

# Flertalsformer af *ari*-ord i den færøske talesprogsbank<sup>1</sup>

Jógvan í Lon Jacobsen

*Fróðskaparsetur Føroya*

## Abstract

The aim of this article is to investigate dialectal variation of plural endings of *ari*-words in Faroese, i.e., masculine words with *ari*-ending in singular. Such words for example *lærari* ‘teacher’ may get different plural endings in spoken Faroese: *ar*, *ir*, *a*, and *R* (*lærarar*, *lærarir*, *lærara*, *læraR*<sup>2</sup>). In the written language there is only one correct plural form which is *ar*: *lærarar*. The empirical material in this article is picked up from a corpus of spoken language, *Føroyskur talumálsbanki*, which is a corpus management and analysis system for annotated corpora. The article is also a study of the usability of the corpus concerning dialectal variation in spoken Faroese. The result of the correlation shows that the non-standardized *ir*-variant is most frequent in the corpus. Here I investigate the variation by correlating them with two non-linguistic variables, place, and age.

Keywords: nomina agentis, sociolinguistics, dialectology, Faroese

## 1. Formål

Formålet med denne artikel er todelt. Dels undersøger jeg variationen af flertalsendelser i ord, som ender på *ari* i ental (ofte omtalt som nomina agentis), fx *lærarar* ‘lærere’ i ubestemt form i nominativ og akkusativ<sup>3</sup> i forhold til to sociale variabler, alder og sted. Dels vurderer jeg, i hvor høj grad datagrundlaget i talesprogsbanken giver et tilfredsstillende billede af denne variation. Den aldersmæssige variation, som bliver omtalt i artiklen, er eksempel på det, man i sociolingvistikken kalder for tilsyneladende tid, som er en metodisk tilgang, hvor man undersøger sprogændring ved at sammenligne folks tale i forskellige aldre. For så vidt sprogændringer foregår, går man ud fra, at de ældre generationer anvender ældre sprogformer, mens de yngre i højere grad anvender nyere former.

## 2. Strukturering af artiklen

Før jeg besvarer spørgsmålet om variation i flertalsendelser, er det nødvendigt først at sige noget om det andet spørgsmål: det sproglige materiale i talesprogsbanken. På grund af det vil rækkefølgen blive omvendt: Først omtales talesprogsbanken, og derefter vil variationen blive diskuteret.

Den følgende tekst er opdelt i otte afsnit (afsnit 3 til afsnit 10). Tredje afsnit drejer sig om *ari*- og *i*-afledninger i færøsk, hvorefter der i fjerde afsnit omtales tre centrale begreber i sociolingvistik, nemlig sprogvariation, variabler og varianter. I det femte afsnit behandles projektet, *Ændringer i færøske dialekter*, som ledte frem til oprettelsen af en færøsk talesprogsbank, som omtales i afsnit seks. I syvende afsnit argumenteres der for fordelene ved en standardiseret retskrivning i transskriptionerne. I afsnit otte omtales nogle udfordringer der er knyttet til talesprogsmateriale, hvorefter distributionen af de forskellige flertalsvarianter af *ari*-ord i talesprogsbanken analyseres i niende afsnit. Det tiende og sidste afsnit består af et sammendrag og diskussion om variantfordelingen og datamaterialet i talesprogsbanken.

<sup>1</sup> Artiklen bygger på en præsentation i forbindelse med 200 års dagen for V.U. Hammershaimbs fødsel (1819-1909).

<sup>2</sup> *R* = phonetic reduction of unstressed vowel in spontaneous speech.

<sup>3</sup> *ari*-ordene har samme form i nominativ og akkusativ i ubestemt flertal.



### 3. *ari*- og *i*-endelser

Nomina agentis refererer til personer, der aktivt udfører eller praktiserer noget, fx *lærari* 'person der underviser andre, især elever i en (folke)skole' og *bakari* 'person der beskæftiger sig erhvervsmæssigt med at bage brød, kager m.m.' (<https://ordnet.dk/ddo>). Disse ord er afledt af tilsvarende verber (*baka* 'bage' > *bakari* 'bager', *læra* 'lære' > *lærari* 'lærer'). Men mange nomina agentis savner det tilsvarende verb, fx *klokkari* 'klokker'. Indholdet af sådanne afledninger er ikke altid transparent. Eksempelvis er *klokkari* ikke en person, der laver klokker, men en, der bestrider erhvervet at ringe med klokkerne ved en kirke. Nogle nomina agentis er afledt af andre substantiver eller navne, f.eks. *átari* 'grovæder' af substantivet *át* 'spisning' og *B36'ari* 'tilhænger af fodboldklubben B36' afledt af navnet på fodboldklubben B36. Ingen tilsvarende verber eksisterer for disse ord. Petersen (2019: 79-80) opdeler nomen agentis i tre funktioner:

- 1) Den gørende (den der udfører en handling): *skrivarin sendi brævið avstað* 'sekretæren sendte brevet afsted'
- 2) Redskab til at udføre en handling med: *upptrekkjari* 'oplukker', *printari* 'printer'
- 3) Den oplevende (den der oplever noget): *lurtari* 'lytter'.<sup>4</sup>

Björn Hagström (1977: 47) siger således om nomina agentis i færøsk:

Många sådana ursprungliga nomina agentis saknar det motsvarande verbet i färöiskan, t. ex. *snikkari*, *skómakari*, *skraddari*. Dermed er förutsättningen given att uppfatta *-ari* som ett suffix för betecknande av yrke utan något verb ligger til grund.

Selvom *ari*-afledningsendelsen er aktiv i moderne færøsk, var der i en periode i 1970'erne en tendens til fra sprogpolitisk side at reducere brugen af denne endelse til fordel for den kortere *i*-endelse, muligvis inspireret af islandsk (jf. fx isl. *blóðgjafi*, fær. *blóðgevi* 'bloddonor' og isl. *sæðisgjafi*, fær. *sáðgevi* 'sæddonor', se [islex.arnastofnun.is](http://islex.arnastofnun.is)). *i*-endelsen er meget tydelig i en liste med geografiske navne med tilsvarende inkolentnavne, udgivet af de nordiske sprognævn i 1974 (jf. *Navne på stater* i litteraturlisten), fx *afgani* 'afganner' i stedet for *afganari* og *amerikani* 'amerikaner' i stedet for *amerikanari*. I oversigten over inkolentnavne på Sprogrådets hjemmeside er denne praksis ændret, således at mange af disse inkolentnavne nu har dobbeltformer, fx *afganari/afgani* og *amerikanari/amerikani* (malrad.fo/orðalistar/lond og tjóðir, 4. januar 2022).

Denne orddannelsesmåde med *i*-endelser i stedet for *ari* er i øvrigt velkendt i dannelsen af afløsningsord i færøsk. En optælling i en nyordsliste viser, at af tyve afløsningsord, som ender på *ari* og *i*, er der ni med *ari*- og elleve med *i*-afledning (Poulsen, 2004: 519-524). Eksempler på *ari* er: *ambætari* 'server', *atstøðari* 'assistent' og *blekkprentari* 'blækprinter'. Og af eksempler på *i* kan nævnes *ostskeri* 'ostehøvl', *floksherji* 'partisan', *hjáseti* 'bisidder', *sjókagi* 'søkkert', *skyni* 'detektor', *vevkagi* 'browser' og *vísi* 'cursor'.

Weyhe (2012a; 2012b; 2012c; 2015: 423) drøfter geografisk variation af flertalsformer af nomina agentis historisk og konkluderer, at *ir*-endelsen breder sig (Denne undersøgelse bekræfter rigtigheden i hans konklusion, især hvis man lægger *ir* og *R* sammen (jf. afsnit 9 nedenfor). Hagström (1977: 47-52) siger, at danske importord, som ender på *er* og *or* får *ari*-endelse i færøsk. Her kan jeg tilføje, at der også findes eksempler på, at danske importord med udlydende *or*, kan få *ur*-endelse, fx *motorur* (da. motor), *traktorur* (da. traktor), *radiatorur* (da. radiator). På den anden side er der også eksempler på, at ord, som refererer til personer, og som på dansk ender på *or*, får tilføjet *ur* i færøsk, fx *senatorur* 'senator' (mask.) og *faktorur* 'opsynsmand, forvalter'<sup>5</sup> (mask.), ikke *\*senatorari*, *\*faktorari* (jf. fx *doktari*, *professari*).

<sup>4</sup> Her drejer det sig kun om egentlige nomina agentis, og derfor er den semantiske kategori indbygget ikke medregnet, fx *kanadiari* 'kanadier'.

<sup>5</sup> Om betydningen af faktor se faktor 1.2. i *Ordbog over det danske sprog*, <https://ordnet.dk/ods>, 4. januar 2022.

#### 4. Sprogvariation, variabler og varianter

Da denne artikel er en deskriptiv sociolingvistisk undersøgelse af sprogvariation i færøsk talesprog, vil jeg kort berøre nogle centrale træk ved sociolingvistik. Formålet med sociolingvistiske variationsundersøgelser er at opnå større bevidsthed om sprogbrug i social sammenhæng. Det betyder, at foruden at undersøge en variants geografiske distribution, relateres sprogbrugen også til sprogksterne (sociale) variabler. Chambers og Trudgill (1990: 54) siger, at alle dialekter både er lokale og sociale: "All dialects are both regional and social, since all speakers have a social background as well as a regional location." Hudson (1999: 3) siger om forskellen mellem sociolingvistik og lingvistik, "that linguistics differs from sociolinguistics in taking account only of the *structure* of language, to the exclusion of the social contexts in which it is learned and used." En sociolingvistisk dialektundersøgelse siger derfor ikke blot noget om geografisk distribution (som traditionelle dialektundersøgelser gør), men afdækker samtidig i hvilken grad distributionen af en sproglig variabel kan knyttes an til sociale variabler, i dette tilfælde alder og sted.

I sociolingvistiske variationsundersøgelser arbejdes der med variabler og varianter. En variabel er en størrelse, som varierer, og varianterne er de forskellige former eller udtryk, som variabelen kan realiseres som. Den sproglige (sproginterne) variabel i denne undersøgelse er flertalsendelser af *ari*-ord (fx *lærarar* 'lærere') med varianterne *lærarar*, *lærarir*, *lærara* og *læraR*. Den sociale (sprogksterne) variabel er en størrelse, der ligger uden for selve sproget, fx alder, geografisk sted, køn osv.

En sociolingvistisk korrelationsstudie gør det muligt at undersøge, i hvilken grad der er sammenhæng mellem de sproglige og sociale variabler. Af pladshensyn har jeg begrænset mig til at undersøge betydningen af to sociale variabler, nemlig alder og sted. Stedsvariablen siger noget om distribution og frekvens af varianterne i de forskellige dialekter, og aldersvariablen afdækker, i hvor høj grad der er sammenhæng mellem distribution af de fire flertalsformer og informanternes alder. Aldersvariablen er naturligvis interessant i en undersøgelse i tilsyneladende tid. I analysen bliver den sproglige variabel korreleret med de sociale variabler for at undersøge, om der er nogen sammenhæng. Formålet er at afdække eventuelle mønstre i variationen. Når vi taler om mønstre i sociolingvistisk forstand, hentyder det til, at variationen af en sproglig variabel knyttes an til en social variabel og kan forklares ud fra den. For eksempel korrelerer udtalevariation ofte med alder.

Traditionelle dialektundersøgelser viser geografisk variation, fx at folk, som bor syd for Skopunar-fjørður, siger [gan̥ga] 'gå', mens folk nord for Skopunarfjørður siger [gen̥ga] (se kort 1). Dette faktum er naturligvis interessant, set i relation til geografien, men det siger intet om, i hvilket omfang sproglig variation forekommer internt i det pågældende dialektområde. Netop sproglig variation er tema i denne artikel, hvor der fokuseres på interindividuel variation i forhold til alder og geografi.

#### 5. Ændringer i færøske dialekter

I 2015 bevilgede Granskingarráð Føroya (det færøske forskningsråd) midler til et treårigt projekt, hvis formål var at undersøge ændringer i færøske dialekter. Projektet blev lavet i samarbejde mellem færøske og norske sprogforskere og var stedfæstet på Føroyamálsdeildin ved Fróðskaparsetur Føroya. Deltagerne i projektet var Hjalmar P. Petersen, Jógvan í Lon Jacobsen og Heðin Jákupsson, Fróðskaparsetur Føroya, og Helge Sandøy, Universitetet i Bergen, og Edit Bugge, Høgskulen på Vestlandet. Paul Meurer, sprogteknolog ved Universitetet i Bergen, havde udviklet programvaren *Corpuscle*, som den norske *Talebanken* var udformet i, og han oprettede også den færøske *Føroyskur talumálsbanki* som et parallelt korpus og lagde det færøske materiale ind i den (jf. afsnit 6 nedenfor). Trond Trosterud, professor i sprogteknologi ved Universitetet i Tromsø, gjorde en grammatisk tagger, som er knyttet til talesprogsbanken. Studerende og ansatte ved instituttet hjalp til med feltarbejdet.

Projektet, som denne artikel bygger på, havde titlen *Ændringer i færøske dialekter gennem to generationer – hvordan, hvor hurtigt og hvorfor?* Der var tale om en sociolingvistisk studie med det formål at få indsigt i relationen mellem samfundsændringer og sprogændringer. For at få den vinkel med var projektet bygget op omkring modsætningen centrum-periferi, og interviewpersonerne var udvalgt efter denne model. Vi udvalgte fem steder spredt over hele landet: Tórshavn, Vágur, Eysturoy (Eiði), Norðoyggjar (Klaksvík, Viðareiði) og Suðuroy (se kort 1 nedenfor). Ifølge dette koncept var Tórshavn tænkt som centrum i forhold til Vágur og Eiði, mens Klaksvík var tænkt som centrum i forhold til Viðareiði,

og Tvøroyri som centrum i forhold til andre bygder på Suðuroy. Denne model er inspireret fra projektet *Dialektendringsprosesser* ved Universitetet i Bergen, og vi syntes, det kunne være interessant at undersøge, om vi kunne se de samme tendenser, som man ser i Norge, nemlig at de såkaldte regionscentre påvirker de mindre dialektområder, som ligger tæt på disse centre (jf. Sandøy, 2008 og Sandøy, 2017). Dette spørgsmål drøftes ikke i denne artikel.



Kort 1: Kort over Færøerne (Kilde: Wikimedia Commons)

Det er naturligvis nødvendigt med en bred repræsentation af informanter, når sociolingvistiske studier skal udføres. Har man en fornuftig spredning med hensyn til alder, geografi, køn osv. kan man få interessante oplysninger om sproglig praksis i sociolingvistisk perspektiv.

Informanterne bestod af tre aldersgrupper. De yngste informanter var 15-årige elever, som gik i 9. klasse. I den mellemste aldersgruppe var informanterne 45-65 år gamle, og i den ældste aldersgruppe var informanterne 70+. Den oprindelige plan var at have mindst fire personer i hver aldersgruppe, dvs. 12 informanter fra hvert dialektområde. Men antallet ligger langt over det, i særdeleshed i Tórshavn, som har et samlet antal informanter på 37. Foruden nyindsamlet materiale brugte vi nogle gamle optagelser. Testdesignet i hele projektet bestod af tre metoder: Interviewer, spørgeskemaer om holdning til dialekter og masketest. Denne artikel bygger kun på interviewene<sup>6</sup>.

Der er i alt 69 interviewere med 103 informanter, 54 kvinder og 49 mænd. Årsagen til at antallet informanter er så stort i forhold til interviewene er, at i de fleste tilfælde sad to og to informanter og talte sammen. Denne metode fungerede godt i de fleste tilfælde, da informanterne kendte hinanden i forvejen.

## 6. Talesprogsbanken

Alle optagelserne er lagt ind i talesprogsbanken som lydfiler sammen med transskriptioner. 30 minutter af hver samtale blev transskriberet med standardretskrivning. Transskriberingen blev gjort i Praat-

<sup>6</sup> Mere kan læses om de færøske holdningsundersøgelser i Bugge (2018) og i Bugge og Jacobsen (2018).

programmet, som egner sig godt til talesprogsdata. Transskriberingen var færdig ved udgangen af 2018. Det samlede antal ord i talesprogsbanken pr. januar 2022 er 471 178.

Når vi nu har adgang til talesprogsmateriale, giver det nye muligheder for at forske i sprogvariation. Talesprogsbanken er et godt arbejdsredskab til sociolingvistiske undersøgelser. Der er mange tekniske funktioner indbygget i programmet, som gør det muligt at lave korrelationsstudier. Den færøske talesprogsbank er en del af den norske talesprogsbank, Clarino (Common Language Resources Infrastructure Norway), som er udviklet ved Universitetet i Bergen.

Alle ord i talesprogsbanken er tagget med grammatiske oplysninger, køn og tal for substantiver, præsens og præteritum for verber osv. Ud over grammatiske oplysninger er der adgang til metaoplysninger om informanterne, fx fødselsår, hjemsted, køn osv. I tabel 1 ser vi et eksempel på tagging og metasproglige oplysninger med et eksempel på skriftsprogsformen *lærarar* (som i øvrigt forekommer 53 gange i talesprogsbanken):

word:	Lærarar
lemma:	Lærari
pos:	N
features:	N Msc Pl Nom Indef
person:	30716
document:	KKG_2TO
age:	G
is-interviewer:	Nei
dialect:	Tórshavn
sex:	K
birthyear:	1945
place:	Tórshavn
admission-year:	2016
interviewer:	90215
admission-place:	Tórshavn
mother-from:	Hvalba
father-from:	Svínoy

Tabel 1: Eksempel på metaoplysninger i talesprogsbanken.

## 7. Standardretskrivning

I transskriptionerne anvendes standardretskrivning og standardbøjning. I udgangspunktet skal transskriptøren skrive det, som bliver sagt, dvs. at transskriptionen skal være en deskriptiv og objektiv gengivelse af talen. Transskriptøren skal koncentrere sig om at forstå og gengive det, som høres på lydfilen uden samtidig at tænke på fonetiske forskelle. I øvrigt er transskriptøren sjældent kendt med alle de forskellige variabler og varianter, som forskere på et senere tidspunkt måtte være interesseret i. Det kan desuden være yderst problematisk at identificere en bestemt udtale af et ord, og den tolkning skal transskriptøren ikke foretage. Derfor valgte vi at transskribere dialektale varianter, fx *lærarir* og *lærara* (flertal af *lærari* 'lærer') med standardformen *lærarar*. Et andet eksempel på standardisering i transskriptionen var, at den dialektale flertalsform *hevði* 'havde' (præteritum af verbet *hava* 'have') blev skrevet *høvdu*, som er den korrekte form ifølge standardgrammatikken. En søgning på formen *høvdu* i søgestrengen giver derfor samtlige belæg på begge udtalevarianter i præteritum flertal. Det betyder dog ikke, at denne transskriberingspraksis er den eneste rigtige eller mest hensigtsmæssige, fordi når man har adgang til en grammatiktagger, kan man nemt lave en avanceret søgning på grammatiske kriterier og opnå de samme søgeresultater. Den største fordel med standardretskrivning er i de tilfælde, hvor der eksisterer mange talesproglige varianter af et enkelt ord. I mange tilfælde forekommer disse talesprogsvarianter aldrig på skrift, og derfor kan det være vanskeligt at gengive sådanne ord med en lydret ortografi. Som eksempel på det vil jeg anføre adverbiet *soleiðis* [so:lai:jts] 'således', som kan have mange forskellige udtaler (i eksemplerne her har jeg bibeholdt *ð*, selvom

det ikke udtales), fx *soloyðis*, *soloyðus*, *soloyðs*, *soleiðsnar*, *soleiðsna*, *soleiðsni*, *soloyðsni*, *soloyðsnar*, *soloyðsna*, *soeiðsna*, *soeiðsnar* osv. Dette eksempel skulle gerne vise fordelene ved at anvende standardretskrivning. At skrive en tilnærmelsesvis lydret gengivelse af udtalen af hvert enkelt ord ville være en umulig opgave. Ved at skrive standardformen ind i søgestrengen får man alle udtaler af vedkommende ord.

## 8. Udfordringer med talesprogsmateriale

Det er spændende men samtidig udfordrende at arbejde med talesprogsmateriale. Det er spændende på den måde, at vi nu for første gang har mulighed for at arbejde med et annoteret korpus med talesprogsmateriale. Udfordringen er bl.a., at andre normer gør sig gældende i talesproget end i skriftsproget. En tilbagevendende udfordring med talesprog er, at udtalen ikke altid er så distinkt, og det går især ud over endelserne, som kan være meget svagt artikuleret og til tider kan være helt væk. Når to personer sidder og taler sammen, som de gjorde i denne undersøgelse, afbryder de ofte hinanden eller taler i munden på hinanden. I enkelte tilfælde afbryder intervieweren samtalen. Ofte er der højlydt latter. Lydkvaliteten på optagelsen kan være dårlig. Der kan være støj i baggrunden. Somme tider bliver informanterne så ivrige, når de fortæller en spændende historie, at taletempoet stiger voldsomt. En spændende historie kan således medføre udistinkt udtale. Nogle personer har en utydelig udtale, andre taler hurtigt osv. Alle disse forhold kan gøre det vanskeligt at høre, hvordan informanterne rent faktisk udtaler ordene. Derfor er der mange tvivlsspørgsmål, når endelserne skal undersøges. Det kommer også tydeligt til udtryk i denne undersøgelse, hvor det kan være vanskeligt at holde *ir*-endelsen og *R*-endelsen ude fra hinanden.

## 9. *ari*-ord i talesprogsbanken

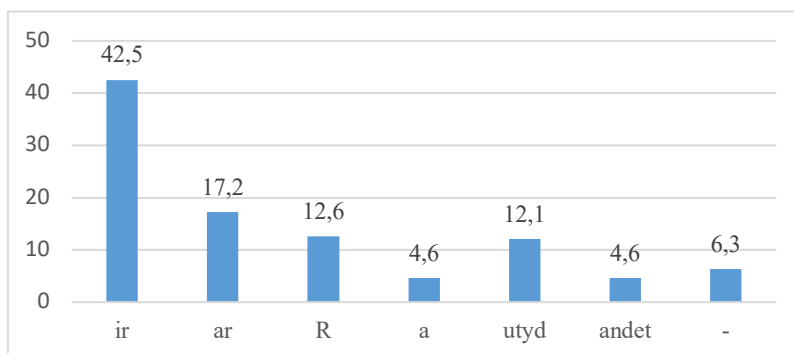
Nu vil jeg undersøge distributionen af de forskellige varianter af *arar*-endelsen i talesprogsbanken. En søgning på skriftsprogsendelsen *arar* (skrevet ”\*arar” i søgestrengen) giver 174 belæg eller forekomster, som opfylder betingelsen for ord, der ender på *arar*. Det burde være tilstrækkeligt til en dialektal variationsundersøgelse. Men det blev hurtigt klart for mig, at der var en del udfordringer i materialet. En fejl i en af lydfilerne gjorde, at lyden ikke kunne høres, og derfor er disse eksempler ikke annoteret. Senere er den blevet lagt ind på ny. I denne lydfil er der 19 belæg på *arar*. Tilbage var der så 155 belæg. Af disse 155 belæg måtte 21 grupperes for sig på grund af utydelig udtale (markeret som *utyd*). Tilbage var der nu 134 belæg, som jeg havde til rådighed i min analyse. Men yderligere to belæg måtte fjernes, fordi de blev brugt af intervieweren. Derfor består analysen af 132 belæg (jf. tabel 2). Når jeg gennemlyttede alle belæggene, blev det klart for mig, at nogle personer brugte en endelse, som hverken var *ar*, *ir* eller *a*. Her kunne jeg rent faktisk ikke høre nogen endelse i det hele taget. Den tryksvage vokal var bortsynkoperet. Denne udtale mener jeg er eksempel på stavelsesbærende *r* (skrevet *R*), fx *danskaR*, *skutaR*. Sandøy (1994: 15) har undersøgt stavelsesbærende *n* i færøsk og siger, at den lyd, som falder ud, enten er *i* eller *u* og at det rammer den sidste stavelse i ordet. I dette tilfælde er det også den sidste stavelse, som bliver bortsynkoperet, fx *skutarir* > *skutaR*. Jeg er enig med den ene fagfællebedømmer, som siger, at hvis *r’et* er stavelsesbærende, skal man have en fornemmelse af, at rytmen forudsætter tre stavelser.

Det kan være vanskeligt at høre forskel på *R* og *ir*, især når folk taler hurtigt. Siger de fx *danskaR* eller *danskarir*? Og jeg skal erkende, at jeg ofte var i tvivl og stadigvæk er det, men ved at lytte igen og igen er jeg kommet frem til 22 belæg (12,6%) på *R* (stavelsesbærende *r*). Hvis jeg skulle klassificere *R*-formerne i en af de andre grupper, ville de komme i *ir*-gruppen, fordi *R*-formerne fonetisk ligger nærmere *ir* end *ar* (*a*-kategorien er udelukket). En sådan artikulatorisk forklaring er vel ikke utænkelig. Afstanden mellem gane og tunge er meget lille ved de høje vokaler *i* og *u*, og det samme gør sig gældende ved udtalen af konsonanten *r*. Derved kan der opstå en slags ”assimilation”.

En foreløbig akustisk undersøgelse af de forskellige endelsers (suffiksers) varighed viser, at *R*-varianten i gennemsnit har en kortere varighed end *ir*- og *ar*-endelserne, hhv. 155 ms for *R* og 282 ms for *ir* og *ar*. Resultaterne skal ses med forbehold grundet lydets kvalitet og analysens foreløbige grovhed. Men målingerne viser en gennemgående tendens til, at *R*-varianterne udtales med kortere varighed end *ir*- og *ar*-varianterne, hvilket underbygger argumentet for kategoriseringen<sup>7</sup>.

<sup>7</sup> Tak til Iben Nyholm Debess for den akustiske undersøgelse, som blev lavet i Praat.

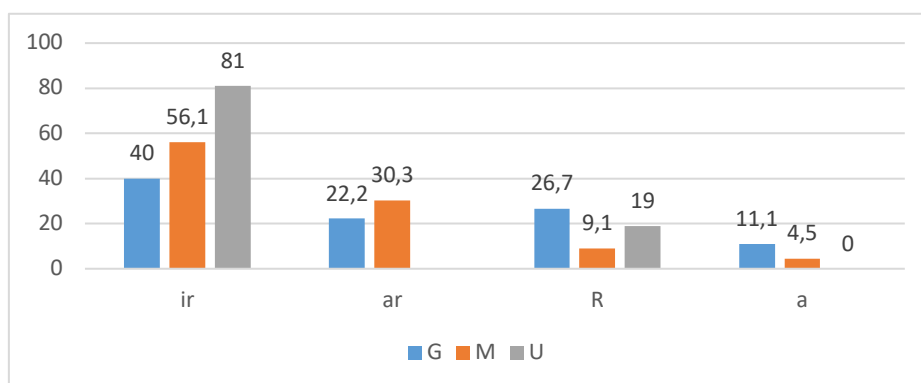
Søjlediagrammet i figur 1 viser den procentuelle distribution af varianterne (medregnet de belæg, der falder uden for varianterne). I tabel 2 nedenfor anføres antallet af belæg sammen med procenttallene fordelt på de tre aldersgrupper.



Figur 1: Den procentuelle distribution af varianterne

I gruppen *andet* placeres fx verber, som ender på *arar*, eksempelvis *svarar* 'svarer'. I gruppen længst til højre med en streg under placeres de uannoterede belæg (pga. en fejl i en lydfil, se ovenfor). Forkortelsen *utyd* betyder utydelig udtale. Distributionen viser meget tydeligt, at *ir*-varianten er den mest frekvente, og hvis man lægger *ir*- og *R*-varianten sammen (hvilket jeg er tilbøjelig til), får man et resultat, som viser, at over halvdelen af samtlige belæg falder ind under *ir*-varianten.

Figur 2 er en visualisering af tabel 2, som viser den procentuelle fordeling i forhold til alder:



Figur 2: Den procentuelle distribution af varianterne i forhold til alder.

	sum	ir	ar	R	a
sum	134 (100,0)	74 (55,2)	30 (22,4)	22 (16,4)	8 (6,0)
	134 (100,0)	74 (69,3)	30 (13,1)	22 (13,7)	8 (3,9)
	2 (100,0)	2 (100,0)			
G	45 (100,0)	18 (40,0)	10 (22,2)	12 (26,7)	5 (11,1)
M	66 (100,0)	37 (56,1)	20 (30,3)	6 (9,1)	3 (4,5)
U	21 (100,0)	17 (81,0)		4 (19,0)	

Tabel 2: Den procentuelle distribution af varianterne i forhold til alder

Tallene i tabel 2 er antal belæg og ikke antal informanter. Distributionen viser, at *ir*-varianten er den hyppigst forekommende i alle aldersgrupper. De yngste bruger udelukkende *ir*- og *R*-varianten. Den høje

frekvens af *ir*-formen betyder vel egentlig, at den har fået status som en overdialektal form, brugt af folk overalt i samfundet.

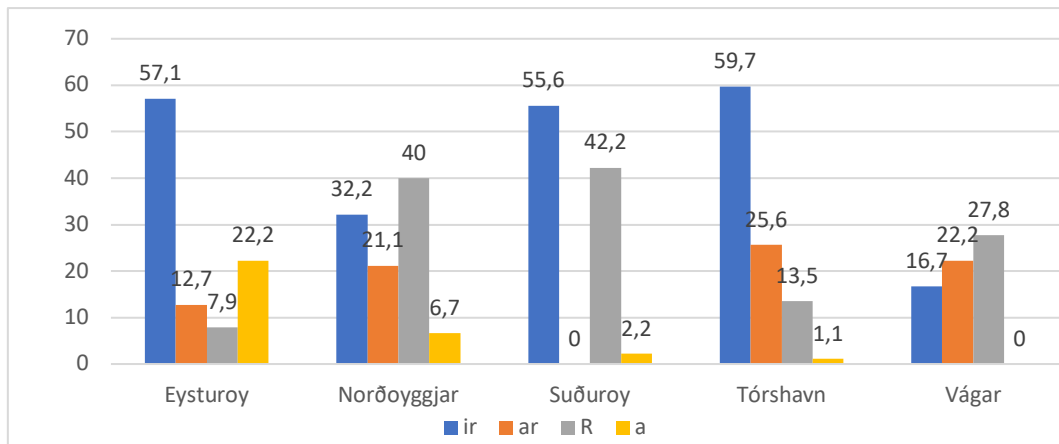
Mens tabel 2 viser distributionen af belæg, viser tabel 3 fordelingen af informanter mellem de forskellige dialektområder. Tabel 3 viser, hvordan de forskellige dialektområder er repræsenteret i forhold til alder og køn:

		Eysturoy	Suðurstreymoy	Norðoyggjar		Suðuroy		I alt
Alder	Vágar	Eiði	Tórshavn	Klaksvík	Viðareiði	Tvøroyri	Andre bygder	
<b>G</b>	2k+2m	2k+2m	7k+6m	2k+2m	2k+2m	2k+3m	2k+1m	37
<b>M</b>	2k+2m	2k+2m	6k+6m	1k+1m	2k	1k+2m	3k	30
<b>U</b>	2k+2m	2k+2m	6k+6m	2k+2m	2k+2m	2k+2m	2k+2m	36
<b>I alt</b>	<b>12</b>	<b>12</b>	<b>37</b>	<b>10</b>	<b>10</b>	<b>12</b>	<b>10</b>	<b>103</b>

Tabel 3: Antal informanter fordelt på alder og køn i de undersøgte dialektområder (k=kvinder; m=mænd)

Kolonnen yderst til højre viser, at antallet informanter i de tre aldersgrupper ligger på nogenlunde samme niveau, mellem 30 og 37. Kønsfordelingen er også rimelig (54 kvinder og 49 mænd). Til gengæld er antallet af informanter fra Tórshavn langt større end fra de andre steder. Årsagen til det forholdsvis store antal er, at vi i et underprojekt ville få mulighed for at undersøge, hvorvidt tilflytning til Tórshavn påvirker folks sproglige praksis. Derfor blev informanterne fra Tórshavn udvalgt efter andre kriterier og opdelt i tre undergrupper: (I) Informanten selv var tilflytter, (II) Informanten havde to tilflytterforældre, og (III) Begge informantens forældre var fra Tórshavn.

Figur 3 viser variantdistributionen mellem de fem dialektområder:



Figur 3: Den procentuelle distribution af varianter i de fem dialektområder

Diagrammet viser, at frekvensen af *ir*-formerne er særlig høj på Eysturoy, Suðuroy og i Tórshavn. På den anden side ligger skriftformen *ar* relativt lavt. Tabel 4 viser en distributionsanalyse, hvor jeg krydser sprogvariant, alder og sted. Tallene i parentes er procenttal af forekomster i hele gruppen, og summen af dem giver 100 procent på den vandrette akse. På den vandrette akse øverst i tabellen står de sproginterne varianter og på den lodrette akse i venstre side står de sprogksterne variabler, sted og alder. Øverst i den første talkolonne står det samlede antal belæg på den undersøgte variabel og nedenfor står de samlede belæg for hvert dialektområde, fordelt på de tre aldersgrupper. Derefter kommer belæggene på hver af de fire varianter: *ir*, *ar*, *R* og *a* (rest) som står øverst til venstre i tabel 4 er to eksempler på spørgsmål fra intervieweren.



	Sum	ir	ar	R	a
<b>Sum</b>	<b>134 (100,0)</b>	<b>74 (55,2)</b>	<b>30 (22,4)</b>	<b>22 (16,4)</b>	<b>8 (6,0)</b>
<b>(rest)</b>	2 (100,0)	2 (100,0)			
	2 (100,0)	2 (100,0)			
<b>Eysturoy</b>	<b>22 (100,0)</b>	<b>12 (57,1)</b>	<b>4 (12,7)</b>	<b>2 (7,9)</b>	<b>4 (22,2)</b>
G	6 (100,0)		1 (16,7)	1 (16,7)	4 (66,7)
M	14 (100,0)	10 (71,4)	3 (21,4)	1 (7,1)	
U	2 (100,0)	2 (100,0)			
<b>Norðoyggjar</b>	<b>15 (100,0)</b>	<b>5 (32,2)</b>	<b>4 (21,1)</b>	<b>4 (40,0)</b>	<b>2 (6,7)</b>
G	3 (100,0)	2 (66,7)	1 (33,3)		
M	10 (100,0)	3 (30,0)	3 (30,0)	2 (20,0)	2 (20,0)
U	2 (100,0)			2 (100,0)	
<b>Suðuroy</b>	<b>21 (100,0)</b>	<b>15 (55,6)</b>		<b>5 (42,2)</b>	<b>1 (2,2)</b>
G	15 (100,0)	10 (66,7)		4 (26,7)	1 (6,7)
M	5 (100,0)	5 (100,0)			
U	1 (100,0)			1 (100,0)	
<b>Tórshavn</b>	<b>65 (100,0)</b>	<b>38 (59,7)</b>	<b>18 (25,6)</b>	<b>8 (13,5)</b>	<b>1 (1,1)</b>
G	18 (100,0)	5 (27,8)	8 (44,4)	5 (27,8)	
M	31 (100,0)	18 (58,1)	10 (32,3)	2 (6,5)	1 (3,2)
U	16 (100,0)	15 (93,7)		1 (6,2)	
<b>Vágar</b>	<b>9 (100,0)</b>	<b>2 (16,7)</b>	<b>4 (22,2)</b>	<b>3 (27,8)</b>	
G	3 (100,0)	1 (33,3)		2 (66,7)	
M	6 (100,0)	1 (16,7)	4 (66,7)	1 (16,7)	
U	0 (0,0)				

Tabel 4: Antal belæg af de fire varianter i forhold til alder og sted.

G = gamle; M = midaldrende; U = unge.

Som det fremgår af tabel 4, er der 74 belæg på *ir*-varianten (svarende til 55,2% af samtlige belæg). Dermed er *ir*-varianten den mest frekvente i hele korpusset. Til sammen udgør *ir*- og *R*-varianten 71,6%. Det er ikke usandsynligt, at *ir*-varianten vil brede sig endnu mere i fremtiden som en slags statusform (jf. Weyhe, 2012b: 368). Den næststørste kategori er *ar*-varianten (skriftsprogvarianten), hvis brugsfrekvens dog ligger meget lavere end *ir*-varianten med 30 belæg (svarende til 22,4%). Mindst frekvent er *a*-varianten med otte belæg (svarende til 6,0%). Af de otte *a*-former, findes fire hos gamle mennesker på Eysturoy og to hos midaldrende fra Norðoyggjar. Der er altså ingen forekomst af *a*-former hos de yngste. Den før så udbredte *a*-form på Eysturoy og i Norðoyggjar er stærkt på retur og afløses af *ir*-former, som i øvrigt forekommer i alle aldersgrupper over hele landet. *R*-formerne forekommer i alle aldersgrupper, dog med en lille overvægt i Tórshavn og på Suðuroy<sup>8</sup>.

## 10. Sammenlæg og diskussion

Formålet med denne artikel er dels at undersøge en morfologisk variabel og dens manifestationer i talesproget, dels at undersøge anvendeligheden af talesprogsbanken som sprogteknologisk værktøj til sociolingvistiske variantundersøgelser.

Den undersøgte variabel er flertalsformer af *ari*-ord i nominativ og akkusativ i ubestemt form. Skriftsprogvarianten er *ar*, mens talesproget desuden har varianterne *ir*, *a* og *R* (fx *lærarar*, *lærarir*, *lærarar*, *læraR*). Undersøgelsen viser meget tydeligt, at *ir*-varianten er den mest frekvente i hele materialet med 74 belæg (55,2%), derefter kommer skriftsprogvarianten *ar* med 30 belæg (22,4%), *R*-varianten med 22 belæg (16,4%) og *a*-varianten med otte belæg (6,0%).

<sup>8</sup> Suðuroy-varianten med en *o*- eller *ø*-lignende lyd tolker jeg som variant af *i*.

I den yngste aldersgruppe på Eysturoy er der kun to belæg på den undersøgte variabel, og på Suðuroy er der kun et enkelt belæg i den yngste aldersgruppe. På Vágar er der intet belæg blandt de yngste af et samlet antal på ni belæg på Vágar. Det siger sig selv, at når man nedbryder ni belæg på tre aldersgrupper, bliver tallene meget små og usikre. Datagrundlaget er yderst spinkelt, hvilket påvirker reliabiliteten i negativ retning og gør konklusionerne uvisse. På den anden side er datamængden for Tórshavn betydeligt større, hvilket øger undersøgelsens reliabilitet. Derfor er det nødvendigt at udvide talesprogsbanken med nyt materiale, især fra områderne uden for Tórshavn for på den måde at få et mere balanceret korpus. Datagrundlaget må altså øges betragteligt, før vi kan få fuldt udbytte af alle de tekniske muligheder, der ligger i analyseprogrammet. Trods den begrænsede datamængde viser materialet nogle tendenser i variationen i forhold til alder og sted.

Undersøgelsen viser, at *ir*-varianten står stærkt blandt de unge. Det er derfor ikke usandsynligt, at denne variant vil brede sig endnu mere i fremtiden som en overdialektal form (jf. Weyhe, 2012b: 368).

Distributionen med *ir*-varianten i top åbner op for en helt anden diskussion, nemlig hvorvidt tiden ikke er inde til at acceptere *ir*-formen som tilladt skriftsprogsform ved siden af *ar*-formen. Det kan virke lidt besynderligt, at den mest frekvente talesprogsvariant ikke er tilladt i skrift. Den afgørelse ligger hos Málráðið (Sprográdet).

Erfaringerne fra dette studie er, at rent teknisk er talesprogsbanken et fortræffeligt redskab til sociolingvistiske dialektstudier og variationsstudier i øvrigt. Nu er det muligt at udføre kvantitative dialekt- og dialektforandringsstudier med udgangspunkt i et dialektkorpus, som er designet til netop dette formål.

Ulempen ved det færøske dialektkorpus er dens begrænsede størrelse med knap en halv million ord. Når man undersøger en variabel med alder og sted, bliver tallene i de forskellige celler meget små. Derfor må man være forsigtig i sine konklusioner. Datagrundlaget må derfor øges betragteligt, før vi kan få fuldt udbytte af alle de tekniske muligheder, der ligger i analyseprogrammet.

## Tak

Tak til de to anonyme fagfæller for yderst relevante kommentarer til første udkast af artiklen. Jeg vil også takke redaktorerne af festskriftet for deres tålmodighed og velvilje. Tak til to af mine kolleger på instituttet, Iben Nyholm Debess og Knút Háberg Eysturstein, for teknisk assistance, og Inger Smærup Sørensen skal have tak for sproglige kommentarer.

## Litteraturliste

- Bugge, Edit. 2018. Attitudes to Variation in Spoken Faroese. I *Journal of Sociolinguistics* 22:3, 1–19. <https://doi.org/10.1111/josl.12283>.
- Bugge, Edit og Jógvan í Lon Jacobsen. 2018. Hugburður til variatión í føroyskum talumáli. I *Íslenskt mál og almenn málfræði*, 40. árg., 97–117. Íslenska málfræðifélagið, Reykjavík.
- Chambers, Jack K. og Peter Trudgill. 1990. *Dialectology*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.
- Hagström, Björn. 1977. Hvi hevur nekarin fepur? I *Fróðskaparrit*, 25. bók, 26–56. Mentunargrunnur Føroya Løgtings, Tórshavn.
- Hudson, R.A. 1999. *Sociolinguistics*. Second edition. Cambridge University Press. Cambridge
- Navne på stater*. Nationalitetsbetegnelser. Dansk-færøsk-islandsk. 1974. *Sprog i Norden*, 81–113. Årsskrift for de nordiske sprognævn. Gyldendalske boghandel. Nordisk forlag, Keypmannahavn.
- Petersen, Hjalmar P. 2019. *Føroysk mállæra 1. Kyn, orðmyndan og bending*. Nám, Tórshavn.
- Poulsen, Jóhan Hendrik Winther. 2004. *Mál í mæti. Greinasavn eftir Jóhan Hendrik W. Poulsen*. Útgivið í sambandi við sjeiti ára føðingardag hansara 20. juni 2004, redigeret af Anfinnur Johansen og Hans Joensen. Føroya Fróðskaparfelag, Tórshavn.
- Sandøy, Helge. 2017. Kjelder og årsaker til dialektendringar i norsk. I *Bók Jógvan. Heiðursrit til Jógvan í Lon Jacobsen á 60 ára degnum*, redigeret af Zakaris Svabo Hansen, Anfinnur Johansen, Hjalmar P. Petersen og Lena Reinert, 371-384. Fróðskapur. Faroe University Press, Tórshavn.

- Sandøy, Helge. 2008. Kontakt og spreining. I *Språkøte. Innføring i sociolingvistikk*, redigeret af Brit Mæhlum, Gunnstein Akselberg, Unn Røyneland og Helge Sandøy, 221-240. Cappelen akademisk forlag.
- Sandøy, Helge. 1994. Stavilsisberandi N í føroyskum. *Málting* 11 (nr. 2, 4. årg.), 12–19.
- Weyhe, Eivind. 2015. Færøsk gennem to hundrede år. I *Talemål etter 1800*, redigeret af Helge Sandøy, 409–431. Novus, Oslo.
- Weyhe, Eivind. 2012a. Pluralis af nomina agentis på *-ari* i færøsk. I *Eivindaródn. Greinar 1979-2011*, redigeret af Anfinnur Johansen, Turið Sigurðardóttir og Tóta Árnadóttir, 355–365. Fróðskapur, Faroe University Press, Tórshavn.
- Weyhe, Eivind. 2012b. Nøkur orð um bendingarmun í føroyskum bygdamálum. I *Eivindaródn. Greinar 1979-2011*, redigeret af Anfinnur Johansen, Turið Sigurðardóttir og Tóta Árnadóttir, 366–373. Fróðskapur. Faroe University Press, Tórshavn.
- Weyhe, Eivind. 2012c. Bendingarmunur í føroyskum málførum. I *Eivindaródn. Greinar 1979-2011*, redigeret af Anfinnur Johansen, Turið Sigurðardóttir og Tóta Árnadóttir, 374-419. Fróðskapur. Faroe University Press, Tórshavn.

**Internetadresser:**

<https://ordnet.dk/ods>

<https://ordnet.dk/ddo>

[islex.hi.is](http://islex.hi.is)

[malrad.fo/orðalistar/lond](http://malrad.fo/orðalistar/lond) og [tjóðir](http://malrad.fo/orðalistar/tjóðir)