

Findes morfemer? Om morfologiens rolle i sprogprocessering

LAURA WINTHER BALLING

Denne artikel giver et overblik over morfemets rolle i menneskelig sprogprocessering på basis af teoretisk morfologi, psykolingvistik og undersøgelser af dansk ordgenkendelse. Jeg beskriver først hvordan teoretisk morfologi arbejder ud fra enten morfembaserede eller ordbaserede modeller, trods problemer med både morfemet og ordet som teoretiske størrelser. Efterfølgende ser jeg på to relaterede men dog adskilte psykolingvistiske traditioner der begge fokuserer på morfologi: for det første den såkaldte datidsdebat der står mellem på den ene side modeller der skelner skarpt mellem regelmæssige og uregelmæssige ord, og på den anden side modeller der postulerer den samme processeringsmekanisme for alle typer ord; for det andet en bredere ordgenkendelseslitteratur der, i modsætning til datidsdebatten og dens fokus på engelske bøjningsformer, undersøger forskellige sprog og forskellige typer morfologi. Dernæst opsummeres resultaterne fra en række eksperimenter om dansk ordgenkendelse som overordnet set tyder på en funktionel rolle for morfologien. Samtidig viser forskelle mellem forskellige typer ord, komplekse frekvens effekter og effekter af det lydige signals overlap med andre ord i sproget et komplekst billede af morfologiens rolle. Til slut opsummeres morfologiens funktionelle rolle i sprogprocesseringen, og de videre perspektiver opridses med en opfordring til et kritisk blik på psykolingvistikens traditionelle metoder og på både ordet og morfemet som teoretiske størrelser.

EMNEORD: morfologi, psykolingvistik, ordforråd, ordgenkendelse

1 INDLEDNING

Nogle vil nok mene at det er at foregribe begivenhedernes gang, men som service til den type læser der gerne springer frem til knaldroma-

nens sidste sider for at se hvordan det hele ender, kan overskriftens spørgsmål besvares med et rungende *nja*. Mest nej, fordi morfemet er problematisk som teoretisk størrelse (se afsnit 2) og uholdbar som grundlæggende enhed for mentale processer, som det vil fremgå af de psykolingvistiske data fra forskellige sprog jeg diskuterer i afsnit 3. Men også ja, fordi morfemet stadig er en nyttig deskriptiv størrelse, og fordi morfologiske strukturer og relationer i ordforrådet spiller en betydelig rolle for hvordan vi forstår og husker ord, som beskrevet på basis af undersøgelser af dansk i afsnit 4. Perspektiverne i dette og mulige videre undersøgelser opridses i afsnit 5.

Men først to primære grunde til at titlens spørgsmål er vigtigt. For det første er morfemet en deskriptiv enhed som spiller en fremtrædende rolle når vi skal beskrive og lære nye sprog: Lægfolk kan tænke på folkeskolens tyskundervisning, i hvert fald som den har udspillet sig i store dele af det 20. århundrede; lingvister for eksempel på beskrivelsen af ukendte sprog. Når et begreb står centralt i sprogbeskrivelsen, bliver det også et interessant spørgsmål om det har psykologisk realitet for (naive) sprogbrugere.

For det andet har spørgsmålet om morfemets rolle i mentale processer udgjort en vigtig kampplads i psykolingvistikken og de kognitive videnskaber mere generelt, siden i hvert fald Rumelhart og McClelland (1986). Accept af morfemet som grundlæggende enhed for sprogprocessering er – mere eller mindre eksplicit og med mere eller mindre god grund – blevet forbundet med accept af medfødt universal grammatik, regelbaseret sprogforståelse og -produktion, symbolsk, modulær kognition og videnskabelig rationalisme, jf. for eksempel Pinker (1999). Omvendt er afvisning af morfemet som selvstændig enhed forbundet med frekvensbaseret sprogprocessering, associativ, analogibaseret kognition og videnskabelig empirisme. Selvom en så skarp opdeling nok kan virke besnærende, viser forskningen at en binær opsplitning er forfejlet, og mere nuancerede forståelser af morfologien som psykologisk fænomen er derfor nødvendige. Titlens binære spørgsmål er derfor ment som afsæt for en overblikartikel der præsenterer en mere nuanceret forståelse af morfologiens rolle i menneskelig kognition, baseret på teorier og data fra teoretisk morfologi og i særdeleshed fra psykolingvistik.

2 MORFEMER OG ORD I TEORETISK MORFOLOGI

En grundlæggende skelnen i teoretisk morfologi er mellem morfembaserede og ordbaserede teorier der ser hhv. morfemet og ordet som den grundlæggende enhed i sproget. Denne skelnen er helt fundamental for forståelsen af morfologi og ordforråd og har præget feltet i mange år, med det resultat at spørgsmålet om ord kontra morfemer i nogen grad har fået præg af trossag snarere end videnskabeligt spørgsmål (Plag & Balling under udgivelse). Et fyldestgørende overblik over så lang og central en debat kan jeg ikke give her og henviser i stedet til hånd- og lærebøger i morfologi som f.eks. Spencer & Zwicky (1998) og Haspelmath & Sims (2010). I stedet vil jeg kort opridse de forskellige positioner og noget af den evidens der bruges for og imod i teoretisk morfologi, som basis for diskussionen af morfemer og ord i sprogprocessering i afsnit 3. En alternativ relateret skelnen er mellem Hocketts (1954) tre mulige modeller *Item-and-Arrangement*, *Item-and-Process*, som begge er morfembaserede, og *Word-and-Paradigm*, som er ordbaseret, men i hvert fald i et psykolingvistisk perspektiv er den binære kontrast langt mere fremtrædende.

2.1 Morfembaserede teorier

Morfemet er en intuitiv størrelse, navnlig i forbindelse med bøjning, hvor det forekommer indlysende at et ord som *ro-ede* består af to klart adskillelige dele, *ro-* og *-ede*, med hver sin betydning, hhv. *benæge en ro-båd ved brug af årer* og *datid*. Det kan derfor forekomme naturligt at se morfemer som sprogets grundlæggende enheder der manipuleres af morfologiske og syntaktiske mekanismer for at forme ytringer. Det forekommer særligt indlysende for agglutinerende sprog med rig morfologi som f.eks. finsk, hvor den enkelte rod kan komme til udtryk i hundredvis af bøjningsformer.

Denne naturlighed holder imidlertid kun så længe det enkelte morfem repræsenterer en klart afgrænset form med en klart afgrænset betydning, og den udfordres derfor af afvigelser fra et sådant en-til-en-forhold mellem form og betydning. Det gælder ændringer i betydning, der typisk signaleres af et morfem, men i visse tilfælde sker uden ændring i form, f.eks. umarkerede flertalsformer som *får*; tilfælde hvor

en morfologisk form optræder uden tilsvarende betydning, f.eks. de danske fugeelementer *-e-* og *-s-*; og tilfælde hvor form og betydning ikke tydeligt kan adskilles, som f.eks. umlaut i danske flertalsformer som *gæs*. Disse eksempler udgør mestendels sproglige undtagelser, i hvert fald på moderne dansk, men problemet gør sig også gældende for mere udbredte fænomener som det at afledte ord ofte antager betydninger der ikke helt kan forudsiges fra de morfemer de består af, som f.eks. *jord-isk* som nok refererer til noget der har med *jord* at gøre, men kun i ret snæver forstand. Denne form for semantisk opacitet (uigennemsigthed) er i sig selv problematisk for morfembaserede teorier og bliver kun mere problematisk fordi opaciteten er graderet, så komplekse ord varierer på et spektrum fra semantisk helt gennemsigtig til helt uigennemsigtig. Et andet ret udbredt fænomen er bundne rod morfemer, på engelsk kaldet *cranberry morphemes*, fordi et skoleeksempel er *cran-* der må postuleres i en morfembaseret analyse af *cran-berry* (tranebær) men som ikke synkront har nogen selvstændig betydning. Et dansk eksempel på dette kunne være *edder-fugl* hvor *fugl* er helt genkendeligt, men betydningen af *edder-* ikke er synkront tilgængelig.

Med en morfembaseret teori følger typisk regler der manipulerer morfemer for at danne bøjede og afledte ordformer, og dermed nødvendigvis en forklaring af undtagelserne fra reglerne; undtagelserne oplystes i det mentale leksikon (ordforrådet), mens reglen anvendes i alle ulistede tilfælde, som 'default' eller efter 'the Elsewhere principle' (Kiparsky 1973), dvs. i alle tilfælde hvor der ikke gælder en mere specifik regel. Empirisk kan det dog være problematisk, fordi undtagelser i mange tilfælde udviser delvise regelmæssigheder som en simpel oplystning ikke dækker tilstrækkeligt, f.eks. de fonologiske ligheder mellem uregelmæssige datidsformer for engelske verber, som er et særlig velundersøgt fænomen. Der findes forskellige psykolingvistiske svar på hvordan regelmæssigheder, delvise regelmæssigheder og undtagelser håndteres; de diskuteres nærmere i afsnit 3.

2.2 Ordbaserede teorier

I kontrast til morfembaserede teorier står ordbaserede teorier, der sætter ord og de morfologiske paradigmer de indgår i, som centrale enheder i sproget og for morfologien. Her forstås morfologien som para-

digmatiske relationer mellem morfologisk komplekse ord, snarere end en proces der agglutinerer eller på anden måde manipulerer morfemer. Sådanne tilgange har været populære i psykolingvistikken i de senere år, med afsæt i Schreuder & Baayen (1997) og Moscoso del Prado Martín et al. (2004), som en måde at forstå morfologiske effekter i psykolingvistiske eksperimenter uden at postulere morfemet som den grundlæggende enhed i det mentale leksikon.

Givet en kontrast mellem morfem- og ordbaserede teorier og den række af problemer med morfemet som jeg pegede på ovenfor, kan det forekomme indlysende at postulere ordet som sprogets grundlæggende enhed. Et sådant postulat bliver yderligere indlysende for brugere af skriftsprog der har mellemrum mellem ord. Desuden bliver ordet snart sagt altid brugt som enhed i psykolingvistiske eksperimenter på dette område, også dem der søger evidens for morfemet som den grundlæggende kognitive enhed. Ordet antager dermed en selvfølgelighed som gør det svært at undgå som mere eller mindre grundlæggende enhed i sprogbeskrivelse og sprogprocessering.

Trods denne selvfølgelighed er problemerne med at definere og afgrænse ordet ved nærmere eftersyn nok ikke mindre end de problemer der er med morfemet. Bloomfield (1933) beskrev ordet som 'the minimal free form' og opstillede en række kriterier for ordstatus, men selv tilsammen kan disse og andre kriterier der er blevet foreslået, ikke skelne systematisk på tværs af sprog og kontekster mellem ord og på den ene side affikser og på den anden side fraser (Haspelmath 2011). Et klassisk kriterie, som Haspelmath også diskuterer, er ordets mulighed for at udgøre en ytring i sig selv, i modsætning til affikser, men dette gælder ikke for alle de enheder vi intuitivt eller ifølge andre kriterier gerne vil kategorisere som ord, f.eks. funktionsord og transitive verber. Mange funktionsord kan vi nok leve med ikke at kunne kategorisere som fuldgældige ord, men vores ordbegreb giver næppe mening hvis transitive verber ikke defineres som ord.

2.3 Modeller uden ord og morfemer

En måde at undgå problemerne med såvel ord som morfemer er at postulere modeller hvor ord og morfemer ikke spiller nogen rolle som selvstændige enheder, men hvor det vi sædvanligvis omtaler som ord

og morfemer opstår i kraft af formmæssige (lydlige eller ortografiske) og semantiske regelmæssigheder. Sådanne modeller er typisk mere generelle modeller fra kognitiv psykologi, som f.eks. distribuerede konnektionistiske modeller (Rumelhart et al. 1986) og *Naive Discriminative Learning* (NDL, Wagner & Rescorla 1972), som finder anvendelse i den mere psykolingvistiske ende af sprogvidenskaben, jf. f.eks. Harm & Seidenbergs (2004) og Norris' (1994) konnektionistiske modeller og Baayen et al.s (2011) model baseret på NDL. Eksemplarmodeller (f.eks. Goldinger 1998) der opererer uden abstrakte repræsentationer af sproglige enheder, men med omfattende lagring af sprogligt input falder i samme kategori.

De ovenfor nævnte modeller koder input ved hjælp af bogstaver eller lyde, gruppevis, parvis eller enkeltvis, men valget af repræsentation tillægges ikke nogen særlig vægt, bortset fra at det er centralt at der ikke er tale om morfemer. Ikke desto mindre skal der dog vælges en repræsentation af inputtet, og det bliver noget ad hoc: i Baayen og kollegers NDL-model kodes inputtet som bogstavpar, dvs. et ord som *best* kodes som /\$h/, /he/, /es/, /st/ og /t\$/ (hvor \$ repræsenterer ordgrænser), som antages at være tilgængelige for sprogbrugeren uden videre begrundelse. Som det fremgår, opereres der altså her med en form for ordgrænse og derudover også med såkaldte 'leksomer' (Baayen et al. 2015) som muligvis har visse ligheder med ord, men ses som mere dynamiske begreber og i øvrigt er for vagt defineret til at det rigtig kan afgøres.

3 MORFEMER OG ORD I SPROGPROCESSERING

Det grundlæggende spørgsmål om morfologi i sprogprocessering er hvorvidt morfologisk komplekse ord huskes, forstås og produceres på basis af ord eller morfemer. Denne formulering af spørgsmålet kan nok siges at være lidt forældet, men er ikke desto mindre et nyttigt udgangspunkt for diskussionen, og et udgangspunkt der knytter an til de morfologiske modeller ovenfor. Meget af debatten har fokuseret på bøjningsformer, mere specifikt på engelske datidsformer i den såkaldte datidsdebat (*past-tense debate*). Valget af engelsk morfologi som kampplads kan forekomme besynderligt i lyset af det engelske sprogs ret begrænsede bøjningsmorfologi, men fordelingen af regelmæssige og

uregelmæssige former og lighederne mellem – eller endda regelmæssighederne i – de uregelmæssige former giver alligevel dette fokus en vis mening.

Datidsdebattens grundlag er tre mulige måder at repræsentere og processere bøjede ord på: med morfemer og regler, med ord og associationer eller med en kombination af de to, så regelmæssige former processeres på basis af morfemer og uregelmæssige former associativt. Den første position repræsenteres af generativ fonologi (Chomsky & Halle 1968; Halle & Mohanan 1985), en teoretisk model som ikke er holdbar i en psykolingvistisk sammenhæng af i hvert fald to grunde: For det første er den aldrig blevet direkte implementeret i psykolingvistiske eksperimenter. For det andet opstiller generativ fonologi regler for uregelmæssige bøjningsformer der falder i lydige grupper, dvs. regler der dækker undtagelser fra den generelle regel, men må så også have regler for de uregelmæssige bøjningsformer der ikke falder inden for en forventet gruppering, dvs. undtagelser fra undtagelserne. Dette meget komplekse system af regler stemmer dårligere overens med psykolingvistiske data end de simple forklaringer der gives af de andre to typer af modeller (Pinker & Ullman 2002). Datidsdebatten kommer derfor til at stå mellem de to sidstnævnte positioner: associative, typisk konnektionistiske, ordbaserede modeller på den ene side og på den anden side dobbeltmekanismemodeller, der skelner skarpt mellem regelmæssige og uregelmæssige former.

Ifølge de associative modeller, med udgangspunkt i Rumelhart og McClelland (1986), kan en enkelt associativ mekanisme redegøre for processeringen af både regelmæssige og uregelmæssige (bøjnings)former, eksemplificeret ved engelske verber. De fonologiske ligheder mellem uregelmæssige verber, f.eks. mønstre som *drink-drank*, *sink-sank*, *stink-stank*, håndteres indlysende godt af en model hvor datidsformen læres og processeres på basis af formmæssigt overlap med kendte verber og deres datidsformer. Den regelagtige generalisering af de regelmæssige datidsformers *-ed* opstår i dette system som følge af disse formers meget høje typefrekvens, og af at *-ed*-former netop ikke danner klynger af former med lydige overlap.

Over for de associative modeller står forskellige dobbeltmekanisme-modeller, mest prominent Ullmans (2001) *Declarative/Procedural Model*.

Tilhængere af disse modeller (f.eks. Pinker 1999) bruger undertiden generativ fonologi som en stråmand for at fremstå som den gyldne middelvej mellem to ekstreme positioner, men dobbeltmekanisme-modellerne repræsenterer i realiteten det ene yderpunkt i debatten. Den stærkeste evidens for denne type model er nok den neurolingvistiske, der viser forskelle mellem regelmæssige og uregelmæssige former i hjernescanninger (Ullman 2004), hos afasipatienter (Ullman et al. 2005) og hos børn med udviklingsforstyrrelser (Ullman & Pullman 2015). Generelt opfører de uregelmæssige former sig mere som ord og de regelmæssige former mere som syntaktiske strukturer, men der er dog mange tilfælde hvor den skarpe skelnen mellem regelmæssige og uregelmæssige former ikke holder, f.eks. de Vaan et al. (2007; se derudover referencerne der). Historisk startede denne type model som et forsvar for regelbaseret kognition og sprogprocessering efter introduktionen af konnektionistiske netværk og deres associative simulering af regelbaseret processering, med særligt fokus på engelske datidsformer, og det er også derfor den nævnes her. Imidlertid skal det bemærkes at Ullmans model langt fra primært er en model for morfologisk processering eller ordgenkendelse, men derimod bruger regelmæssige og uregelmæssige former som diagnostiske værktøjer for hhv. procedurale og deklaratve kognitive processer.

Mens datidsdebatten altså fokuserer ret snævert på processeringen af bøjningsformer, findes der også en psykolingvistisk litteratur der undersøger ordgenkendelse mere generelt, inklusiv både forskellige morfologiske processer og forskellige sprog. De grundlæggende spørgsmål, og de forskellige mulige positioner, er i nogen grad de samme, men den bredere ordgenkendelseslitteratur drager almindeligvis mindre vidtgående konklusioner om sprog og kognition end de studier der hører hjemme inden for datidsdebatten.

Inden for ordgenkendelsesforskningen er der ligesom i datidsdebatten og i teoretisk morfologi yderpunkter der postulerer at enten ordet eller morfemet er den grundlæggende og eneste enhed for processeringen, traditionelt repræsenteret af hhv. Butterworths (1983) *full-listing*-model og Tafts (1979) *prefix-stripping*-model. Herimellem findes igen forskellige dobbeltmodeller som opererer med både ord og morfemer i en eller anden kombination; her mener nogle at den morfembaserede

'vej' til ordgenkendelse er sekundær (Caramazza et al. 1988), men det har været bredere accepteret at to 'veje' er aktive samtidig (Frauenfelder & Schreuder 1992; Schreuder og Baayen 1995). En anden diskussion fokuserer på forskellige niveauer i processeringen hvor morfemer kan være relevante i de indledende (Taft 1994, 2004) eller de senere (Giraudo & Grainger 2003) stadier af ordgenkendelsen. Uanset hvad, er det interessant at notere sig at ordet også her er uundgåeligt: Selv i Tafts morfembaserede model handler den *morfembaserede* processering om at nå frem til genkendelsen af det rigtige *ord*, og der stilles ikke spørgsmålstegn ved ordet som enhed.

De vigtigste metoder i ordgenkendelsesforskningen er priming og eksperimentelle manipulationer baseret på ords og morfemers korpusfrekvenser. Logikken bag manipulationen af frekvens er at enheder med høj frekvens genkendes hurtigere end enheder med lav frekvens; hvis noget udgør en enhed der er relevant for processeringen, vil vi altså se frekvens effekter. Hvis ordgenkendelseseksperimenter således viser effekter af morfemers, især rødders, frekvenser, ses det som evidens for morfembaseret processering; hvis der derimod konstateres effekter af hele komplekse ords frekvenser, ses det som evidens for ordbaseret processering. Hvis altså et ord som *sam-tanke* genkendes hurtigere end et ord som *gen-digte*, der har samme helordsfrekvens (i Korpus 90 og 2000, se nærmere i Balling 2008, kapitel 3), men langt lavere rodfrekvens, ses det som evidens for morfembaseret processering af disse ord. Og hvis en sammenligning af ord med samme rodfrekvens som f.eks. *gen-skin* og *und-skyld* viser hurtigere processering af *und-skyld*, som har langt højere helordsfrekvens, ses det som evidens for ordbaseret processering. Ordparrene her tjener kun til illustration; eksperimenterne omfatter altid mange ord i hver kategori eller, i de nyere regressionsbaserede¹ eksperimenter, slet ikke kategorier, men ord med kontinuert varierende frekvenser. Desuden må eksperimenterne på den ene eller den anden måde tage højde for andre faktorer der påvirker ordgenkendelsen, som f.eks. ordenes længde, lighed med andre ord og semantiske egenskaber. Resultaterne af disse eksperimenter

1 Regression er en særlig statistisk teknik der har vundet frem inden for lingvistikken de senere år. For en teknisk præsentation af den mest fremherskende teknik, se f.eks. Baayen et al. 2008, for en mindre teknisk introduktion, se Balling og Hvelplund 2015. Se også afsnit 4.2.

viser både helords- og morfemfrekvenser, og undertiden begge dele på en gang, se overblik i Balling (2008) eller Plag & Balling (under udgivelse) og flere detaljer om frekvenseffekter på dansk i afsnit 4. Effekterne af ordfrekvens er dog både større og langt mere systematisk forekommende, og peger igen i retning af ordet som den mest uundgåelige størrelse.

Den anden hovedmanipulation i ordgenkendelseeksperimenter er priming. Her undersøger man hvordan genkendelsen af et ord, det såkaldte *target*, påvirkes af et ord der præsenteres før, en *prime*. Hvis der er en påvirkning af prime på target, kaldes det en primingeffekt; den kan være enten negativ eller positiv. Positive effekter, altså hurtigere genkendelse af target efter præsentation af en prime, bruges som evidens for en kognitiv relation mellem de to ord. I forbindelse med undersøgelser af morfologisk komplekse ord sammenligner man typisk primingeffekter mellem morfologiske komplekse ord og deres rødder, sammenlignet med den priming der opstår mellem ord der kun er formmæssigt overlappende, og ord der kun er semantisk relaterede. På dansk kunne man for eksempel sammenligne effekten på genkendelse af et target som *vind* af præsentationen af det morfologisk overlappende *medvind*, det formoverlappende *sind* og det semantisk relaterede *storm*. De morfologiske relaterede ord er per definition både semantisk og formmæssigt overlappende i større eller mindre grad, og priming-effekterne sammenlignes derfor mellem de tre typer. Hvis primingeffekten fra den morfologiske prime er større end de formmæssige og semantiske primingeffekter tilsammen, argumenterer nogle for at det er evidens for morfemers rolle i ordgenkendelsen (f.eks. Marslen-Wilson et al. 1994); andre (bort)forklarer sådanne resultater ved at sige at morfologiske effekter skyldes ikke-additive men interagerende effekter af semantisk og formmæssigt overlap (Gonnerman et al. 2007).

Resultaterne af primingstudier er blandede, afhængig både af grader af semantisk, formmæssigt og morfologisk overlap og af hvor lang tid der går mellem præsentationen af prime og target. Overordnet set er morfologisk priming mere robust på tværs af forskellige tidsintervaller end både semantiske og formmæssige primingeffekter. Sidstnævnte er desuden ret usystematiske, men er typisk forsinkende (jf. f.eks. Gonnerman et al. 2007 og Marslen-Wilson et al. 1994). Tidsintervallet er

særlig vigtigt i forhold til om morfologiske primingeffekter afhænger af graden af semantisk relation mellem de morfologisk relaterede ord; ved almindelig primepræsentation giver morfologisk relaterede ord med høj grad af semantisk overlap stærkere primingeffekter end dem med lavt semantisk overlap, f.eks. ville man forvente stærkere priming mellem *ironi* og *ironiske* end mellem *jord* og *jordiske*. Disse gradvise effekter giver, som nævnt i afsnit 2, problemer for en morfembaseret model. Når det kommer til subliminal priming, dvs. så kort præsentation af en prime at den ikke processeres bevidst, er det dog uafklaret om den semantiske relation spiller en rolle, hvilket er emne for megen debat (Feldman et al. 2015 kontra Rastle & Davis 2008).

Morfologiens rolle i menneskelig sprogprocessing er således et fokusområde for megen forskning og debat, med tre primære positioner: processing baseret på ord, på morfemer eller på en blanding. Disse tre positioner dukker op både i datidsdebatten om associativ og symbolsk kognition og i den bredere ordgenkendelsesforskning. Dette afsnit har dækket de vigtigste argumenter og evidens for de forskellige positioner, mens vi kommer tættere på en konklusion efter næste afsnits gennemgang af resultater fra dansk.

4 MORFOLOGIENS ROLLE I DANSK ORDGENKENDELSE

4.1 Dansk morfologi

Dansk morfologi hører ikke til blandt de mest komplekse, men der er ikke desto mindre nogle fænomener der er interessante i forhold til at undersøge morfologiens rolle i ordgenkendelsen. Blandt disse er den meget produktive kompositumsdannelse som giver meget store såkaldte morfologiske familier, dvs. grupper af ord der indeholder den samme rod. For *bruge* består den morfologiske familie for eksempel af 2103 separate afledninger og sammensætninger i Korpus 90 og 2000, inklusiv *brugsforening*, *brugbar* og *misbrug*. Interessant nok er de morfologiske familier på dansk meget større end hvad vi ser i tilsvarende korpora for nært beslægtede sprog som engelsk og hollandsk (se sammenligning i Balling 2008, s. 88).

Et andet fænomen der er interessant at undersøge, er forskellene mellem forskellige morfologiske operationer, ikke fordi sådanne forskelle er specifikke for dansk, men fordi dansk udgør en udmærket case. En mulig sammenligning er mellem bøjning, afledning og sammensætning, hvor nogle teoretikere postulerer kategoriske forskelle, navnlig mellem bøjning og afledning, der skulle høre til henholdsvis syntaks og leksikon (Chomsky 1970). En anden sammenligning er mellem præfigering og suffigering, hvor morfemernes rækkefølge kan have en effekt på hvordan de processeres.

4.2 Dansk ordgenkendelse

Selvom dansk ikke er blandt de mest velundersøgte sprog når det kommer til ordgenkendelse hos voksne, normale sprogbrugere, er der dog nogle resultater der er interessante for vores spørgsmål om morfologiens rolle i sprogprocesseringen. De er rapporteret i detaljer i Balling (2008) og i Balling & Baayen (2008, 2012); her vil jeg ikke gå ind i detaljerne, men derimod beskrive nogle hovedtendenser i eksperimenterne og bruge dem til yderligere at belyse de modeller vi så på i afsnit 3.

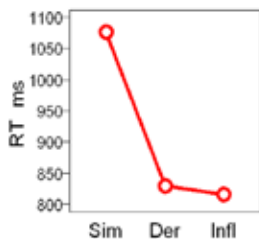
Resultaterne stammer alle fra såkaldt *lexical decision*-eksperimenter, hvor sprogbrugere læser eller lytter til en blanding af rigtige ord og nonsensord der er fonotaktisk korrekte men ikke findes. For hvert ord trykker deltagerne på en knap for at indikere om det er et ord de genkender på dansk eller ej; hvor hurtigt deltagerne trykker på knappen for et rigtigt ord, er en indikation af hvor nemt eller svært det ord er. Hvert eksperiment omfattede 200-250 danske ord og lige så mange nonsensord; for ordene var en række variable manipuleret så effekterne af de variable på ordgenkendelsen kunne undersøges. Eksperimenterne fokuserede primært på afledte ord og auditiv ordgenkendelse, men omfattede også andre typer ord og et enkelt visuelt eksperiment. Jeg vil her diskutere resultaterne for de tre vigtigste grupper af variable: for det første ordenes morfologiske type, for det andet ordenes frekvenser og for det tredje forskellige variable der har at gøre med det enkelte lydige signals ligheder med og forskelle fra andre ord i ordforrådet.

Eksperimenter der undersøger morfologiens rolle i sproget, fokuserer typisk – naturligt nok – på morfologisk komplekse ord, primært dem der består af to morfemer, og undersøger variable der har spe-

cifikt med den type ord at gøre. Men det er faktisk også interessant at sammenligne processeringen af morfologisk komplekse ord med morfologisk simple ord. Her opstår imidlertid den praktiske vanskelighed at andre forskelle mellem simple og komplekse ord end deres morfologiske type påvirker processeringen af dem; først og fremmest er simple ord generelt både kortere og hyppigere forekommende end komplekse ord, og det er begge forhold der trækker i retning af hurtigere processering. Længde og korpusfrekvens bliver dermed variable der skal kontrolleres: den traditionelle tilgang til det er at matche en gruppe af komplekse ord med en gruppe af simple ord i forhold til deres længde og frekvens, men det giver potentielt urepræsentative ord i begge kategorier, nemlig for lange og sjældne simple ord og for korte og hyppige komplekse ord, i forhold til deres kategori mere generelt. Her kan moderne statistik hjælpe: Med en regressionsanalyse (se fodnote 1 ovenfor) kan man bruge ord med forskellig længde og frekvens, inkludere længde og frekvens i den statistiske analyse og se på effekten af for eksempel morfologisk type når der er taget højde for længde og frekvens (samt andre relevante variable).

Med en sådan tilgang viser der sig noget interessant i de danske eksperimenter, nemlig at komplekse ord alt andet lige processeres hurtigere end simple ord. Det vil sige: når vi tager højde for deres længde og frekvens, genkendes komplekse ord hurtigere end simple ord. Effekten er stærkest for de suffigerede bøjede ("Infl" i figur 1) og afledte ("Der" i figur 1) ord hos Balling & Baayen (2008), som kan ses i figur 1, men findes også for sammensatte og præfigerede afledte ord (Balling & Baayen 2012), se figur 2. Det går imod ideer om at kombinatorisk processering af morfemer tager tid (Bertram et al. 1999; Lehtonen et al. 2006), og også imod en intuition om at komplekse ord nødvendigvis må være komplekse at processere. Derimod stemmer resultatet fint overens med ideen om *maximisation of opportunity* (Libben 2006): at sprogbrugeren søger at optimere processeringen ved at bruge alle tilgængelige typer information, i dette tilfælde både viden om hele komplekse ord og deres morfologiske struktur. Den hjælp som lytteren får fra den morfologiske struktur, lader til at være størst for ord hvor roden kommer først, altså i de suffigerede snarere end de præfigerede ord. Denne fortolkning ser ikke morfemet som den centrale kognitive enhed, men

FIGUR 1. REAKTIONSTID

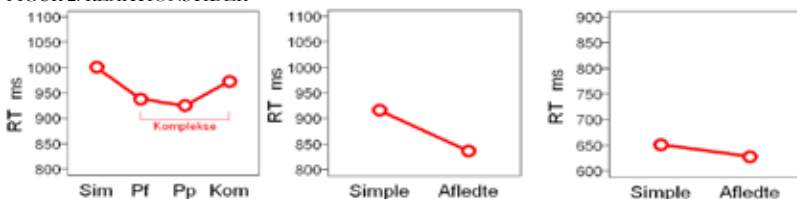


Reaktionstid for simple (Sim), afledte (Der) og bøjede (Infl) ord fra det auditive eksperiment hos Balling & Baayen (2008). Figuren viser forskellen på de tre typer ord med alle andre relevante variable holdt konstante.

giver en forståelse af hvordan morfologiske effekter alligevel kan opstå. Resultatet kan også forstås i forhold til den kritik af læsbarhedsformler (af samme type som Lix) som kommer til udtryk f.eks. hos Bailin & Grafstein (2016) der argumenterer mod disse formlers fokus på ordlængde med henvisning til at lange ord der består af kendte morfemer, ikke som udgangspunkt er så svære som deres længde indikerer

Udover forskellen mellem simple og komplekse ord kan typeforskelle mellem komplekse ord også være interessante. Inden for teoretisk morfologi drages der undertiden meget skarpe grænser mellem forskellige morfologiske operationer (som f.eks. under *the Lexicalist Hypothesis*, Chomsky 1970). Derfor sammenligner de danske eksperimenter afledte ord med bøjede former (Balling & Baayen 2008) og med sammensatte ord (Balling & Baayen 2012), men finder ingen forskel. Det taler imod en kategorisk forskel på forskellige typer morfologi, selvom resultaterne i sig selv ikke betyder at en sådan kategorisk forskel absolut kan afvises.

FIGUR 2. REAKTIONSTIDER



Figur 2: Reaktionstid for simple og komplekse ord i tre eksperimenter hos Balling & Baayen (2012): Den højre figur viser det auditive eksperiment 1, den midterste og højre figur viser eksperiment 2 og 3 som er hhv. auditiv og visuel version af det samme eksperiment. Figureerne viser forskellen på ordtyperne med alle andre relevante variable holdt konstante.

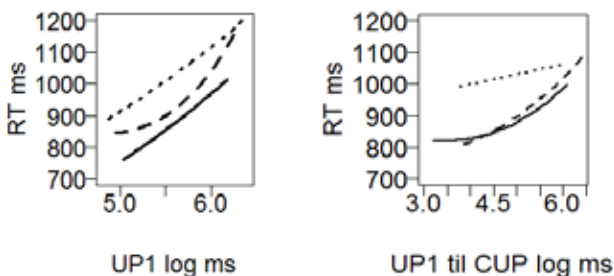
Det andet hovedresultat har at gøre med frekvens effekter. Som vi så i afsnit 3, er det en hovedmanipulation i ordgenkendelsesforskningen og dermed også interessant at undersøge for dansk, ikke mindst i den type regressionsdesigns som vi brugte i de danske eksperimenter. Det handler igen – blandt andet – om repræsentativitet: hvis man skal matche ord præcist på helordsfrekvens, men variere deres rodfrekvens, eller vice versa, som i eksemplerne i afsnit 3, kan man hurtigt ende med meget få og meget atypiske ord, fordi de to typer frekvenser på tværs af ordforrådet har en tendens til at være korreleret. I et regressionsdesign kan begge variable undersøges fordi den ene ikke skal holdes konstant, og det giver både flere og mere typiske ord og mulighed for at sammenligne effekten af de to slags frekvenser. Billedet på tværs af de forskellige eksperimenter er ganske komplekst, men der er dog nogle overordnede tendenser. For det første er helordsfrekvensen klart dominerende, med effekter i alle eksperimenter og større effekter end morfemfrekvenserne. For det andet ses effekter af morfemfrekvenser ikke i alle eksperimenter, og når de ses, er det i interaktion med helordsfrekvenserne, altså er morfemfrekvenserne kun relevante for nogle værdier af helordsfrekvens. Endelig er det interessant at helordsfrekvens effekter der traditionelt udgør en indikation af ikkemorfembaseret processering, optræder samtidig med forskellige morfologiske effekter. Det kan fortolkes sådan at helordsfrekvenserne i virkeligheden ikke bare er en indikation af hvor bekendt en form er for sprogbrugere, men også er en indikation af den kombinatoriske sandsynlighed af at ordets to morfemer optræder sammen (Baayen et al. 2007).

Det tredje og sidste hovedresultat jeg vil diskutere, har at gøre med en variabel der er særligt relevant for auditiv ordgenkendelse, nemlig det såkaldte *Uniqueness Point* (UP). UP'et blev introduceret af Marslen-Wilson & Welsh i 1978 og er et centralt begreb i Marslen-Wilsons (1984) Kohortemodell. Den grundlæggende ide i den model er at der på ethvert givet tidspunkt i den auditive ordgenkendelse er en gruppe – en kohorte – af ord der passer med inputtet; kohorten bliver mindre og mindre efterhånden som mere og mere af ordet er hørt, og UP'et er det fonem hvor det hørte ord adskiller sig fra alle konkurrerende ord i kohorten, bortset fra bøjede, afledte og sammensatte udgaver af ordet selv. UP'et er altså det teoretisk tidligste punkt hvor det faktiske ord

kan identificeres når det optræder ude af kontekst, og i adfærdseksperimenter viser det sig at ord med senere UP genkendes senere. Efterhånden som Kohortemodellen viste sig at være utilstrækkelig, blev UP'er umoderne som fokus for forskning, men de stærke effekter har dog nødvendiggjort kontrol af UP-variablen i auditive ordgenkendelseseksperimenter.

Dette var også i første omgang grunden til at inkludere UP'er i de danske auditive eksperimenter. Det viste sig imidlertid at den traditionelle definition af UP'et var utilstrækkelig for de komplekse ord i disse eksperimenter, og vi introducerede derfor en omfortolkning og udvidelse af begrebet: et første UP – UP1 – som det punkt hvor ordets første morfem bliver unikt ved at adskille sig fra alle ord der ikke har dette morfem i første position, og et komplekst UP – CUP – hvor hele det komplekse ord adskiller sig fra andre ord med det samme første morfem. For et sammensat ord som *fodbold* er UP1 det første *d* hvor ordet adskiller sig fra konkurrenter som *foto*; CUP optræder på det andet *o* hvor *fodbold* adskiller sig fra andre ord med *fod* som første morfem, f.eks. *fodbad*. Begge UP'er har solide effekter på ordgenkendelsen med meget langsommere genkendelse af ord med senere UP1 og noget langsommere genkendelse af ord med senere CUP, som illustreret i figur 3. Med andre ord er ord sværere at genkende jo længere tid konkurrencen fra både urelaterede og relaterede ord varer, som målt af hhv. UP1 og CUP (eller afstanden fra UP1 til CUP).

FIGUR 3.



Reaktionstid som funktion af UP1 til venstre og afstand fra UP1 til CUP til højre i tre auditive eksperimenter: de fuldt optrukne linjer viser effekterne hos Balling og Baayen (2008), de to stiplede er effekterne hos Balling og Baayen (2012). Hver figur viser effekten af det givne UP med alle de andre relevante variable boldet konstante.

De to UP'er skal forstås samlet, og CUP bliver dermed et udtryk for betinget processering af hele ordet, givet at det første morfem allerede er identificerbart. På den anden side er UP'erne også udtryk for to forskellige typer leksikalsk konkurrence: UP1 er konkurrence mellem morfologiske urelaterede ord, mens CUP er et indeks for konkurrence mellem morfologisk relaterede ord. Sidstnævnte er interessant i forhold til en anden morfologisk variabel, nemlig den morfologiske familie (se afsnit 4.1). Denne variabel har typisk en faciliterende effekt: jo flere familiemedlemmer et ord har, jo nemmere er det alt andet lige at forstå (Schreuder & Baayen 1997). Men i de danske eksperimenter er effekterne af morfologisk familie blandede: de meget store familier hæmmer typisk genkendelsen af det enkelte ord. Disse hæmmende effekter af specifikke morfologiske kohorter, det være sig familien der inkluderer alle ord hvor en bestemt rod optræder, eller CUP-kohorten der kun indeholder ord med den samme rod i første position, nuancerer billedet af morfemernes positive effekt som vi så i fordelene for de komplekse ord ovenfor.

5 PERSPEKTIVER OG VIDERE UNDERSØGELSER

Som lovet er det samlede billede af morfologiens rolle i sprogprocesseringen altså adskilligt mere komplekst end overskriftens binære spørgsmål indikerer, og det er i sig selv en pointe: Den binære kontrast mellem ord og morfemer – og de modeller der accepterer og bygger på den kontrast – er forsimplet i forhold til den adfærd vi kan observere hos lyttende og læsende sprogbrugere. Det lader nemlig til at sprogbrugerne trækker på alle tilgængelige informationskilder i processeringen; i større eller mindre grad afhængigt af omstændighederne, men dog sjældent eller aldrig et spørgsmål om *enten* ord *eller* morfemer. Denne ide om *maximisation of opportunity* giver også en forklaring på hvordan morfologien kan spille en rolle, selvom selvstændige repræsentationer af morfemer er problematiske.

I denne forståelsesramme er morfologien funktionel: den morfologiske struktur gør vores ordgenkendelse mere effektiv og mere robust. Vi har set i afsnit 4 hvordan dette gør sig gældende for voksne sprogbrugere i ordgenkendelseseksperimenter. Derudover er der en større

litteratur om hvordan implicit eller eksplicit morfologisk viden støtter børns læsning, stavning og skrivning af ord og tekster (jf. f.eks. Northey et al. 2016 for et overblik). For ordblinde, fremmedsprogslearere og andre svagere læsere kan denne funktionalitet være endnu vigtigere (Goodwin & Ahn 2010).

Jeg har vovet den påstand at den adfærd vi observerer hos sprogbrugere er kompleks, men i en anden forstand kan den kritiseres for at være for simpel. Langt de fleste undersøgelser af ordgenkendelsen bygger nemlig på de samme forsimplende paradigmer: Vi undersøger genkendelsen af enkelte ord uden kontekst, vi kalder ord komplekse når de består af to morfemer og arbejder sjældent med mere indviklede strukturer end det, og vi bruger paradigmer der ikke findes i naturlig sprogbrug, som for eksempel *lexical decision*. Jeg vil stadig argumentere for at resultaterne siger noget om ordgenkendelsen, men vi bør samtidig holde os for øje at vi langt fra kan regne med at de giver det fulde billede. Vi har set i afsnit 4 hvordan moderne statistiske teknikker kan ændre vores forståelse af morfologiens rolle i ordgenkendelsen, men mere naturalistiske eksperimenter hvor ord processeres i kontekst, er nødvendige for en fuld forståelse af morfologiens rolle i den samlede sprogprocessering.

Som lovet i overskriften har jeg prøvet at tegne et billede af morfemets og morfologiens rolle som teoretisk og psykologisk begreb, men det viser sig at diskussionen af det er uløseligt forbundet med en diskussion af ordet som begreb. Hver gang vi afviser eller problematiserer morfemet, bliver alternativet per default ordet, også i de modeller der principielt ikke bekender sig til noget bestemt repræsentativt niveau. Alligevel er ordet ikke teoretisk mindre problematisk end morfemet, men nok i højere grad praktisk uundgåeligt. En mulig måde at forstå ordets rolle på kunne være som et andet niveau der bidrager til *maximisation of opportunity*: Såvel ord som morfemer kan ses som uperfekte abstraktioner, der dog abstraherer over en type information der er central for sprogprocesseringen.

Laura Winther Balling

Department of International Business Communication

Copenhagen Business School

lwb.ibr@cbs.dk

LITTERATUR

- Baayen, R.H. et al. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118. 438-81. DOI: 10.1037/a0023851.
- Baayen, R.H. et al. 2015. Comprehension without segmentation: A proof of concept with naive discriminative learning. *Language, Cognition, and Neuroscience* 31. 106-128. DOI: 10.1080/23273798.2015.1065336.
- Baayen, R.H., L.H. Wurm & J. Aycocock. 2007. Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon* 2. 419-463. DOI: 10.1075/ml.2.3.06baa.
- Baayen, R.H., Davidson, D.J. & Bates, D.M. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390-412. DOI: 10.1016/j.jml.2007.12.005.
- Bailin, A. & A. Grafstein. 2016. *Readability: Text and context*. Houndmills: Palgrave Macmillan.
- Balling, L.W. 2008. *Morphological effects in Danish auditory word recognition*. Ph.d.-afhandling. Aarhus Universitet.
- Balling, L.W. & R.H. Baayen. 2008. Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes* 23. 1159-1190. DOI: 10.1080/01690960802201010.
- Balling, L.W. & R.H. Baayen. 2012. Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition* 125. 80-106. DOI: 10.1016/j.cognition.2012.06.003.
- Balling, L.W. & K.T. Hvelplund. 2015. Design and statistics in quantitative translation (process) research. *Translation Spaces* 4. 169-186. DOI: 10.1075/ts.4.1.08bal.
- Bertram, R., M. Laine & K. Karvinen. 1999. The interplay of word formation type, affixal homonymy, and productivity in lexical processing evidence from a morphologically rich language. *Journal of Psycholinguistic Research* 28. 213-226. DOI: 10.1023/A:1023200313787.
- Bloomfield, L. 1933. *Language*. New York: Henry Holt.
- Butterworth, B. 1983. Lexical representation. B. Butterworth (red.), *Language production (Vol. II): Development, writing and other language processes*, 257-294. London: Academic Press.
- Caramazza, A., A. Laudanna & C. Romani. 1988. Lexical access and inflectional morphology. *Cognition* 28. 297-332. DOI: 10.1016/0010-0277(88)90017-0.

- Chomsky, N. 1970. Remarks on nominalization. R. Jacobs & P. Rosenbaum (red.), *Readings in transformational grammar*, 184-221. Waltham, MA: Ginn & Co.
- Chomsky, N. & M. Halle. 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.
- Feldman, L.B. et al. 2015. Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in human neuroscience* 9, Article 111. DOI: 10.3389/fnhum.2015.00111.
- Frauenfelder, U.H. & R. Schreuder. 1992. Constraining psycholinguistic models of morphological processing and representation: The role of productivity. G.E. Booij & J. van Marle (red.), *Yearbook of morphology 1991*, 165-183. Dordrecht: Kluwer. DOI: 10.1007/978-94-011-2516-1_10.
- Giraud, H. & J. Grainger. 2003. A supralexic model for French derivational morphology. E.M.H. Assink & D. Sandra (red.), *Reading Complex Words. Cross-Language Studies*, 139-156. New York: Kluwer. DOI: 10.1007/978-1-4757-3720-2_7.
- Goldinger, S.D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105. 251-279. DOI: 10.1037/0033-295X.105.2.251.
- Gonnerman, L.M., M.S. Seidenberg & E.S. Andersen. 2007. Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General* 136. 323-345. DOI: 10.1037/0096-3445.136.2.323.
- Goodwin, A.P. & S. Ahn. 2010. A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia* 60. 183-208. DOI: 10.1007/s11881-010-0041-x.
- Halle, M. & K.P. Mohanan. 1985. Segmental phonology in Modern English. *Linguistic Inquiry* 16. 57-116.
- Harm, M.W. & M.S. Seidenberg. 2004. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review* 111. 662-720. DOI: 10.1037/0033-295X.111.3.662.
- Haspelmath, M. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45. 31-80. DOI: 10.1515/flin.2011.002.
- Haspelmath, M. & A. Sims. 2010. *Understanding morphology*, 2. udg. London: Routledge.
- Hockett, C.F. 1954. Two models of grammatical description. *Word* 10. 210-234.
- Kiparsky, P. 1973. "Elsewhere" in phonology. S. Anderson & P. Kiparsky (red.), *A festschrift for Morris Halle*, 93-106. New York: Holt, Rinehart & Winston.
- Korpus 90, Korpus 2000*. Tilgængelig fra <http://korpus.dsl.dk/resources.html>. Tilgæet 1. juni 2016.

- Lehtonen, M. et al. 2006. Recognition of inflected words in a morphologically limited language: Frequency effects in monolinguals and bilinguals. *Journal of Psycholinguistic Research* 35. 121-146. DOI: 10.1007/s10936-005-9008-1.
- Libben, G. 2006. Why study compound processing? An overview of the issues. G. Libben & G. Jarema (red.), *The representation and processing of compound words*, 1-22. Oxford, UK: Oxford University Press. DOI: 10.1093/acprof:oso/9780199228911.003.0001.
- Marslen-Wilson, W.D. & A. Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10. 29-63.
- Marslen-Wilson, W.D. 1984. Function and process in spoken word recognition. H. Bouma & D.G. Bouwhuis (red.), *Attention and performance X: Control of language processes*, 125-150. Hillsdale, NJ: Lawrence Erlbaum.
- Marslen-Wilson, W.D. et al. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review* 191. 3-33. DOI: 10.1037/0033-295X.101.1.3.
- Moscoso del Prado Martín, F., A. Kostić & R.H. Baayen. 2004. Putting the bits together: an information theoretical perspective on morphological processing. *Cognition* 94. 1-18. DOI: 10.1016/j.cognition.2003.10.015.
- Norris, D. 1994. Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52. 189-234. DOI: 10.1016/0010-0277(94)90043-4.
- Northey, M., D. McCutchen & E.A. Sanders. 2016. Contributions of morphological skill to children's essay writing. *Reading and Writing* 29. 47-68. DOI: 10.1007/s11145-015-9579-7.
- Pinker, S. 1999. *Words and rules: The ingredients of language*. London: Phoenix.
- Pinker, S. & M.T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6. 456-463. DOI: 10.1016/S1364-6613(02)01990-3.
- Plag, I. & L.W. Balling. (Under udgivelse). Derivational morphology: An integrative perspective on some fundamental questions. V. Pirrelli, I. Plag, & W. U. Dressler (red.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*. Berlin: De Gruyter.
- Rastle, K. & M.H. Davis. 2008. Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes* 23. 942-971. DOI: 10.1080/01690960802069730.
- Rumelhart, D.E. & J.L. McClelland. 1986. On learning the past tenses of English verbs. J.L. McClelland, D.E. Rumelhart & the PDP Research Group (red.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*, 216-271. Cambridge, MA: MIT Press.

- Rumelhart, D.E., J.L. McClelland & the PDP Research Group (red.). 1986. *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Schreuder, R. & R.H. Baayen. 1995. Modeling morphological processing. L.B. Feldman (red.), *Morphological aspects of language processing*, 131-154. Hillsdale, NJ: Lawrence Erlbaum.
- Schreuder, R. & R.H. Baayen. 1997. How complex simplex words can be. *Journal of Memory and Language* 37. 118-139. DOI: 10.1006/jmla.1997.2510.
- Spencer, A. & A.M. Zwicky (red.). 1998. *The handbook of morphology*. Oxford: Blackwell.
- Taft, M. 1979. Recognition of affixed words and the word frequency effect. *Memory and Cognition* 7. 263-272. DOI: 10.3758/BF03197599.
- Taft, M. 1994. Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes* 9. 271-294. DOI: 10.1080/01690969408402120.
- Taft, M. 2004. Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology* 57. 745-765. DOI: 10.1080/02724980343000477.
- Ullman, M.T. 2001. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research* 30. 37-66. DOI: 10.1023/A:1005204207369.
- Ullman, M.T. 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92. 231-270. DOI: 10.1016/j.cognition.2003.10.008.
- Ullman, M.T. et al. 2005. Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language* 93. 185-238. DOI: 10.1016/j.bandl.2004.10.001.
- Ullman, M.T. & M.Y. Pullman. 2015. A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews* 51. 205-222. DOI: 10.1016/j.neubiorev.2015.01.008.
- Vaan, L. de, R. Schreuder & R.H. Baayen. 2007. Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2. 1-24. DOI: 10.1075/ml.2.1.02vaa.
- Wagner, A. & R. Rescorla. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. A.H. Black & W.F. Prokasy (red.), *Classical conditioning II: Current research and theory*, 64-99. New York: Appleton-Century-Crofts.