

*Hjælp! Jeg har mistede min
yndlings rød taske.*
Et studie af grammatikafvigelser

Katrine Falcon Søby
& Line Burholt Kristensen

Abstract

Grammar anomalies are frequent in texts produced by second language learners. When describing these anomalies, two main issues arise: What is anomalous use of grammar? And how are grammar anomalies distinct from orthographic and lexical anomalies?

We review earlier error-definitions and suggest defining grammar anomalies according to an explicit norm instead of L1 usage. We propose a broader definition of grammar than in previous studies, based on Boye & Harder (2012).

The distinction between grammar, lexicon and orthography is illustrated with data from 28 adult L1 English learners of L2 Danish. In the corpus, 55.9 % of the anomalies were related to grammar. Finally, we discuss how definitions and procedures can be used in future studies of naturally occurring grammar anomalies.

Nøgleord

grammatik, afvigelse, sprognormer, dansk som andetsprog

1. Introduktion

Når man lærer dansk som andetsprog, er kendskab til leksikon og ortografi ikke tilstrækkeligt. For at kommunikere klart og effektivt på skriftligt dansk er det også vigtigt at beherske den danske grammatik.

I dansk som andetsprog forekommer en lang række forskelligartede grammatikafvigelser, fx afvigende brug af grammatiske ord, ordstilling og bøjning. En central del af forskningsprojektet Broken Grammar and

Beyond (2019) er at undersøge hvilke typer af grammatikafvigelser der forekommer i dansk som andetsprog. Et øget kendskab til kursisternes typiske grammatikafvigelser kan give underviserne på landets sprogskoler et bedre grundlag for at tilrettelægge undervisningen i grammatik med disse afvigelser for øje.

Det er dog ikke ukompliceret at opgøre grammatikafvigelser i kursiststile. I denne artikel peger vi på to væsentlige problemer: at definere hvad der tæller som en afvigelse, og at definere hvad der tæller som grammatik. Afvigelserne i *jeg har mistede min yndlings rød taske* er relativt lette at identificere, men i kraft af sprogudviklinger, lektale forskelle, variationsmuligheder mv. kan det i andre tilfælde være svært at skelne norm fra normbrud. Ligeledes kan det være svært at afgøre om et normbrud er relateret til grammatik, ortografi eller leksikon (jf. Jervelund 2007). Et ord som *områden* (for *området*) indeholder en kongruensafvigelse, men afvigelsen kunne også anskues som en taste- eller stavefejl.

I denne artikel præsenterer vi vores bud på hvordan man kan koble teoretisk funderede definitioner af *norm* vs. *normbrud* og *grammatik* vs. *ikke-grammatik*. Vi viser også hvordan definitionerne kan anvendes i praksis ved studier af skriftligt dansk. Målet er at belyse hvor stor en andel af afvigelser der reelt har med grammatik at gøre, hvilket ikke har været i fokus i tidligere stavefejlsundersøgelser. Vi illustrerer dilemmaer og konsekvenser med autentiske eksempler fra stile skrevet af voksne sprogskolekursister med engelsk som L1 og dansk som L2. Definitioner og fremgangsmåder vil også kunne anvendes ved undersøgelser af grammatikafvigelser i skriftligt dansk for skribenter med andre sproglige baggrunde.

I de følgende afsnit præsenteres undersøgelsens datamateriale og fremgangsmåden for opmærkning af de indsamlede kursisttekster. Med udgangspunkt i dette datamateriale diskuterer vi dernæst hvad en afvigelse er, og hvilke afvigelser der kan kategoriseres som grammatiske. Derpå følger et kort resultat afsnit med et overblik over de hyppigste underkategorier af grammatikafvigelser i materialet. Til slut skitseres mulighederne for fremtidige undersøgelser.

2. Deltagere og data

Indsamlingen af tekster fra sprogskolekursister foregik fra september 2017 til marts 2018 på Københavns Sprogcenter, som på det tidspunkt var en af de største udbydere af danskundervisning i København. Skolens kursister er indplaceret på én af tre danskuddannelser afhængigt af deres uddan-

nessesbaggrund. Vores informanter er fra DU2 (for kursister med kort skolegang) og DU3 (for kursister med en mellemlang eller lang uddannelse).

Hver danskuddannelse er inddelt i 5-6 moduler, og efter hvert modul skal kursisterne bestå en modultest for at kunne fortsætte. Ifm. de skriftlige test udleverede Katrine, som dengang var lærer på skolen, og hendes kolleger en samtykkeerklæring og et spørgeskema og indsamlede således skriftlige testbesvarelser fra 217 kursister, primært fra DU3, med over 50 forskellige nationaliteter og modersmål. Besvarelserne blev anonymiseret efter indsamlingen. De skriftlige tests er forskellige på hvert modul (og findes ikke på alle). De består af 1-2 bundne skriveopgaver, typisk opslag, e-mails og debatindlæg. Et begrænset antal opgavesæt per modul genbruges på skift til prøverne. Opgaverne skrives i hånden, og kursisterne må benytte ordbøger, men ikke elektroniske hjælpemidler.

Det mest udbredte modersmål i datasættet er engelsk. Vi startede derfor med at undersøge afvigelserne hos denne gruppe learnere. I alt havde 38 kursister angivet engelsk som deres L1, men for at få en mere homogen gruppe inkluderede vi kun kursister fra lande som ifølge den britiske regering er “majority English speaking” (University of Sheffield 2019). Vi frasorterede hermed 10 kursister med andre nationaliteter, fx græsk eller nepalesisk. Tilbage var 28 deltagere i alderen 22-63 år (gns. 32,6 år, $\pm 7,1$), heraf 13 amerikanere, 12 briter, 2 australiere og en irer (i alt 17 mænd og 11 kvinder).

2.1. Korpus og tekststopmærkning

De 28 kursister har hver skrevet 43-621 ord, i alt 5.685 ord. Den store spredning kan bl.a. skyldes at der er stor forskel på hvor meget tekst der skal produceres, og hvor lang tid testdeltagerne har til opgaverne, på modul 2 over for modul 4.

Til at opmærke afvigelserne brugte vi XML. Først grovsorterede vi afvigelserne i de provisoriske kategorier ortografi, morfologi, syntaks og ordvalg. For hver afvigelse har vi foreslået en form der er i overensstemmelse med normen, og angivet hvilken underkategori der er tale om. Hvor det er relevant, har vi angivet ordklasse og evt. om stavemåden er autoriseret. Til sidst har vi vurderet om afvigelsen var grammatisk eller leksikalsk¹. Fremgangsmåden er udviklet i samarbejde med Jørgen Schack og Philip Diderichsen fra Dansk Sprognævn.

Da begge artiklens forfattere har opmærket tekster, har vi for at sikre

1 Tak til Kasper Boye for hjælp til vanskelige tilfælde.

ensartethed i opmærkningen udregnet vores *interrater reliability score*. Vi opmærkede tekster fra seks deltagere hver for sig og undersøgte i hvor høj grad opmærkningen var ens. Vi forholdt os dels til om vi var enige om hvad der talte som en afvigelse, og dels til om vi var enige om kategoriseringen. I alt var vi enige om 255 afvigelser. Derudover havde Line markeret 16 afvigelser, og Katrine 32 andre afvigelser. Dvs. vi var enige i 85 % af opmærkningen. Det var især inden for kategorien ordvalg vi havde opmærket forskelligt; mest enige var vi inden for kategorierne ortografi og morfologi. Beregningen af scoren blev efterfulgt af en kvalitativ gennemgang af teksterne hvor vi blev enige om antal afvigelser og justerede underkategorierne hvor de var brugt forskelligt.

Inden gennemgangen af undersøgelsens resultater i afsnit 5 gennemgår vi først hvad vi forstår ved en afvigelse (afsnit 3), og hvilke afvigelser vi ser som grammatikrelatede (afsnit 4).

3. Hvad er en afvigelse?

Inden for andetsprogsforskningen er der ikke enighed om hvordan begrebet *error* skal defineres. De definitioner som er blevet brugt i *error analysis*-studier gennem årene, er med Pawlaks (2013: 3) ord “far from satisfactory”. Lennon (1991a) sætter spørgsmålstegn ved om begrebet overhovedet kan defineres da mange faktorer har indflydelse på hvad der ses som en afvigelse, fx situationen, samtalepartneren, stiltniveauet og produktionspresset på taleren. Thewissen (2015: 65) beskriver grænsen mellem egentlige afvigelser og *infelicities* (mindre acceptabel sprogbrug) som “a notoriously grey area for anyone who has ever got their hands dirty doing error analysis work.”

Den mest udbredte måde at definere afvigelser på er ifølge Pawlak (2013) at sammenligne L2-ytringerne med hvad en modersmålstalende ville producere i en lignende situation. Et eksempel på denne tilgang findes i Lennons (1991b) udbredte arbejdsdefinition af *error*:

“a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers’ native speaker counterparts” (Lennon 1991b: 182).

Det er dog problematisk at sammenligne med modersmålstalendes *usus*, da de også producerer afvigelser (jf. bl.a. Jervelund 2007 og Blom et al. 2017): Skulle manglende præsens-*r* så ikke tælle som en afvigelse hos lea-

rnere fordi manglende præsens-*r* også er udbredt hos L1-talere? Fremfor at definere afvigelser ift. modersmålstalendes praksis definerer vi i denne undersøgelse afvigelser ift. gældende normer, hvor normer forstås som *langue* hos Harder (2010):

“*Langue* is not a set of actually attested regularities, but the set of options/constraints that attested regularities have generated as the context of any future potential linguistic acts[...]” (Harder 2010: 302).

Ved at definere afvigelser ift. en norm kan L2-former ses som afvigende, selvom modersmålstalende laver tilsvarende afvigelser. Udfordringen ved at sammenligne med en norm er dog at den kan ændre sig eller være lektalt betinget. Derfor må vi vælge og eksplicitere hvilken norm vi følger.

3.1. Afvigelser fra et modtagerperspektiv

Vi kan ikke afgøre hvordan afsenderens sprognormer eller *langue* er. Vi kan kun forholde os til hvilke sprognormer vi som modtagere har. Når vi beskriver en afvigelse, forholder vi os til hvordan vi som modtagere opfatter afvigelsen. Vi kan ikke vide hvordan eller hvorfor en afvigelse er opstået, og hvad afsenderen intenderede at skrive. Vi kan kun forholde os til hvordan vi som modtagere opfatter afvigelsen og vil kategorisere den.

Denne introspektive fremgangsmåde er ikke uproblematisk; vi har ikke altid den fornødne selvindsigt til fx at afgøre om en norm er ved at rykke sig, eller om vi bedømmer tekster skrevet af personer med dansk som L2 hårdere end tekster fra modersmålstalende. For at kunne kategorisere afvigelser må vi også i opmærkningen skrive et forslag til en korrekt målform, dvs. vi kan ikke undgå at skulle gætte på hvad afsenderen ville skrive.

Når vi har lokaliseret afvigelser i praksis, har vi set på om der var afvigelser ift. Dansk Sprognævns normer (fx *bedede* for *bad*), ift. danske grammatikker og ift. vores egen sproglige intuition. Nogle gange er der dog ikke overensstemmelse mellem disse kilder. Fx bruger mange kursister *ny* uden endelse i eksempler som *min ny lejlighed*. Brugen af *ny* fremfor det mere udbredte *nye* kan opfattes som afvigende, men da Retskrivningsordbogen godtager begge former, er *ny* ikke markeret som en afvigelse.

I mange tilfælde var det let at identificere afvigelser, især inden for morfologi og syntaks som i (1):

- (1) I 1948 Maria begyndte at studere medicin (3.4.G²)

2 Henvisningen læses som: danskuddannelse 3, modul 4, kursist G

I andre tilfælde er det mere vanskeligt, særligt hvor der er forskellige normer for tale- og skriftsprog, eller hvor normen måske er ved at ændre sig, som i *Venligst ring*. Konstruktionen er ikke beskrevet i standardgrammatikker og falder uden for normen ift. Hansen & Heltofts (2011) beskrivelse af at kun *så* og *da* kan stå på fundamentpladsen før imperativer. Christensen (2012: 26) skriver at det i moderne dansk virker “mest rimeligt at placere ordet *venligst* der, hvor man placerer andre småord, der påvirker, hvordan imperativen skal forstås – nemlig til højre for den.” *Venligst ring* er derfor tagget som en afvigelse, selvom der kunne være en normændring undervejs, og konstruktionen tilsyneladende er udbredt – i hvert fald i et kursusvurderingsskema fra Det Humanistiske Fakultets ph.d.-skole på Københavns Universitet hvor den optræder flere gange, fx *Venligst husk at skrive dit navn øverst på side 1*.

Endelig ser vi flere eksempler som (2) med helsætningsordstilling i ledsætninger. Her kan man jf. Christensen & Jensen (2014) diskutere om ordstillingen i ledsætningen er afvigende eller ej da ledsætningen kan ses som forgrund.

(2) Men de tror at der er ikke så meget de kan at gøre (3.3.C)

Udbredelsen af helsætningsordstilling i ledsætninger er os bekendt ikke undersøgt for skriftsprog, men vi formoder at det forekommer langt mindre hyppigt på skrift end i tale. Derudover undervises kursister på sprogskoler i – og bedømmes på evnen til – at lave ledsætningsordstilling i ledsætninger. Af disse to grunde er ledstillingen i (2) medregnet som en afvigelse.

3.2. Tidligere undersøgelser af dansk skriftsprog

Vores undersøgelse har fokus på afvigelser i grammatik, ikke i ortografi eller ordvalg. Den sigter mod et bredt og samlet overblik over grammatikafvigelser. I tidligere skriftsprogsundersøgelser har rapporteringen af grammatikafvigelser ikke været fremtrædende, og begrebet grammatik har været snævrere defineret. I det følgende gennemgås hvordan vores fokus adskiller sig fra tidligere undersøgelser.

Det grundlæggende skel i vores undersøgelse er skellet mellem grammatikafvigelser og ikke-grammatikafvigelser. Jervelunds (2007) udførlige gennemgang af tidligere normbrudsundersøgelser for dansk skriftsprog viser at fokus oftest har været retstavning. Det vigtigste skel i disse undersøgelser har været mellem stavfejl og ikke-stavfejl, mens grammatikafvigelser enten ikke rapporteres eller ses som en restkategori. Ofte regnes forvekslingen af

springe/sprænge og *lærere* for *lærerne* for bogstavforvekslinger, snarere end for grammatisk betingede (Jervelund 2007).

Senere skriftsprogsundersøgelser har haft særlig opmærksomhed på fonologiens rolle ved normbrud i skriftsprog, bl.a. Johannsen (2012) og Brink, Elbro & Johansen (2014) der skelner mellem *tegnsetsningsfejl*, *udtaleoverensstemmende fejl* og *ikke-udtaleoverensstemmende fejl*. Med sit fokus på udtale forholder denne sortering sig ikke til skellet mellem grammatik og ikke-grammatik. Tegnsætningsfejlene kan dog hovedsageligt regnes for grammatikrelaterede. Både i Johannsens (2012) undersøgelse af danske gymnasiestile og i en senere undersøgelse af universitetsopgaver (Blom et al. 2017) er tegnsætningsfejl (især kommafejl) den hyppigste afvigelsestype. Selvom afvigelser i tegnsætning er grammatikrelaterede, ser vi i vores undersøgelse helt bort fra dem. På danske sprogskoler undervises der først i tegnsætning på sene moduler, så ikke alle kursister i vores undersøgelse har forudsætningerne for produktion af korrekt dansk tegnsætning. Desuden er tegnsætningsfejlene i Johannsen (2012) så talrige (90 % af alle afvigelser) at en opgørelse af tegnsætning i vores undersøgelse må formodes at blive unødigt tidskrævende ift. den begrænsede teoretiske indsigt de kunne give i learneres beherskelse af dansk grammatik og i modersmålets indflydelse.

I modsætning til tidligere undersøgelser synes studiet af sproglige afvigelser i danske universitetsopgaver (Blom et al. 2017) ikke at opstille et primært skel med udgangspunkt i ortografi eller fonologi. Studiet har til formål både at beskrive ortografi, tegnsætning og grammatik, og det forholder sig til en bred kam af afvigelsestyper: tegnsætning, stavning, layout, semantik, idiomer, reference, kongruens og syntaks. Foruden tegnsætning kan de tre sidstnævnte kategorier ses som grammatikafvigelser, men mange af de øvrige kategorier indeholder også afvigelser, som vi, med en bredere definition af grammatik, ser som grammatiske. Fx optælles *af vide* (for *at vide*) som en konsonantforveksling, og ikke som grammatikrelateret, til trods for at infinitivmarkøren *at* er et grammatisk ord. Også forveksling af *synes/syntes* opgøres som en stavfejl (dog af typen bøjningsafvigelse). Afvigende brug af et ord, manglende ord og overflødige ord placeres i kategorien semantik. Hvis ordet er grammatisk, fx en infinitivmarkør, opgøres det i vores undersøgelse som en grammatikafvigelse, jf. (3) - (5).

- (3) Afvigende grammatisk ord: vi har masse ting til taler om (3.3.B)
- (4) Manglende grammatisk ord: god til (at) pas på mig selv (3.4.E)
- (5) Overflødig grammatisk ord: der er ikke så meget de kan at gøre (3.3.C)

I forhold til tidligere undersøgelser af skriftsprog har vores undersøgelse primært fokus på grammatik. Først undersøges hvad der kunne være relateret til grammatik, og dernæst placeres øvrige afvigelser i restkategorierne ortografi-afvigelser og leksikalske afvigelser.

4. Hvad er grammatik?

I det følgende gennemgås hvordan den brede tilgang til grammatik udmønter sig, og hvilke praktiske problemer der opstår ved afgrænsningen af grammatik fra ortografi og leksikon.

4.1. En bred tilgang til grammatik

Vores brede forståelse af grammatik bunder i ProGram-teorien (Boye & Harder 2012). If. denne teori kan grammatik tage mange former. Foruden syntaks og grammatiske morfemer inkluderer teoriens grammatikbegreb også grammatiske ord: både ord der traditionelt er beskrevet som funktionsord (fx konjunktioner, artikler og hjælpeverber) og, mere utraditionelt, visse former for pronominer og præpositioner. Særligt præpositionerne kan dog være vanskelige at kategorisere (jf. Messerschmidt et al. 2018). If. ProGram-teorien har de forskellige former for grammatik tilfælles at de per konvention udgør den diskursive baggrund: Grammatiske elementer som bøjningsendelser og grammatiske ord kan ikke udgøre det primære budskab i en ytring. Kun leksikalske elementer (som *ko* og *spis-*) kan udtrykke det primære budskab. Vha. fokuseringstest kan man afgøre om et ord er grammatisk eller leksikalsk (jf. Kristensen & Boye 2016).

Med afsæt i ProGram-teorien kan vi både beskrive rent kombinatoriske grammatikafvigelser som i (6) og afvigelser ved grammatiske items som i (3) - (5).

(6) imens giftede sig hun (3.4.F)

De fleste afvigelser relaterer sig både til kombinatoriksiden og itemsiden, som bøjningsafvigelsen i (7), hvor kombinationen af alkvantoren *hele* og substantivet *nat* afviger da bestemthedssuffikset, et grammatisk item, mangler.

(7) Vi skal danse hele nat (3.2.K)

Med udgangspunkt i ProGram-teorien har vi afgjort om et item er grammatisk eller leksikalsk for også at fange de itemrelaterede afvigelser der typisk ikke opgøres som grammatikrelaterede i andre undersøgelser. Vi definerer altså en afvigelse som en grammatikafvigelse hvis den relaterer sig til kombinatorik eller til brug eller stavning af grammatiske items (grammatiske ord eller grammatiske affikser). Afgrænsningen til ortografi og leksikon byder dog i nogle tilfælde på udfordringer.

4.2. *Grammatik vs. ortografi*

Afgrænsninger af stavfejl og grammatikafvigelser er en kendt problematik i skriftsprogsundersøgelser. Jervelund (2007: 34) diskuterer kategoriseringen af *bedte* for *bad* og *nogle* for *nogen*. Begge typer afvigelser regnes traditionelt for stavfejl selvom *bedte* er en bøjningsafvigelse, og selvom det er brugen, ikke stavningen, af *nogle* der afviger. Når afvigelserne traditionelt regnes for stavfejl, kan det skyldes det praktiske hensyn at “[d]en korrekte fordeling af *nogen* og *nogle* fremgår [...] af Retskrivningsordbogen, ligesom bøjningsformer fremgår af ordbogen” (Jervelund 2007: 34).

I vores undersøgelse er skellet mellem grammatik og ikke-grammatik teoretisk, snarere end praktisk, funderet. Afvigelser i brug eller stavning af grammatiske elementer (jf. Boye & Harder 2012) kategoriseres som grammatikafvigelser uanset hvad Retskrivningsordbogen måtte indeholde af oplysninger. Afvigelser i bøjning (som i *bedte* for *bad* eller i *har mistede* for *har mistet*) og afvigelser i brugen af grammatiske ord (*nogen/nogle*) kategoriseres alle som grammatikafvigelser. Dermed undgår vi at grammatikafvigelser kamoufleres som en del af andre kategorier. Kun afvigelser i leksikalske elementers ortografi (i rødder eller afledningsmorfemer) regnes for rent ortografiske, såsom *tyskland* for *Tyskland*, *tet* for *tæt*, *solbrinner* for *solbriller* og *skatejage* for *skattejagt*.

Denne fremgangsmåde er både teoretisk funderet og operationaliserbar. Den høje interrater reliability i vores undersøgelse for kategorisering af ortografiske elementer viser at skellet er overvejende klart (88,2 %). Der er dog fortsat enkelte forekomster som ligger på grænsen mellem ortografi og grammatik. Eksemplerne (8) og (9) indeholder begge afvigelser i rod-morfemerne for *boger/bøger* og *stør/stor*.

- (8) Boger, Kuglepen, telefon (3.3.E)
 (9) in stør hund (2.3.A)

Afvigelserne kan vidne om kursisternes kendskab til regelmæssige bøjningsmønstre og dermed tolkes som grammatikafvigelser. *Boger* kan ses som regelmæssigt dannet fra *en bog*. *Stør* kan ses som regelmæssigt dannet gennem mønsteret *stør, større, størst*. Dermed minder de om grammatikafvigelsen *bedte* for *bad*. Vi har dog kategoriseret (8) og (9) som ortografiske afvigelser ud fra det princip at afvigelserne forekommer i rodmorfemer, ikke i grammatiske morfemer.

4.3. Grammatik vs. leksikon

Forveksling af ord eller afvigende brug af ord optræder også. Hvis afvigelsen berører et grammatisk ord, som i (3) - (5), ses den som grammatisk. Hvis ordet er leksikalsk (fx *plads* for *sted*), er afvigelsen som udgangspunkt kategoriseret som leksikalsk. Der er dog 15 særlige tilfælde i kategorien ordvalg i vores korpus. I eksemplerne (10) og (11) forveksles to leksikalske elementer, men i begge tilfælde er der tale om forveksling af afledninger (*interessere* for *interessant* og *glæde* for *glad*). Det afvigende valg af ordklasse resulterer i en afvigende kombinatorik. Derfor har vi kategoriseret afvigelserne som grammatiske.

(10) DET ER MEGET INTERESSERE EMNE (3.4.A)

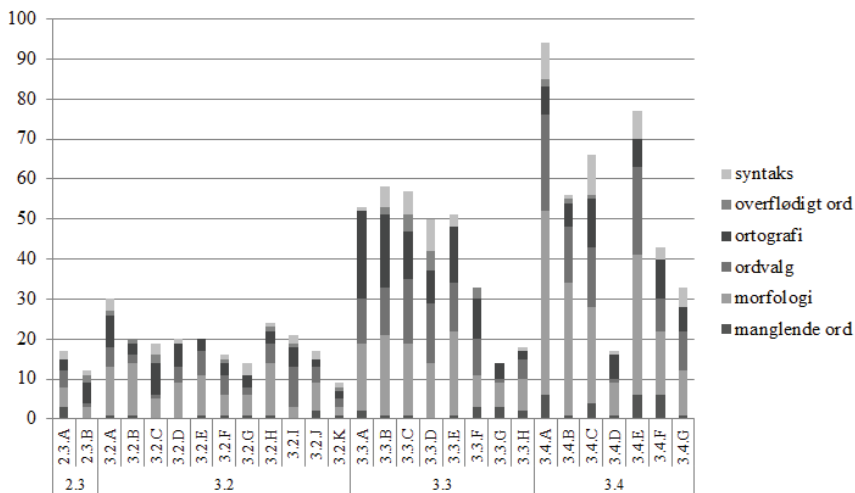
(11) Jeg er glæde for at høre fra dig (3.3.D)

Afvigelsen i (11) med to forvekslede ord der udtales næsten ens, er særligt vanskelig da den også kunne anses for ortografisk.

5. Kort om resultater

Alle 28 kursister har afvigelser i deres tekster, både grammatiske og ikke-grammatiske. Som det fremgår af figur 1 har alle kursister nemlig forekomster af morfologiafvigelser og af ortografiafvigelser, men derudover er der stor individuel variation. Fx har ikke alle kursister syntaksafvigelser (jf. 3.2.B, 3.2.E, 3.3.F og 3.3.G i figur 1).

Umiddelbart ser det ud til at kursisterne på de høje moduler producerer flest afvigelser, men det skyldes at de skriver længere tekster. Fx har 3.4.D flere afvigelser per 100 løbende ord end 3.4.E der har skrevet en længere tekst. Når man opgør hyppigheden af grammatikafvigelser, kan man med fordel inddrage flere forskellige frekvensmål. Man kan som nævnt opgøre frekvensen ift. hvor tit afvigelsen forekommer i en tekst: som absolut frekvens eller som antal afvigelser per 100 løbende ord i korpusset. Man



Figur 1: Antal afvigelsestyper per kursist. X-aksen viser kursistnumre ordnet efter danskuddannelse og modul (2.2., 3.2, 3.3 og 3.4). Inden for hvert modul er kursisterne rangordnet, så kursisterne med flest afvigelser per 100 løbende ord står mod venstre. Y-aksen viser antal afvigelser.

	grammatiske	ikke-grammatiske	uafklarede præpositioner	total	andel i procent
ortografi	-	198	-	198	20,7 %
ordvalg	56	125	45	226	23,6 %
overflødig ord	7	7	14	28	2,9 %
manglende ord	16	13	21	50	5,2 %
morfologi	377	-	-	377	39,3 %
syntaks	80	-	-	80	8,3 %
total	536	343	80	959	100 %
andel i procent	55,9 %	35,8 %	8,3 %	100 %	

Tabel 1: Fordeling af undersøgelsens afvigelser på overordnet type.

kan undersøge kursisternes individuelle variation: Hvor mange kursister producerer afvigelsen, og hvor mange afvigelser producerer hver kursist i snit? I undersøgelser af *learner errors* er det også oplagt at rapportere andelen af potentielle afvigelser (jf. Thewissen 2015). Ortografiafvingelser kan forekomme i alle slags ord, men grammatikafvingelser er kontekstbestemte. Der er kun potentiale for en afvigelse i verbalmorfologi hvis der er et verbum. I alt opmærkede vi 131 verbalafvingelser i korpussets 1050 verber. Andelen af potentielle verbalafvingelser er således 12,5 %.

5.1. *Grammatikafvingelser samlet set*

Tabel 1 giver et overblik over det totale antal afvingelser samt fordelingen mellem grammatikafvingelser og ikke-grammatikafvingelser. Grammatikafvingelserne udgør i alt 55,9 % af det samlede antal afvingelser, mens ikke-grammatikafvingelser udgør 35,8 %. En kontrastiv undersøgelse af vores korpus vs. et tekstkorpus for skribenter med dansk som L1 vil kunne afgøre om den høje andel grammatikafvingelser udelukkende kan tilskrives det brede grammatikbegreb, eller om den kan skyldes forskelle mellem dansk som L1 og L2.

Kategorien ortografi indeholder ikke grammatikafvingelser (jf. afsnit 4.3), men de øvrige kategorier i tabel 1 gennemgår vi herunder. Ordvalgsafvingelser og manglende og overflødige ord kan være grammatiske. Syntaksafvingelser og morfologiafvingelser er altid grammatiske.

5.2. *Ordvalg*

Med 226 forekomster udgør ordvalgsafvingelser 23,6 % af det totale antal afvingelser. Heraf er 125 ikke-grammatiske, fx leksikalsk forveksling af *synes* for *tror* eller *folk* for *mennesker*. Derudover er der 56 grammatikafvingelser, fx forveksling af grammatiske ord som hjælpeverber (*har* blev *venner med*) og konjunktioner (12). Desuden er der forveksling på tværs af ordklasser, som i (3), (10) og (11).

- (12) De kom i tirsdags omkring klokken to i aften når vi var ikke hjem
(3.3.C)

Endelig er der 45 eksempler på forveksling af præpositioner, hvis grammatiske status er uafklaret, jf. afsnit 4.1.

5.3. Manglende og overflødige ord

Manglende og overflødige ord i korpusset er oftest præpositioner, fx *Ring mig!*, *Hils til familien*. Manglende ord er grammatiske i 16 af 50 tilfælde, mens overflødige ord er grammatiske i 7 af 28 tilfælde. De grammatiske ord er typisk infinitivmarkøren *at*, jf. eksemplerne (4) og (5), konjunktioner og artikler: *jeg har en fødselsdag*. Leksikalske ord er typisk pronominer og substantiver, fx *et par stykke øl*.

5.4. Syntaksafvigelser

Syntaksafvigelser drejer sig hyppigst om placering af sætningsadverbial (27 afvigelser) og om manglende inversion af subjekt og verbal (23 afvigelser). Placeringen af sætningsadverbialet afviger hos 15 kursister: i 17 ledsætninger og 10 helsætninger. I 14 ud af de 17 ledsætninger er der valgt helsætningsordstilling, i resten er adverbialet placeret andre steder i ledsætningen. I 8 af ledsætningerne med helsætningsordstilling kan man jf. afsnit 3.1. diskutere om ordstillingen afviger. I de resterende 6 er ordstillingen afvigende som i den adverbiale ledsætning i (13).

- (13) Hvis parken er ikke derhen, må vi cykle 5 km til næsten park (3.4.C)

Vi forventede at manglende inversion af subjekt og verbal ville være udbredt blandt kursisterne. Afvigelsen forekommer dog kun hos 10 kursister, måske fordi de undgår at producere sætninger hvor der skal være inversion. Når sådanne sætninger produceres, er andelen af afvigelser til gengæld høj: 25 % af alle deklarativer sætninger med adverbial, ledsætning eller objekt i forfeltet mangler inversion. Kursisterne har tilsyneladende godt styr på inversion ved spørgsmål og topikaliseret objekt. Udfordringen ligger efter adverbialer og ledsætninger i forfeltet:

- (14) Derefter jeg håber vi skal danse (3.2.C)

- (15) Hvis jeg kunne sige en ting til om mit opdragelse, det er at jeg er glad for det (3.4.E)

5.5. Morfologiafvigelser

De mest fremtrædende morfologiafvigelser i korpusset vedrører verbalformer (34,7 %), adjektivbøjning (22,5 %) samt nominalers numerus (15,6 %) og genus (10,9 %). Alle 28 kursister har afvigelser i verbalformer (4,7 i gns.). Som nævnt afviger hele 12,5 % af verberne i korpusset. Især infinitiv er vanskelig (57 tilfælde): I 22,5 % af de tilfælde hvor infinitiv er den norme-

rede form, bruges en anden form i stedet: præsens (39 tilfælde), imperativ, præteritum eller perfektum participium:

- (16) Glæder mig til at fortæller mere (3.3.F).
 (17) Derfor skal vi find en løsninger (3.4.C)

Afvigelsen findes typisk efter modalverber (40 tilfælde), som i (17). En mulig forklaring på brugen af præsens og imperativ i stedet for infinitiv kunne være at de tre former ofte er sammenfaldende på engelsk, undtagen 3.sg. På dansk bruger kursisterne tilsyneladende præsens som default. Kun sjældent bruges infinitiv for præsens: Afvigelsen udgør 3,2 % af alle verber i præsens.

Adjektivernes bøjning afviger hos hele 23 kursister (3,7 i gns.). Afvigende brug af nøgen form (for enten e- eller t-form) udgør 60 % af de samlede adjektivafvigelser, fx *offenlig torve* (3.3.E) og *vejret var varm* (3.3.A). Samme mønster ses i Janiks (2017: 98) undersøgelse af adjektivbøjning hos learnere med engelsk som modersmål og norsk som andetsprog. Adjektivafvigelser har både at gøre med genus, numerus og bestemthed, men det er svært at skelne mellem afvigelser i bestemthed vs. numerus da e-formen kan dække over begge dele.

Afvigelser i genus og numerus forekommer også ved andre ordklasser. Der er 59 numerusafvigelser hos 20 kursister på alle niveauer (3,1 afvigelser i gns.). De findes mest i substantiver, men også i pronominer og artikler. Det kan være uautoriserede former (*gode tings*), singularis for pluralis (*min kollegaer*) og, mindre hyppigt, pluralis for singularis (*mine kone*).

I substantiver, pronominer og artikler ses også genusafvigelser hos 17 kursister (2,4 afvigelser hver). Genusafvigelser ses hyppigere i ubestemte artikler (13,6 %, fx *en budget*) end for suffikser (2,5 %, fx *togen*), når vi ser på andelen af potentielle afvigelser. Utrum bruges oftere for neutrum end omvendt (26 vs. 15 forekomster). Eneste undtagelse fra dette mønster er de fritstillede pronominer: Her er der kun afvigelser, hvor *det* er brugt for *den*.

6. Konklusion og perspektiver

Korpusundersøgelsen viser at det ikke er let at definere hvad en grammatikafvigelse er. Eftersom L1-talere også producerer afvigelser, kan de ikke bruges som sammenligningsgrundlag. I stedet sammenlignes med en norm. Desuden kan det være vanskeligt at afgøre om et normbrud er relateret til grammatik, ortografi eller leksikon.

Til trods for disse udfordringer har vi præsenteret teoretisk funderede

definitioner der i praksis muliggør en skelnen mellem grammatik og ikke-grammatik. Vi har identificeret dele af dansk grammatik som er særligt vanskelige når engelsk er L1, herunder adjektivbøjning, brug af infinitiv og inversion efter adverbialer og ledsætninger. Vores undersøgelse har en høj interrater reliability og kun få uafklarede tilfælde: de svært klassificerbare præpositioner.

Da tidligere danske sprogfejlsundersøgelser har kategoriseret en del grammatikafvigelser som stavfejl, er det ikke klart om andelen af grammatikafvigelser i vores kursiststile afviger fra andelen af grammatikafvigelser hos fx gymnasieelever (Johannsen 2012) og universitetsstuderende (Blom et al. 2017). Men fremtidige kontrastive undersøgelser af dansk som L1 og L2 kan afgøre om det er tilfældet.

Sammenligninger med andre kursistgrupper kan besvare en række spørgsmål relateret til andetsprogsforskning: Hvilke afvigelser kan skyldes tværspørgs indflydelse fra modersmålet eller andre sprog, og hvilke er fælles for alle lærnere uanset L1? Skal grammatikundervisningen tilrettelægges forskelligt afhængigt af kursistens modersmål?

Vores grammatikdefinition har den fordel at den dækker en bred vifte af grammatisk fænomener som i andre undersøgelser ville blive kategoriseret som ortografiske eller leksikalske. Dermed får vi et overblik over forskellige typer grammatik som fremover kan bruges til at undersøge hvordan forskellige typer grammatikafvigelser i en tekst påvirker læsning og forståelse. Fremadrettet håber vi vha. læseforsøg med naturalistiske grammatikafvigelser at kunne opstille en model for processering og forståelse af naturligt forekommende grammatikafvigelser.

Om forfatterne

Katrine Falcon Søby, ph.d.-studerende, Broken Grammar and Beyond, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.
Line Burholt Kristensen, ph.d., Broken Grammar and Beyond, Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet.

Litteratur

Blom, J.N., Rathje, M., Jakobsen, B. le E., Holsting, A., Hansen, K.R., Svendsen, J.T., Vildhøj, T.W., Lindø, A.V. (2017): Linguistic Deviations in the Written Academic Register of Danish University Students. *Oslo Studies in Language* 9 (3), 169-190.

- Boye, K. & Harder, P. (2012): A usage-based theory of grammatical status and grammaticalization. *Language* 88 (1), 1-44.
- Brink, E. T., Elbro, C. & Johannsen, S. (2014): Sproglige fejl og karakterer i danske gymnasiestyle. *Mål og Måle* 36 (2), 5-9.
- Broken Grammar and Beyond (<http://www.nors.ku.dk/forskning/projekter/broken-grammar-and-beyond/>) (lokaliseret d. 5. januar 2019).
- Christensen, T. (2012): Læs venligst her - om at tale venligt. *Danes, Magasinet for danskere i udlandet* 6, 26-27.
- Christensen, T. & Jensen, T. (2014): Det er klart at ledsætninger har da ledsætningsordstilling – eller har de? *Nyt fra Sprognævnet* 2014 (1).
- Hansen, E. & Heltoft, L. (2011): *Grammatik over det Danske Sprog*. Det Danske Sprog- og Litteraturselskab, København: Syddansk Universitetsforlag.
- Harder, P. (2010): *Meaning in Mind and Society. A Functional Contribution to the Social Turn in Cognitive Linguistics*. Berlin: Mouton de Gruyter.
- Janik, M. (2017): Positive and Negative Transfer in the L2 Adjective Inflection of English-, German- and Polish-speaking Learners of L2 Norwegian. I: *Crosslinguistic Influence and Distinctive Patterns of Language Learning. Findings and Insights from a Learner Corpus*. Golden, A., Jarvis, S. & Tenfjord, K. (red.). Bergen: Multilingual Matters.
- Jervelund, A. Å. (2007): *Sådan staver vi, om ortografi og stavfejl*. Frederiksberg: Dansk Sprognævn/Dansklærerforeningen.
- Johannsen, S. (2012): Fejl og karaktergivning i danske gymnasiestyle. *Mål og Måle* 35 (1), 18-23.
- Kristensen, L. B. & Boye, K. (2016): Svært ved Grammatikken – del 3. *Logos* 77, 27-31.
- Messerschmidt, M., Boye, K., Overmark, M., Kristensen, S. & Harder, P. (2018): Sondringen mellem grammatiske og leksikalske præpositioner. *Ny forskning i grammatik* 25, 89-106.
- Lennon, P. (1991a): Error and the Very Advanced Learner. *International Review of Applied Linguistics in Language Teaching* 29 (1), 31-44.
- Lennon, P. (1991b): Error: Some Problems of Definition, Identification, and Distinction. *Applied Linguistics* 12 (2), 180-196.
- Pawlak, M. (2013): *Error correction in the foreign language classroom, reconsidering the issues*. Heidelberg: Springer.
- Thewissen, J. (2015): *Accuracy across Proficiency Levels. A Learner Corpus Approach*. Louvain-LaNeuve: Presses Universitaires de Louvain.
- University of Sheffield (<http://www.sheffield.ac.uk/international/english-speaking-countries>) (lokaliseret d. 4. januar 2019).