

Moderne importord over tid i høj opløsning – foreløbige resultater

Philip Diderichsen
Dansk Sprognævn

1. Introduktion

Hvor mange udenlandske ord er der i dansk? Kommer der flere eller færre af dem? Hvordan ændrer de sig når de kommer ind i det danske sprog? Hvad synes folk om dem?

Spørgsmål som disse kan nyordsforskningen være med til at besvare. Spørgsmålet om mængden og udviklingstendenserne, som er i fokus i denne artikel, har været genstand for videnskabelig interesse i hvert fald siden 1980'erne, hvor der begyndte at komme korpusbaserede undersøgelser af sagen i de nordiske lande – især Norge og Sverige. I første omgang foregik undersøgelserne af ordimporten i de nordiske sprog i spredt fægtning hvad angår undersøgelsesernes omfang, metode osv. Den generelle konklusion fra disse tidlige undersøgelser er at engelske og andre importerede ord er i klart undertal med forekomstfrekvenser på mellem 0,1 og 3,4 % afhængigt af tekstarten, med mere alment sprogligt materiale såsom blandet avistekst generelt i den lave ende af intervallet og mere specialiserede tekster/genrer (herunder også talesproglige) generelt i den høje ende (Graedler, 2017).

I 2000'erne kom resultaterne af den store fællesnordiske undersøgelse Moderne importord i språka i Norden (forkortet MIN-projektet), som var en ambitiøs, koordineret indsats for at undersøge ordimporten i de nordiske sprogsamfund på en sammenlignelig måde. Visse stereotype forestillinger blev her gjort til skamme idet man fx kunne vise at Norge ikke som bredt antaget var et af de mest puristiske

sprogsamfund, men tværtimod var det sprog der havde den største ordimport i redaktionel avistekst (som projektet brugte som bedste repræsentant for alment sprog). Projektet havde også en tidsdimension og kunne vise at der i samtlige nordiske sprogsamfund kunne registreres større andele af moderne importord i 2000 end i 1975 (hvilket dog til dels kan forklares af at definitionen af moderne importord er ord der er kommet ind i et givet sprog efter anden verdenskrig, altså efter 1945). (Se Brunstad (2003), Graedler (2017), Sandøy (2007), Sandøy (under udg.)).

Forskningsprojektet *Yes, det er coolt. Om påvirkning af dansk fra andre sprog* (bedre kendt som Moderne importord i dansk, forkortet MID-projektet), som denne artikel er en del af, lægger sig i forlængelse af MIN-projektet med det formål at undersøge hvordan det er gået med ordimporten i dansk i den mellemliggende periode, bl.a. mht. importordenes hyppighed (jf. Heidemann Andersen (2019) og Diderichsen & Schack (2019)).

Heidemann Andersen har undersøgt hyppighedsspørgsmålet ved at gentage MIN-projektets excerpering af moderne importord i aviser. Hun finder, med visse forbehold for sammenlignelighed og repræsentativitet af data, et overordnet fald i importordsfrekvens i et materiale fra 2016 (et eksemplar af hhv. Politiken, Århus Stiftstidende og BT) i forhold til et (nogenlunde) tilsvarende materiale fra MIN-projektet (et eksemplar af hhv. Politiken, Århus Stiftstidende og Ekstra Bladet fra år 2000). I 2000-materialet er importordsfrekvensen på 0,97 moderne importord pr. 100 løbende ord, mens den i 2016-materialet er på 0,58 – inklusive importord i annoncer. I det redaktionelle tekstmateriale alene er der en lignende forskel (Heidemann Andersen, under udg.).

Disse resultater indikerer hvordan den relative hyppighed af importord samlet set har bevæget sig siden 2000 – her altså tilsyneladende i nedadgående retning. Den foreliggende artikel er en mere detaljeret undersøgelse af de moderne importords frekvens over tid. Det nye her er at man kan se hvordan brugen af det enkelte importord bevæger sig over tid, ikke kun hele gruppen af importord på en gang. Som Graedler bemærker, kræver korpusundersøgelser af enkeltord ”svært store” korpusser (Graedler, 2017:167). Det gælder selvfølgelig i endnu højere grad over tid, hvor der er brug for flere

store korpusser efter hinanden. En sådan serie af korpusser har vi til rådighed i Dansk Sprognævn sammen med et særligt udviklet værktøj, Tidslinjen, der gør undersøgelser af ordfrekvens over tid enkle og ligetil, jf. Diderichsen (2019b). Søgningerne i det større korpus kan også eliminere visse skævheder. Fx kan nogle ord være overrepræsenteret i en excerptundersøgelse fordi de bogstaveligt talt er dagsaktuelle og optræder hyppigt en enkelt eller nogle få dage og så stort set aldrig mere. Med vores tidslinjeverktøj har vi således mulighed for at få et retvisende billede af de enkelte importords frekvens over tid.

Jeg undersøger i denne artikel et større udvalg af de importord der er excerperet i MID-projektet. Spørgsmålet er hvordan det fald Heidemann Andersen (under udg.) observerer, ser ud når man undersøger de excerperede ord lidt nærmere. I sidste ende er målet at få klarlagt om brugen af moderne importord virkelig er faldende i det danske skriftlige almensprog, og i så fald hvor meget. Som det vil fremgå, kan de resultater der præsenteres her, kun i begrænset grad bidrage til dette da hele materialet endnu ikke er undersøgt. Men der kan dog dokumenteres relevante tendenser indenfor nogle udvalgte emnesfærer, og man kan desuden få et indtryk af hvordan den komplette undersøgelse vil kunne give et bredere, skarpere, mere nuanceret billede af importordenes karriere i ordforrådet.

2. Korpus

Undersøgelsen er baseret på to forskellige tekstmaterialer: dels de tre aviseksemplarer der blev excerperet i MID-projektet, og dels et stort aviskorpus. Aviseksemplarenes rolle er at levere den stikprøve af moderne importord der undersøges; korpusset udgør den totalitet af ord som de excerperede importords mere nøjagtige frekvens over tid regnes ud i forhold til.

Korpusset består af store dele af teksterne fra de landsdækkende aviser Berlingske, Information, Jyllands-Posten, Kristeligt Dagblad, Politiken og Weekendavisen i perioden fra og med december 2009 til og med maj 2018. Tidsspændet er således mindre end spændet mellem den seneste excerpering i MIN-projektet fra 2000 og MID-projektets excerpter fra 2016. Til gengæld er korpusset langt mere omfangsrigt end de excerperede aviser, og resultaterne kan dermed også blive langt mere nuancerede. Korpusset består i alt af 280.366.288 løbende ord

fordelt på 104 måneder – således altså godt 100 månedskorpusser a knap 3 mio. løbende ord.

Tekstmaterialet i korpuset er anderledes end de excerperede aviseksemplarer på måder man må holde sig for øje i det følgende. Korpuset indeholder fx ikke tekstmateriale fra annoncer. Desuden er kilderne anderledes: Korpuset indeholder kun tekst fra store, landsdækkende aviser minus formiddagsavisene, hvorimod de excerperede aviser var et eksemplar af hhv. Politiken, Århus Stiftstidende og BT (som blev valgt for at sikre sammenligneligheden med de tre tilsvarende aviseksemplarer i MIN-projektet).

Den vigtigste konsekvens af disse forskelle har med generaliserbarhed at gøre. Korpuset består af redaktionel tekst fra landsdækkende ikke-tabloidaviser, og de fundne resultater kan derfor heller ikke uden videre generaliseres til andre typer tekst end netop denne afgrænsede type.

3. Materiale

Det her undersøgte datamateriale består af ord der er excerperet fra tre aviser (et eksemplar af hhv. Politiken, Århus Stiftstidende og BT) i overensstemmelse med Selback (2007) – se ordene i afsnittene nedenfor. Ordene er indtil videre klassificeret i kategorierne ”ordform”, ”lemma”, ”ordklasse”, ”fra”, ”oprindelig fra”, ”hybrid”, ”teksttype”, ”genre”, ”emne”, ”år”, ”avis”, ”sektion”, ”side” og ”spalte”. Korpusundersøgelsen er afgrænset i forhold til disse klassifikationer som følger. Navne er udeladt for at kunne sammenligne med MIN-projektets tal, hvor navne også er udeladt. Programoversigter for TV og radio er også udeladt, igen for at sikre sammenligneligheden med MIN-projektet. Excerpter fra annoncetekst sorteres fra eftersom det korpus der er til rådighed, som nævnt ikke indeholder denne teksttype. Endelig ser jeg bort fra flerordsexcerpter da konkrete flerordsforbindelser erfaringsmæssigt er sjældne og derfor antagelig med fordel kan behandles kvalitativt.

Eftersom tidslinjeværktøjet ikke kan aftvetydiggøre ord vha. grammatiske annotationer (fx ordklasser), er de excerpter der søges efter principielt nødt til at være entydige. Her opererer jeg dog med omtrentlig entydighed således at jeg accepterer ord som entydige, så længe deres homografer har samme semantiske kerne. Fx kan *blog*

være et verbum (infinitiv af at *blogge*) eller et substantiv, men formen betragtes her som entydig fordi en forestilling om det at blogge i begge tilfælde aktiveres.

Der findes selvfølgelig også ord i datamaterialet der ikke er entydige. Nogle af excerpterne måtte undersøges nærmere i korpuset, og af disse accepterede jeg nogle som tilstrækkeligt entydige, mens jeg forkastede andre. Et par eksempler kan ses her. Bemærk at excerpter konsekvent skrives med versaler for at signalere at der søges versalblindt i korpuset, dvs. uden forskel på små og store bogstaver.

- Et *medie* kan iflg. Den Danske Ordbog (DDO, se ordnet.dk/ddo) være et kommunikationsmiddel, en spiritist m.m. I korpuset dominerer kommunikationsmiddelbetydningen totalt: Over 99 % af forekomsterne har denne reference. *MEDIER* og *MEDIERNE* tages derfor med i undersøgelsen.
- Et *net* kan iflg. DDO være diverse typer af snore- eller trådnæt, netværk af fx veje m.v. og endelig en forkortelse for *internet*. Ordet *NETTET* har i omkring 90 % af tilfældene den sidste betydning og beholdes derfor.
- Derimod må ordet *REFORMER* forkastes da den excerperede betydning (en særlig træningsform) langt overskygges af det danske ord *reform* i pluralis, der udgør over 99 % af forekomsterne.

Af i alt 1824 excerpter efterlader disse afgrænsninger 549, nogenlunde jævnt fordelt på de tre excerperede aviser: I BT er der 197 excerpter, i Politiken 213 og i Århus Stiftstidende 139.

Antallet af forskellige ordformer blandt excerpterne er 399. Det er begrænset hvor mange gange de enkelte ordformer forekommer – langt de fleste forekommer kun en gang.

Jeg vælger i denne undersøgelse at holde mig indenfor netop de ordformer der er excerperet, og ikke fx inddrage de øvrige bøjningsformer fra det enkelte leksem, fordi jeg på den måde kommer tæt på et tilfældigt udvalg (en stikprøve) af moderne importord i redaktionel avistekst samtidig med at den direkte sammenlignelighed med Heidemann Andersens excerpter (Heidemann Andersen 2019). Endnu bedre ville det have været at undersøge en stikprøve excerperet

fra et antal korte tekstprøver trukket tilfældigt fra en længere periode (jf. fx Diderichsen & Schack (2019) og Ruus (1995)) – hvilket på den anden side ville være gået ud over sammenligneligheden med excerptundersøgelsen.

4. Moderne importord over tid i udvalgte domæner

I det følgende undersøges de udvalgte avisexcerpter fra MID-projektet for hvordan deres forekomstfrekvens har ændret sig måned for måned over den undersøgte 8-årsperiode. For at kunne sige noget om udviklingen der på den ene side er mere nuanceret end alle ordenes samlede hyppighedsudvikling og på den anden side mere generelt end hvert enkelt ords udvikling, har jeg inddelt ordene i domæner. Samtidig har jeg på denne måde undgået at måtte gøre alting på én gang, og nedenfor fremlægger jeg således ikke en foreløbig analyse af hele excerptmateriale, men derimod komplette delanalyser af tre sådanne domæner.

Ifølge Graedler (2017:169) blev importord i mange ældre studier klassificeret efter hvilke emner de refererer til, hvor nyere studier snarere inddeler hele tekster efter emner og derefter opgør antallet af importord indenfor disse. I MIN-projektets analyser af afløsningsord blev der også udvalgt et antal domæner til særlig undersøgelse (Jarvad 2007:87; Kvaran 2007b:14), nemlig *IT*, *mad*, *boldspil* og *ungdomsmusik*, simpelthen med det formål at finde mere materiale, idet man i disse domæner formodede ”at der allerede eksisterede en del afløsningsord i alle seks sprog” (Kvaran 2007a:14).

Min opdeling følger samme princip, nemlig at inddele selve ordene i semantiske domæner snarere end at måle det samlede importordsindhold indenfor visse stofområder. Det gør at de resulterende ordgrupper kan ses som en generalisering af de enkelte ords betydning, som så kan følges over tid.

Opdelingen er foretaget således at ord jeg syntes hørte til de samme domæner, blev grupperet sammen. Mit kriterium for at inkludere de forskellige ord i de forskellige domæner var at de var nogenlunde specifikke og karakteristiske for de således fremkomne grupper. De ord der efter min mening ikke kunne opfylde dette kriterium, placerede jeg i gruppen *Uspecifik*. Min kategorisering af ordene blev efterprøvet ved at en kollega kategoriserede en stikprøve af ordene i de samme

domæner jeg havde fundet frem til, og *intercoder agreement*-tests viste substantiel overensstemmelse mellem de to kodninger, hvilket giver kategoriseringen en rimelig validitet (Diderichsen 2019a). Samtlige domæner jeg kom frem til, er anført her:

Anmelderord, Computere/ny teknologi, Film-/tv-branchen, Fritid, Gambling, Helse, Identitet, Kriminalitet, Mad/restauranter, Managementjargon, Mode, Musik, Nye medier, Politik/økonomi, Rusmidler, Shopping, Sladder, Slang, Sport, Traditionelle medier, Transport/rejser, Uddannelse, Uspecifik.

Nedenfor analyserer jeg domænet *Nye medier* med *Traditionelle medier* som sidegevinst samt domænet *Uspecifik*.

4.1 Nye medier

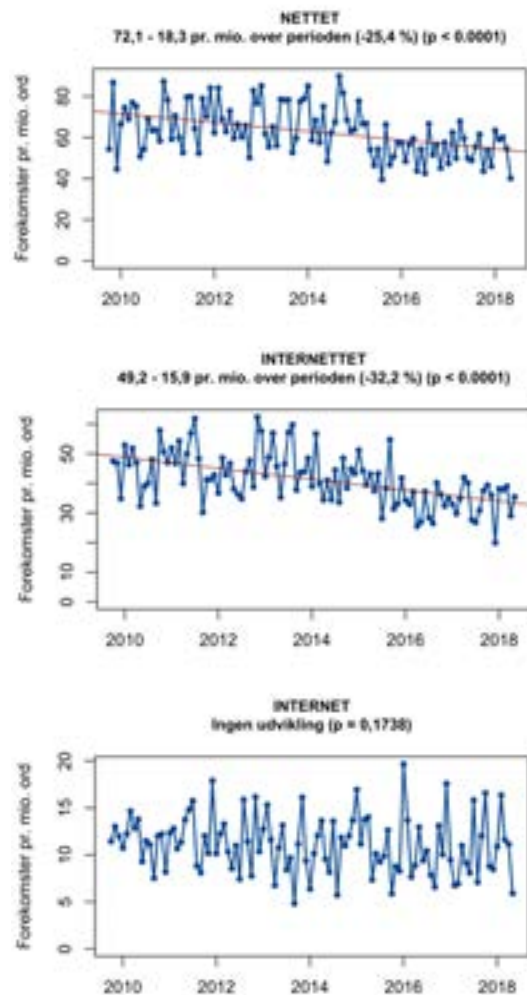
Herunder vises tendenser for ord i domænet ”nye medier”. Ved hver graf vises en p-værdi som angiver om tendensen er statistisk signifikant. P-værdien er baseret på en lineær regression. Hvis p-værdien er statistisk signifikant, vises der også en lineær tendenslinje, og der vises et tal for hvor mange forekomster pr. million tendensen er udtryk for. Hvis tendenslinjens startpunkt ikke er negativt, vises der også et tal for stigningen eller faldet i procent af tendenslinjens startværdi. Man skal i øvrigt være opmærksom på at minimumværdien i de tilgrundsiggende data er 2 (ord med en forekomst på under 2 i de enkelte månedskorpusser er udeladt). Det betyder at punkterne i graferne altid ligger et stykke over 0.

Der er excerperet 51 ord fra domænet ”nye medier”. Heraf er der tidslinjer til de 37. Resten er ikke hyppige nok til at der kan laves en tidslinje (dvs. ordene forekommer ikke mere end én gang i noget månedskorpus). Alle excerperterne er vist her:

BLOG, BLOGGEREN, BLOGGES (ikke fundet), CHATTET, E-MAIL, E-MAILS, EMAIL, FACEBOOK-PROFIL, FACEBOOKEDE, FACERAPE (ikke fundet), FITNESSBLOGS (ikke fundet), HOLLYWOOD-BLOGGEREN (ikke fundet), INSTAGRAM-OPSLAG (ikke fundet), INTERNET, INTERNET-ABONNENTER (ikke fundet), INTERNET-TØJBUTIK (ikke fundet), INTERNETGIGANTEN, INTERNETTET, KATTEVIDEOER, LINK, LINKS, MAIL, MAILEN, MMS, MOBIL, MOBILTELEFONEN, MODEBLOGGER, NETMEDIET, NETTET, ONLINE, PHOTOBOMBEDE (ikke fundet), PHOTOBOMBING (ikke fundet), PODCAST, ROAMING, SELFIE, SKYPE, SMARTPHONE, SMARTPHONEEJERE,

SMARTPHONES, SMS, SMS'ES (ikke fundet), SNAPCHAT-MESSAGE (ikke fundet), STREAM (flertydigt, udgår), STREAME, STREAMET, TWEET, TWEETED, TWEETES (ikke fundet), TWEETS, VIDEO, VIDEOFORBINDELSEN (ikke fundet).

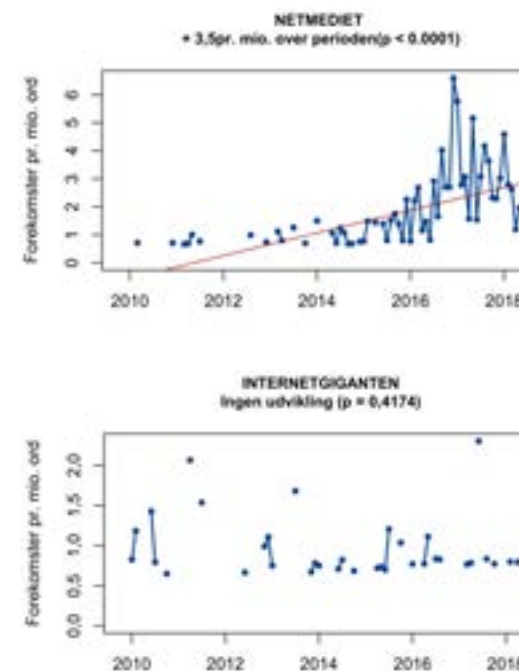
De første ord – med hyppigste først – refererer alle til internettet: NETTET, INTERNETTET, INTERNET.



Figur 1. Tidslinjer for ord der refererer til internettet

INTERNETTET er signifikant faldende over perioden. Og det ser ikke ud til at være fordi den kortere form NETTET er ved at overtage – denne form er også signifikant faldende. Den ubestemte form INTERNET ser ud til at være let faldende, men faldet er ikke statistisk signifikant.

INTERNETGIGANTEN er temmelig sjældent og samtidig ret sporadisk observeret og viser ingen signifikant tendens. Det eneste der er i stigning, er NETMEDIET, som går fra at være ikke-observeret til at vise kraftig vækst indenfor det seneste par år. Gængse ord for det ”gode gamle” internet kunne altså, selv om de vel at mærke stadig er klart de hyppigste i domænet ”nye medier”, se ud til at have haft deres klimaks i avissproget.



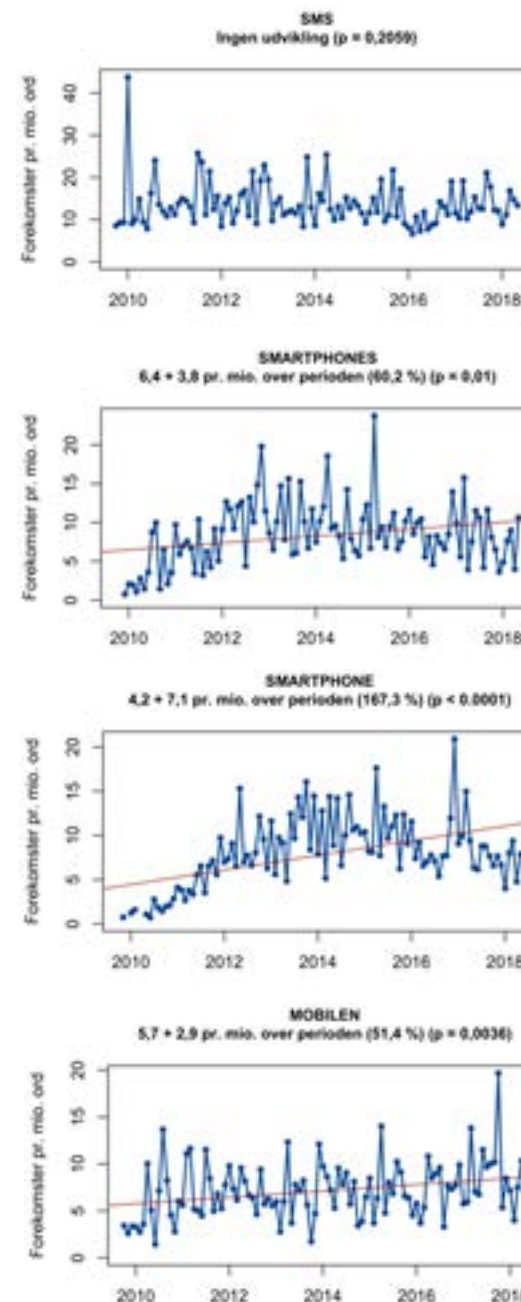
Figur 2. Tidslinjer for NETMEDIET og INTERNETGIGANTEN

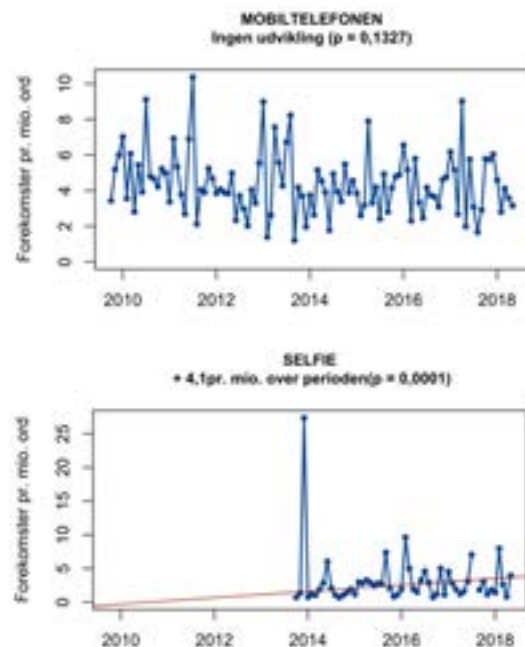
De internetrelaterede ord LINK og LINKS er forholdsvis sjældne, og hhv. uden tendens og let faldende (LINK: ingen udvikling ($p = 0,757$), LINKS: $-1,7$ forekomster pr. mio. over perioden ($-55,5\%$) ($p = 0,0032$)).

MAIL, E-MAIL, E-MAILS, MAILEN og EMAIL viser et lidt blandet billede. Eftersom jeg holder mig til ordformer der faktisk er excerperet, får vi ikke det fulde billede af variationen mellem former med og uden "E-". Jeg har dog smugkigget og kan afsløre at "E"-formerne konsekvent ligger lavere end "ikke-E"-formerne – også hvis "E"-former uden bindestreg lægges sammen med "E"-former med bindestreg. Samlet set synes omtalen af e-mails, som jo kan siges at tilhøre samme "æra" som internettet, ikke ud til at være stagnerende i avissproget (MAIL ($25 + 15,8$ /mio. ($63,5\%$), $p < 0,0001$), E-MAIL ($17,3 - 7,8$ /mio. ($-44,9\%$), $p < 0,0001$), E-MAILS ($2,9 + 3,8$ /mio. ($132,5\%$), $p < 0,0001$), MAILEN ($2,5 + 2,5$ /mio. ($100,3\%$), $p = 0,0001$), EMAIL ($2,7 - 0,7$ /mio. ($-26,8\%$), $p = 0,0651$)).

ONLINE er i stigning over perioden ONLINE: $12,3 + 6,4$ pr. mio. over perioden ($51,9\%$) ($p < 0,0001$). Det kunne tyde på at man måske i højere grad er begyndt at tale om at *være online* end at *være på (inter) nettet*.

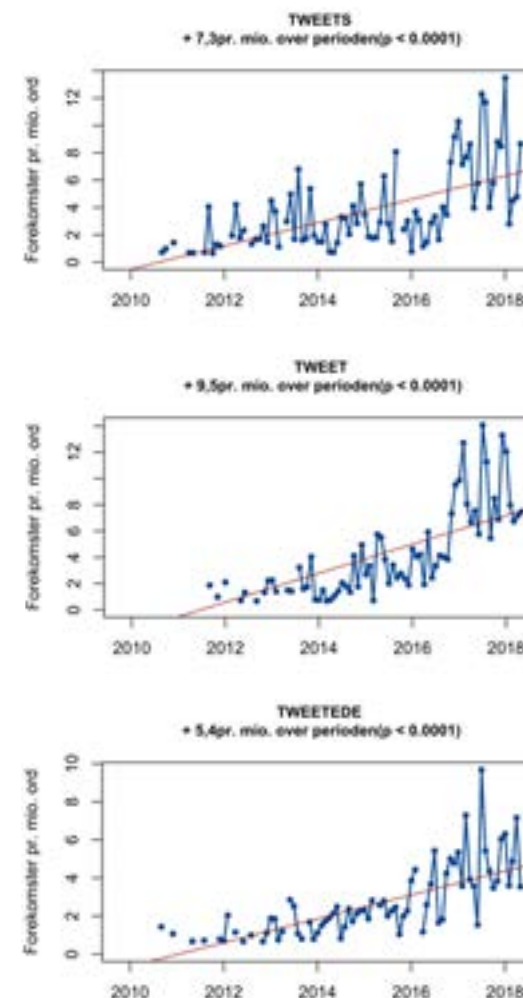
Der refereres signifikant mere til mobiltelefoner i løbet af perioden, især med betegnelsen SMARTPHONE. Dog er det ret tydeligt at der for SMARTPHONE og SMARTPHONES ikke er tale om en lineær stigning. Begge ord synes at nå et klimaks ca. i midten af perioden. Til gengæld er MOBILEN, der ligger på ca. samme hyppighedsniveau, tilsyneladende i mere stabil vækst. Det ligger lige for at forestille sig at det her er nyhedsværdien af smartphonen som sådan vi her ser fordulte og en ny normaltilstand (hvor de fleste mobil(telefon)er er smartphones) indtræffe. Det ældre MOBILTELEFONEN er ikke i stigning, og det er den veletablerede og efterhånden gamle teknologi SMS heller ikke (men det er det hyppigste ord i denne gruppe). Endelig ser vi selfien dukke op i midten af perioden – meget passende netop på det tidspunkt hvor betegnelsen SMARTPHONE er på sit højeste.





Figur 3. Tidslinjer for ord der har at gøre med mobiltelefoner

Brugen af ordet BLOG falder signifikant i perioden, og BLOGGEREN er ikke i stigning. Ordet MODEBLOGGER forekommer sjældent og sporadisk og mest i første halvdel af den undersøgte periode. Bloggen er altså en af de ting der synes at være på vej ud af avisernes spalter: BLOG: 20,9 - 16,7 pr. mio. over perioden (-80,1 %) ($p < 0,0001$), BLOGGEREN: 1 - 0,3 pr. mio. over perioden (-33,5 %) ($p = 0,2626$). Twitter er til gengæld ikke færdig med at vokse i sprogbrugen. Både substantivet TWEETS, det ordklasseflertydige TWEET og ikke mindst verbalformen TWEETEDE er i lettere eksplosiv vækst i perioden.

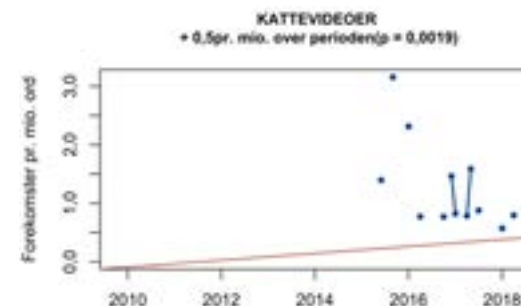
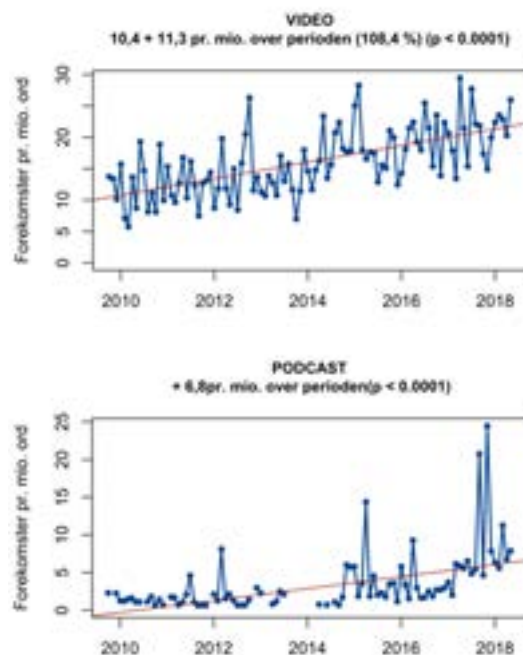


Figur 4. Tidslinjer for ord der har at gøre med det at tweete

FACEBOOK-PROFIL er i moderat vækst over perioden: FACEBOOK-PROFIL: 1,3 + 0,9 pr. mio. over perioden (67,5 %) ($p = 0,0129$). Verbalformen FACEBOOKEDE dukker også op, men kun som et enkelt lille blip på radaren i løbet af 2011.

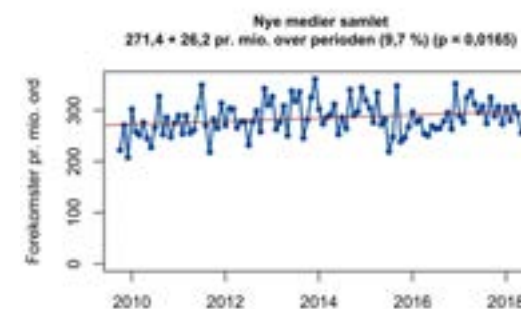
Brugen af verbalformerne STREAME og STREAMET er sjælden, men vokser signifikant over perioden, i begge tilfælde til over det dobbelte: STREAME: 0,7 + 0,6 pr. mio. over perioden (82,4 %) ($p = 0,0655$), STREAMET: + 0,9 pr. mio. over perioden ($p < 0.0001$).

VIDEO er solidt stigende i perioden, og PODCAST (der også kan være en videodistributionsform (jf. DDO), men i korpusset overvejende refererer til lyd) viser en kraftig stigning i slutningen af perioden. SKYPE (der er flertydigt mellem videofoniapplikationen og det at bruge den) ligger stabilt på et lavt niveau (omkring 3 forekomster pr. mio.) over hele perioden: SKYPE: ingen udvikling ($p = 0,4509$). Endelig er KATTEVIDEOER nævnt her fordi kattevideoer notorisk udgør meget af indholdet på netop *videoer* distribueret på *streamingtjenester* som YouTube m.fl. Ordet har tilsyneladende haft sin foreløbige storhedstid ved årsskiftet 2015/2016. (VIDEO er i øvrigt blevet tjekket for om det er den digitale eller den analoge betydning der dominerer i korpusset, og af 100 tilfældige forekomster i korpusset har kun én entydigt med det gamle analoge videomedie at gøre (i konstruktionen ”optage noget på video”).



Figur 5. Tidslinjer for VIDEO, PODCAST og KATTEVIDEOER

Ordene ovenfor er ordformer som faktisk er blevet excerperet i aviser fra en tilfældig dato i perioden, som beskrevet. De udgør på den måde en stikprøve af moderne importord i domænet *Nye medier* fra perioden. Jeg vil her slå de 37 ord sammen og vise en samlet tendens for ordene. Lægger man alle forekomsttallene sammen og dividerer med korpusstørrelsen pr. måned for at få en relativ frekvens, ser man følgende samlede udvikling i netop disse 37 ords andel af ordstoffet i nationale avistekster. Samlet set stiger denne andel fra omkring 271 til omkring 298 forekomster pr. million over perioden, svarende til 9,7 %.



Figur 6. Samlet tidslinje for ordene i domænet *Nye medier*

I tal der måske er lidt nemmere at forholde sig til, svarer udviklingen til at et typisk aviseksemplar med 50.000 løbende ord går fra at indeholde ca. 14 forekomster af ”nye medier”-ordene i starten af perioden til at indeholde ca. 15 i slutningen.

En anden måde at anskue det på er at ekstrapolere til den samlede andel af moderne importord i avistekst (annoncer undtaget). Hvis man antager at denne andel er på ca. 0,5 % i starten af perioden (jf. Heidemann Andersen 2019), og stigningen på 9,7 % gælder for alle moderne importord samlet set, ender den samlede andel med at være på ca. 0,55 % i slutningen af perioden. Anskuet på den måde er det klart at det er en ret ubetydelige stigninger vi har at gøre med. At denne stigningstakt dog ikke nødvendigvis er generel, vil vise sig nedenfor.

4.2 Sociale og andre medier

I dette afsnit analyserer jeg domænet ”traditionelle medier” med udtrykket *sociale medier* som indfaldsvinkel.

Sociale medier er et vigtigt udtryk fra domænet ”nye medier”, der også forekommer blandt excerpterne. Det er ikke med i den ovenstående gennemgang da ordforbindelser jo er sorteret fra. Det er dog så væsentligt et nyt udtryk at det alligevel skal med. Det trækker samtidig domænet ”traditionelle medier” med sig.

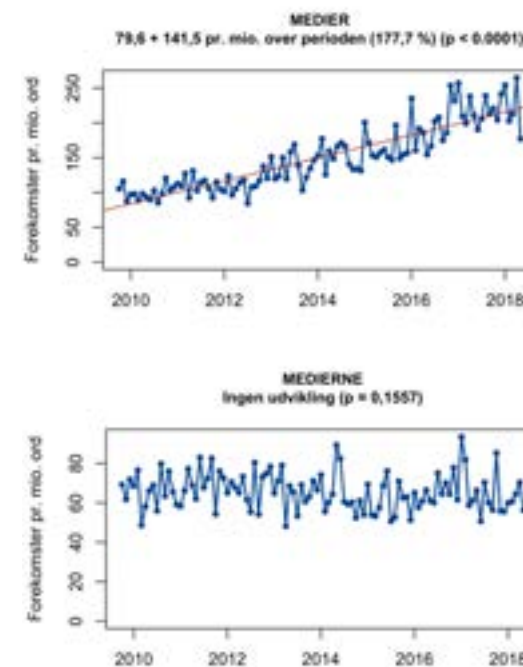
SOCIALE MEDIER er i kraftig vækst, hvilket ikke er mærkeligt eftersom sociale medier først begyndte at dukke op i midten af 00’erne.



Figur 7. Tidslinje for udtrykket SOCIALE MEDIER

Antallet af forekomster er steget støt fra et nulpunkt i starten af perioden til omkring 100 forekomster pr. mio. i slutningen af perioden. Tilvæksten er tre gange så stor som den samlede tilvækst for alle ordene i domænet ”nye medier” ovenfor.

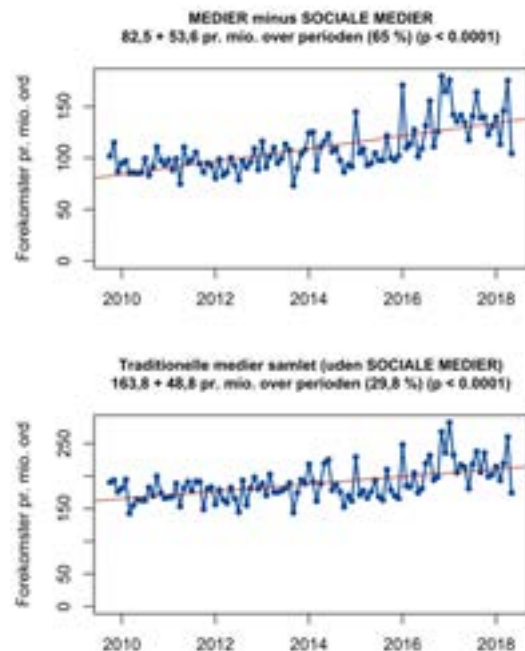
Denne meget betydelige stigning skal man også være opmærksom på når man ser på ord for medier i det hele taget. Jeg har kategoriseret et lille antal ord som tilhørende domænet ”traditionelle medier”, herunder netop ordet MEDIER. MEDIER ser ud til at være i kraftig vækst hvorimod formen MEDIERNE ikke viser nogen signifikant udvikling, som man kan se i graferne nedenfor.



Figur 8. Tidslinjer for MEDIER og MEDIERNE før filtrering for SOCIALE MEDIER

En søgning på MEDIER vil imidlertid også finde alle forekomster af SOCIALE MEDIER. Derfor viser kurven for MEDIER ikke kun udviklingen for traditionelle medier, tværtimod: Det er især udviklingen

i SOCIALE MEDIER der præger kurven. Hvis man simpelthen trækker forekomsterne af SOCIALE MEDIER fra forekomsterne af MEDIER, får man derfor et mere reelt billede af udviklingen i domænet ”traditionelle medier”. Men kurven for MEDIER fratrasket SOCIALE MEDIER viser trods alt en anseelig stigning på 65 %:



Figur 9. Tidslinjer for MEDIER fratrasket SOCIALE MEDIER samt tidslinje for hele domænet *Traditionelle medier* fratrasket SOCIALE MEDIER

MEDIER dominerer i domænet ”traditionelle medier” idet det både er klart det hyppigste og er det eneste af ordene der er i signifikant stigning. De andre ord i domænet er DEADLINE, LAYOUT og LAYOUTET, foruden MEDIERNE. Den ganske pæne stigning man kan se i domænet ”traditionelle medier”, skyldes således udelukkende formen MEDIER.

4.3 Uspecifikt domæne

I det følgende gennemgår jeg et andet domæne – nemlig det uspecifikke domæne hvor ordene ikke kun tilhører et enkelt domæne, men flere, og evt. alle (dvs. ingen specifikke domæner). Et ord der tilhører mere end ét domæne kunne være *Bladerunneren*, som både hører til i para-atletik (hvor en løber med såkaldte blades i stedet for underben kan omtales sådan) og i et bestemt science fiction-univers (hvor en Blade Runner er en der jager menneskelignende robotter). Et ord der kan fungere i mange domæner er fx *teenager*.

Jeg har kategoriseret 54 ord som hørende til dette domæne. Heraf er 39 fundet og analyseret – kun 39 pga. de ovennævnte huller i korpusset og fordi flertydige ord udgår:

ACTION, BLADERUNNEREN (ikke fundet), BODYGUARD, BOOKET, CO-PRODUKTION, DUMPER (flertydigt, udgår), ESTABLISHMENT, EURO-ATLANTISKE, EXIT, FACTS, FANS, FEEDBACK, FLASHY, GAFFATAPE, HITTER, HITTET, IMAGE, IN-STEDER, JACKPOT-COMEBACK (ikke fundet), KICK, KULTURFACILITETER (ikke fundet), KUNDECENTER, LIGHTUDGAVE, MEGA-BUTIK (ikke fundet), MEGA-INTERESSANT (ikke fundet), METRO-VOGN (ikke fundet), MISSE, MORFE (ikke fundet), NEOPREN, NYLON, POLYMERER, POWER, PREMATUR (ikke fundet), PRODUCERE (flertydigt, udgår), PRODUCERET (flertydigt, udgår), REFORMER (flertydigt, udgår), SIKKERHEDS-TEAM (ikke fundet), SKANNET, SMACK, T-SHIRTS, TACKLE, TACKLER, TAPE, TEENAGER, TEENAGERE, TREND, TWIST, VIDENSCENTER, VIRUS, VIRUS-BOMBE (ikke fundet), WHAT, ÜBER-PERFEKTIONISTISK (ikke fundet), YES, ZOOME

De hyppigste seks ord i det uspecifikke domæne er ganske vist moderne importord, men ingen af dem er ligefrem flunkende nye i dansk længere. *Fan* og *teenager* ses første gang i Retskrivningsordbogen i 1986, og de andre ord er alle belagt i 1970’erne eller tidligere (se Nye ord i dansk – fra 1955 til i dag, dsn.dk/noid): FANS (RO 1986), IMAGE (1963), TEENAGER (RO 1986), POWER (1973), TEENAGERE og TREND (omkring 1965).

Denne gruppe er en noget mere blandet landhandel end ordene i domænet *Nye medier*. Nogle ord er i vækst, andre det modsatte. De fleste ord er hverken signifikant stigende eller faldende.

De ord der er signifikant stigende, er:

TEENAGER (+3,9/mio., 42,2 %, $p = 0,0002$), TREND (+1,8/mio., 23,3 %, $p = 0,0331$), FEEDBACK (+3,1/mio., 121,3 %, $p = 0,0052$), ESTABLISHMENT (+3,0/mio., 304,1 %, $p < 0,0001$), VIDENSCENTER (+2,8/mio., 329,8 %, $p < 0,0001$), ZOOM (+0,6/mio., 94,6 %, $p = 0,0188$), MISSE (+0,5/mio., 101,8 %, $p = 0,0297$).

De ord der er signifikant faldende, er:

FANS (-4,5/mio., -17,2 %, $p = 0,0254$), IMAGE (-6,9/mio., -38,9 %, $p < 0,0001$), ACTION (-3,6/mio., -45,6 %, $p = 0,0008$), TWIST (-1,5/mio., -45,2 %, $p = 0,0014$), YES (-1,8/mio., -59,3 %, $p = 0,0261$), TAPE (-0,9/mio., -39,1 %, $p = 0,0385$), KICK (-1,3/mio., -55,5 %, $p = 0,0013$), NYLON (-0,4/mio., -78,6 %, $p = 0,0342$).

Resten er hverken i signifikant stigning eller fald:

BODYGUARD, BOOKET, CO-PRODUKTION, EURO-ATLANTISKE, EXIT, FACTS, FLASHY, GAFFATAPE, HITTER, HITTET, IN-STEDER, KUNDECENTER, LIGHTUDGAVE, NEOPREN, POLYMERER, POWER, SKANNET, SMACK, SPOTTET, T-SHIRTS, TACKLE, TACKLER, TEENAGERE, VIRUS, WHAT.

Samlet set ender det da også med en ikke-signifikant tendens, som man kan se i den følgende graf.



Figur 10. Samlet tidslinje for det uspecifikke domæne

Der kan være en vis logik i at det er sådan for disse mere almene ord. Almen udbredelse vil ofte hænge sammen med at ordene ikke er helt nye mere, eftersom nye ord formentlig oftest vil dukke op i ret specifikke sammenhænge først og derefter evt. blive udbredt til det mere almene sprog.

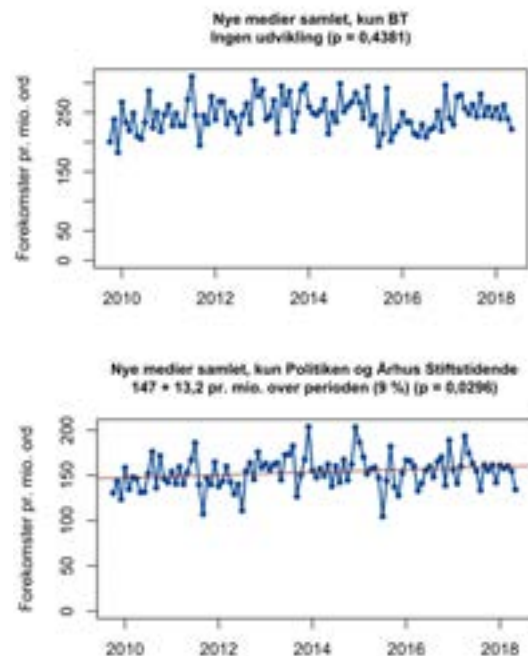
5. Diskussion og konklusioner

I det ovenstående har jeg foretaget tidsserieanalyser af et antal excerpere moderne importord fra tre aviseksemplarer fra 2016 i tre forskellige domæner: to specifikke med ord der har med medier at gøre, og en uspecifik med diverse ord som jeg ikke har opfattet som knyttet til noget specifikt domæne.

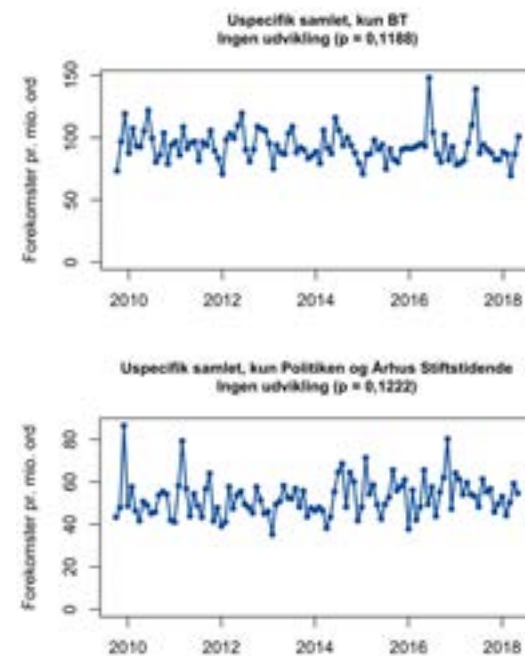
Først og fremmest kan det konkluderes at ikke alle domæner udviser vækst i brugen af moderne importord. Af de tre undersøgte domæner er det kun mediedomænerne der er i samlet stigning. Og i domænet ”nye medier” er det tydeligt betegnelserne for de nyeste ting og fænomener der driver den overordnede vækst der ses. Nuværende hverdagsteknologier som internet, mail, mobiltelefoner og sms viser ikke på samme måde konsekvent stigende tendenser i avisomtalen.

En mulig fejlkilde i korpusundersøgelsen er at der ikke indgår tabloidaviser i korpusset, eftersom sproget i tabloidaviser meget vel kan tænkes at være anderledes end i de store aviser, også med hensyn til brugen af moderne importord. Noget lignende gør sig gældende for de regionale aviser, men antagelig i mindre grad (hvilket dog selvfølgelig er et empirisk spørgsmål).

Skiller man BT-ordene fra, kan man se at domænet ”nye medier” her faktisk ikke er i signifikant vækst, hvilket det til gengæld er i de to andre aviser. (Her skal man være opmærksom på at nogle af ordene vil gå igen i de to optællinger, og at gengangerne i sagens natur vil være nogle af de hyppigste). Til gengæld ser ”nye medier”-ordene fundet i BT ud til generelt at være hyppigere. Samme øvelse udført på domænet ”uspecifik” viser ingen signifikante stigninger eller fald, men ordene fra BT er igen generelt hyppigere.



Figur 11. Samlet tidslinje for ord fra domænet *Nye medier* i BT og samlet tidslinje for ord fra domænet *Nye medier* i Politiken og Århus Stiftstidende



Figur 12. Samlet tidslinje for ord fra det uspecifikke domæne i BT og samlet tidslinje for ord fra det uspecifikke domæne i Politiken og Århus Stiftstidende

BT kunne således se ud til at være en anelse mere konservativ i brugen af moderne importord. Om det gælder tabloidpressen generelt, og om der er forskel på landsdækkende og regionale aviser, er spørgsmål der er værd at undersøge.

En anden konklusion er at mange af de moderne importord der er undersøgt her, er overordentligt sjældne. Adskillige af ordene forekommer knap nok overhovedet i det anvendte korpus, som ellers er af en anseelig størrelse. Kun *teenager* kvalificerer sig lige netop til at blive et såkaldt kerneord, dvs. meget højfrekvent (vel at mærke i ugebladsdata, se Diderichsen & Schack 2019). Vi kigger altså her ind i en overvejende lavfrekvent del af ordforrådet. Selv indenfor de domæner hvor der ses en stigende tendens, ser moderne importord altså generelt ud til at være sjældne.

Det kan desværre ikke konkluderes inden for rammerne af denne undersøgelse om brugen af moderne importord mere generelt er i stigning eller ej. Men flere resultater er på vej fra undersøgelser af de resterende domæner.

Litteratur

- Brunstad, E. (2003) Omfanget av moderne importord i nordiske aviser. *Moderne Importord I Språka I Norden* 1.
- Diderichsen, P. (2019a) Moderne importord over tid i høj opløsning. M. Heidemann Andersen & P. Jarvad (red.) *Yes, det er coolt. moderne importord i dansk*. København: Dansk Sprognævn.
- Diderichsen, P. (2019b) Tidslinjen: Dansk sprognævns n-gram-viewer. M. Heidemann Andersen & P. Jarvad (red.) *Yes, det er coolt. moderne importord i dansk*. København: Dansk Sprognævn.
- Diderichsen, P. & J. Schack (2019) Nye danske kerneord – kerneordskorpus 2.0. Y.U. Goldshtein, I. Schoonderbeek Hansen & T. Thode Hougaard (red.) *17. Møde om Udforskningen af Dansk Sprog*. Århus: Aarhus Universitet:257-275.
- Graedler, A.L. (2017) MIN-projektet jamført med tidligere studier. *Moderne Importord I Språka I Norden* 3:157-173.

- Heidemann Andersen, M. (2019) Yes, det er coolt. Moderne importord i danske aviser. Y.U. Goldshtein, I. Schoonderbeek Hansen & T. Thode Hougaard (red.) *17. Møde om Udforskningen af Dansk Sprog*. Århus: Aarhus Universitet:45-57.
- Jarvad, P. (2007) Afløsningsord i dansk. *Moderne Importord I Språka I Norden* 6.
- Kvaran, G. (2007a) Brug af afløsningsord i de nordiske sprog. sammenligning og konklusioner. G. Kvaran (red.) *Udenlandske eller hjemlige ord? En undersøgelse af sprogene i Norden*. Oslo: Novus Forlag:169-187.
- Kvaran, G. (2007b) Undersøgelse af afløsningsord i de nordiske sprog. indledning. G. Kvaran (red.) *Udenlandske eller hjemlige ord? En undersøgelse af sprogene i Norden*. Oslo: Novus Forlag:9-18.
- Ruus, H. (1995) *Danske kerneord. centrale dele af den danske leksikalske norm*, bd. I. Museum Tusculanum.
- Sandøy, H. (2007) Avissspråket i Norden – ei jamføring. *Moderne Importord I Språka I Norden* 3:127-155.
- Sandøy, H. (under udg.) Moderne importord i danske aviser. M. Heidemann Andersen & P. Jarvad (red.) *Yes, det er coolt. moderne importord i dansk*. København: Dansk Sprognævn.
- Selback, B. (2007). Dansk. *Moderne Importord I Språka I Norden* 3:67-86.