

Hvad kan nonsensord fortælle os om stødets fonologi?

Præsentation af et ph.d.-projekt

Annika Høeg

Stockholms Universitet

1. Indledning

Stødet er et prosodisk fænomen i dansk, som forekommer på visse stavelser i visse ord, fx i ordene *bil* ['bi:ʔ] og *april* [a'pɕi:ʔ]. For at et ord kan have stød, antages det, at den trykstærke stavelse skal have stødbasis, hvilket vil sige, at den enten har en lang vokal eller en kort vokal fulgt af en sonor konsonant (Grønnum 2005:62). At et ord kan have stød, betyder dog ikke nødvendigvis, at det har det. Stødet har en kompleks distribution, men der findes nogle generelle mønstre; ord med tryk på sidste stavelse har ofte stød, som *bil* ['bi:ʔ] og *april* [a'pɕi:ʔ], og ord med tryk på næstsidste stavelse har sjældent stød, som *måned* ['mɔ:nəð] og *vise* ['vi:sə] (vb.). Denne generalisering indgår i Basbølls ikke-stødprincip, som forudsiger forekomsten af stød ud fra trykkets placering og endelsers produktivitet (Basbøll 2005). Alle undtagelser til dette antages at have en leksikalsk markering for +stød, fx i *hummer* ['hɔmʔʌ], –stød, fx i *vært* ['væɾd], eller en ekstraprosodisk sonorant¹, fx i *ven* ['ven]. En anden generalisering er, at enstavede ord som slutter på lang vokal altid har stød, fx *bi* ['bi:ʔ], hvorimod det varierer mere ved enstavede ord med kort vokal fulgt af en sonor konsonant, fx *hal* ['halʔ] og *øl* ['øl] (Grønnum 2005:221-223).

I denne artikel vil jeg præsentere mit igangværende ph.d.-projekt, som handler om generaliseringer af stødets forekomst i dansk og deres produktivitet. I projektet vil jeg se på, hvornår talere udtaler ikkeeksisterende ord med stød, såkaldte nonsensord (eller pseudoord). Dette gøres for at kunne sige noget om, hvilke generaliseringer fra eksisterende danske ord, som talere overfører til nonsensord. Eftersom projektet er i et tidligt stadie, og der endnu ikke findes noget resultat, vil fokus for denne artikel være en overgribende præsentation af projektet samt nonsensordene og fremgangsmåden i det første eksperiment.

¹En ekstraprosodisk sonorant kan, kort forklaret, siges at "svæve frit" i slutningen af ordet og er derfor ikke med til at give ordet stødbasis. Et eksempel på dette er sonoranten /n/ i ordet *ven*. Dette er en leksikalsk undtagelse ligesom –stød, men med en anden forklaring til undtagelsen.

2. Formål

Formålet med projektet er at undersøge, hvilke generaliseringer for stødets forekomst som talere overfører til nonsensord, for at undersøge stødets produktivitet. Dette bliver gjort ved at sammenligne generaliseringer ud fra eksisterende ord med taleres udtale af forskellige nonsensord. Disse nonsensord er mulige danske ord som ikke betyder noget. Det antages, at nonsensord kan sige noget om stødets produktivitet, eftersom de per definition ikke har nogen leksikalsk information. En forventning er, at dette bl.a. vil kunne give empirisk støtte for, hvorvidt det kun er trykkets placering der afgør om et ord har stød, eller om der er andre faktorer som kan påvirke.

Projektet kommer til at bestå af tre eksperimenter, der alle indebærer at informanternes udtale af forskellige nonsensord bliver spillet ind på en lydoptagelse. Nonsensordene i eksperimenterne bliver udformet ud fra hvilken generalisering som skal testes. I eksperiment 1 er det den første del af Basbølls ikke-stødprincip som er i fokus: ”The penultimate syllable of the min-word has non-stød” (Basbøll 2005:380).² Dvs. at trykstærke stavelser (med stødbasis), som er tredjesidst eller sidst i et ord, har stød, og trykstærke stavelser, som er næstsidst i et ord, har ikke stød. Alle undtagelser til dette er leksikalske. Eftersom nonsensord ikke har leksikalsk information, burde stødforekomsten i nonsensord følge denne generalisering uden undtagelser. I eksperiment 1 er det derfor nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse, som vil blive testet. Formålet med eksperiment 2 er at undersøge forekomsten af stød i monosyllabiske ord, ud fra hvilken sonorant der er i codaen. Derfor vil det være monosyllabiske nonsensord med et rim³ bestående af en kort vokal fulgt af en sonor konsonant, som bliver testet. I eksperiment 3 er fokus på forekomsten af stød ud fra morfologiske alternationer. Her vil både bøjnings- og afledningsmorfemer blive kombineret med forskellige nonsensord for at se nærmere på, hvordan det påvirker forekomsten eller fraværet af stød. En del af eksperiment 3 vil også være at teste den anden del af Basbølls ikke-stødprincip: ”A monosyllabic stem has non-stød before a syllable (domain: Basic Word)” (Basbøll 2005:380).⁴ I denne del af ikke-stødprincippet afhænger stødets forekomst af antallet stavelser i stammen, og hvorvidt endelsen inkluderes i et vist domæne i ordstrukturen. Også her er der leksikalske undtagelser til reglen (se Grønnum 2005:228–236 for en gennemgang).

² Lidt forsimplet kan man sige, at ”min-word” er monomorfematiske ord.

³ Rim er stavelers kerne og coda, dvs. vokalen og efterfølgende konsonanter.

⁴ Se Basbøll (2005:377–379) for en gennemgang af domænerne *min-word* og *Basic Word*.

3. Nonsensord som metode

I dette afsnit vil jeg gå nærmere ind på, hvad der menes med nonsensord, og hvordan de kan anvendes til at teste forskellige fonologiske generaliseringer.

Kort sagt kan man sige, at nonsensord er opdigtede ord, som til formen (mere eller mindre) ligner rigtige ord, men ikke har nogen betydning eller funktion i det aktuelle sprog.

Nonsensord kan udformes på forskellige måder og med forskellige grader af nonsens, alt efter hvad det er, man vil teste. Det kan fx være nonsenssammensætninger, hvor eksisterende ord kombineres til en ny sammensætning, som fx *taxigris* i Bruces (1974) undersøgelse af toneaccent i sammensatte ord i sydsvensk. Det kan også være sammensætninger, hvor et eksisterende ord kombineres med et nonsensord, som fx *korstamse* i Grønnum m.fl.s (2020) undersøgelse af tilkomst af stød i sammensætninger. Et andet eksempel er nonsensord, eller nonsensrødder, som testes i en eksisterende morfologisk kontekst, hvilket gøres i Bennett og Bravers (2020) undersøgelse af palatalisering i Xhosa med nonsensrødder som *-fam-* i konteksten *iya-fam-ba* og *iya-fam-wa*. Endvidere kan nonsensord også bestå af fuldstændigt nonsens, forstået som at der ikke indgår nogle eksisterende morfemer i ordet. Et eksempel på fuldstændigt nonsens er *galeaft* fra Turcsan og Herments (2015) undersøgelse af tildeling af tryk i engelsk. I mit projekt består nonsensordene i det første eksperiment, som er i fokus i denne artikel, også af fuldstændigt nonsens.

Selv når nonsensord udelukkende består af nonsenselementer, er de for det meste mulige ord i det aktuelle sprog, hvilket vil sige, at de følger sprogets fonotaks. I det følgende, og resten af artiklen, vil udgangspunktet være danske nonsensord. Selvom nonsensord skal være mulige eller "tilladte" ord i sproget, kan de have en struktur, som påmindrer om fremmedord.⁵ Fx vil polysyllabiske nonsensord med mere end én fuldvokal, som *gnimpov* eller *suvulga*, virke mindre hjemlige, eller måske til og med mindre danske, end monosyllabiske nonsensord og polysyllabiske nonsensord med schwa, som *fril* eller *slærbe*. Det kan dog næppe påstås, at to (eller flere) fuldvokaler kun forekommer i usædvanlige eller eksotiske ord i dansk, jf. fx *tomat*, *ananas*, *sofa* og *cola*. Desuden kan stødet sagtens forekomme i låneord, som ikke ligner danske ord, fx *gymnasium* [gym'næ:ʔɔm]. At nonsensordene har en struktur, som ikke forekommer i danske arveord, er ikke nødvendigvis et problem, så længe informanterne udtaler ordene som om det er dansk. En anden grund til at have "mindre danske" nonsensord med i undersøgelsen er, at ord med tryk på før-næstsidste stavelse, som *kamera*, er mindre sædvanlige i dansk, og denne gruppe er ganske lille blandt eksisterende ord. Derfor ser ingen af nonsensordene med

⁵ Alt efter hvad det er, man vil undersøge. Nonsensord kan fx også bruges til at undersøge, hvor grænsen går for, hvad talere af et vist sprog bedømmer som acceptabelt.

tryk på den før-næstsidste stavelse særligt danske ud, men de er nødvendige for at kunne sige noget om, hvorvidt nonsensord med tryk på før-næstsidste stavelse altid udtales med stød, som det antages i Basbølls ikke-stødprincip.

Pointen med nonsensord er, at de er nye for taleren, og dermed ikke er memorerede med nogen leksikalsk information. De kan derfor bruges til at sige noget om sprogsystemet, i det her tilfælde fonologien, fx gennem at undersøge, hvordan informanter reagerer på, opfatter eller vurderer forskellige nonsensord, som de præsenteres for auditivt eller visuelt. I mit projekt anvendes nonsensordene til at undersøge produktion (udtale), hvor informanter bliver præsenteret for nonsensord i skrift og læser dem højt.

I det informanterne bliver præsenteret for nonsensordene i skrift, giver ortografien visse begrænsninger. Eksempelvis markeres vokallængde eller trykstærke stavelser generelt ikke i dansk ortografi. Hvis formålet fx er at teste nonsensord, hvor rimet består af en kort vokal fulgt af en sonor konsonant, bliver det et problem, hvis informanterne i stedet udtaler nonsensordene med en lang vokal. En anden udfordring er, hvordan man får informanterne til at lægge trykket på en specifik stavelse i polysyllabiske ord, men samtidigt undgår at de udtaler stavelsen med emfase. En mulighed er at skrive den trykstærke stavelse med versaler, som det gøres i en studie med nonsensord af Braver (2014). Jeg vil komme nærmere ind på hvordan det bliver gjort i det aktuelle projekt i næste afsnit.

Nonsensord er altså potentielle ord, som ikke har nogen betydning, og kan bruges til at teste forskellige generaliseringer, når der ikke er nogen leksikalsk information. I de følgende afsnit vil jeg præsentere projektets første eksperiment, med fokus på forsøgsdesignet.

4. Eksperiment 1 – trykkets placering

I dette eksperiment er formålet at undersøge, hvilken udstrækning taleres udtale af forskellige nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse stemmer overens med Basbølls ikke-stødprincip. For at teste dette har jeg genereret tre kategorier af nonsensord, ud fra hvilken stavelse der er trykstærk. I tabel 1 kan der ses eksempler på forskellige nonsensord, som testes i eksperimentet. Fremgangsmåden for genereringen og udvalget af nonsensordene vil blive uddybet nærmere i det næste afsnit.

Nonsensord med tryk på sidste stavelse	Nonsensord med tryk på næstsidste stavelse	Nonsensord med tryk på før-næstsidste stavelse
SPLU	PRULbe	NILdefe
PRILP	BÅpa	VÅNlupe
fluNIL	saKRÆLje	HÆJgopi
stjeKRÆL	fyKVILru	FImeFa

Tabel 1. Eksempler på nonsensord som testes i eksperiment 1.

I eksperimentet læser informanterne i alt 336 nonsensord højt. Nonsensordene er fordelt på fire blokke med 84 ord i hver. Rækkefølgen på blokkene varierer fra gang til gang, og rækkefølgen på nonsensordene i hver blok er randomiseret. Inden oplæsningen får informanterne bl.a. følgende instruktion: ”Det her går som sagt ud på, at du skal læse forskellige ord højt på dansk. Det er ord som du ikke har set før. Du kan forestille dig, at du spiller et nyt spil, og at det her er ting som du kan købe i den verden.” Dette gøres for at få informanterne til at tænke på ordene som substantiver, men i øvrigt er der ikke nogen information, som peger mod at ordene har en bestemt ordklasse. Informanterne ser ordene et efter et i midten af en grå skærm og trykker videre, når de har læst ordet højt.



Figur 1. Eksempel fra dataindsamlingen til eksperiment 1.

I Figur 1 kan der ses et eksempel på, hvordan nonsensordene vises for informanterne.⁶ Som det kan ses i figuren, er den næstsidste stavelse skrevet med versaler. Dette anvendes for at angive, hvilken stavelse som trykket skal lægges på (jf. Braver 2014). Inden oplæsningen instrueres informanterne i at lægge trykket på den stavelse, som har store bogstaver. De ser forskellige minimale par på skærmen, som fx skrives FORMel og forMEL, og derefter ord hvor trykket flyttes, som CARdigan og carDIgan, for at sikre at de har forstået, hvordan de skal gøre, og at de kan flytte trykket. Denne markering af den trykstærke stavelse behøves for at få alle informanterne til at lægge trykket på samme stavelse, og fordi det ellers er svært at få dem til

⁶ Eksperimentet er bygget i programmet PsychoPy (se Peirce m.fl. 2019).

at lægge trykket på den før-næstsidste stavelse. Jeg har udført pilotundersøgelser, hvor nonsensordene enten blev vist med versaler i den trykstærke stavelse eller uden nogen markering af tryk, for at sikre at markeringen ikke medførte, at informanterne udtalte flere ord med stød. Versalerne virkede ikke til at få informanterne i pilotundersøgelserne til at udtale flere af nonsensordene med stød.

5. Generering af nonsensord til eksperiment 1

Under arbejdet med udformningen af nonsensordene er der meget at tage hensyn til. I det følgende vil jeg redegøre for, hvilke overvejelser jeg har gjort om udformningen af nonsensordene, deres struktur, hvilke specifikke kombinationsmuligheder der kan indgå, hvordan de er blevet genereret samt hvordan udvælgelsen blandt de genererede nonsensord er gået til.

I eksperimentet er trykkets placering den uafhængige variabel, som bliver manipuleret for at undersøge den afhængige variabel, som er forekomst eller fravær af stød. Nonsensordene udformes dermed så trykkets placering varierer. Dog er trykkets placering ikke det eneste, som varierer; bl.a. vokalkvalitet og antal stavelser i nonsensordene varierer også. Dette er variabler, som ikke skal testes, men som eventuelt kan påvirke resultatet. Denne type af uvedkommende variabler (eng. *extraneous variables*) kan bl.a. håndteres med hjælp af randomisering eller afbalancering (Barr m.fl. 2013:257). De uvedkommende variabler kan balanceres på tværs af den uafhængige variabel, så at der fx er lige mange af en bestemt vokal eller sonorant i nonsensordene.

De uvedkommende variabler, som jeg stræber efter at kontrollere i udformningen af nonsensordene, er stavelsesantal, vokalkvalitet i den toniske (trykstærke) stavelse, codaens struktur i den toniske stavelse samt rimets struktur i den posttoniske stavelse for nonsensordene med tryk på næstsidste stavelse. Disse variabler kontrolleres gennem en balancering af dem i de tre kategorier af nonsensord, således at der ikke er en specifik vokal eller sonorant, som dominerer i en af kategorierne, fx at en stor andel af nonsensordene med tryk på næstsidste stavelse har en åben trykstærk stavelse (dvs. ingen coda), hvorimod en stor andel af nonsensordene med tryk på sidste stavelse har sonoranten /n/ i codaen i den trykstærke stavelse. For at mindske sandsynligheden for at en sådan forskel kan påvirke, hvor ofte nonsensordene i kategorien får eller ikke får stød, har jeg forsøgt at balancere de ovennævnte uvedkommende variabler i hver kategori af nonsensord.

For den uvedkommende variabel, *stavelsesantal*, er det hvorvidt nonsensordet har en prætonisk stavelse eller ej, som jeg vil kontrollere, snarere end selve stavelsesantallet. Dette er kun relevant for nonsensordene med tryk på sidste eller næstsidste stavelse, eftersom mono-

morfematiske nonsensord med fire stavelser bliver for unaturlige i dansk tale.⁷ I kategorien med tryk på sidste stavelse er der altså en underkategori med monosyllabiske nonsensord og en med bisyllabiske nonsensord, ligesom der i kategorien med tryk på næstsidste stavelse er en underkategori med bisyllabiske nonsensord og en med trisyllabiske nonsensord.

Den uvedkommende variabel *vokalkvalitet* kontrolleres kun i den trykstærke stavelse. For vokalkvalitet har jeg forsøgt at få så stor spredning i vokaldiagrammet som muligt, men samtidigt begrænset det til fire vokaler for at kunne balancere dem på tværs af kategorierne uden at behøve for mange nonsensord. Disse fire vokaler er <i>, <æ>, <u> og <å>. Eftersom nonsensordene præsenteres i normal skrift, begrænses kontrollen over vokalkvaliteten, og dermed balanceringen, af ortografien. Jeg er vidende om, at den ortografiske repræsentation ikke nødvendigvis vil blive udtalt med samme fonem i alle kontekster, eller endda kan variere fra gang til gang, men jeg håber, at de fire grafemer kan give en nogenlunde spredning og afbalancering af vokalkvaliteten.

Den næste uvedkommende variabel er *codaens struktur i den trykstærke stavelse*. Eftersom det er forekomsten af stød, som undersøges, skal nonsensordene kunne have stød, dvs. udtales med stødbasis. I alle tre kategorier af nonsensord er konsonanten direkte efter vokalen derfor en af sonoranterne /m/, /n/, /l/, /ð/, /ɹ/, /ʊ/ eller /ɸ/.⁸ Codaen kan også være tom, hvilket angives med Ø i genereringen af nonsensordene. Ved en tom coda, altså åbne stavelser, er forventningen, at informanterne udtaler vokalen som en lang vokal. Nonsensordene ville også have stødbasis, hvis det var en lang vokal fulgt af en obstruent, men da der er risiko for, at det bliver en kort vokal, og dermed ingen stødbasis, og for at begrænse mængden kombinationer, som skal testes, indgår obstruenterne ikke i codaen i nonsensordenes trykstærke stavelser (med mindre de foregår af en sonorant). I nonsensordene med tryk på sidste stavelse kan sonoranten enten være final eller fulgt af endnu en konsonant.⁹ I nonsensordene med tryk på næstsidste eller før-næstsidste stavelse består codaen af maksimalt én sonorant, eftersom konsonanten efter bliver ansat i den efterfølgende stavelse. Desuden vil for lange interludes¹⁰ medføre større risiko for, at nonsensordene opfattes som sammensætninger.

Den sidste uvedkommende variabel er *rimets struktur i den posttoniske stavelse* og er kun relevant for nonsensordene med tryk på næstsidste stavelse. Anledningen til dette er, at der i Grønnums inddeling (2005:222) er forskel på, hvorvidt forekomsten af stød er forudsigeligt i

⁷ Nonsensordene med tryk på før-næstsidste stavelse er allerede på (eller måske over) grænsen. Se foregående afsnit for en diskussion om dette.

⁸ Ortografisk repræsenteret som <m>, <n>, <l>, <d>, <j>, <v> og <r>.

⁹ Hvis vokalen er kort, falder disse under det Grønnum (2005:222) kalder kategori III respektive kategori IV.

¹⁰ Konsonanterne mellem to stavelser, dvs. både coda og ansat. Termen er hentet fra Basbøll (2005).

ord med tryk på næstsidste stavelse, ud fra hvorvidt den sidste stavelse indeholder en fuldvokal eller et schwa, og hvorvidt stavelsen med schwa er åben eller lukket. Derfor deles de bisyllabiske nonsensord med tryk på næstsidste stavelse ind i undergrupperne *åben eller lukket posttonisk stavelse med fuldvokal*, *åben posttonisk stavelse med schwa* og *lukket posttonisk stavelse med schwa*. De trisyllabiske nonsensord med tryk på næstsidste stavelse deles kun ind i undergrupperne *åben posttonisk stavelse med schwa* og *åben posttonisk stavelse med fuldvokal*.¹¹ At de to underkategorier ikke deles ind i lige mange undergrupper, resulterer i, at der indgår flere bisyllabiske nonsensord med tryk på næstsidste stavelse i eksperimentet end trisyllabiske.

Vokallængde er egentlig også en uvedkommende variabel, som burde balanceres, men da det er svært at styre, hvilken vokallængde informanterne udtaler vokaler med i stavelser, som har en enkelt sonorant i codaen, valgte jeg ikke at lave en opdeling baseret på dette. I stedet vil antallet lange og korte vokaler blive opregnet ved databehandlingen, for at eventuelt kunne tage hensyn til det i analysen.

Antallet nonsensord i hver af de tre kategorier, *tryk på sidste stavelse*, *tryk på næstsidste stavelse* og *tryk på før-næstsidste stavelse*, bestemmes af antallet kombinationer af de ovennævnte uvedkommende variabler. Nonsensord med tryk på sidste stavelse har to muligheder for stavelsesantal, fire muligheder for vokal og 14 muligheder for coda¹². Dette giver 112 kombinationsmuligheder. I eksperimentet testes ét nonsensord for hver kombinationsmulighed og dermed 112 nonsensord med tryk på sidste stavelse (56 monosyllabiske og 56 bisyllabiske). For de bisyllabiske nonsensord med tryk på næstsidste stavelse er der fire muligheder for vokal, otte muligheder for coda og tre muligheder for den posttoniske stavelse, hvilket giver 96 kombinationsmuligheder. For trisyllabiske nonsensord med tryk på næstsidste stavelse er der fire muligheder for vokal, otte muligheder for coda og to muligheder for den posttoniske stavelse, hvilket giver 64 kombinationsmuligheder. I eksperimentet bliver der altså testet 160 nonsensord med tryk på næstsidste stavelse. For nonsensordene med tryk på før-næstsidste stavelse er der fire muligheder for vokal og otte muligheder for coda, hvilket giver 32 kombinationsmuligheder. For at der ikke skal blive for få nonsensord i denne kategori, indgår der to nonsensord for hver kombinationsmulighed, hvilket giver 64 nonsensord.

Øvrige dele af nonsensordene, dvs. ansatsen, vokalen i den prætoniske stavelse og inter-

¹¹ Mit indtryk er, at den posttoniske stavelse i trisyllabiske ord med tryk på næstsidste stavelse altid er åben. Dette er baseret på en gennemgang af en ordliste bestående af Retskrivningsordbogen (Dansk Sprognævn 2012) kombineret med transskriptioner fra NST-leksikonet (Andersen 2011).

¹² /ð/ forekommer kun finalt i disse nonsensord, eftersom dette segment ikke virker til at forekomme ordfinalt med efterfølgende konsonanter. Dette er baseret på en gennemgang af en ordliste bestående af Retskrivningsordbogen (Dansk Sprognævn 2012) kombineret med transskriptioner fra NST-leksikonet (Andersen 2011).

ludes som ikke er efter den toniske stavelse, antages ikke at påvirke forekomsten af stød. Disse vælges derfor tilfældigt ud fra mulige kombinationer. For dette har jeg gået ud fra beskrivelser af dansk fonotaks af Schachtenhaufen (2022) og Grønnum (2007), samt Retskrivningsordbogen som elektronisk ordliste (Dansk Sprognævn 2012) kombineret med fonetiske transskriptioner fra NST-leksikonet (Andersen 2011). I tabel 2 angives grafemerne, som kan forekomme i hver del af stavelsen i de bisyllabiske nonsensord med tryk på sidste stavelse. I den første kolonne angives hvilken del af stavelsen grafemerne kan udgøre, og i den anden kolonne angives grafemerne. Nogle af grafemerne angives i parentes, hvilket betyder at disse udgør en af de ovennævnte uvedkommende variabler. Her bliver der altså valgt tilfældigt inden for hver parentes.

At det er grafemer, og ikke fonemer, som angives i tabellen skyldes at nonsensordene er genereret ud fra ortografien. Jeg har udgået fra fonemerne i begrænsningerne for fonotaksen, og valgte derefter et grafem, som kan modsvare hvert fonem. Som det blev nævnt i forbindelse med vokalerne, kan dette resultere i, at ordene udtales anderledes end jeg havde tænkt. Dog er den eksakte kvalitet på fonemerne ikke afgørende, eftersom det er trykkets placering, som er i fokus, men det kan have betydning for balanceringen og mindsker kontrollen (Jf. Schmitz m.fl. 2021:582–583).

Del af stavelsen	Grafemer som kan indgå
Ansats, prætonisk	b t d k g f s sj h m n l v j Ø st sp sk tj spl spr spj str stj skr skj sv sm sn sl pl pr pj tr kn kl kr kj bl br bj dr dj gn gl gr gj fn fl fr fj vr
Kerne, prætonisk	a e i o u y
Interlude, prætonisk– tonisk	r b j k l m n t v s f br bl d fr fl kr kn kv p pl sj sp st sk sl sv tr bs bt pt ft gs gt sm sf
Kerne, tonisk	(i) (u) (æ) (å)
Coda, tonisk	(l) (lp lt lk lf ls) (m) (mp mt ms) (n) (nt nk ns) (j) (jn jl jt js) (v) (vn vl vs) (r) (rm rn rp rt rk rs) (d) (Ø)

Tabel 2. Mulige grafemer i de forskellige dele af stavelserne i bisyllabiske nonsensord med tryk på sidste stavelse. Grafemer, som angives i parentes, udgør en gruppe. Ø betyder, at pladsen kan være tom.

I tabellen angives altså de grafemer, som bliver anvendt til at generere nonsensordene, hvilket bliver gjort med hjælp af programmeringssproget Python, version 3.10.11 (van Rossum &

Drake 2009).^{13,14} Jeg har anvendt en kode for hver kategori af nonsensord og underkategorierne ud fra antal stavelser.¹⁵ I koden defineres de grafemer, som kan stå på de forskellige plader i stavelserne, og i hvilken rækkefølge de skal kombineres. Koden for de bisyllabiske nonsensord med tryk på sidste stavelse giver 56 grupper, hvor hver gruppe er en liste med alle kombinationsmuligheder.¹⁶ Længden på hver liste varierer alt efter, hvor mange forskellige grafemer der kan være i codaen (se tabel 2). De korteste lister består af 11448 nonsensord, og de længste af 68688 nonsensord. I alt genereres 1.465344 bisyllabiske nonsensord med tryk på sidste stavelse.¹⁷ Af disse er det dog kun ét ord fra hver liste, som skal testes i eksperimentet. For at gøre udvælgelsen af disse 56 ord mere overskuelig, og for at minimere bias i udvælgelsen, anvendes *random.seed*, som er en funktion i modulet *random*, til at vælge et tilfældigt nonsensord fra hver af de 56 lister. Med *random.seed* bliver det randomiserede udvalg desuden det samme, hver gang man kører koden.

Koden for de bisyllabiske nonsensord med tryk på sidste stavelse giver altså 56 tilfældigt udvalgte nonsensord, med ét ord fra hver gruppe. En manuel gennemgang er dog nødvendig, inden de kan testes i eksperimentet. Nogle af kombinationerne er allerede eksisterende ord eller homonymer, eller stavelser, som er identiske med eksisterende affikser. Nonsensordene kan også indeholde uønskede kombinationer, som fx <vr> + <u> i den trykstærke stavelse. For at forsøge at minimere bias i valget af nonsensord har jeg opsat kriterier for, hvilke ord der ikke vælges. Fx fravælges nonsensord med samme konsonant inden og efter vokalen, som *krirk*, og nonsensord med samme ansats to gange i træk, som *flefliv*. Det er altså det første tilfældige nonsensord, som ikke fravælges på grund af kriterier som de ovenstående, som testes, også selvom et senere nonsensord måske lyder bedre. De nonsensord som denne proces resulterer i, testes derefter i en pilotundersøgelse. I tilfælde hvor informanterne havde svært ved at læse ordene højt, hvor flere informanter tøvede eller hvor de ofte læste noget andet, valgtes nye nonsensord ud fra samme proces.

¹³ Se Hammond (2020:50–52), for et eksempel på hvordan monosyllabiske nonsensord kan genereres. Fremgangsmåden for polysyllabiske nonsensord er nogenlunde den samme, bortset fra at konsonanter og vokaler for hver del af stavelsen skal defineres.

¹⁴ For generering og randomiseret udvalg i Python anvendes pakken *numpy* (Harris m.fl. 2020), og det indbyggede modul *random* (van Rossum & Drake 2009).

¹⁵ Tak til Carl Hedman for hjælp med at skrive koden. Koden for generering, sortering og tilfældigt udvalg af nonsensordene vil blive gjort tilgængelig på GitHub. Kontakt forfatteren for at få koden, inden den bliver tilgængelig.

¹⁶ Hver gruppe udgøres af en kombination af de uvedkommende variabler *vokalkvalitet* (fire muligheder) og *codaens struktur i den trykstærke stavelse* (18 muligheder). Dvs. for de bisyllabiske nonsensord med tryk på sidste stavelse består en gruppe af alle kombinationer, som slutter på <il>, en anden gruppe består af alle kombinationer, som slutter på <ilp>, <ilt>, <ilk>, <ilf> eller <ils>, en tredje gruppe består af alle kombinationer, som slutter på osv.

¹⁷ Listerne indeholder også eksisterende ord, som indgår i det som her kaldes ”genererede nonsensord”.

At vælge ord fra som er 'for mærkelige' uden at vælge med bias, er en svær balancegang. Her har jeg vægtet at undgå bias tungere. Man kan diskutere, hvorvidt det er den bedste beslutning. Et udvalg hvor nonsensordene i stedet vælges ud fra intuition, ville formodentlig give mere 'naturligt danske' nonsensord, men det ville også øge risikoen for, at der ubevidst vælges nonsensord, som udtales med eller uden stød.

6. Afslutning

I denne artikel har jeg givet en præsentation af mit ph.d.-projekt, som har til formål at undersøge stødets produktivitet med hjælp af nonsensord. Artiklen har fokuseret på projektets første eksperiment, hvor 336 nonsensord med tryk på sidste, næstsidste eller før-næstsidste stavelse skal læses højt af forskellige informanter. Ved eksperimentdesign er der mange hensyn at tage, og som Arunachalam (2013:221) pointerer; "designing a carefully controlled study is not trivial." I artiklen har jeg præsenteret nogle af de beslutninger, jeg har taget under udformningen af eksperimentet, fx i forhold til hvordan nonsensordene præsenteres for informanterne, og i forhold til udformningen af nonsensordene. Under arbejdet med nonsensordene har en svær balancegang været at få nonsensord, som virker "naturligt danske", men samtidigt holde kontrollen og validiteten så høj som muligt, og desuden undgå den bias det ville give at håndplukke de bedste (mest plausibelt danske) nonsensord.

Litteratur

- Andersen, Gisle (2011) *Leksikalsk database for dansk*. Technical report, Nasjonalbiblioteket.
<https://www.nb.no/sprakbanken/en/resource-catalogue/oai-nb-no-sbr-26/>.
- Arunachalam, Sudha (2013) Experimental Methods for Linguists. *Language and Linguistics Compass*, 7(4), 221–232. <https://doi.org/10.1111/lnc3.12021>
- Barr, Dale J., Levy, Roger, Scheepers, Christoph & Tily, Harry J. (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
<https://doi.org/10.1016/j.jml.2012.11.001>
- Basbøll, Hans (2005) *The phonology of Danish*. Oxford: Oxford University Press.
- Bennett, William & Braver, Aaron (2020) Different speakers, different grammars: Productivity and representation of Xhosa labial palatalization. *Phonological Data & Analysis*, 2(6), 1–29.
<https://doi.org/10.3765/pda.v2art6.9>
- Braver, Aaron (2014) Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua* (152), 24–44.
<https://doi.org/10.1016/j.lingua.2014.09.004>
- Bruce, Gösta (1974) Tonaccentregler för sammansatta ord i några sydsvenska stadsmål. I C. Platzack (red.). *Svenskans beskrivning* 8, 62–75.

- Dansk Sprognævn (2012) *Retskrivningsordbogen* (4. udg.) <https://dsn.dk/ordboeger/retskrivningsordbogen/ro-elektronisk-og-som-bog/>
- Grønnum, Nina (2005) *Fonetik og fonologi: almen og dansk*. 3 udg. København: Akademisk forlag.
- Grønnum, Nina (2007) *Rødgrød med fløde, en lille bog om dansk fonetik*. København: Akademisk Forlag.
- Grønnum, Nina, Pharaoh, Nicolai & Basbøll, Hans (2020) Stød in unexpected morphological contexts in Standard Danish: An experimental approach to sound change in progress. *Nordic Journal of Linguistics*, 43(2), 147–180. <https://doi.org/10.1017/S0332586520000074>
- Hammond, Michael (2020) *Python for Linguists*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108642408>
- Harris, Charles R., Millman, K. Jarrod, van der Walt, Stéfan J., Gommers, Ralf, Virtanen, Pauli, Cournapeau, David, Wieser, Eric, Taylor, Julian, Berg, Sebastian, Smith, Nathaniel J., Kern, Robert, Picus, Matti, Hoyer, Stephan, van Kerkwijk, Marten H., Brett, Matthew, Haldane, Allan, Fernández del Río, Jaime, Wiebe, Mark, Peterson, Pearu, Gérard-Marchant, Pierre, Sheppard, Kevin, Reddy, Tyler, Weckesser, Warren, Abbasi, Hameer, Gohlke, Christoph, & Oliphant, Travis. E., (2020) Array programming with NumPy. *Nature*, 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Peirce, Jonathan, Gray, Jeremy R., Simpson, Sol, MacAskill, Michael, Höchenberger, Richard, Sogo, Hiroyuki, Kastman, Erik & Lindeløv, Jonas (2019) PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Schachtenhaufen, Ruben (2022) *Ny dansk fonetik*. Hellerup: BoD.
- Schmitz, Dominic, Baer-Henney, Dinah & Plag, Ingo (2021) The duration of word-final /s/ differs across morphological categories in English: evidence from pseudowords. *Phonetica*, 78(5-6), 571-616. <https://doi.org/10.1515/phon-2021-2013>
- Turcsan, Gabor & Herment, Sophie (2015) Making Sense of Nonce Word Stress in English. I Mompean, J. A., & Fouz-González, J. (red.). *Investigating english pronunciation: Trends and directions*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137509437_2
- van Rossum, Guido, & Drake, Fred L., (2009) *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.