledes enorme udfordring der ligger i at finde ud af hvordan ordene hænger sammen.

For det første er sprogets samlede ordforråd enormt. Det er også uhyre svært at fastslå cirka hvor stort en enkelt persons ordforråd er, for det kommer an på hvordan man definerer ord; om man tæller bøjede former, slangord og tekniske udtryk med (Ringgaard 1988). Hertil kommer at sproget varierer og konstant forandrer sig. Det er dog blevet anslået at en gennemsnitlig 20-årig kender et sted mellem 17.000 og 42.000 ord plus slang, navne, polysemer, idiomatiske udtryk og bøjede former (Brysbaert et al. 2016, D'Anna et al. 1991). Mange voksne har et ordforråd på 50-70.000 (Kristensen 2016), nogle sågar kæmpemæssige ordforråd på over 100.000 ord (Bloom 2001: 1098). Det er således ikke usandsynligt at en gennemsnitlig voksen sprogbruger kender ca. 20.000 opslagsord.

---

1 Det er selvfølgelig en tilsnigelse, for der er teoretisk set et uendeligt antal mulige sætninger.

Lad os tage udgangspunkt i følgende sætning:

- (1) Hendes kat jagtede den lille mus.

Hvis vi antager at fordelingen mellem ordklasserne er nogenlunde den samme som i engelsk, er ca. hvert syvende et verbum, hvert andet et substantiv og hvert fjerde et adjektiv (Oxford Dictionaries 2016)<sup>2</sup>. Lad os ydermere for nemheds skyld kun medregne 18 artikler: *en, et, den, det, de, denne, dette, disse, min, mine, din, dine, hans, hendes, dens, vores, jeres, deres*.<sup>3</sup> Selv hvis vi antager at det kun er én ud af en million potentielt mulige sætninger der er grammatisk korrekt og giver mening, og at man kun udsættes for sådanne velformede sætninger – dvs. ingen ’performans-fejl’ såsom talefejl, grammatiske fejl, langstrakte pauser, etc. (hvilket er umuligt, jf. Aitchison 2008: kapitel 11, Harley 2010: kapitel 8) – så er antallet af logisk mulige sætninger af typen i (1) ca. 24,4 millioner milliarder. Et astronomisk antal der ville tage lidt over 1,5 milliard år at lære hvis man lærte en sætning hvert andet sekund 24 timer i døgnet. Det er naturligvis ikke sådan sprogtilegnelsen forløber, men det sætter problemet i perspektiv. Når man så tager med i betragtningerne at sproget er kombinatorisk (lyde, morfemer og ord kan kombineres, og måden de kombineres på er betydningsbærende, jf. *Han kyssede hende* vs. *Hun kyssede ham*) og rekursivt (fx kan sætninger indeholde andre sætninger, såsom *De sagde at de mente at hun så at han kyssede hende*), og dermed har uendeligt mange kombinationsmuligheder (Chomsky 2005: 11), ja, så er opgaven absolut umulig. Medmindre, selvfølgelig, at sprogtilegnelsen, altså det at lære sit modersmål, på en eller anden måde er styret således at vi ikke behøver at lære sproget én sætning ad gangen. Vi er nødt til for det første at finde ordene, gruppere dem og generalisere til en struktur (eller konstruktion).

---

2 Jeg har ikke kunnet finde fordelings tallene for dansk, men ifølge Sproget.dk (2016) er ca. 70 % af ordene i Retskrivningsordbogen substantiver, mens det er ca. 50 % i Oxford Dictionaries. Denne forskel kan være tilfældig og skyldes at ordbøgerne bare ikke er lige store. Der er i hvert fald ikke nogen umiddelbar grund til at antage at der skulle være nogen betydelig forskel.

3 Denne liste mangler fx den tomme artikel for ubestemt flertal, som i heste, de interrogative artikler, såsom *hvilken* eller *hvafforen*, og kvantorerne, fx *alle, ingen, nogle, mange* og *få*; vi ignorerer også alle talordene der ville gøre antallet uendeligt.

## 2.2 En skov af træer

En af de enormt store udfordringer i sprogtilegnelse og læring (udover de uendeligt mange mulige sammenstillinger af ord) er at finde ud af hvordan de sproglige elementer skal grupperes. Hvordan man kan opdele en streng i mindre enheder (konstituenten), det såkaldte *bracketing problem* (Weisstein 1999a). Problemet er altså hvordan man kommer til den rette analyse der passer til den intenderede mening, blandt alle de andre analyser – eller hvordan man finder det rigtige træ i en skov af potentielt mulige træer. Dette problem skal løses før det er muligt at generalisere (eller ”analogisere”) til en eller anden konstruktion.<sup>4</sup> For nemheds skyld vil jeg igen holde mig til ordniveauet, selvom det som nævnt kan være svært at definere præcist hvad et ord er. Jeg ignorerer desuden den åbenlyse kompleksitet at der i mange sætninger kan være diskontinuerte konstituenten, det der i generativ grammatik analyseres som ’flytning’. I sætning (2) er objektet delt op, og den ene del (*en bog*) står på objektets ’normale’ plads lige efter verbet, mens den anden (understregede) del er flyttet ud til venstre i ekstraposition efter adverbialt *sidste år*. Sådanne ordstillingsvariationer, der også ses i fremflytninger såsom *hv-spørgsmål* (3) og fokalisering/topikalisering (4), får antallet af mulige strukturelle analyser til at eksplodere, og som jeg vil vise, er antallet af muligheder i forvejen stort.

- (2) Jeg læste [en bog \_\_\_\_] sidste år der helt klart beskrev hvordan problemet skulle løses.
- (3) Hvem har du snakket med \_\_\_\_ i dag?
- (4) Min gamle computer må du da godt låne \_\_\_\_ igen.

For et givet antal ord (eller andre symboler) kan man beregne antallet af mulige konstituentanalyser, dvs. antallet af mulige kombinationsmuligheder uden at bytte om på ordenes rækkefølge. Tag fx sætningerne i (5) og (6), der begge synes at være helt almindelige og hverdagsagtige.

- (5) Tror du at det kommer til at sne igen i morgen?

---

<sup>4</sup> Bemærk at det samme problem gør sig gældende uanset om man anvender et syntaktisk eller et semantisk perspektiv, og om man foretrækker at lave kassedigrammer eller feltanalyser fremfor træstrukturer (Vikner og Jørgensen 2017).

- (6) Hvad med at vi smutter ud i køkkenet og ser om der er flere af de der lækre kager?

I sætning (5) er der 11 ord, og det betyder at der er 518.859 potentielle måder at tilskrive den en struktur på hvis vi antager at ord danner konstituenten der består af to eller flere ord, og at konstituenten kan indgå i større konstituenten der selv består af to eller flere konstituenten. Sætning (6) består af 18 ord hvilket giver mere end 300 milliarder mulige analyser!<sup>5</sup> Til sammenligning er der 'kun' omkring 100 milliarder stjerner i vores galakse, Mælkevejen (ESA 2016).

Hvis vi antager princippet om binær forgrening (engelsk *binary branching*), dvs. hvor ordene eller symbolerne grupperes rekursivt to og to, som er standardantagelsen inden for de fleste versioner af generativ grammatik (fx Adger 2003: 55, Carnie 2010: 120, Radford 2004: 70), reduceres antallet radikalt<sup>6</sup>. Sætning (5) har så ca. 17.000 mulige analyser, mens (6) har lige under 130 millioner – stadigvæk enormt mange muligheder.

Det er altså ikke bare ligetil at lære simple sætninger og finde ud af hvordan de enkelte ord hænger sammen. Mængden af mulige men forkerte analyser er nærmest ufattelig, og disse bogstaveligt talt astronomiske tal gør det klart at sprogindlæringen (og sprogforarbejdning, *processing*, generelt) nødvendigvis må være styret og dermed afgrænset.

Her kunne man måske fristes til at sige at en stor del af fejlanalyserne undgås pga. hyppighed (frekvens) og kollokation. Det er vigtigt at bemærke at selvom der argumenteres for at der er behov for styring og afgrænsning, dvs. en slags 'medfødt' grammatik, så betyder det ikke at input eller stimuli og dermed hyppighed ikke spiller en vigtig rolle. Der er ingen tvivl om at hyppighed spiller en stor rolle i sproglæring såvel som i sprogbrug og sprogforandring, men det betyder ikke at hyppighed skal eller kan være det eneste styrende princip eller det mest centrale (Yang 2004, 2015). Som Jerry Fodor skrev i 1985:

---

5 Antallet af mulige strukturer beregnes ved hjælp af det såkaldte "super Catalan number". For  $n$  symboler er der  $S_n = (3(2n-3)S_{n-1} - (n-3)S_{n-2})/n$  muligheder (Weisstein 1999b).

6 Spørgsmålet om antallet af binære strukturer for  $n$  symboler kaldes Catalans problem, efter Eugène Charles Catalan (1814-1894) der løste problemet matematisk. For  $n$  symboler er der  $C(n-1)$  mulige strukturer, hvor  $C = (2n)!/(n+1)!n!$ , således at der for fire symboler er fem binære strukturer (Stanley og Weisstein 1999).



No nativist has ever supposed that innate capacities are unaffected or unformed by environmental interactions. [...] How many times has this point been made in the last 30 years? How many times in the last three centuries? How many times is one going to have to make it again? (Fodor 1985: 36).

Enhver organisme, mekanisme eller algoritme der skal kunne lære noget, skal nødvendigvis være udstyret med en passende repræsentation af den relevante type data (Yang 2004: 451). Uden en sådan indbygget styring (afgrænsning) ville der ikke være nogen målrettet læring, for der ville ikke være mulighed for at fokusere på hvad der er relevant, og dermed drage de rette konklusioner (generaliseringer) for at lære:

... all learning must be constrained. Without a severely restricted hypothesis space, the learner cannot even get started (Lidz og Waxman 2004: 165).

Det ville Tomasello formodentlig være enig i. Uenigheden går på hvad den indbyggede styring består af:

The problem, as with all problems of induction, is that an infinite number of generalizations are consistent with any finite sample of data. Many curves can be drawn through a set of points, many laws are consistent with a set of observations, and many grammars are consistent with a set of sentences. Therefore any learner who correctly induces a function, theory, or grammar must respect prior ('innate') constraints on its hypothesis space; the data alone are insufficient (Pinker 2004: 451).

Vi må således være udstyret med den rette læringsmekanisme inklusive et kategorisystem af den rette slags der adskiller sig fra generel læring, kontra Tomasello (2003), ellers er vi tilbage hvor vi startede. Sprogtilægnelsen er baseret på et kombinatorisk og rekursivt system af kategorier. Indenfor generativ grammatik kaldes dette system *universal grammar* (UG) (Bolender 2010, Chomsky 1972: 27, Everaert et al. 2015, Pinker 1994: 22, Yang 2004). UG-tilgangen er således fuldstændig i tråd med den ovenfor nævnte udbredte antagelse i kognitionsvidenskaben at hjernen er en forudsigelsesmaskine der konstant konstruerer og opdaterer modeller af verden omkring os. Ligesom enhver form for

læring forudsætter en form for begrænsning eller styring, forudsætter forudsigelse også styring, dvs. en model (eller hypotese).

### 3 PSYKOLINGVISTISKE UNDERSØGELSER

#### 3.1 *En uventet hyppighedseffekt*

Som forklaret spiller hyppighed, dvs. hvor ofte et sprogligt element (fx lyd eller ord) eller en streng af elementer høres, naturligvis også en vigtig rolle i sprogtilegnelsen (selvom jeg argumenterer for at den også er styret af underliggende principper for sproglig form). I dette afsnit vil jeg præsentere en række psykolingvistiske forsøg der viser hvordan hyppighed nogle gange også spiller sammen med syntaksen i sprogforståelsen. Ydermere viser de at form og funktion kan adskilles, og at formen er en forudsætning for funktionen, dvs. fortolkning og hyppighed.

Det antages ofte at flytning ud af en relativsætning (en speciel slags sætningsknode) er universelt umulig (fx Kush et al. 2013, Phillips 2013, Ross 1967), selvom fx Erteschik-Shir (1982) og Engdahl (1982) har argumenteret for at det i hvert fald er muligt i de skandinaviske sprog. I følgende forsøg (Christensen og Nyvad 2014) undersøgte vi flytning ud af relativsætninger på dansk som i (7) hvor objektet til verbet *havde* i relativsætningen er flyttet ud og helt frem i starten af matrixsætningen (de understregede ord er flyttet fra den understregede tomme plads).

(7) Sådan en hund har Pia engang **mødt/set/kendt/kysset** en pensionist [**der** havde \_\_\_\_].

Forsøget bestod af en spørgeundersøgelse der indeholdt 16 forskellige matrixverber, fx *mødt/set*, fordelt ligeligt på i alt 64 sætninger: 32 med flytning og 32 uden, halvdelen med relativsætningen indledt med *der*, den andel halvdel med *som*, plus 18 såkaldte fillers. Der var 112 deltagere i forsøget. Sætningerne var fordelt på fire pseudorandomiserede lister sådan at hver deltager kun så det enkelte matrixverb i en sætning én gang. Deltagerne bedømte sætningerne på en skala fra 1 til 7 (1=helt uacceptabel, 7=helt acceptabel).

Forsøget viste to ting af relevans for denne artikel: For det første viste vores resultater at acceptabiliteten af sætninger som (7) var signifikant lavere end af sætninger uden fremflytning, men det er ikke overraskende. Tidligere undersøgelser har vist at denne slags flytning har en statistisk signifikant negativ virkning på acceptabiliteten pga. belastningen af arbejdshukommelsen, og at det generelt er tilfældet for fuldt ud grammatiske sætninger med fremflytning, såsom topikaliseringer og *hw*-spørgsmål (Christensen et al. 2013a, 2013b, Fanselow og Frisch 2006, Phillips et al. 2005). For det andet viste forsøget at acceptabiliteten afhænger af den relative hyppighed af matrixverbet (hovedverbet i hovedsætningen), hvor hyppighed er målt som antal forekomster af de specifikke ordformer i KorpusDK (<http://ordnet.dk/korpusdk>). Fx er *møde* mere frekvent end *se*, der er mere frekvent end *kende*, som igen er hyppigere end *kysse*: *møde* > *se* > *kende* > *kysse*. Sætningen i (9) vurderedes af testpersonerne som mest acceptabel med *møde* som matrixverbum, i faldende grad med hhv. *se* og *kende*, og mindst med *kysse*. Forsøget havde 16 forskellige matrixverber, men den eneste faktor med signifikant effekt udover flytning var matrixverbets hyppighed. Det er dog mildest talt uklart hvad der teoretisk set skulle forklare dette mønster. Mig bekendt er der ingen lingvistisk teori der forudsiger at hvis et nominalt objekt indeholder en relativsætning, så er det bedst hvis matrixverbet er højfrekvent, men dog ikke umuligt selvom det er lavfrekvent. Alt andet lige er det sådan at jo oftere vi ser eller hører noget, jo hurtige er vi til at genkende det og til at acceptere det (Zajonc 1968); det er også det der kaldes *priming* (Harley 2014: 175). Det gælder også på det syntaktiske niveau – dog kun så længe der er tale om grammatiske sætninger, for det gælder ikke for ugrammatiske sætninger (Christensen et al. 2013a, Sprouse 2007).

### 3.2 Hyppighed, grammatikalitet og anomali

Et andet eksempel på at hyppighed har indflydelse på syntaktisk forarbejdning, er at vi opbygger en vis forventning om hvilke ord der vil optræde sammen, ud fra hvilke ord der oftest optræder sammen: fx går *drikke* fint med *vand* (der jo kan drikkes), mindre godt med *sovs* (der dog godt kan drikkes, men som oftest spises) og endnu mindre med *støv* (der normalt ikke kan drikkes), så hvis man hører/læser verbet *drikke*

forventer man *vand* snarere end *sovs* eller *støv*. Når forventningen ikke opfyldes, udløser det en anomalieffekt i hjernen. Som vist i (8), er det mere sandsynligt at *Hun drak et dejligt glas...* efterfølges af *vand*, mens *sovs*, *støv* og *kunne* i stigende grad er usandsynlige (?/??) og/eller umulige (\* betyder ugrammatisk):

(8) Hun drak et dejligt glas vand/?sovs/??støv/\*kunne.

Anomalieffekten viser sig ved EEG-målinger som en negativ ændring i det elektriske spændingsfelt omkring hjernen 400 millisekunder efter man ser det underlige eller usandsynlige ord (Dudschig et al. 2016, Kutas og Hillyard 1980), og ved fMRI-målinger som øget blodgen-nemstrømning i den del af pandelappen i hjernen der kaldes Brocas område (Christensen og Wallentin 2011). Det at lav acceptabilitet udløser en anomalieffekt, gør sig også gældende på sætnings- eller konstruktionsniveau. I et neurolingvistisk hjerneskanningsforsøg med fMRI-skanning undersøgte vi den såkaldte lokativalternation og sprogforståelse på dansk (Christensen og Wallentin 2011). Nogle verber er kompatible med begge varianter i alternationen, (9), mens nogle kun fungerer med den ene (10), og andre kun med den anden (11):

- (9) a. Han fylder jord i hullet.  
b. Han fylder hullet med jord.
- (10) a. Han kaster græs i vandet.  
b. \*Han kaster vandet med græs.
- (11) a. \*Han blokerer skrald i afløbet.  
b. Han blokerer afløbet med skrald.

I forsøget blev 22 forsøgsparticipanter præsenteret for 96 sætninger (16 af hver af de tre par i (9)-(11)) som de skulle vurdere om gav mening eller ej.

Resultaterne viste som forventet at forsøgsparticipanterne fandt de ugrammatiske sætninger, (10b) og (11a), uacceptable, og at sådanne ugrammatiske sætninger udløste en anomalieffekt i form af forøget

hjerneaktivitet i Brocas område. I dette forsøgsresultat hænger det uventede, dvs. meget lavfrekvente, sammen med ugrammatikalitet, og det er måske fristende at tænke at det er en hyppigheds- eller (u) sandsynlighedseffekt. Andre forsøg viser dog at der laves temmeligt usandsynlige (midlertidige) fortolkninger undervejs i sprogforståelsen, og at det påvirker den overordnede acceptabilitet af ellers fuldt ud grammatiske sætninger (Christensen et al. 2013a, 2013b, Kizach et al. 2013, Nyvad et al. 2014).

### 3.3 Midlertidige (mis)fortolkninger

At den overordnede acceptabilitet påvirkes af midlertidige fortolkninger, kan fx ses ved *hv*-flytning ud af en indlejret sætning (Christensen et al. 2013a), som illustreret i (12) og (13) nedenfor. Dette forsøg bestod af 144 sætninger, alle med *vide* som matrixverbum, mens der var 16 forskellige hovedverber i de indlejrede sætninger (der var 8 forskellige slags sætninger, herunder de to i (12a-b), men det er ikke centralt for indeværende). Deltagerne (60 personer) skulle vurdere sætningerne en efter en i randomiseret rækkefølge på en skala fra 1 til 5 (1=helt uacceptabel, 5=helt acceptabel). (Forsøget havde desuden et replikations-eksperiment med sætninger svarende til (13a-b) med 30 forsøgspersoner og med samme resultater.)

(12) a. Hvad ved hun godt [at man kan leje \_\_\_\_ dér]?

b. Hvor ved hun godt [at man kan leje noget \_\_\_\_]?

(13) a. Hvilken båd foreslog naboen [at vi skulle sælge \_\_\_\_ ret billigt]?

b. Hvor billigt foreslog naboen [at vi skulle sælge vores båd \_\_\_\_]?

Selvom både (12a) og (12b) er grammatisk korrekte, syntes respondenterne/forsøgspersonerne at (12a) var væsentligt mere acceptabel end (12b). Det samme gør sig gældende for (13a) over for (13b). Det skyldes at når vi læser (eller hører) en sætning, sker det fra venstre mod højre, og at vores grammatiske forståelsesapparat, vores *parser*, forsøger at integrere fremflyttede elementer så hurtigt som muligt ved førstkommande hovedverbum. I (12a) er det uproblematisk: *hvad* fortolkes midlertidigt som objekt for *ved*, og det er fuldt ud kompatibelt med

verbets syntaks og semantik. Efterfølgende bliver det klart at det skal være objekt for det indlejrede verbum, *leje*. I (12b) analyseres *hvor* midlertidigt som adverbial for *ved*, hvilket er en umulig situation (jf. *\*Hvor ved hun det? \*Hun ved det i stuen*). Efterfølgende bliver det hurtigt klart at det hører hjemme i den indlejrede sætning som adverbial for *leje*. Denne midlertidige anomali gør at helsætningen, der ellers er grammatisk korrekt, vurderes som mindre acceptabel.

Resultaterne viser altså at sprogforståelsen ikke kun styres af meningen (funktionen). Vi laver visse fortolkninger undervejs i sprogforståelsen uanset om de giver mening eller.

### 3.4 Syntaksen har førsteprioritet

Det bliver endnu mere interessant når vi kigger på hvordan vi parser verber der ikke kan tage en nominalfrase som objekt. I dette forsøg (Kizach et al. 2013) undersøgte vi sætninger med verber som *bemærke* og *formode*, der adskiller sig fra hinanden ved at *bemærke* kan tage enten en nominalfrase eller en sætning som objekt, mens *formode* kun kan tage en sætning, jf. (14) og (15). Der er med andre ord en syntaktisk forskel i verbernes valens, deres specifikke krav til deres komplement. Spørgsmålet er om vi også laver absurde og lavfrekvente midlertidige fortolkninger i (15b) på samme måde som i (12) og (13) ovenfor, eller om syntaksen så at sige styrer udenom og blokerer for den potentielle fejlfortolkning.

- (14) a. Mia **bemærkede** grisen.
- b. \*Mia **formodede** grisen.
  
- (15) a. Mia **bemærkede** grisen i stalden manglede vand.
- b. Mia **formodede** grisen i stalden manglede vand.

I dette forsøg brugte vi G-Maze (Forster et al. 2009), som er et program der viser ordpar, fx *grisen* og *anså*, på computerskærmen, og forsøgspersonen skal så hurtigt som muligt (ved at trykke højre eller venstre på keyboardet) vælge det ord der passer bedst som fortsættelse på den sætning der er ved at blive opbygget, fx *Mia bemærkede* som i (14a) eller (15a). Ved at præsentere sætningerne ord for ord, er det muligt at

måle præcist hvor lang tid det tager (og dermed hvor svært det er) at integrere de enkelte ord i konteksten, og om der er statistisk signifikante forskelle. Vores eksperiment havde 60 deltagere og 16 sætninger (8 sætningspar svarende til det i (15) ovenfor) plus 26 fillers. Det viste sig at det tager længere tid at behandle *formodede* end *bemærkede* umiddelbart efter ordet *Mia*, hvilket formentligt skyldes at parseren (det grammatiske 'forståelsesapparat') forudsiger og konstruerer en sætningsstruktur i stedet for en nominalfrase som objekt, og at en sætningsstruktur alt andet lige er mere kompleks end en nominalfrase. En sætningsstruktur har pladser til subjekt, verbum (eller rettere prædikat), og potentielt et objekt (og evt. adverbialer), hvoraf både subjekt og objekt begge kan være og i dette tilfælde er nominalfraser, henholdsvis *grisen* og *vand*. Det interessante er at det tager signifikant længere tid at behandle ordet *manglede* i (15a) end i (15b) (bemærk at det er det samme ord i begge sætninger). Forklaringen er at i (15a) 'tror' parseren at objektet for *bemærkede* er en nominalfrase, for en nominalfrase er alt andet lige simplere end en sætning og passer fint som objekt til *bemærke*. Men når man kommer til ordet *manglede*, bliver det klart at det var forkert, og at strukturen skal reanalyseres sådan at *grisen i stalden* er subjekt i en indlejret sætning. I (15b), derimod, er det ikke muligt at have en nominalfrase som objekt for *formodede* (verbet kræver en sætning som objekt, jf. (14b)), og derfor analyseres *grisen i stalden* med det samme som subjekt i en indlejret sætning, og der er derfor ikke reanalyse ved ordet *manglede*.

Vi har også anvendt G-Maze til at undersøge *hv*-flytning ud af en indlejret sætning, som i (12) og (13) ovenfor (Nyvad et al. 2014). Resultatet viser igen at vi laver denne midlertidige analyse hvor et fremflyttet element analyseres som komplement eller adjunkt til hovedverbet i hovedsætningen, uanset om det giver mening eller ej.

Forsøgene understøtter den gængse model for sprogforståelsen hvor vi, når vi hører eller læser ord, gradvist opbygger en sætningsrepræsentation, dvs. et syntaktisk træ (eller tilsvarende) (Frazier 1987, Harley 2014: 291-293). Ordene føjes ind i den struktur som parseren er ved at konstruere, uanset om det umiddelbart giver mening eller ej (Fodor og Inoue 1998). Vores forsøg viser desuden at det kun sker hvis denne struktur er kompatibel med verbets subkategorisering. Med andre ord: Vi accepterer det semantisk eller pragmatisk usandsynlige, men ikke

det syntaktisk umulige. Den midlertidige misfortolkning man laver i (12b), *\*Hvor ved hun det?*, er rasende usandsynlig, mens misfortolkningen i (15b), *\*Mia formodede grisen*, er fuldstændig umulig, hvorfor man ikke laver den. Eftersom begge må formodes at have en hyppighed på nul, er det meget svært at se hvordan en hyppighedsbaseret model ville kunne redegøre for hvorfor der er en forskel.

### 3.5 Grammatisk illusion og (mis)forståelse

Sandsynlighed (semantisk plausibilitet) og hyppighed (brugsfrekvens) forklarer heller ikke hvorfor eller hvordan vi tilskriver mening til sætninger der ikke har en fast betydning, eller hvorfor eller hvordan vi konsekvent misforstår visse andre sætninger. Sådanne sætningstyper der konsekvent misforstås, kaldes grammatiske illusioner (Christensen 2010, 2016; Kizach et al. 2015; Phillips et al. 2011). Sammenlign sætningerne i (16) og (17). Mens der formodentlig ikke er nogen tvivl om at (16) er tvetydig, og enten betyder at det var manden med tasken der blev slået, eller at det var med tasken hun slog ham, kan folk ikke blive enige om hvad sætningerne i (17) betyder (Christensen 2011, 2016):

(16) Hun slog manden med tasken.

(17) a. ?Flere folk har været i Paris end jeg har.

b. ?Flere kvinder har været på ferie end mænd har i år.

Jeg har undersøgt om folk finder sætninger som dem i (17) meningsfulde, og i givet fald hvad sætningerne betyder, i en række undersøgelser og eksperimenter, herunder et fMRI-studie med 19 deltagere (Christensen 2010), en uformel spørgeundersøgelse med 63 deltagere (Christensen 2011), en internetspørgeundersøgelse med 545 deltagere og to eksperimenter med henholdsvis 32 og 60 deltagere (Christensen 2016). Resultaterne viser gang på gang at de fleste mennesker bliver narret, dvs. de finder sætningerne acceptable og meningsfulde, men de er ikke enige om betydningen. Sætningerne i (19) er grammatiske illusioner. Jeg har kaldt dem 'blindgyder' (Christensen 2011) eller *dead ends* på engelsk (Christensen 2010, 2016); andre kalder dem 'komparative illusioner' (Phillips et al. 2011).



For at se hvordan illusionen virker, så tag fx (17a). Den ser umiddelbart grammatisk og meningsfuld ud (i hvert fald for de fleste), men den har ikke nogen entydig sammenhængende betydning. For det første ser det ud som om *end jeg har i* (17a) er en elliptisk ledsætning, men det kan ikke være tilfældet, for der er ikke noget i den forudgående sætningskonstekst der potentielt kunne gentages, såsom i (18) nedenfor hvor *kommer* er udeladt og kan gentages på pladsen markeret med "[...]”:

(18) Den ene **kommer** fra Norge, den anden [...] fra Sverige.

Udover at (17a) ser ud som om der er ellipse (men det er der ikke, det er pseudoellipse), så ser det også ud som om at *end jeg har* er sat i ekstraposition. Dvs. det ser ud til at være flyttet ud til højre, og at det derfor skal fortolkes (rekonstrueres) som hørende til et sted inde i sætningen – på samme måde som i (19) hvor relativsætningen, *der handler om sprog og statistik*, lægger sig til (modificerer) *bog*, ikke *loppemarkedet*:

(19) Jeg fandt en bog \_\_\_\_ på loppemarkedet der handler om sprog og statistik.

Men rekonstruktion virker heller ikke i (17a). Sammenlign (17a), gengivet i (20a), med sætningen i (20b). Hvis man ophæver ellipsen (21) og derefter rekonstruerer i begge sætninger (22), bliver det tydeligt at kun den ene er meningsfuld og grammatisk, mens den anden er noget vås:

(20) a. ?Flere folk har været i Paris end jeg har.

b. Flere folk har været i Paris end i Køge.

(21) a. \*Flere folk har været i Paris end jeg har (været i Paris).

(pseudo-ellipse ophævet)

b. Flere folk har været i Paris end (der har været) i Køge.

(ellipse ophævet)

(22) c. \*Flere folk [end jeg har (været i Paris)], har været i Paris \_\_\_\_.

(rekonstrueret)

- d. Flere folk [end (der har været) i Køge], har været i Paris \_\_\_\_.  
(rekonstrueret)

Eftersom sætninger af typen i (17a)/(20a) ikke har nogen sammenhængende betydning, er det interessant at mange ikke opdager det. I et forsøg med hjerneskaning med fMRI (Christensen 2010) viser det sig at der ikke er nogen anomalieffekt ved denne sætningstype, som man ellers ville forvente i stil med effekten for (8)-(11) beskrevet ovenfor. I stedet er folk meget hurtige til at vurdere dem som grammatiske og meningsfulde (Christensen 2016). Det er også bemærkelsesværdigt at folk ikke bare gætter på en fortolkning. I stedet tilskriver folk sætningerne én ud af et lille sæt af forskellige og gensidigt inkompatible fortolkninger (hvoraf én dog er at sætningen er meningsløs), som vist i (23). Bemærk at det fx varierer om subjektet *jeg* overhovedet har været i Paris. Det er her også værd at bemærke at det er meget anderledes end tvetydigheden i (16) ovenfor (*Hun slog manden med tasken*).

(23) Flere folk har været i Paris...

- a. ... undtagen mig.
- b. ... end bare mig.
- c. ... oftere end jeg har.

Sætningen i (17a) er som sagt en grammatisk illusion: Det syntaktiske forudsigelsesapparat (parseren) bliver 'snydt' og leverer en semantisk fortolkning der ikke er kompatibel med ordene og strukturen. Det er et tilfælde af "overfladisk fortolkning", eller *shallow processing* (Sanford og Sturt 2002), men bemærk at det kun er den semantiske fortolkning der er overfladisk, ikke den syntaktiske. Som beskrevet ovenfor i forbindelse med (12)-(15), tillader parseren at vi laver absurde midlertidige fortolkninger undervejs i sprogforståelsen, men kun så længe at det er kompatibelt med den syntaktiske struktur. At fortolkningen er afhængig af at der er en syntaktisk struktur at hænge den op på, kan bl.a. også ses i de såkaldte *garden path*-sætninger (Christensen 2010, Frazier 1987, Pritchett 1992; jeg har kaldt dem 'vildspor' på dansk, Christensen 2011). En *garden path*-sætning, fx (24) nedenfor, er grammatisk korrekt, selvom folk finder den uacceptabel og meget svær at forstå – parse-

ren bliver ført på vildspor og kan derfor ikke levere en grammatisk velformet syntaktisk struktur. Det viser sig altså at parseren følger én syntaktisk analyse indtil der er grund til at ændre den, og sommetider er denne reanalyse så svær at man synes at sætningen ikke er grammatisk.

- (24) De mest overbebyrdede offentligt ansatte møder læser også bøger.  
(sml.: De mest overbebyrdede folk som de offentligt ansatte møder, læser også bøger.)

Et andet typisk eksempel på en grammatisk men uacceptabel sætning er (25), hvor der er en subjektsrelativsætning indlejret i en anden subjektsrelativ – en såkaldt multipel centerindlejring eller 'løgsætning' ((24) og (25) er taget fra Christensen 2011: 114-115):

- (25) Biler [mænd [kvinder hader], køber], larmer og ruster.

Problemet med (25) er at den overbelaster arbejdshukommelsen, ikke at der er multipel indlejring som sådan. Hvis vi ændrer en lille smule på den inderste sætning ved at gøre den til en subjektsrelativ i stedet for en objektsrelativ, bliver den meget lettere at forstå og er ikke længere uacceptabel: *Biler mænd der hader kvinder, køber, larmer og ruster.* (Betydningen er altså også ændret her, men pointen er at konstruktionen som sådan er acceptabel.)

### 3.5 En overdosis af kompleksitet og anomali

De ovenstående forsøg viser at det er muligt at adskille form og funktion, struktur og mening. Så længe de syntaktiske kriterier for en velformet, grammatisk sætning er opfyldt, vil sprogsystemet levere en semantisk fortolkning, også selvom den er meget usandsynlig eller decideret absurd, som i (12b) og (13b). Talere kan sågar være helt uenige om betydningen af tilsyneladende simple sætninger som dem i (17), hvilket er meget ulig hvordan vi ser på resten af sproget. Dertil kommer at hvis den strukturelle kompleksitet er tilstrækkelig høj, som i (25), daler acceptabiliteten, til tider så langt at vi finder sætningen ugrammatisk selvom den ikke er det. Ydermere viser forsøgene at så længe syntaksen er i orden, er der åbent for forskellige hyppighedseffekter, jf. hvordan

(15a) og (15b) parses forskelligt. Men hvad sker der hvis nu vi i stedet skruer yderligere op for den semantiske kompleksitet?

Eksempel (26) nedenfor er også en illusion, men den er hverken tvetydig som (16) eller ugrammatisk som (17). De fleste læser (26) som ensbetydende med (27), mens den i virkeligheden betyder det stik modsatte (Wason og Reich 1979, Natsopoulos 1985, Kizach et al. 2015). Sætninger som (26) kaldes 'dybdebomber' (engelsk *depth charges*).

(26) Ingen hjerneskode er for ubetydelig til at blive ignoreret.

(27) Ingen hjerneskode er for ubetydelig til at blive **behandlet**.

Illusionen skyldes en sammensværgelse af tre faktorer: en semantisk anomali, en pragmatisk anomali og for mange negationer. Udgangspunktet, eller 'baseline', er (28) nedenfor, der er velformet og let at fortolke. I (29) er der en overtrædelse af den semantiske 'regel' eller generalisering at *for X til Y* betyder *jo mere X, jo mindre Y*, som fx i *Jo mere dumt et spørgsmål er, jo mindre er vi tilbøjelige til at besvare det* i (28). Men det fungerer ikke godt med (29): *Jo mere velmagende en lagkage er, jo mindre er vi tilbøjelige til at spise den*, og det er ret let at se at den er en semantisk anomali.

(28) Ingen spørgsmål er for dumme til at blive besvaret.  
(baseline)

(29) Ingen lagkage er for velmagende til at blive spist.  
(faktor: semantik)

(30) Ingen medicin er for nyttig til at blive hemmeligholdt.  
(faktor: pragmatik)

(31) Intet benbrud er for ufarligt til at blive undgået.  
(faktor: negation)

(32) Ingen advarsel er for ubetydelig til at blive overhørt.  
(dybdebombe)

I (30) er der en åbenlys pragmatisk anomali da det bryder med vores viden om verden at tale om at hemmeligholde medicin. Sætning (31), derimod, indeholder ingen anomalier, hverken semantiske eller pragmatiske. Den er bare meget kompliceret på grund af det høje antal

negative elementer: *intet* (sætningsnegation), *ufarligt* (morfologisk negation) og *undgå* (semantisk negation). I (32) er alle tre faktorer til stede, og sætningen er en dydbombe ligesom (26).

I vores forsøg (Kizach et al. 2015) blev 29 forsøgspersoner præsenteret for 150 sætninger på en skærm en ad gangen, og hver sætning blev efterfulgt af en kort fortsættelse der enten passede eller ikke passede til den, fx kunne (28) blive fulgt af *Derfor besvarer vi ofte/sjældent spørgsmål*. Opgaven bestod i at svare på om det efterfølgende gav mening, og på om forsøgspersonen var sikker på sit svar. Resultatet kan beskrives med hierarkiet i (33) hvor '>' betyder 'udløste signifikant flere fejl end' (forkellen mellem (30) og (31) var ikke statistisk signifikant, derfor er der komma mellem dem):

(33) dydbombe (32) > {pragmatik (30), negation (31)} > semantik (29) > baseline (28).

Som forventet udløste dydbomber som (32) flest fejl. Faktisk svarede folk forkert over 60 % af gangene (dvs. at de angav en fortsættelse som meningsfuld selvom den logisk set ikke var det). Interessant nok var forsøgspersonerne generelt sikre på deres svar, med undtagelse af (31). Det er meget svært at forstå flere negationer (Sherman 1976), og det gjorde folk signifikant mindre sikre på deres svar. Der var dog ingen forskel på de andre fire typer. Selvom folk svarede forkert over 60 % af gangene med dydbomber som (32), var de overbevist om at de svarede rigtigt.

## 4 KONKLUSIONER

Jeg har argumenteret for sprogtilegnelsen nødvendigvis må være styret af slags grammatik (fx universalgrammatikken UG) der også indeholder ordkategorierne. En model for sprogtilegnelse udelukkende baseret på hyppighed (stimulus) og analogidannelse, er utilstrækkelig. Uden en form for styring vil det ikke være muligt at generalisere til andre konstruktioner og den sprogspecifikke grammatik og dens anvendelse. Der er simpelthen for mange potentielle men forkerte generaliseringer der skal undgås.

Denne artikel har diskuteret resultaterne fra en række neuro- og psykolingvistiske undersøgelser der understøtter den formelle lingvistik grundtese, nemlig at det er muligt at adskille sproglig form og sproglig betydning/funktion. Sprogforståelsen (i betydningen parsing) må nødvendigvis være styret af visse principper, ligesom enhver anden form for søge- eller læringsalgoritme. Beregninger viser at hvis hjernen kun tager udgangspunkt i at sproglige symboler, fx morfemer eller ord, kan kombineres, er mængden af mulige fejlanalyser ufattelig stor, bogstaveligt talt astronomisk. Man kunne måske fristes til at sige at en stor del af fejlanalyserne undgås pga. hyppighed (frekvens) og kollokation idet vi opbygger en vis forventning om hvilke ord der optræder sammen. Når forventningerne så ikke opfyldes, udløser det en anomalieffekt i hjernen. De nævnte forsøg viser dog at vi i visse kontekster laver usandsynlige fortolkninger i real-time undervejs i sprogforståelsen, og at sandsynlighed og hyppighed heller ikke forklarer hvorfor eller hvordan vi misforstår visse sætningstyper, eller hvorfor vi tilskriver mening til sætninger der ikke har en fast betydning. De psykolingvistiske eksperimenter viser på den ene side at visse semantiske og syntaktiske kriterier skal være opfyldt, hvilket passer med at form og funktion går hånd i hånd. På den anden side viser forsøgene også at det er muligt at kortslutte det sproglige fortolkningssystem og dermed adskille de to. Alt tyder således på at der er behov for såvel en formel beskrivelse som en funktionel beskrivelse af sproget, og at de to tilgange komplementerer hinanden. Der er dog tilfælde hvor den funktionelle beskrivelse må give fortabt, som fx hvor den semantiske og/eller pragmatiske fortolkning kan være (endda meget) overfladisk så længe den syntaktiske struktur er i orden.

Ken Ramshøj Christensen  
Institut for Kommunikation og Kultur  
Aarhus Universitet  
krc@cc.au.dk

## LITTERATUR

- Adger, D. 2003. *Core syntax: a minimalist approach*. Oxford: Oxford University Press.
- Adger, D. 2015a. Mythical myths: Comments on Vyvyan Evans' "The Language Myth." *Lingua* 158. 76–80. DOI:10.1016/j.lingua.2015.02.006.
- Adger, D. 2015b. More misrepresentation: A response to Behme and Evans 2015. *Lingua* 162. 160–166. DOI:10.1016/j.lingua.2015.05.005.
- Aitchison, J. 2008. *The articulate mammal. An introduction to psycholinguistics*. 5. udg. London: Routledge.
- Allott, N. & G. Rey. 2017. The many errors of Vyvyan Evans' The Language Myth. *The Linguistic Review* 34(3). 1–20. DOI:10.1515/tlr-2017-0011.
- Bloom, P. 2001. Précis of how children learn the meanings of words. *Behavioral and Brain Sciences* 24(6). 1095–1103. DOI:10.1017/S0140525X01000139.
- Bolender, J. 2010. Universal grammar as more than a programmatic label. *Lingua* 120(12). 2661–63. DOI:10.1016/j.lingua.2010.03.010.
- Brysbaert, M., M. Stevens, P. Mander & E. Keuleers. 2016. How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology* 7 (artikel nr. 1116). 1–11. DOI:10.3389/fpsyg.2016.01116.
- Bubic, A., D.Y. von Cramon & R.I. Schubotz. 2010. Prediction, cognition and the brain. *Frontiers in Human Neuroscience* 4. 1–15. DOI:10.3389/fnhum.2010.00025.
- Carnie, A. 2010. *Constituent structure*. 2. udg. Oxford: Oxford University Press.
- Chomsky, N. 1959. Review: Verbal Behavior by B. F. Skinner. *Language* 35(1).: 26–57. DOI:10.2307/411334.
- Chomsky, N. 1972. *Language and mind. Enlarged edition*. San Diego: Harcourt Brace Jovanovich.
- Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI:10.1162/0024389052993655.
- Christensen, K.R. & A.M. Nyvad. 2014. On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37(1). 29–45. DOI:10.1017/S0332586514000055.
- Christensen, K.R. & M. Wallentin. 2011. The locative alternation: distinguishing linguistic processing cost from error signals in Broca's region. *NeuroImage* 56(3). 1622–31. DOI:10.1016/j.neuroimage.2011.02.081.
- Christensen, K.R. 2010. Syntactic reconstruction and reanalysis, semantic dead ends, and prefrontal cortex. *Brain and Cognition* 73(1). 41–50. DOI:10.1016/j.bandc.2010.02.001.

- Christensen, K.R. 2011. Flere folk har været i Paris end jeg har. I.S. Hansen & P. Widell (red.), *13. Møde om Udforskningen af Dansk Sprog*. 113–36. Aarhus: Aarhus Universitet. [http://muds.dk/rapporter/MUDS\\_13.pdf](http://muds.dk/rapporter/MUDS_13.pdf).
- Christensen, K.R. 2016. The dead ends of language: the (mis)interpretation of a grammatical illusion. S. Vikner, H. Jørgensen & E. Van Gelderen (red.), *Let Us Have Articles Betwixt Us – Papers in Historical and Comparative Linguistics in Honour of Johanna L. Wood*, 129–60. Aarhus: Aarhus Universitet. <http://ebooks.au.dk/index.php/aul/catalog/download/119/107/466-1?inline=1>.
- Christensen, K.R., J. Kizach & A.M. Nyvad. 2013a. Escape from the Island: grammaticality and (reduced) acceptability of *wh-island violations in danish*. *Journal of Psycholinguistic Research* 42(1). 51–70. DOI:10.1007/s10936-012-9210-x.
- Christensen, K.R., J. Kizach & A.M. Nyvad. 2013b. The processing of syntactic islands – an fMRI study. *Journal of Neurolinguistics* 26(2). 239–51. DOI:10.1016/j.jneuroling.2012.08.002.
- Clark, A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36(03). 181–204. DOI:10.1017/S0140525X12000477.
- D’Anna, C., E. Zechmeister & J. Hall. 1991. Toward a meaningful definition of vocabulary size. *Journal of Literacy Research* 23(1). 109–22 DOI:10.1080/10862969-109547729.
- Dudschig, C., C. Maienborn & B. Kaup. 2016. Is there a difference between stripy journeys and stripy ladybirds? The N400 response to semantic and world-knowledge violations during sentence processing. *Brain and Cognition* 103. 38–49. DOI:10.1016/j.bandc.2016.01.001.
- Ellis, N.C. 2002. Frequency Effects in Language Processing. *Studies in Second Language Acquisition* 24(2): 143–188. DOI:10.1017/S0272263102002024.
- Elman, J.L., A. Karmiloff-Smith, E. Bates, M. Johnson, D. Parisi, & K. Plunkett. 1996. Rethinking innateness: *A connectionist perspective on development*. *Neural network modeling and connectionism*. Cambridge, Mass.: MIT Press.
- Engdahl, E. 1982. Restrictions on unbounded dependencies in Swedish. E. Engdahl & E. Ejerhed (red.), *Readings on Unbounded Dependencies in Scandinavian Languages*, 151–174. Stockholm: Almqvist & Wiksell International.
- Erteschik-Shir, N. 1982. Extractability in Danish and the pragmatic principle of dominance. E. Engdahl & E. Ejerhed (red.), 1982. *Readings on Unbounded Dependencies in Scandinavian Languages*, 175–191. Stockholm: Almqvist & Wiksell International.



- ESA. 2016. How many stars are there in the universe? *European Space Agency*. (Tilgået 11. november 2016). [http://www.esa.int/Our\\_Activities/Space\\_Science/Herchel/How\\_many\\_stars\\_are\\_there\\_in\\_the\\_Universe](http://www.esa.int/Our_Activities/Space_Science/Herchel/How_many_stars_are_there_in_the_Universe).
- Evans, Vyvyan. 2014. *The language myth: why language is not an instinct*. Cambridge: Cambridge University Press.
- Everaert, M.B.H., M.A.C. Huybregts, N. Chomsky, R.C. Berwick & J.J. Bolhuis. 2015. Structures, not strings: linguistics as part of the cognitive sciences. *Trends in Cognitive Sciences* 19(12). 729–43. DOI:10.1016/j.tics.2015.09.008.
- Fodor, J.A. 1985. Précis of the modularity of mind. *Behavioral and Brain Sciences* 8(1). 1–42. DOI:10.1017/S0140525X0001921X.
- Fodor, J.D., & A. Inoue. 1998. Attach anyway. J.D. Fodor & F. Ferreira (red.), *Reanalysis in Sentence Processing*, 101–41. Springer Netherlands. DOI:10.1007/978-94-015-9070-9\_4.
- Forster, K.I., C. Guerrero & L. Elliot. 2009. The maze task: measuring forced incremental sentence processing time. *Behavior Research Methods* 41(1). 163–71. DOI:10.3758/BRM.41.1.163.
- Frazier, L. 1987. Sentence processing: A tutorial review. M. Coltheart (red.), *Attention and Performance XII*, 559–86. Hove: Lawrence Erlbaum Associates.
- Fuster, Joaquin. 2015. *The prefrontal cortex*. 5. udg. Amsterdam: Academic Press/Elsevier.
- Global Language Monitor. 2014. Number of words in the English language: 1,025,109.8. (Tilgået 4. januar 2014). <http://www.languagemonitor.com/number-of-words/number-of-words-in-the-english-language-1008879/>.
- Harley, T.A. 2010. *Talking the talk: language, psychology and science*. Hove: Psychology Press.
- Harley, T.A. 2014. *The psychology of language: from data to theory*. 4. udg. Hove: Psychology Press.
- Kizach, J, A.M. Nyvad & K.R. Christensen. 2013. Structure before meaning: sentence processing, plausibility, and subcategorization. *PLoS ONE* 8(10). e76326. DOI:10.1371/journal.pone.0076326.
- Kizach, J, K.R. Christensen & E. Weed. 2015. A verbal illusion: now in three languages. *Journal of Psycholinguistic Research*, 1–16. DOI:10.1007/s10936-015-9370-6.
- Kristensen, K. 2016. Lidtomordforråd. Sproget.dk. Accessed October 11. <http://sproget.dk/temaer/ord-og-bogstaver/kjeld-kristensen-lidt-om-ordforraad.html>.
- Kush, D., A. Omaki og & N. Hornstein. 2013. Microvariation in islands? J. Sprouss & N. Hornstein (red.), *Experimental syntax and island effects*, 239–64. Cambridge: Cambridge University Press.

- Kutas, M., & S.A. Hillyard. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207(4427). 203–205. DOI:10.1126/science.7350657.
- Lidz, J., & S. Waxman. 2004. Reaffirming the poverty of the stimulus argument: A reply to the replies. *Cognition* 93 (2): 157–65. DOI:10.1016/j.cognition.2004.02.001.
- Natsopoulos, D. 1985. A verbal illusion in two languages. *Journal of Psycholinguistic Research* 14(4). 385–397. DOI:10.1007/BF01067882.
- Newmeyer, F.J. 2003. Grammar is grammar and usage is usage. *Language* 79(4). 682–707. DOI:10.1353/lan.2003.0260.
- Newmeyer, F.J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. New York: Oxford University Press.
- Newmeyer, F.J. 2010. Formalism and functionalism in linguistics. *Wiley Interdisciplinary Reviews: Cognitive Science* 1(3). 301–7. DOI:10.1002/wcs.6.
- Nyvad, A.M., J. Kizach & K.R. Christensen. 2014. (Non-)arguments in long-distance extractions. *Journal of Psycholinguistic Research*, 1–13. DOI:10.1007/s10936-014-9300-z.
- Oxford Dictionaries. 2016. How many words are there in the English language? (Tilgæt 20. september 2016). <https://en.oxforddictionaries.com/explore/how-many-words-are-there-in-the-english-language>.
- Phillips, C. 2013. Some arguments and nonarguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 28. 156–187. DOI:10.1080/01690965.2010.530960.
- Phillips, C., M.W. Wagers & E.F. Lau. 2011. Grammatical illusions and selective fallibility in real-time language comprehension. J.T. Runner (red.), *Experiments at the Interfaces*, 147–80. Bingley: Emerald Group Publishing.
- Phillips, C., N. Kazanina & S.H. Abada. 2005. ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22.3, 407–428. DOI:10.1016/j.cogbrainres.2004.09.012.
- Pinker, S. 1994. *The language instinct: the new science of language and mind*. New York: Penguin.
- Pinker, S. 1997. *How the mind works*. New York: Norton.
- Pinker, S. 2002. *The blank slate: The modern denial of human nature*. New York: Viking.
- Pinker, S. 2004. Clarifying the logical problem of language acquisition. *Journal of Child Language* 31(4). 949–53. DOI:10.1017/S0305000904006439.

- Pritchett, B.L. 1992. *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.
- Radford, A. 2004. *Minimalist syntax: exploring the structure of English*. Cambridge: Cambridge University Press.
- Reilly, R.G., & and N.E. Sharkey (red.). 1992. *Connectionist approaches to natural language processing*. Hove: Lawrence Erlbaum Associates Ltd.
- Ringgaard, K. 1988. Hvor mange ord er der i dansk? *Nyt Fra Sprognavnet*, 1. 1–6.
- Ross, J.R. 1967. *Constraints on variables in syntax*. Ph.d.-afhandling. Massachusetts Institute of Technology.
- Sanford, A.J. & P. Sturt. 2002. Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences* 6(9). 382–386. DOI:10.1016/S1364-6613(02)01958-7.
- Sherman, M.A. 1976. Adjectival negation and the comprehension of multiply negated sentences. *Journal of Verbal Learning and Verbal Behavior* 15(2). 143–57. DOI:10.1016/0022-5371(76)90015-3.
- Skinner, B.F. 1957. *Verbal behavior*. New York: Appleton-Century-Crofts, Inc.
- Sproget.dk. 2016. Substantiver. (Tilgået 11. oktober 2016). <http://sproget.dk/raad-og-regler/typiske-problemer/substantiver>.
- Sprouse, J. 2007. Continuous acceptability, Categorical grammaticality, and Experimental syntax. *Biolinguistics* 1. 123–34.
- Stanley, R. & E.W. Weisstein. 1999. Catalan number. *MathWorld – A Wolfram Web Resource* (Tilgået 22. september 2016). <http://mathworld.wolfram.com/Catalan-Number.html>.
- Togeby, O. 2017. Formel og funktionel grammatik. *NyS: Nydanske Sprogstudier* 52–53. 107–134.
- Tomasello, M. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.
- Vikner, S. & H. Jørgensen. 2017. En formel vs. en funktionel tilgang til dansk sætningsstruktur. *NyS: Nydanske Sprogstudier* 52–53. 135–168.
- Wason, P.C. & S.S. Reich. 1979. A verbal illusion. *Quarterly Journal of Experimental Psychology* 31(4). 591–97. DOI:10.1080/14640747908400750.
- Weisstein, E.W. 1999a. Bracketing. *MathWorld – A Wolfram Web Resource* (Tilgået 4. november 2015). <http://mathworld.wolfram.com/Bracketing.html>.
- Weisstein, E.W. 1999b. Super Catalan number. *MathWorld – A Wolfram Web Resource* (Tilgået 22. september 2016). <http://mathworld.wolfram.com/SuperCatalan-Number.html>.

- Yang, C.D. 2004. Universal grammar, statistics or both? *Trends in Cognitive Sciences* 8(10). 451–56. DOI:10.1016/j.tics.2004.08.006.
- Yang, C.D. 2015. For and against frequencies. *Journal of Child Language* 42(2). 287–93. DOI:10.1017/S0305000914000683.
- Zajonc, R.B. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology* 9(2, Pt.2). 1–27. DOI:10.1037/h0025848.