

Ordvarighed og grammatisk og pragmatisk funktion. Tilfældet *egentlig*.

JAN HEEGÅRD

Artiklen afrapporterer en undersøgelse af sammenhængen mellem topologi, pragmatisk betydning og variation i varigheden af udtalen af adverbialt *egentlig*, og artiklen lader variation i varighed være en afspejling af graden af fonetisk reduktion henholdsvis ekspansion i udtalen af *egentlig*. Undersøgelsen viser at der er en stor variation i udtalen af *egentlig* i henhold til adverbialets placering i sætningsskemaet, der traditionelt anvendes for dansk syntaks. På alle andre pladser end 'lille a' har *egentlig* som oftest lang varighed, dvs. udtales ekspanderet, mens ordet på lille a udviser stor variation, fra optimalt ekspanderet til kraftigt reduceret. Undersøgelsen viser også at der er stor forskel på de enkelte talere mht. varighed, og at der er en vis men ikke afgørende sammenhæng mellem *egentlig*s varighed og tale tempo. I en multifaktoriel *mixed-effect*-regressionsanalyse sammenholdes disse faktorer med faktorer der i andre studier har vist sig at spille ind på fonetisk reduktion. Analysen viser at topologisk placering, talehastighed og final eller ikke final placering i ytringen påvirker *egentlig*s varighed i signifikant grad, men at også andre faktorer spiller ind, fx talers køn og alder og om *egentlig* står foran et andet adverbial eller ej. Resultaterne bliver i diskussionen sat i perspektiv til andre, danske og internationale undersøgelser. Artiklen afrundes med bud på hvilke konsekvenser resultaterne kan have for fremtidig udforskning af årsager til fonetisk reduktion og af sammenhængen mellem sprogets udtryksside og dets pragmatiske funktion.

Emneord: fonetisk reduktion, topologi, individuel variation, multifaktoriel regressionsanalyse, subjektivitet

1 INTRODUKTION

Det er en uomgængelig del af talesproget at man kan bruge forskellige grader af fonetisk distinkthed til at fremhæve eller nedtone vigtigheden af det man vil sige. Vil man betone et eller flere ord i en ytring, er en

måde at gøre det på at artikulere det eller dem tydeligere. Omvendt kan man have tendens til netop ikke at artikulere tydeligt i de passager i en ytring der ikke har så stor betydning. De mekanismer og faktorer der spiller ind på denne måde at udnytte ords lydlig elasticitet på, og de udslag dette har i form af faktisk fonetisk reduceret eller ekspanderet ('distinkt') udtale, har i de senere år fået mere opmærksomhed i forskningsverdenen, både herhjemme og internationalt, jf. fx oversigter i Ernestus & Warner 2011 og Schachtenhaufen 2013 samt temasektionen i *Danske Talesprog* 12 (2012). Dette har uden tvivl blandt andet at gøre med dels at vi nu har store databaser for naturlig tale til rådighed der gør det muligt at lave generaliseringer på et solidt kvantitativt grundlag, dels at sprogteknologien er ovre sin første fase og nu stræber efter at efterligne og frembringe naturligt lignende tale frem for kunstigt lydende tale. Forskningen i fonetiske reduktionsprocesser har fokuseret både på nogle af de mekanismer der spiller ind på hvornår henholdsvis fonetisk reduktion og fonetisk ekspansion finder sted, og på de faktiske fonetiske realisationer af reducerede ord og segmenter. Denne artikel er et bidrag til førstnævnte fokus.

1.1 Årsager til fonetisk reduktion

Dokumentation for at nogle faktorer frem for andre bidrager til den fonetiske reduktion henholdsvis ekspansion, finder man bl.a. i en lang række undersøgelser af især engelsk og hollandsk (for hvilke sprog der er større talesprogsdatabaser til rådighed). Flere studier viser fx at højfrekvente ord kan have mere eller hyppigere reduktion end lavfrekvente ord (fx Bush 2001, Pluymaekers m.fl. 2005, Bell m.fl. 2009; jf. også omfattende litteraturliste i sidstnævnte). Gregory m.fl. (1999), Jurafsky m.fl. (2001), Schuppler m.fl. (2009), Hanique m.fl. (2010), Hanique & Ernestus (2011) og Schuppler m.fl. (2012) viser at et syntaktisk forudsigeligt ord, fx et hovedverbum i en participiumform forudgået af et finit hjælpeverbum, har større tendens til reduktion end et ikkeforudsigeligt ord. Andre forskere har peget på en 'iboende' forskel med hensyn til reduktionstendens for indholdsord (fx substantiver, adjektiver og hovedverber) over for funktionsord (fx præpositioner, konjunktioner og hjælpeverber) (fx Jurafsky m.fl. 2002, Adda-Decker m.fl. 2005 og Bell m.fl. 2009).

På hjemlig grund har Pharao (2009, 2010) vist at bl.a. ordfrekvens har indflydelse på i hvor høj grad intervokaliske konsonanter undergår reduktion (fx af typen *liget*, *livet* og *lejet*). Heegård (2012, 2013a) har vist at verber i datidsformer på *-te* og *-ede* oftere mister den finale schwavokal end adjektiver med homofone *-te* og *-ede*. Heegård (2013b) og Heegård & Mortensen (2014) har vist at varierende udtaler af adverbialet *faktisk*, fx fra [faʊs]¹ til [ˈfagdisg] kan relateres til ordets syntaktiske og pragmatiske funktion. Heegård & Thøgersen (2012, 2014) har ligeledes påpeget at sproglige såvel som ikkesproglige faktorer spiller ind på reduktionen af endelserne *-te* og *-ede*.

Ruben Schachtenhaufen har i en række arbejder, men først og fremmest i sin ph.d.-afhandling (Schachtenhaufen 2013), dokumenteret at visse sproglige og ikkesproglige faktorer indvirker på reduktion henholdsvis ekspansion, fx funktionsord vs. indholdsord, fonologiske og syntaktiske omgivelser, enkeltsegmentets placering i stavelsen, ordklasse, (tendens til integration med) naboord, ny vs. kendt information, gentagelse af ord, samtalepartner (akkommodation), talesituation (monolog vs. dialog), tryk, tøven, pauser samt individuelle præferencer vedrørende artikulation og taleproduktion (se også Schachtenhaufen 2010 og 2012). Det er en pointe hos Schachtenhaufen, og også pointeret af Heegård & Thøgersen (2012, 2014), at fonetisk reduktion ofte for ikke at sige altid er et resultat af et samspil af sådanne forskelligartede faktorer, sproglige såvel som kontekstuelle (fx Schachtenhaufen 2013: 52).²

1.2 Artikelens fokus

Denne artikel er et bidrag til denne forskning, men et bidrag der specifikt undersøger sammenhængen mellem udtalevariation og visse grammatiske, pragmatiske, produktionsmæssige og sociolingvistiske faktorer. Artiklen fokuserer på adverbialet *egentlig* og dets varierende

1 Danske udtaler er gengivet med IPA som defineret for dansk af Grønnum (2005).

2 Enkeltordsudtaler i distinkte henholdsvis reducerede udtaler i det danske ordforråd er solidt belagt i *Den Store Danske Udtaleordbog* (Brink m.fl. 1991), som angiver udtaleformer med bemærkninger som ”distinkt”, ”mere distinkt”, ”mindre distinkt” osv. Lars Brink har selv andetsteds, fx Brink (2013), illustreret hvordan et ord kan have forskellige, ’hierarkisk’ ordnede udtaler.

varighed i et stort talesprogskorpus. Med denne indsnævring af undersøgelsesgenstanden kan faktorer som forskel i frekvens, indholdsord vs. funktionsord og grammatisk forudsigelighed (som sætningsadverbial er *egentlig* jo ikke syntaktisk obligatorisk) udelukkes som determinerende for variationen i varighed. Risikoen ved afgrænsningen af undersøgelsesgenstanden er så at de faktorer der kan påvises at have en indflydelse på variationen af udtalen af *egentlig*, kun gælder for dette ord. Dette lader jeg være op til fremtidige studier at undersøge.

Ved at fokusere på et epistemisk adverbial som *egentlig* vil undersøgelsen også kunne relatere sig til det forskningsfelt der undersøger subjektivitet og sprogforandring (Jensen 2001; Traugott 1989, 1995). Som det vil fremgå nedenfor, har *egentlig* et betydningsdomæne der omfatter både en mere konkret betydning og en mere generaliseret, subjektiv eller epistemisk betydning. Sådanne betydningsovergange, eller polyseme forhold, er velkendte i verdens sprog (Traugott 1980, 1995; Traugott & Dasher 2002) og har ofte været diskuteret i forbindelse med begrebet grammatikalisering (Traugott 1988; Hopper & Traugott 2003; Boye & Harder 2012), men disse betydningsskifts relation til fonetikken og korrelation med fonetisk reduktion henholdsvis ekspansion er endnu underbelyst. I diskussionsafsnittet vil undersøgelsens resultater blive forholdt til dele af dette forskningsfelt.

Artiklen viser nemlig at en række sproglige såvel som ikkesproglige faktorer korrelerer med den variation i varighed som udtalen af *egentlig* er præget af, og at disse derfor bør inddrages eller i det mindste bør tages i betragtning fremover. En meget vigtig faktor er *egentlig*s syntaktiske, eller topologiske, placering, som her påvises at være medbestemmende for ordets varighed. Med påvisningen af topologi som relevant parameter for udtalevariation er resultatet samstemmende med Heegårds (2013a) resultater for adverbialet *faktisk*, om end ikke direkte med Schachtenhaufens (2013) undersøgelser som ikke inddrager dette parameter. Således er artiklen et argument for at topologi er et vigtigt parameter for udtalevariation, og at dette parameter bør indgå i fremtidige undersøgelser af sammenhængen mellem ords udtalevariation og grammatiske og (skiftende) pragmatiske funktion.

Som nævnt tager undersøgelsen, måske noget uventet, ikke udgangspunkt i *egentlig*s udtale i form af segmentrealisering, altså om ordet fx

udtales ekspanderet [^he:^henli], hvad det sjældent gør, eller, som oftest, reduceret [^he:^hni] eller, ikke sjældent, endnu mere reduceret [çi]. En ulempe ved dette metodiske valg er at den undersøgte variation ikke direkte kan relateres til den del af den øvrige forskning som netop tager udgangspunkt i den fonetiske realisation af et ords segmenter. Undersøgelsen vil derfor ikke kunne give evidens for om eller i hvor høj grad et ords varighed korrelerer med antal og art af realiserede segmenter, eller for at det er graden eller arten af segmentreduktion der korrelerer med de undersøgte faktorer. En væsentlig fordel ved metoden er dog at den giver en tidsbesparende mulighed for at undersøge udtalevariation i et stort korpus for naturligt talesprog som ikke er transskriberet fonetisk, men hvor der er mange belæg på den undersøgte variabel i forskellige syntaktiske og diskursive kontekster. Det vil være ønskeligt og gøre undersøgelsen mere pålidelig hvis resultaterne af undersøgelsen kan underbygges af dokumentation af faktisk udtale, men det er en tidskrævende opgave at lydskrive i halvfin transskription de 900+ forekomster der er indgået i undersøgelsen, og jeg har derfor afstået fra dette.³

1.3 Artikelens struktur

Artiklen er struktureret således: Først gives en kort sammenfatning af ordbøgernes fremstilling af *egentlig* og af beskrivelser af *egentlig*s semantik og pragmatik. Dernæst beskrives data, herunder de distributionelle tendenser der kan spores. Efter dette diskuteres målemetoden, varighed, og herefter redegøres der for de faktorer der indgår i undersøgelsen. Efter dette fremlægges først indsigter fra den deskriptive statistik, og derpå beskrives anvendelsen af den statistiske metode der måler de forskellige faktorerers indflydelse på *egentlig*s varighed. Efter dette følger et resultatafsnit og herpå et perspektiverende afsnit. Artiklen rundes af med et konkluderende afsnit.

3 DanPASS-korpusset, som Schachtenhaufen har anvendt i sine analyser, er fuldt annoteret for bl.a. faktisk udtale, og transskriptionen er af høj kvalitet, men korpusset rummer kun tre forekomster af *egentlig* (http://schwa.dk/filer/udtaleordbog_danpass; se Grønnum 2009 for en nærmere beskrivelse af DanPASS-korpusset).

2 OM EGENTLIG

Den brug af *egentlig* der undersøges her, er funktionen som sætningsadverbial, dvs. *egentligs* anvendelse på et ikkepropositionelt niveau. Som adjektiv og som mådesadverbium vil *egentlig* bidrage til sætningens propositionelle niveau (Jensen 2006: 1), dvs. sætningens egentlige betydning. Som sætningsadverbium henviser *egentlig* til hvad der bredt kan karakteriseres som 'forhold i kommunikationssituationen', dvs. i en brug der ifølge Hansen & Heltoft (2011: 1033) i en vis forstand er "kommentarled", dvs. at det angiver "den talendes kommentar til resten af sætningens beskrivende indhold", og derved, ifølge Andersen (1982: 86), på forskellig måde inddrager forholdet mellem afsender og modtager.

2.1 Egentligs betydning

Den Danske Ordbog (DDO) giver to betydningsbeskrivelser af adverbiet *egentlig* der lægger sig tæt op ad mere konkrete adjektiviske betydninger: (1) 'af en bestemt oprindelse eller af en grundlæggende natur', fx *hun er egentlig fra Finland, men bor i Oslo*, og (2) 'i virkeligheden (når der ses bort fra de ydre omstændigheder)', fx *han var egentlig flink nok*. Synonymudtryk til denne betydning er *i grunden*, *ret beset*, *inderst inde*. Hertil kommer en mere abstrakt betydning, der i DDO har fået en brugsdefinition hvad der indikerer at selve betydningen ikke lader sig bestemme semantisk præcist: 'bruges for at nedtone et synspunkt eller (tøvende) indrømme noget', fx *det tror jeg egentlig nok* eller *det er egentlig underligt*. Til denne brug angives *faktisk* som synonym. I det analyserede datasæt er denne generaliserede, eller mere subjektiverede, betydning langt den almindeligste, hvad der meget fint afspejler det angivne brugsdomæne i DDO, "især talesprog".

2.2 Egentligs pragmatiske funktion

Hos Jensen (2006) analyseres *egentlig* sammen med *faktisk* ud fra et polyfont perspektiv. I dette tilfælde vil det sige at adverbialerne så at sige udtrykker to synspunkter på én gang. Med *faktisk* i fx *han spiller faktisk godt klaver* udtrykker man dels at noget er gældende, 'at han spiller godt klaver', og at der er et *modsynspunkt* til dette, at nogen mener, eller at det i almindelighed menes at han ikke spiller godt klaver. I Jensens analyse

præsupponerer *egentlig* ikke et modsynspunkt, men derimod et synspunkt som nogen har, eller som i almindelighed bare findes, og som taler erklærer sig enig i. I *han spiller egentlig godt klaver* udtrykker afsender sin holdning, *han spiller godt klaver*, og samtidig det (samstemmende) synspunkt hos nogen, eller 'ude i verden', at *han spiller godt klaver*. Med *egentlig* erklærer man sig altså enig i et synspunkt, og *egentlig* har dermed en "bekræftelsesfunktion, og er for så vidt redundant, men ... bibringer noget subjektivt ved at afsenders engagement fremhæves" (Jensen 2006: 12). Denne analyse lægger sig stærkt op ad den indrømmende, generaliserede og subjektiverede funktion som *egentlig* ifølge DDO især har i talesprog.

2.3 Egentligs udtale

Den Store Danske Udtaleordbog (Brink et al. 1991: 385) giver kun udtaler til adjektivet *egentlig*: ['e:[?]jnli]⁴ (distinkt form), ['e:[?]nli] (hovedform), ['e:[?]enli] (mere distinkt), ['e:[?]jnd,li] (mere distinkt). Ordbogen oplyser at adverbiet *egentlig* har de samme udtaler "men reduceres oftere som adv., hvor det – undt. på sætningens 1.-plads – også ofte står med bi-tryk istf. hovedtryk". Som nævnt ovenfor er datasættet med de 900+ forekomster af *egentlig* ikke som helhed blevet transskriberet fonetisk, men ud fra umiddelbare gennemlytninger og selektive transskriptioner kan ordbogens hovedform bekræftes, og lige sådan kan tendensen til ikkereduktion på sætningens førsteplads hvorimod udtalerne med mærkatet "mere distinkt" er så godt som aldrig forekommende. Tabel 1 giver eksempler på udtaler i datasættet med syntaktisk kontekst og varighed i sekunder.

TABEL 1: EKSEMPLER PÅ EGENTLIGS UDTALE- OG VARIGHEDSVARIATION; TALER-ID63

	Udtale	Eksempel	Varighed i sek.
(1)	[ən]	og jeg tror egentlig jeg havde ham ...	0,123
(2)	['e: [?] nə]	så vi øh har så ... øh og egentlig opgivet tanken om at ...	0,255
(3)	['e: [?] ne]	jeg havde en ... dansklærer som også havde det der kulturprojekt egentlig	0,32

4 *Den Store Danske Udtaleordbog* anvender lydskriftalfabetet Dania. I denne artikel er udtalerne angivet i Dania translittereret til IPA i henhold til Grønnum (2005; jf. note 1 ovenfor).

(4)	[² e:ni]	men egentlig så er det tit noget med at hvis der kommer ...	0,340
(5)	[² e:nli]	.. hvor vi så blev kærester sådan egentlig lidt ved et tilfælde	0,314
(6)	[² e:ndnə]	så det forekom ret nemt egentlig umiddelbart ikke at gøre det	0,46

Fra gennemlytning, men altså ikke ud fra en fonetisk transskription af alle eksemplerne, er det mit klare indtryk at de hyppigste udtaler er mellem (lidt) reduceret [²e:ɲə] til [²e:ni], [²e:li], [²ẽ:ni]. Hvis der ekspanderes segmentalt, kan udtaler med [l] efter [n] forekomme, som i (5). Udtaler med tre stavelser er så godt som fraværende. Men mere reducerede udtaler som i (1) eller [₁e:ɲ] er heller ikke sjældne. (Bemærk at stødet er sejlivet i reducerede udtaler af *egentlig*; undertiden er stødet bedre transskriberet som en laryngal modifikation af vokalen, fx [²eni]). Det er mit indtryk, hvilket bør efterprøves med mere systematisk transskription og akustiske målinger, at variationen i udtalen af *egentlig* i væsentlig grad skal tilskrives varigheden af de realiserede segmenter i en typisk udtale som [²e:ni], i særdeleshed varigheden af det initiale og undertiden trykstærke [e]. Bemærk at det segmentelt mere fyldige [²e:ɲli], (5), har kortere varighed end det lidt mere reducerede [²e:ni], (4).

Hverken analysen eller det umiddelbare indtryk af segmentrealiseringen kan be- eller afkræfte den påståede sammenhæng mellem trykforhold og reduktion. Jeg har i denne undersøgelse fundet det uoverkommeligt på rimelig konsistent vis at bestemme hoved-, bi- og nultryk i den løbende talestrøm hvorfra forekomsterne er taget. Bestemmelsen af graden af reduktion i forhold til tryk, eller forekomsten af trykreduktion i forhold til øvrig fonetisk reduktion og den syntaktiske – og pragmatiske – funktion som *egentlig* måtte have, overlades således til fremtidige undersøgelser.

3 DATA

De forekomster af *egentlig* der er genstand for analysen i denne artikel, stammer fra sociolingvistiske interviews der er udført med og af københavnsktalende informanter og interviewpersoner i København

i perioden 2005–2010. Denne variant af dansk kan kaldes moderne standardkøbenhavnsk. Interviewene udgør den del af det *nye bysocio-lingvistik-korpus* (NyBySoc), der indgår i det mere end 8 millioner ord store LANCHART-korpus (Gregersen 2009). Delkorpusset NyBySoc indeholder ca. 789.000 ord, næsten ti gange flere end i DanPASS, og det er transskriberet ortografisk (og desuden annoteret for en række andre grammatiske og diskursive træk der ikke er relevante for denne undersøgelse). Hver forekomst af *egentlig* er først blevet lokaliseret vha. af det søgeprogram der er udviklet til LANCHART-korpusset, og efterfølgende vha. analyseprogrammet *Praat* (Boersma & Weenink 2015), der sammenholder en indlæst lydfil (fx vha. et spektrogram) med ord for ord-transskriptionen i en tilhørende tekstfil. I NyBySoc-transskriptionerne er ord for ord-tilpasningen med lydfilerne imidlertid ikke præcis på ordniveau, og derfor er alle forekomster af *egentlig* i forbindelse med denne analyse manuelt blevet tilpasset i en særlig tekststreng. Denne tilpasning har dannet grundlag for den varighed af *egentlig* der er denne analyses omdrejningspunkt.⁵

Der er i alt 1003 forekomster af *egentlig* i det undersøgte korpus. 67 forekomster er ekskluderet fra datasættet, der således omfatter 936 forekomster.⁶

4 VARIGHED SOM MÅL FOR REDUKTION

Varighed er ikke en ukontroversiel målestok for reduktion, men den ses anvendt i den internationale litteratur (fx Fosler-Lussier & Morgan

5 Ved søgning i LANCHART genereres der en csv-fil der kan indlæses i et regneark til videre analyse. For hver forekomst af ens søgeord giver csv-filen en række yderligere oplysninger, fx om taler, talers alder, køn og socialklasse, samt, efter ønske, den syntaktiske kontekst, dvs. *egentligs* placering i en given talers ytring. Nogle af disse ikkesproglige oplysninger er blevet anvendt i denne analyse, og oplysningerne om *egentligs* syntaktiske kontekst og placering i helytringer har dannet udgangspunkt for de sproglige analyser der er indgået som de sproglige variable i den statistiske analyse.

6 I 39 forekomster optræder *egentlig* i en afbrudt sætning eller i en sætning der er ubestemmelig med hensyn til syntaktisk struktur, og i 23 forekomster kan *egentlig* ikke afgrænses præcist på grund af overlap, hosten, latter, mumlen eller kontraktion med naboord. Blot to forekomster af *egentlig* er adjektiver, og tre *outliers*, dvs. forekomster der afviger ekstremt fra de øvrige og den gennemsnitlige varighed, er ligeledes ekskluderet fra datasættet på grund af varigheder på 0,89, 0,93 og 1,33 sek. (Den gennemsnitlige varighed for de 936 forekomster er 0,241 sek.).

1999, Gahl 2008, Bell m.fl. 2009, Schubotz m.fl. 2015). Varighed som målestok for reduktion kan anfægtes af tre grunde. For det første er ikke alle segmenter iboende lige lange; lavere vokaler er alt andet lige længere end høje vokaler (Keating 1985: 118; Lehiste 1970), og en velær hæmmelyd [x], som reduceret manifestation af /k/ eller /g/, kan være længere end [k] eller [g]. Dette vil give længere varighed til et ord der består af fx fem iboende lange segmenter over for et ord med fem iboende korte segmenter. For det andet kan man sagtens forestille sig en segmentalt reduceret udgave af *egentlig* blive udtalt med vokalfør-længelse [eːʔji] og som sådan være længere i varighed end en segmentalt mere fyldig og mindre reduceret udtale: [eʔnli]. For det tredje, og måske vigtigst i forbindelse med denne undersøgelse, kan et ords varighed afhænge af talehastigheden, fx målt som ord per sekund, hvorimod et ords segmentrealisering ikke nødvendigvis, eller ikke for alle talere, vil afhænge af talehastigheden; man kan godt tale hurtigt og distinkt, med tydelig artikulation af alle et ords segmenter, og man kan tale langsomt men med ikketydelig, reduceret realisation af segmenter (Zwicky 1972; Rischel 1992; Brink 2013: 14). Hertil kommer at individuelle tilbøjeligheder og nogle ords iboende tendens til eller iboende modstandsdygtighed over for reduktion kan medføre afvigelser fra et ellers overensstemmende forhold mellem tydelig artikulation og talehastighed. Man bør dog forvente en vis form for sammenhæng mellem talehastighed og reduktion på ord- og segmentniveau. Dette finder man også til en vis grad i det undersøgte datasæt.

Det vil være ideelt at kunne supplere varighed med en transskription af det faktisk udtalte og lade en sådan være et alternativt mål for at se om eller i hvor høj grad et kortvarende *egentlig* også er segmentalt reduceret, og om eller i hvor høj grad et længerevarende *egentlig* er segmentalt ekspanderet. Men en sådan målestok er heller ikke ukontroversiel, for man kan med rimelighed spørge hvornår et realiseret segment skal regnes som en reduceret udgave af et forventet segment, og hvilke kriterier man skal følge for at fastlægge kvaliteten af et forventet segment. Desuden er et sådant mål følsomt for finhed i transskriptionen, jo finere man transskriberer fonetisk, jo større chance er der for at man transskriberer et segment som reduceret i forhold til det forventede.⁷

⁷ Schachtenhaufen (2013) og Heegård (2013) bruger deres egne skalaer for 'reduktionsværdi' som beregnes ud fra en halvfins transskription sat i forhold til en forventet distinkt udtale. Hos

Der er både en fonetisk og en praktisk grund til at *egentlig*s tendens til reduktion ikke er målt ud fra transskriptioner af faktisk udtale. Dels vil det gå ud over undersøgelsens format at lave en halvfin auditivt baseret transskription af alle 936 forekomster, dels er udtalen af *egentlig* i spontan tale undertiden så reduceret og så fonetisk afvigende fra de lydværdier som et fonetisk alfabet som IPA stiller til rådighed, at en given udtale kun tilnærmelsesvist kan repræsenteres. Dette medfører stor brug af diakritiske tegn, og det vil være omstændeligt at håndtere og gøre den enkelte transskription svært læselig (og tilsvarende indviklet at værdisætte med hensyn til reduktionsgrad).

4.1 Egentligs varighed og talehastighed

LANCHART-korpusset giver mulighed for en rimelig om end halvgrov automatisk estimering af talehastighed. I udskrivningen af lydfilerne har udskriverne sat grænser ved pauser i talestrømmen, mellem hvilke der så er et stræk bestående af et antal ord. Disse stræk behøver ikke være selvstændige eller velformede ytringer, idet man jo kan holde en pause midt i en sætning, eller undlade at holde pause mellem to sætninger. Ved hjælp af et script har det kunnet lade sig gøre at fastslå talehastigheden af hvert af disse stræk af ord som indeholder *egentlig*.⁸ Strækkene kan indeholde ét ord, nemlig *egentlig*, sagt prosodisk men ikke nødvendigvis syntaktisk isoleret, flere ord svarende til en (kort) sætning eller mange ord, svarende til flere sætninger. Det anvendte script har målt både antallet af ord og, ved hjælp af den automatiske

Schachtenhaufen fastsættes reduktionsværdien for et givent fonem eller et givent ord, eller dets 'reduktionstilbøjelighed', ud fra hvor mange gange ud af 100 et fonem eller ord udtales reduceret. Hos Heegård får adverbialet *faktisk* tildelt et antal point baseret på hvor mange segmenter det er udtalt med, og om disse er reducerede udgaver eller ej af et forventet distinkt udtalt segment.

8 Når målet er groft, skyldes det at det ikke tager hensyn til om der i strækket ytres rigtige eller velformede ord eller ej. Fx er selvfærdelser i LANCHART skrevet med bindestreg, fx *fone-*, hvis en taler har villet sige *fonetik* men afbrudt sig selv. Også tøvelyde, som i henhold til LANCHART-konventionerne er skrevet *øh*, er talt med som et ord. Derudover kan nogle stræk være særdeles lange, fx over 70 ord, og inden for selv kortere ytringer end disse er det velkendt at talehastighed kan variere. Det ville have været ønskeligt at kunne anvende talehastighed inden for den prosodiske frase som mål, som fx anvendt i Schubotz m.fl. (2015), og som det er muligt at gøre i analyser baseret på DanPASS-korpusset, men dette kræver en opmærkning som LANCHART-korpusset i skrivende stund ikke tilbyder. Det her anvendte mål for talehastighed er dog finere end de meget grove mål for talehastighed som i en lydfil blot tæller antal ord og dividerer dette tal med tallet for lydfilens længde uden hensyntagen til fx pauser.

fonologiske opmærkning i LANCHART-korpusset, antallet af fonologiske stavelser i det stræk som *egentlig* optræder i,⁹ og det har ved hjælp af tidskoder målt hvor langt strækket er i sekunder, og derpå givet et tal for hvor mange ord henholdsvis stavelser der er per sekund i målestrækket. Disse automatisk udtrukne mål giver på denne måde mulighed for at beregne i hvor høj grad der er sammenhæng, *korrelation*, mellem *egentlig*s varighed og talehastighed målt både som ord per sekund og (fonologisk) stavelse (= vokal) per sekund, dvs. om *egentlig* har kort varighed når ordet indgår i et stræk af ord der globalt set har høj talehastighed (mange ord eller stavelser per sekund), eller omvendt, om *egentlig* har lang varighed når det indgår i et stræk af ord med lav talehastighed (få ord eller stavelser per sekund).

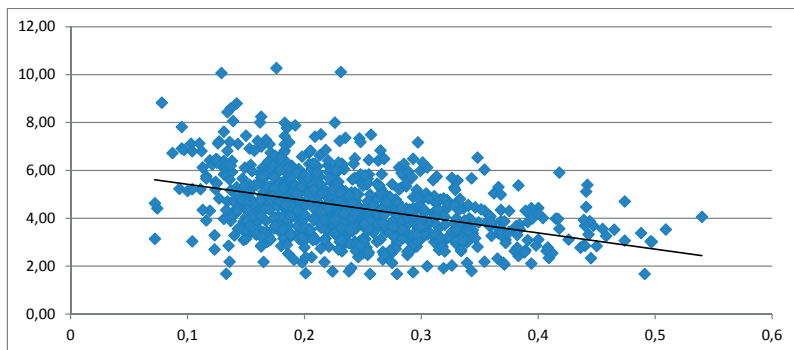
Regressionslinjen i grafen i figur 1 viser at der i datasættet er en overensstemmelse mellem *egentlig*s varighed og talehastigheden i det ordstræk som det indgår i. Regressionen har den forventede retning: Jo højere talehastighed et ordstræk med *egentlig* har (y-aksen), jo hurtigere bliver *egentlig* udtalt (x-aksen). Med de signifikante p-værdier bekræfter Spearmans rangkorrelationstest denne tendens, men testens r-værdier fortæller os at regressionen ikke er overvældende stor (en perfekt korrelation hvor alle forekomster præcist følger regressionslinjen, vil have værdien -1), hvad vi også kan se ud fra den relativt flade vinkel som regressionslinjen har.¹⁰ Det er altså ikke givet at høj talehastighed automatisk altid medfører kort varighed af *egentlig*, og omvendt at en kortvarig udtale af *egentlig* ikke er et entydigt indeks for høj talehastighed. Dette peger på at varighed og talehastighed i et vist omfang kan anvendes som to selvstændige mål i en fonetisk undersøgelse. Men at

9 Målet for antal fonologiske stavelser er baseret på en strengt konservativ udtale, bl.a. er der regnet med fem fonologiske stavelser i et ord som *materialer* trods etableret udtale med fire, [matri'æ:lɐ], med to stavelser i *ikke*, trods en hyppig udtale [eg], og, ikke mindst, med tre stavelser i *egentlig* [e:'jɛndli], selvom om de hyppigste udtaler i datasættet kun har to.

10 En test for normalfordeling viser en lille tendens til to pukler, dvs. en ikkeperfekt normalfordeling, og derfor er Spearmans rangkorrelationstest anvendt (Rasinger 2013: 161–167, 221–225). (En test med Pearsons korrelationstest for normalfordelte data giver et tilsvarende resultat.) Overensstemmelsen mellem Varighed og Ord per sekund er en lille smule bedre end mellem Varighed og Stavelser per sekund, mellem -0,04 og -0,07, og derfor bruges målet Ord per sekund i denne undersøgelse. Regressionen er beregnet med alle 936 forekomster. Hvis outliers ekskluderes, forbedres r-værdierne kun minimalt, fra 0,02 til 0,07 – i retning mod mere overensstemmelse, og da der af denne grund ikke er et væsentligt argument for at ekskludere outliers, er disse bibeholdt i analysen nedenfor.

der trods alt er en (mindre og signifikant) sammenhæng mellem de to mål, betyder at talehastighed ikke blot kan ignoreres, og det er derfor inkluderet i den statistiske model der afdækker de indvirkende faktorer på varigheden af *egentlig*.

FIGUR 1: FORHOLDET MELLEM VARIGHED I SEKUNDER OG TALEHASTIGHED



Grafen viser regressionslinjen af forholdet mellem varighed i sekunder (x-akse) og talehastighed målt som antal ord per sekund (y-akse); $n = 936$; Spearmans rangkorrelationstest: $r = -0,4159217$, p -værdi = $2,501e-09$.

5 FAKTORER

Variationen af *egentlig*s varighed er holdt op imod en række sproglige såvel som ikkesproglige faktorer der i andre reduktions- og variationsstudier har vist sig at have en indflydelse på de undersøgte fænomener (jf. indledningen), eller som af informationsstrukturelle grunde kunne tænkes at have en indflydelse på ytringers varighed eller reduktionstendens.¹¹

5.1 Ikkesproglige faktorer

De ikkesproglige faktorer der indgår i analysen, er traditionelle sociolingvistiske variabler: Generation (Ung, Mellem og Ældre), Køn,

11 Forekomsterne er ikke blevet opmærket for dialogisk funktion, fx i henhold til Jensens (2006) eller Mortensens (2010) analyser, og som i Plugs (2005) analyse af hollandsk *eigenlijk* og i Heegård & Mortensens (2014) analyse af *faktisk*. Dette ville have været en særdeles tidskrævende opgave. Til sammenligning analyser Plug blot ”a small subset” af 269 forekomster i et hollandsk talesprogskorpus (Plug 2005: 132) og Heegård & Mortensen (2014) kun 138 sætninger. Forekomsterne er derimod blevet undersøgt for betydning i henhold til DDO’s betydningsbeskrivelse af *egentlig*, men da ca. 1/3 af forekomsterne ikke entydigt lader sig bestemme betydningsmæssigt, indgår denne opmærkning ikke i den statistiske analyse.

Socialklasse (arbejderklasse, AK, og middelklasse, MK)¹² og desuden Rolle i interviewet (interviewer eller informant). Generationen Ung omfatter talere under 25 på optagetidspunktet, generationen Ældre talere over 45 på optagetidspunktet. På nær en er alle interviewere i Mellem-gruppen. Generation er blevet undersøgt som en trevejsfaktor, som to tovejsfaktorer, Ældre vs. Ikkeældre og Ung vs. Ikkeung, og som en seksvejsfaktor i henhold til fødselsår, 1940–1990’erne. Alle interviewere, som alle var ansatte ved Københavns Universitet på interviewtidspunktet, er tilskrevet socialklassen MK.

5.2 Sproglige faktorer

De sproglige faktorer kan deles op i hvad der kan kaldes kollokationelle variabler, og variabler der har at gøre med syntaks eller informationsstruktur. Det er velkendt fra den internationale litteratur at et givent ords tendens til reduktion kan hænge sammen med hvilke ord det optræder i umiddelbar sammenhæng med (Bybee & Scheibmann 1999; Bell m.fl. 2009; Tily m.fl. 2009). Dette er der opmærket for i datasættet i henhold til om *egentlig* er placeret lige før eller lige efter et andet sætningsadverbium, *egentlig_ADV* (fx *egentlig gerne* eller *egentlig ikke*) og *ADV_egentlig* (fx *sgu egentlig* eller *vel egentlig*), om *egentlig* knytter sig til et kognitionsverbium som fx *tro* eller *synes*, og om *egentlig* efterfølger den modsigende dialogiske partikel *jamen*. Datasættet er også opmærket for om *egentlig* optræder i nærheden af en tøvemarkør, skrevet ”øh” i korpusset, jf. Bell m.fl. 2003 mht. fonetisk reduktion og ikkeflydende tale).

Med afsæt i at fonetisk reduktion henholdsvis ekspansion kan relateres til principper for præsentation af kendt versus ukendt information (Schachtenhaufen 2010, 2013) er datasættet opmærket for syntaktiske konstruktioner hvis funktion på forskellig vis er at fremhæve særlige led: fundamentfordobling, sætningsknode, sætningskløvning og topikaliserings, dvs. om der er et andet led i fundamentfeltet end subjektet.¹³

12 Alle informanterne i LANCHART-korpusset tilskrives socialklasserne MK eller AK ud fra kriterierne uddannelse, ledelsesfunktion på informantens arbejdsplads og indholdet af informantens arbejde (Gregersen 2009: 9–11).

13 Når *egentlig* optræder i ledsætninger, er der ikke markeret for fundamentfordobling, sætningskløvning, sætningsknode og andet led i fundamentfeltet end subjekt.

Torben Juel Jensen og Tanya Karoli Christensen har i en række artikler (fx Christensen & Jensen 2013 og Jensen & Christensen 2014) påvist at ledstillingsrækkefølge i ledsætninger kan korrelere med semantisk vigtighed, i deres terminologi forgrundsinformation over for baggrundsinformation. Jeg har i analysen ladet mig inspirere heraf og opmærket alle forekomster for om de optræder i hovedsætninger eller i ledsætninger.¹⁴ Alle forekomster af *egentlig* er også opmærket for om ordet optræder som absolut sidste ord i sætningen, dvs. foran en pause, en placering der i andre undersøgelser er påvist at korrelere positivt med mindre reduktion (Heegård 2013: 110–113; Schachtenhausen 2013: 182–184; Schubotz m.fl. 2015: 366). Endvidere er alle forekomster noteret for om de forekommer i spørgsmål eller ej, ud fra den hypotese at spørgsmål retter fokus på en bestemt type information.

Endelig er alle forekomster noteret for deres topologiske placering, dvs. deres plads i sætningsskemaet anvendt for dansk. Sætningerne (7)–

14 Ved placering i højre interjektionalfelt efter ledsætninger kan det være svært at afgøre om *egentlig* knytter sig til udsagnet i ledsætningen eller i hovedsætningen. I sådanne tilfælde er *egentlig* ikke tilskrevet forekomst i hovedsætning eller ledsætning. Endvidere er alle ledsætninger opmærket for om de er adverbelle, nominale eller attributive. Se også Schubotz m.fl. (2015) for undersøgelse af om en skelnen mellem placering i hovedsætninger over for ledsætninger, og forskellige ledsætningstyper, har en effekt på varigheden af *you know* over for diskursmarkøren *y'know*, baseret på antagelsen at talesprogsledsætninger "are frequently used as statements and have none of the parenthetical character they are associated with in written language" (Schubotz m.fl. 2015: 377).

(12) i sætningsskemaet nedenfor viser hvilke pladser *egentlig* per default optræder på i datasættet.

	Venstre Interjek- tional	Venstre Ekstra- position	Funda- ment	v	n	a	V	N	A	Højre Interjek- tional	
(7)	egentlig		det	er	løn	nok			i sig selv		(ID9)
(8)		egentlig	så	er	det	tit		noget	med at hvis der kommer ...		(ID63)
(9)		egent- lig	har	min mor			været	rimelig large			(ID6)
(10)			jeg	tror		egent- lig			jeg havde ham ...		(ID63)
(11)			vi	blev			kære- ster	sådan egentlig nærmest lidt ved et tilfælde			(ID63)
(12)			der	boede	børn				i alle op- gange	egent- lig	(ID3)

I nogle sætninger optræder *egentlig* på ikkedefault position, fx mellem N og A, (13), eller mellem V og N, (14).

- (13)

så det forekom ret [nemt (N)] **egentlig** [umiddelbart ikke at gøre det (A)]

(ID63)
- (14)

jeg kan heller ikke engang [huske (V)] **egentlig** [hvor meget (N)]

(ID22)

Endelig optræder *egentlig* også i sætningsækvivalenter som i ID34's svar i (15).

- (15)

ID60

er der mange fra klassen der ryger

ID34

nej **egentlig** ikke

6 OVERORDNEDE TENDENSER

6.1 DISTRIBUTION OG GENNEMSNITSVARIGHED

Undersøgelser af reduktionen af *faktisk* (Heegård 2013; Heegård & Mortensen 2014) viser at dette adverbial er signifikant kortere – og mere reduceret – på den topologisk centrale a-plads end på de ikke-centrale Venstre Interjektional, Venstre Ekstraposition, Fundament og Højre Interjektional.¹⁵ Dette gælder også for *egentlig*, som det ses i tabel 1.

TABEL 1: DISTRIBUTION OG GENNEMSNITLIG VARIGHED AF *EGENTLIG* PÅ FEM TOPOLOGISKE PLADSER

	Lille-a	Ikke default	Sætnings-ækvivalent	Højre Interjektional	Fundament, Venstre Ekstraposition og Venstre Interjektional	Samlet
Antal	738	34	36	107	21	936
Gennemsnit	0,225	0,293	0,298	0,303	0,315	0,241

Tabellen viser en ret stor overensstemmelse mellem *egentlig*s varighed i de topologiske kategorier der ikke er Lille-a, over for den noget lavere varighedsværdi på Lille-a-pladsen. Parvise signifikanstest mellem de positioner der ikke er Lille-a, viser at der ikke er signifikante forskelle mellem disse. Det er der derimod mellem Lille-a og Ikke-default, dvs. mellem det laveste og det næstlaveste gennemsnit. Den topologiske placering er således en vigtig faktor for *egentlig*s varighed. Topologiens vigtighed bekræftes hvis de topologiske kategorier Ikkedefault, Højre-Interjektional og Fundament og Venstre-Ekstraposition slås sammen og stilles over for Lille-a (og den ytringsafvigende ’Sætningsækvivalent’ udelades), som vist i tabel 2.

TABEL 2: DISTRIBUTION OG GENNEMSNITLIG VARIGHED AF *EGENTLIG* I DE TO TOPOLOGISKE KATEGORIER LILLE-A OG IKKE-LILLE-A

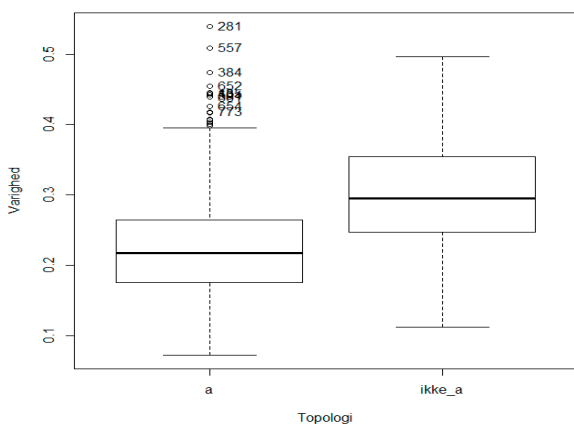
	Lille-a	Ikke-Lille-a	Samlet
Antal	738	162	900
Gennemsnit	0,225	0,302	0,239

15 Der er kun tre forekomster på adverbialfeltets Store-A, og disse er slået sammen med kategorien Højre-Interjektional.

Forskellen mellem de to kategorier er signifikant ($p < 2e-16$; f-værdi 153,8, df 1), og vi kan dermed konkludere at når *egentlig* optræder på Lille-a, er ordet gennemsnitligt kortere end når det optræder andre steder i sætningskemaet.

Modifikationen 'gennemsnitligt' er vigtig, for det er ikke sådan at *egentlig* ikke kan have langvarig udtale på Lille-a, eller at *egentlig* kun udtales med stor varighed på en anden plads. Dette fremgår af boksplottet i figur 2.

FIGUR 2: BOKSPLOT AF EGENTLIGS GENNEMSNIITSVARIGHEDER I DE TOPOLOGISKE KATEGORIER LILLE-A OG IKKE-LILLE-A



Boksplottet viser med de fede vandrette streger 50 %-medianen for de to kategorier. De to bokse omfatter forekomsterne fra 25 %- til 75 %-kvartilerne, og de to *whiskers* oven for og neden for boksene angiver højeste og laveste værdier i kategorierne, eksklusive de forekomster som den statistiske software definerer som *outliere*; vi ser oven for den øvre whisker for Lille-a at denne kategori har nogle stykker af disse. Boksplottet viser tydeligt at gennemsnitsvarigheden for *egentlig* er lavere i Lille-a-kategorien, men også at der er en pænt stor spredning i varighed inden for begge topologiske kategorier.

Som opsummering på den deskriptive statistik kan vi fastslå at langt de hyppigste forekomster i materialet er på Lille-a-pladsen, at forekomsterne på denne plads gennemsnitligt er signifikant kortere end på andre pladser, at forekomsterne på de andre topologiske pladser er

nogenlunde ensartede med hensyn til varighed, og at der er forholdsvis stor variation på både Lille-a og Ikke-Lille-a. Det er denne variation der vil blive undersøgt i det efterfølgende.

7 HVORNÅR EN FAKTOR ER SIGNIFIKANT ELLER IKKE-SIGNIFIKANT

Spørgsmålet om faktorerers signifikans i forhold til hinanden kan undersøges statistisk vha. den type regressionsanalyse der kaldes *mixed-model-multipel-linear-regressionsanalyse*. Dette statistiske værktøj giver som output analytikeren en såkaldt *model*, dvs. en fremstilling af hvordan og i hvilken grad en eller flere uafhængige faktorer påvirker eller kan forudsige udfaldet af eller variationen i den afhængige faktor, her den varighed som udtalen af *egentlig* har. Afhængigt af den specifikke software der anvendes, vil det af modellen ligeledes fremgå om en given faktor har en signifikant indflydelse på den afhængige faktor. Alle mål beregnes under hensyntagen til den påvirkning de andre faktorer man vælger at inkludere i modellen, også måtte have på den afhængige faktor.¹⁶

7.1 Taler som tilfældig faktor

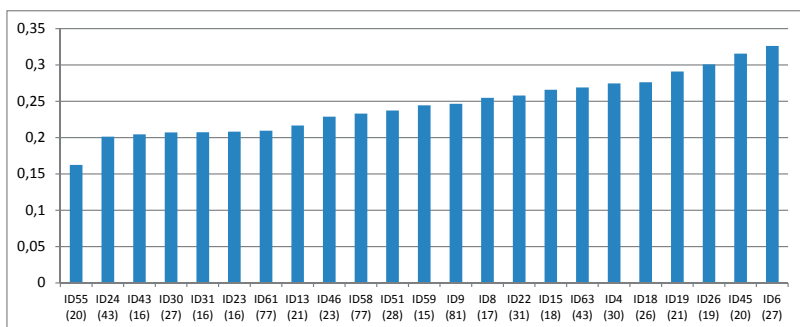
Den anvendte regressionsanalyse har den fordel over for en simpel regressionsanalyse at den kan kontrollere for den effekt som *tilfældige faktorer* (eng. *random factors*) kan have på den afhængige variabel. En tilfældig faktor i en stikprøve er typisk variable faktorer der er tilfældige i forhold til en samlet population, og som ikke vil gå igen i en replikation af eksperimentet (Baayen 2008: 241–242; Winther Balling 2008: 181; Pharao 2010: 149–150). I denne undersøgelse er Taler, dvs. den person der udtaler en given forekomst af *egentlig*, en oplagt vilkårlig faktor. I en replikation af eksperimentet vil man formentlig ikke vælge de samme talere, men, selvfølgelig, mere eller mindre samme sæt af uafhængige variable. Man ville i stedet undersøge om resultaterne passer med hvad man kan analysere sig frem til på basis af et materiale med andre

¹⁶ Analysen er udført vha. *The R Project for Statistical Computing* (R Core Team 2014), R i kortform, og funktionen *lmer* i pakken *lme4* (version 1.1-8; Bates et al. 2014) og analyseværktøjerne i funktionen *Remdr*. Visse af plottene er udført vha. funktionerne *plotLMER.fnc* i pakken *languageR* (version 1.0; Baayen 2007) og *plotmeans* i pakken *plots* (Warnes m.fl. 2014).

taleres udtale af *egentlig*. Med ‘kontrollere for’ menes at den statistiske analyse tager højde for den påvirkning den individuelle taler måtte have på *egentlig*s varighed, og at de værdier der gives for de ikketilfældige faktorer, de *faste faktorer* (‘fixed factors’), er beregnet ud fra den påvirkning som den tilfældige faktor måtte have.

Ud over den metodisk-principielle grund er der to andre gode grunde til at inkludere Taler som en tilfældig faktor i denne undersøgelse. For det første er der stor spredning med hensyn til hvor mange forekomster af *egentlig* den enkelte taler har: 18 talere har fem eller færre forekomster, fem talere har mere end 40 forekomster; gennemsnitligt er der i datasættet 15 forekomster per taler. For det andet er der netop en stor – og signifikant – forskel med hensyn til hvor lang tid udtalen af *egentlig* varer internt i gruppen af talere, jf. figur 3 og efterfølgende tekst.

FIGUR 3: GENNEMSITLIG VARIGHED FOR *EGENTLIG*



Figuren medtager kun talere med over 15 belæg (antal i parentes); alle topologiske pladser; $n = 712$, gennemsnit 0,245 sek.

Af figur 3 fremgår det at den taler der gennemsnitligt har den korteste varighed, ID55, er omtrent dobbelt så hurtig om at artikulere *egentlig* som den taler der har den længste gennemsnitlige varighed, ID6, nemlig 0,162 vs. 0,32 sek. En test for variansens størrelse blandt disse 23 talere viser at denne er signifikant ($p < 2e-16$; f -værdi 7,965, df 22). Det er altså i høj grad berettiget at lade Taler være en tilfældig faktor i de videre statistiske tests.

7.2 Modelbygning

De uafhængige variabler præsenteret ovenfor blev i analysen inkluderet som faste faktorer. I den procedure der førte til den endelige model, blev enhver faktor der ikke viste sig at have en signifikant indflydelse på den afhængige faktor, udeladt fra modellen. Kun faktorer der kunne påvises at have en signifikant indflydelse på *egentlig*s varighed, forblev i modellen til videre analyse.¹⁷

Den model man på denne måde ender med at have, er ikke nødvendigvis den eneste model med signifikante faktorer, dvs. ikke den eneste repræsentation af de faktorer der på signifikant vis kan redegøre for variationen i ens data. Men af de gode modeller der kan det, er det den der gør det bedst. Dette betyder så igen at der kan være faktorer eller kombinationer af faktorer (*interaktioner*) der isoleret set eller i andre gode modeller har en væsentlig indflydelse på den uafhængige faktor.

8 ÅRSAGER TIL VARIATIONEN I *EGENTLIG*S VARIGHED

Dette afsnit afrapporterer resultaterne af de inferentielle statistiske analyser af *egentlig*s varighed. Først præsenteres den model som det statistiske software har givet som output som den bedste af flere mulige gode modeller.

I henhold til den individuelle talers mulige indflydelse på *egentlig*s gennemsnitlige varighed (jf. tabel 2) er Taler inkluderet i modellerne som *tilfældig faktor*. For faktoren Generation gav tovejsinddelingen Ældre vs. Ikkeældre, inkl. interviewere, den mest sigende model. Også faktoren Socialklasse er blevet afprøvet med og uden interviewere, og også her gav inklusionen af interviewere det mest sigende resultat. En række faktorer er blevet afprøvet i henhold til procedurebeskrivelsen oven-

17 Modellen er trinvist opbygget, 'step-up-procedure', således at faktorerne en for en blev testet i modellen. Hvis en faktor viste sig at have en signifikant indflydelse på *egentlig*s varighed, forblev den i modellen, hvis ikke, blev den ladet ude af betragtning i den videre modelopbygning. Hver signifikant faktor blev testet i alternative modeller for tovejsinteraktioner med de faktorer der i forvejen var i modellen. Hvis en interaktion viste sig signifikant, blev den alternative model som interaktionen indgik i, sammenlignet med 'grundmodellen' og eventuelt andre alternative modeller vha. funktionen ANOVA. Hvis en mere kompleks model, dvs. en model med en interaktion, ikke viste sig signifikant stærkere end en mere simpel model (fx uden interaktion), blev den mest komplekse model forkastet ifølge princippet om simpelhed (Crawley 2007: 325).

for, men har ikke været stærke nok til at kunne indgå i den endelige model som enkeltfaktorer eller faktorer i interaktioner. Disse er Rolle, ADV_egentlig (dvs. samforekomst med et foranstående adverbial, fx i *han er jo egentlig god nok*), Jamen-kontekst, placering i Hovedsætning versus Ledsætning samt de faktorer der vedrører informationsstruktur: Topikalisering, Sætningskløvning, Sætningsknode og Fundamentfeltfordobling. Tilbage som signifikante faktorer i den bedste mulige model står de faktorer der præsenteres i tabel 3.

TABEL 3: MODEL FOR VARIATION I EGENTLIGS VARIGHED (N=900); TILFÆLDIG FAKTOR: TALER (N=62)

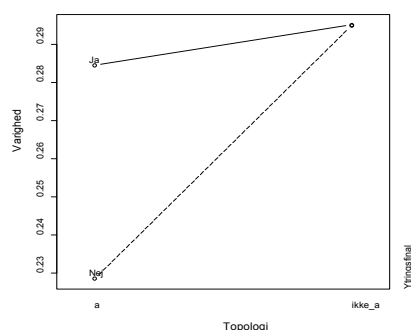
	Estimat	Standard- fejl	t-værdi	Signifikansni- veau
(Skæringspunkt)	0,304	0,011	28,395	<2e-16***
Topologi: Ikke_Lille_a	0,066	0,007	9,755	<2e-16***
Ord per sekund	-0,017	0,002	-9,468	<2e-16***
Ytringsfinal: Ja	0,056	0,014	3,866	0,0001***
Generation: Ældre	0,031	0,008	3,634	0,0007***
Øh_kontekst: Ja	0,025	0,008	3,191	0,0015**
Køn: Mand	-0,021	0,008	-2,707	0,0092**
Spørgsmål: Ja	0,015	0,009	1,628	0,1039
egentlig_ADV: Ja	-0,003	0,005	-0,672	0,5018
Topologi: Ikke_Lille_a * Ytringsfinal: Ja	-0,056	0,018	-3,150	0,0017**
egentlig_ADV: Ja * Øh_kontekst: Ja	-0,040	0,013	-3,139	0,0018**
Generation: Ældre * Spørgsmål: Ja	-0,051	0,021	-2,488	0,0130*

I aflæsningen af modellen skal vi først og fremmest lægge mærke til tallene der angiver værdierne for estimaterne og signifikansniveauerne (stjernerne ud for signifikanstallene angiver grad af signifikans). Estimaterne for hver faktor og hvert faktorniveau viser os hvor meget den enkelte faktor påvirker *egentlig*s varighed og i hvilken retning. Værdierne for standardfejl giver ikke anledning til bekymring, og t-værdierne er de værdier som signifikansniveauet er beregnet ud fra.

Der er seks enkeltfaktorer og tre interaktioner der på signifikant vis påvirker variationen i *egentlig*s varighed. Fire af de seks enkeltfaktorer indgår også i interaktionerne og vil blive omtalt nedenfor med udgangspunkt i figur 3–5. For faktoren Ord per sekund fortæller modellen at jo flere ord der siges per sekund, altså jo hurtigere taletempo, jo kortere bliver *egentlig* (jf. minustegnet foran estimatværdien). For faktoren Køn ses det at *egentlig* er kortere når det bliver sagt af mænd end af kvinder.

Faktorerne Topologi og Ytringsfinals påvirkning af *egentlig*s varighed illustreres i figur 3.

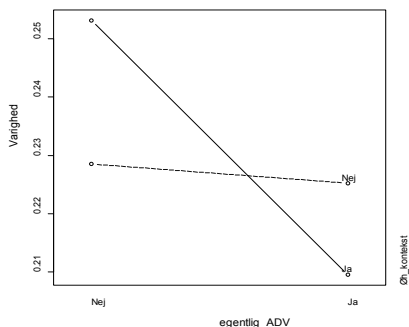
FIGUR 3: PLOT AF INTERAKTIONEN TOPOLOGI * YTRINGSFINAL



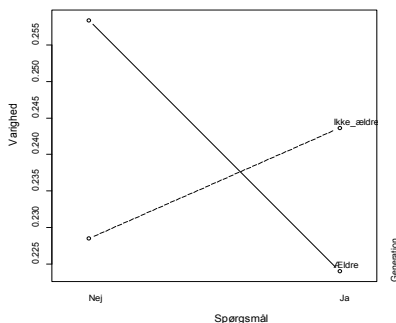
I plottet ser vi at når *egentlig* står på Lille-a, er det væsentligt kortere når det *ikke* er ytringsfinalt end når det er (linjerne forbinder blot fikspunkterne, de angiver ikke forekomst på en skala). Optræder det både i kategorien Lille-a og er ytringsfinalt, er det længere ("Ja"-punktet inde i grafen) end hvis det ikke er ytringsfinalt ("Nej"-punktet inde i grafen). Ytringsfinal position giver altså en længere varighed (for kategorien Lille-a). Når *egentlig* ikke optræder ytringsfinalt, er ordet signifikant kortere når det optræder på Lille-a end når det ikke gør. Med andre ord: Den ekspanderende effekt som faktoren Ytringsfinal har, kan altså ikke spores når *egentlig* optræder på anden topologisk plads end Lille-a, for der har *egentlig* under alle omstændigheder relativt lang varighed (jf. tabel 2 og tabel 3).

Figur 4–5 illustrerer faktorerne *egentlig*_ADV og Øh_konteksts henholdsvis Spørgsmål og Generations indflydelse på *egentlig*s varighed.

FIGUR 4: PLOT AF INTERAKTIONEN
EGENTLIG_ADV * ØH_KONTEKST



FIGUR 5: PLOT AF INTERAKTIONEN
SPØRGSMÅL * GENERATION



I kontekster uden tøvesignaler er der ingen forskel på om der forekommer et adverbial efter *egentlig* eller ej. Men i tøvekontekst er der signifikant forskel på om der er et efterfølgende adverbial eller ej. Hvis der er, udtales *egentlig* hurtigere end hvis der ikke er. Når *egentlig* efterfølges af et adverbium, er det kortere når det står i nærheden af en tøve, end når det ikke gør ($p < 0,01$). Dette er et svært forklarligt resultat: I nærheden af en tøve markør bliver *egentlig* længere når der ikke følger et adverbium, men markant kortere når der gør. Man kunne være fristet til at sige at den forkortende faktor kollokationen *egentlig_ADV* så dominerer over den forlængende faktor *Øh_kontekst*, men uden for *Øh_kontekst* ser vi *ikke* en signifikant påvirkning af faktoren *egentlig_ADV* (jf. den relativt flade stiplede linje i figur 4 og den lave estimatværdi 0,003 i tabel 3). Det er ikke indlysende hvorfor det forholder sig sådan, og disse faktorer bør undersøges nærmere i fremtidige studier.

Af figur 5 ser vi at der er en krydseffekt: Ældre forkorter *egentlig* i spørgsmål, men ikke i andre typer sætninger, mens yngre og midaldrende gør det omvendte, nemlig forkorter *egentlig* i fremsættende sætninger, men ikke i spørgsmål. I det analyserede datasæt gør ældre talere altså forskel på varighed af *egentlig* i henhold til sætningstype. Også dette forhold bør undersøges nærmere i fremtidige studier.¹⁸

18 Der er 72 spørgsmål i datasættet, hvad der ikke er faretruende lavt. Men det faktum at kun 15 af dem er ytret af ældre, mens 57 er ytret af ikkeældre, bør give anledning til forsigtighed i fortolkningen.

9 PERSPEKTIVER

9.1 Ikke sproglige faktorer

Vi så i analysen at der var signifikante forskelle inden for de traditionelle sociolingvistiske kategorier Køn og Generation. Enkelte andre, internationale studier af fonetisk reduktion har rapporteret lignende tendenser (Ernestus & Warner 2011; Keune m.fl. 2005), om end de traditionelle sociolingvistiske kategorier ikke synes at være de højst-prioriterede undersøgelsesfaktorer i udforskningen af fonetisk reduktion. Schachtenhaufen (2013: 165) finder ingen effekt overhovedet af traditionelle sociolingvistiske kategorier.¹⁹

At ældre talere taler langsommere, som vist i fremsættende sætninger, eller har mindre fonetisk reduktion, er ikke overraskende ud fra den almene opfattelse. Det er en udbredt antagelse at unge taler hurtigere, og mere 'uartikuleret' eller 'sjusket', og undersøgelsen her understøtter til en vis grad denne opfattelse. Årsagen til denne tilsyneladende generationsforskel kan måske findes i ikke sproglige faktorer, fx i fysiologien, at taleorganerne eller forbindelseslinjerne mellem hjernen og taleorganerne forandres eller sløves, eller i erfaring og afdæmpethed, at man som ældre får en mere rolig eller tilbagelænet personlighed. Der kan også være tale om en lydudvikling hvor den yngre generation taler et dansk der for visse ords og morfemers vedkommende har svækkede eller helt udeladte fonetiske segmenter, et ikke ukendt udviklingstræk i dansk sproghistorie. Dette sidste perspektiv kan med fordel undersøges ved inddragelse af ældre optagelser, fx optagelserne fra senfirserne i LANCHART-korpusset med denne undersøgelses interviewede i Ældre- og Mellem-kategorierne. Disse betragtninger bør dog tages med nogen forbehold; undersøgelsen viste jo også at ældre talere var hurtigere om at udtale *egentlig* i spørgsmål.

Hvad angår mænd som hurtigere talere end kvinder, er det interessant at Heegård (2013) i sin undersøgelse af vokaltab, schwaapokope, i verbale og adjektiviske *te*-endelser finder at mænd oftere har schwa-reduktion end kvinder. Og Keune m.fl. (2005) viser ligeledes at mænd oftere har fonetisk reduktion end kvinder. Årsagen er dunkel. Mange

19 Derimod finder Schachtenhaufen (2013: 165–168) tendens til at talesituation, monolog vs. dialog, og rolle i DanPASS-optagelser, instruerende over for adlydende i kort-rute-beskrivelserne, kan have en effekt for graden af fonetisk reduktion for den individuelle taler.

sociolingvistiske undersøgelser peger på at det er kvinder der fører an i sprogudviklingen fra ikkestandardsprog til standardsprog, og at mænd oftere end kvinder bruger en stigmatiseret variant (Labov 2001: 265 ff; Jensen & Maegaard 2012: 186–187, 190). Hvis en fonetisk reduceret form betragtes som ikkestandard og en ekspanderet form som standard, antyder denne undersøgelse et kønsdifferentieret mønster der ligner det som er fundet i sociolingvistiske undersøgelser: Kvinder taler mere standardnært end mænd. Sammenhænge mellem kønsforskelle, fonetisk reduktion og sprogudvikling er dog endnu et område der påkalder sig mere belysning.

9.2 Signifikante sproglige faktorer

Det var forventeligt at Ord per sekund ville komme ud som en signifikant faktor. Man kan ligefrem sige at det interessante ved undersøgelsen er at denne faktor ikke er den eneste signifikante faktor, men at der trods den åbenlyse sammenhæng mellem *egentlig*s varighed og talehastighed målt på et større stræk 'er plads' til at andre faktorer viser sig som relevante. Der er imidlertid flere måder at måle talehastighed på, ligesom der kan være flere måder at måle fonetisk reduktion på, og det er ikke utænkeligt at inklusionen af andre, mere forfinede talehastighedsmål vil kunne resultere i andre signifikante modeller end den her præsenterede. Jeg har imidlertid fundet det uden for projektets omfang at udforske hvilke typer talehastighedsmål der bedst vil kunne beskrive variationen i talehastighed. Opgaven med at udvikle differentierede mål for talehastighed er dog for undertegnede at se en vigtig fremtidig opgave for den danske fonetikforskning der vil undersøge reduktionsfænomener i naturlig og spontan tale.

At *egentlig* bliver længere når det står (absolut) ytringsfinalt end når det ikke gør, er heller ikke en overraskende erkendelse. Ytringsfinalt er i denne undersøgelse defineret som både sætningsfinal og foran en pause i talestrømmen, og foran en pause finder både Schachtenhaufen (2013: 182–184, ff), Heegård (2013: 110–115) og Heegård & Thøgersen (2012, 2014) markant eller signifikant mindre grad af reduktion. Årsagen må skyldes aspekter ved taleproduktionen – i ens talestrøm er der tendens til at sætte taletempoet ned sidst i en ytring (der ofte efterfølges af en pause), jf. fx Tøndering (2008).

En anden faktor der har at gøre med produktionsprocessen, er Øh_kontekst, det at *egentlig* (uden et andet efterfølgende adverbial) bliver længere når det står i nærheden af tøven, transskriberet i korpusset med ”øh”. Forklaringen må være at når vi tøver, sætter vi også talehastigheden ned, dvs. afvikler færre ord og stavelser per sekund. Pudsigt nok finder undersøgelsen ikke nogen signifikant sammenhæng mellem forekomster i nærheden af ”øh” og en nedsat talehastighed. Dette kan skyldes at målet for talehastigheden er for groft, for ikkelokalt, så at sige, eller det kan skyldes at den kollokationelle faktor *egentlig*_ADV også indvirker på *egentlig*s varighed i denne position. Et større datasæt eller supplerende undersøgelser af disse faktorerers indflydelse på fonetisk reduktion vil kunne belyse dette komplekse samspil.

9.3 Implikationer af topologiens indvirken på *egentlig*s varighed

Denne undersøgelses måske væsentligste resultat er påpegningen af sammenhængen mellem *egentlig*s varighed og topologi, dvs. syntaktisk placering i henhold til sætningsskemaet for dansk. Undersøgelsen af *egentlig* bekræfter Heegårds (2013a) og Heegård & Mortensens (2014) observationer, at *egentlig*s ’søsteradverbial’ *faktisk* (jf. Jensen 2006 og Mortensen 2010) er signifikant kortere (og segmentelt mere reduceret) når det står på pladsen Lille-a end forrest eller sidst i sætningsskemaet eller uden for det centrale sætningsskema (fx på Venstre-Ekstraposition, Højre-Ekstraposition eller interjektionalpladserne). Også variationen inden for de to kategorier Lille-a og Ikke-Lille-a er identisk med tidligere undersøgelser. Dette forhold peger på en sammenhæng mellem syntaktisk-pragmatisk funktion og fonetisk realisation, i hvert fald hvad angår visse adverbialer, en sammenhæng der så vidt undertegnede er orienteret, hverken er dokumenteret på et større empirisk grundlag eller nøjere beskrevet i litteraturen, den danske såvel som den internationale. De følgende kommentarer skal således læses som foreløbige tanker om måder at beskrive dette forhold på og som et muligt afsæt til videre udforskning.

En fortolkning af resultaterne tager udgangspunkt i hypotesen at længere varighed (mindre fonetisk reduktion) er en måde at fremhæve information på i en sætning, og omvendt at kortere varighed (fonetisk reduktion) er en måde ikke at fremhæve information på. Undersøgel-

sen peger på at især pladserne i den topologiske kategori Ikke-Lille-a er pladser der typisk bruges til en sådan optoning, men også, med mindre andre faktorer spiller ind, at placering på Lille-a-pladsen kan bruges til dette (jf. at *egentlig* sagtens kan have lang varighed på Lille a).

En anden fortolkning tager udgangspunkt i Hansen og Heltofts (2011: 1042–1044) inddeling i typer af sætningsadverbialer. Af deres eksempler fremgår det at *egentlig* på Lille-a kan regnes som et vidensadverbial, der angiver grad af talers sikkerhed for ytringen, ”fra sikkerhed over sandsynlighed og mulighed til det lidet sandsynlige” (s. 1042). *Egentlig* i Fundamentfeltet, derimod, er et repræsentativt adverbial der har den funktion at markere engagement og distance. Sammenholdt med nærværende undersøgelse kan man sige at funktionen som repræsentativt adverbial har ’leksikaliseret’ sig i en lang varighed i den fonetiske realisation, hvorimod funktionen som vidensadverbial, på Lille-a, tillader variation i den fonetiske realisation; eller, med andre ord, at Lille-a kan rumme *egentlig* som både vidensadverbial og repræsentativt adverbial. Kun tre ud af 21 forekomster af *egentlig* til venstre for nexus, fx i Fundamentfeltet, har en varighed under gennemsnittet for alle forekomster. Dette er til dels en bekræftelse af oplysningen hos Brink m.fl. (1991) hvor vi får at vide at adverbiet *egentlig* ikke ”ofte” reduceres på ”sætningens førsteplads” (s. 385). (Det er ikke klart fra Hansen & Heltoft (2011) hvordan *egentlig* skal kategoriseres med hensyn til de øvrige topologiske pladser). Sammenhængen mellem ’leksikalisering’ og fonetisk reduktion er også påpeget i internationale undersøgelser af homografiske ordforbindelser, fx af engelsk *you know* som subjekt og verbum over for *you know* som efterstillet diskurspartikel (fx stavet *y’know*), jf. Schubotz m.fl. 2015 (se også Bell m.fl. 2009 og henvisninger heri).

En tredje fortolkning tager udgangspunkt i *egentlig*s betydning og funktion som beskrevet i DDO (jf. afsnit 2.1). Som nævnt har det ikke været muligt entydigt at bestemme hver en forekomst i datasættet ud fra DDO’s tre betydninger, nogle eksempler kan have et vist element af ’i virkeligheden’- eller ’ret beset’-betydning (= den mere konkrete betydning 2) og samtidig også et element af indrømmelse og nedtoning (= den generaliserede betydning 3). Desuden kræver en præcis indsnævring af *egentlig*s betydning meget ofte en analyse af konteksten,

dvs. samtalepartnerens bidrag op til brugen af *egentlig* og undertiden også efterfølgende ytringer, og dette gøremål ville have haft sprængt rammerne for denne undersøgelse. Forsøget på semantisk kategorisering med udgangspunkt i DDO har dog peget på visse tendenser. På de topologiske pladser Fundament, Venstre-Ekstraposition og Venstre-Interjektional finder vi som oftest den mere konkretiserede betydning, fx i (19) hvor 'ret beset'-elementet er kraftigt, og hvor konteksten ikke lægger op til en nedtonende læsning, eller til at taleren forsøger at indrømme et sagsforhold (varighed 3,111 sek.).

- (19) altså de bliver nogle andre personligheder ikke ja **egentlig** er (ID4)
det jo ganske forfærdeligt alt det der alzheimer

Eksemplet i (20) viser en brug af *egentlig* som også svarer til DDO's betydning 2. Også denne betydning er pænt præsenteret blandt de 34 forekomster på en ikkedefaultplads, i eksemplet som modifikator til et adjektiv på N-pladsen. Af eksemplets nærmeste kontekst er det heller ikke her oplagt med en nedtonende og indrømmende fortolkning (varighed 0,331 sek.).

- (20) de er sådan meget åbne og høflige og og **egentlig** positive øh (ID45)
i et i et stort stykke hen ad vejen ikke altså

Anderledes semantisk tvetydig er *egentlig* i højre del af sætningskemaet og på Lille-a. I eksempel (21), hvor taleren karakteriserer sine klassekammerater, og hvor *egentlig* står på Højre-Interjektional, og 'ret beset'/'i virkeligheden'-elementet er kraftigt til stede (varighed 0,354 sek.).

- (21) der er ikke nogen der skiller sig ud øh på en dårlig måde heller (ID45)
ikke rigtig på en god måde **egentlig**

Men i (22), et ja-nej-spørgsmål, er det det nedtonende element der er kraftigt til stede (varighed 0,287 sek.)

- (22) øh har du familie andre steder i landet **egentlig** (ID60)

Tilsvarende på Lille-a. I eksempel (23), fra interviewerens instruktion til informanten før en sprogkarakteristikprøve, er det 'ret beset'-elementet der er kraftigt til stede i begge *egentlig*'er (varighed 0,174 og 0,292, henholdsvis):

- (23) du behøver ikke tænke så meget over ... hvem det er og du (ID55)
behøver **egentlig** heller ikke at tænke over hvad de **egentlig**
snakker om

mens det i eksempel (24) ret oplagt (også) er det indrømmende element der er kraftigt til stede, idet intervieweren erkender at informanten tidligere har givet information der besvarer det spørgsmål intervieweren lige har stillet (varighed 0,142):

- (24) men det har du **egentlig** også fortalt om (ID55)

Denne selektive eksemplificering af *egentlig*s betydning (i henhold til DDO) og topologisk placering har peget på en tendens til en delvis sammenhæng mellem disse to faktorer. Pladserne først i sætningsskemaet (Venstre Interjektional, Venstre-Ekstraposition og Fundament) og ikkedefaultpladserne synes oftest at optræde med en mere konkret betydning af *egentlig*. Eller: En mere konkret, ikkegeneraliseret betydning af *egentlig* udtrykkes gerne (men ikke udelukkende) på disse pladser. De samme pladser kan derfor siges at bidrage til en særlig kraftig pragmatisk effekt. Pladserne før neksus netop fordi de er de første pladser i ytringen og derved er kraftigt medkonstituerende for hvordan ytringen skal opfattes, ikkedefaultplaceringen fordi den netop ikke er default og derfor påfaldende når den er til stede.

Den mere konkrete betydning er tilsyneladende ikke lige så gerne analog med placering på Lille-a og til højre i sætningsskemaet. Her er *egentlig* oftere tvetydig semantisk set, eller den nedtonende, indrømmende funktion er oftere fremhævet. Disse pladser lader det således i højere grad være op til et pragmatisk parameter hvordan *egentlig* skal forstås. Denne analyse af sammenhængen mellem fonetisk reduktion, topologisk placering og pragmatisk funktion lægger sig tæt op ad Eva Skafte Jensens analyse af visse adverbialers funktion og mulighed for placering i sætningsskemaet. I Jensens analyse af *altså*s udvikling re-

degøres der for at adverbialet går fra at have en konkret og objektiv betydning, der kan udtrykkes ved placering både i Fundamentfeltet og i Lille-a, til en mere abstrakt betydning der ”bruges af afsender som subjektivt signal om engagement i det ytre”, og som kun kan udtrykkes på Lille-a (Jensen 2000: 62-77).

10 KONKLUSION

Artiklen har præsenteret resultaterne af en større kvantitativ undersøgelse af adverbialet *egentlig*s varighed. Undersøgelsen har overordnet vist at der er en stor spredning i *egentlig*s varighed, at varigheden, som ventet, til en vis grad er afhængig af talehastigheden, og at der er stor talerspecifik variation med hensyn til varighed: Nogle talere har oftere eller kun korterevarende udtaler, andre oftere eller fortrinsvis længerevarende udtaler. Undersøgelsen viste tillige at der er stor spredning i varighed afhængig af hvor i sætningsskemaet *egentlig* er placeret. I en nærmere statistisk undersøgelse der inddrog disse og andre sproglige og ikkesproglige faktorer, og som tog højde for den påvirkning den enkelte taler har på *egentlig*s varighed, viste det sig at især topologi, talehastighed, placering sidst i sætningen (før pause), generation, placering i en tvekontekst og talers køn har en indvirken på *egentlig*s varighed. Nogle af disse faktorerers signifikante indvirken på fonetisk reduktion er også dokumenteret i andre undersøgelser. Andre, fx topologi, er ikke.

Både den anvendte metode og resultaterne peger i retninger der i den videre forskning vil kunne belyse forholdet mellem fonetisk reduktion, topologi og pragmatisk funktion. Med den form for statistik der vurderer forskellige faktorerers indflydelse i forhold til hinanden, og som kan tage hensyn til individers eventuelle særlige udtalevaner, vil man kunne lade de faktorer ude af betragtning der ikke har (nok) indflydelse på, i dette tilfælde, et givent ords mulighed for ekspansion henholdsvis reduktion. Artiklen anviser således en måde hvorpå man kan udrede hvilke faktorer eller kombinationer af faktorer der for et givent genstandsfelt påvirker i hvilken retning.

Sammenhængen mellem adverbialers funktion og topologi er påpeget af andre, men denne artikel viser relevansen af at inddrage endnu en dimension i udtrykssystemet, nemlig lydsiden. For det første anslår

undersøgelsens resultater at ikke kun tryk men også spændet mellem fonetisk reduktion og ekspansion har syntaktisk relevans, foruden indholdsmæssig, pragmatisk relevans. For det andet vil det være et interessant spor at følge fremover i hvilken grad forskydninger i adverbialers semantik og brug korrelerer med udtalevariation i et synkront snit og udtaleforandring over i et diakront blik. Dette er et forskningsfelt der fortjener større interesse, både hjemme og i udlandet.

TAK

Arbejdet med undersøgelsen og artiklen er kommet i stand gennem en forskningsbevilling fra Carlsbergfondet. En stor tak til Torben Juel Jensen, Jacob Thøgersen, Nicolai Pharaos og Tanya Karoli Christensen for råd og vejledning angående den statistiske undersøgelse. Også en stor tak til deltagerne ved Faglig Frokost på Sprogforandringscentret på Københavns Universitet for konstruktiv feedback, og en tak til Ida Elisabeth Mørch, en anonym fagfællebedømmer og to NyS-redaktører for kommentarer til en tidligere version af artiklen, samt til Tore Kristiansen for sociolingvistisk perspektiv. Endelig, en meget, meget stor tak til Gert Foget Hansen for udarbejdelse af script til undersøgelse af talehastighed. Jeg alene hæfter naturligvis for fejl og mangler i artiklen.

Jan Heegård

Institut for Nordiske Studier og Sprogvidenskab

Københavns Universitet

janhp@hum.ku.dk

LITTERATUR

- Adda-Decker, Martine, Philippe Boula de Maréuil, Gilles Adda & Lori Lamel. 2005. Investigating syllabic structures and their variation in spontaneous French. *Speech Communication* 46(2). 119–139.
- Andersen, Torben. 1982. Modalpartikler og deres funktion i dansk. *Danske Studier* 86–95.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2014. *lme4: Linear Mixed-Effects Models Using Eigen and S4*. R Package Version 1.1-6. <http://CRAN.R-project.org/package=lme4>.
- Bell, Alan, Dan Jurafsky, Eric Fosler-Lusier, Cynthia Girand, Michelle Gregory & Daniel Gildea. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113(2). 1001–1024.
- Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand & Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1). 92–111.
- Boersma, Paul & David Weenink. 2015. *Praat: Doing Phonetics by Computer. Version 5.4.15*. <http://www.praat.org/>.
- Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88(1). 1–44.
- Brink, Lars. 2013. *Phonetic reduction*. Jan Heegård & Peter Juel Henriksen (red.), *New Perspectives on Speech in Action* (Copenhagen Studies in Language 43). 13–22. Frederiksberg: Samfundslitteratur Press.
- Brink, Lars, Jørn Lund, Steffen Heger & Jens Normann Jørgensen. 1991. *Den Store Danske Udtaleordbog*. København: Munksgaard.
- Bush, Nathan. 2001. Frequency effects and word-boundary palatalization in English. Joan L. Bybee & Paul J. Hopper (red.), *Frequency and the Emergence of Linguistic Structure*. 255–280. Amsterdam: John Benjamins Publishing Company.
- Crawley, Michael J. 2007. *The R Book*. Hoboken, NJ: John Wiley & Sons.
- Christensen, Tanya Karoli & Torben Juel Jensen (2014). Sætningsadverbialer i ledsætninger som forgrundssignal. *Ny forskning i grammatik* 21. 37–53.
- Den Danske Ordbog*. <http://ordnet.dk/ddo/>.
- Ernestus, Mirjam & Natasha Warner. 2011. An introduction to reduced pronunciation variants. *Journal of Phonetics, Special Issue, Speech Reduction* 39(3). 253–260.

- Fosler-Lussier, Eric & N. Morgan. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech and Communication* 29(2). 137–158.
- Gahl, Susanne. 2008. 'Thyme' and 'time' are not homophones. Word durations in spontaneous speech. *Language* 84(3). 474–496.
- Gregersen, Frans. 2009. The data and design of the LANCHART study. *Acta Linguistica Hafniensia* 41(1). 3–29.
- Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier & Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS 35. The Panels*. 151–166.
- Grønnum, Nina. 2005. *Fonetik og fonologi, almen og dansk*. København: Akademisk.
- Grønnum, Nina. 2009. A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Communication* 51. 594–603.
- Hanique, Iris & Mirjam Ernestus. 2011. Final /t/ reduction in Dutch past-participles: The role of word predictability and morphological decomposability. *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), Florence, Italy*. 2849–2852.
- Hanique, Iris, Barbara Schuppler & Mirjam Ernestus. 2010. Morphological and predictability effects on schwa reduction: The case of Dutch word-initial syllables. *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan*. 933–936.
- Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog*. København: Det Danske Sprog- og Litteraturselskab.
- Heegård, Jan. 2012. Funktionelt sjusk. Endelserne *-te* og *-ede* i dansk. *Danske Talesprog* 12. 34–61.
- Heegård, Jan. 2013a. Functional phonetic reduction? The case of the Danish adverb *faktisk* ('actually'). Jan Heegård & Peter Juel Henriksen (red.), *New Perspectives on Speech in Action. Proceedings from the 2nd SJUSK Conference on Contemporary Speech Habits* (Copenhagen Studies in Language 43). 59–78. Frederiksberg: Samfundslitteratur Press.
- Heegård, Jan. 2013b. Morphologisation or reduction by context? The *-te* ending on adjectives and preterite verb forms in Standard Copenhagen Danish. *Acta Linguistica Hafniensia* 45(1). 100–125.
- Heegård, Jan & Janus Mortensen. 2014. Fonetisk reduktion og kommunikative kontraster. Tilfældet 'faktisk'. *Ny forskning i grammatik* 21. 87–102. Odense: Institut for Sprog og Kommunikation, Syddansk Universitet.

- Heegård, Jan & Jacob Thøgersen. 2012. 'Her er Pressens Radioavis' – med det minimalt acceptable sjuskedansk. *Danske Talesprog* 12. 62–95.
- Heegård, Jan & Jacob Thøgersen. 2014. Style-dependent rules for reduction: Mixed-model multiple linear regression analysis of reduction phenomena in Danish news readings. *Nordic Journal of Linguistics* 37(3). 325–365.
- Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Jensen, Eva Skaft. 2000. *Danske sætningsadverbialer og topologi i diakron belysning*. Ph.d.-afhandling. Københavns Universitet.
- Jensen, Eva Skaft. 2006. 'Egentlig' og 'faktisk' og to slags spørgsmål polyfoniteoretisk anskuet. Rita Therkelsen, Nina Møller Andersen & Henning Nølle (red.), *Sproglig Polyfoni. Arbejdspapirer 6, Marts 2006*. 27–52.
- Jensen, Torben Juel & Tanya Karoli Christensen. 2013. Promoting the demoted: The distribution and semantics of 'main clause word order' in spoken Danish complement clauses. *Lingua* 137: 38–58.
- Jurafsky, Daniel, Alan Bell & Cynthia Girand. 2002. The role of the lemma in form variation. Carlos Gussenhoven & Natasha Warner (red.), *Papers in Laboratory Phonology VII*. 1–34. New York: Mouton de Gruyter.
- Jurafsky, Daniel, Alan Bell, Michelle Gregory & W.D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. Joan L. Bybee & Paul Hopper (red.), *Frequency and the Emergence of Linguistic Structure*. 229–254. Amsterdam: Benjamins.
- Keating, Patricia. 1985. Universal phonetics and the organization of grammars. Victoria A. Fromkin (red.), *Phonetic Linguistics*. 115–132. New York: Academic Press.
- Keune, Karen, Mirjam Ernestus, Roeland van Hout & R.H. Baayen. 2005. Variation in Dutch: From written MOGELIJK to spoken MOK. *Corpus Linguistics and Linguistic Theory* 1(2). 183–223.
- Labov, William. 2001. *Principles of Linguistic Change. Social Factors*. Malden: Blackwell Publishers.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.
- Mortensen, Janus. 2010. *Epistemic stance marking in the use of English as a lingua franca*. Ph.d.-afhandling. Roskilde Universitet.
- Pharao, Nicolai. 2009. *Consonant Reduction in Copenhagen Danish. A Study of Linguistic and Extra-Linguistic Factors in Phonetic Variation and Change*. Ph.d.-afhandling. Københavns Universitet.
- Pharao, Nicolai. 2010. Ordbrug og udtaleforandringer. *NyS* 39. 137–165.

- Plug, Leendert. 2005. From words to actions: The phonetics of *eigenlijk* in two communicative contexts. *Phonetica* 62. 131–45.
- Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118(4). 2561–2569.
- Rasinger, Sebastian M. 2013. *Quantitative Research in Linguistics. An Introduction*. London: Bloomsbury.
- Rischel, Jørgen. 1992. Formal linguistics and real speech. *Speech Communication* 11. 379–392.
- Schachtenhaufen, Ruben. 2010. Looking for lost syllables in Danish spontaneous speech. Peter Juel Henriksen (red.), *Linguistic Theory and Raw Sound* (Copenhagen Studies in Language 40). 61–88. Frederiksberg: Samfundslitteratur.
- Schachtenhaufen, Ruben. 2012. Reduction of word final schwa in Danish – Rhythmic and syntactic factors. Jan Heegård & Peter Juel Henriksen (red.), *Speech in Action. Proceedings of the 1st SJUSK Conference on Contemporary Speech Habits* (Copenhagen Studies in Language 42). 53–71. Frederiksberg: Samfundslitteratur Press.
- Schachtenhaufen, Ruben. 2013. *Fonetisk reduktion i dansk*. Ph.d.-afhandling. Copenhagen Business School.
- Schubotz, Louise, Nelleke Oostdijk & Mirjam Ernestus. 2015. *Y'know vs you know*: What phonetic reduction can tell us about pragmatic function. Sander Lestrade, Peter de Swart & Lotte Hogeweg (red.) *Addenda. Artikelen voor Ad Foolen*. Nijmegen: Radboud University. 361–380.
- Schuppler, Barbara, Wim A. van Dommelen, Jacques Koreman & Mirjam Ernestus. 2012. How linguistic and probabilistic properties of a word affect the realization of its final /p/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40(5). 595–607.
- Tønndering, John. 2008. *Skitser af prosodi i spontant dansk*. Ph.d.-afhandling. Københavns Universitet.
- Traugott, Elizabeth Closs. 1980. Meaning-change in the development of grammatical markers. *Language Sciences* 2(1). 44–61.
- Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. *Berkeley Linguistics Society: Proceedings of the Fourteenth Annual Meeting, February 13-15, 1988: General Session and Parasession on Grammaticalization*, red. Shelley Axmaker & Helen Singmaster. 406–16.
- Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalisation. Dieter Stein & Susan Wright (red.), *Subjectivity and Subjectivisation. Linguistic Perspectives*. 31–54

- Cambridge: Cambridge University Press.
- Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in Semantic Change* (Cambridge Studies in Linguistics 97). Cambridge: Cambridge University Press.
- Warnes, Gregory R., Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber, Andy Liaw & Thomas Lumley 2014. *Gplots: Various R Programming Tools for Plotting Data*. R Package Version 2.13.0. <http://CRAN.R-project.org/package=gplots>.
- Winther Balling, Laura. 2008. A Brief introduction to regression designs and mixed-effects modelling by a recent convert. Peter Juel Henriksen (red.), *Looking at Eyes. Eye-Tracking Studies of Reading og Translation Processing* (Copenhagen Studies in Language 36). 175–192. København: Samfundslitteratur.
- Zwicky, Arnold. 1972. On Casual Speech. *CLS* 8. 607–115.