



Dynamikken i en språkendringsprosess: Bortfall av hunkjønnsformer i norsk*

Björn Lundquist¹, Rachel Klassen¹ og Marit Westergaard^{1,2}

¹UiT Norges arktiske universitet

²NTNU Norges teknisk-naturvitenskapelige universitet

Det finnes betydelig regional og sociolingvistisk variasjon i genussystemet i norsk, hovedsakelig når det gjelder hunkjønns-kategorien. Mange morsmåls-talere av norsk har et robust tregenussystem (hankjønn, hunkjønn og intetkjønn), mens andre har et system med kun to genus (felleskjønn og intetkjønn), der alle eller de fleste hunkjønnsformene er erstattet av hankjønnsformer. Uavhengig av dialekt er alle morsmålstalere eksponert for både to- og tregenussystemet gjennom forskjellige medier, noe som betyr at det er stor variasjon også i skriftlig og talt input. I denne artikkelen setter vi opp en hypotese om at input og output gjensidig påvirker hverandre; mer spesifikt at variasjonen i input har en effekt på morsmålstaleres output og over tid bidrar til språkendringer. Vi sammenligner frekvensdata for hunkjønnsformer i to korpus (skrift og tale) med resultater fra en genusklassifiseringsstudie som også målte reaksjonstider. Vi finner en tydelig korrelasjon mellom klassifiseringsmønstrene og frekvensdata fra det skriftlige korpuset: Hankjønnsst substantiver som ofte opptrer med hankjønnsformer i korpuset, hovedsakelig bestemthetssuffikset *-en*, blir oftere klassifisert som hankjønnsord og har lengre reaksjonstider.

Nøkkelord: genus, hunkjønn, språkforandring, norsk, variasjon

* Vi har satt stor pris på nyttige kommentarer fra to konsulenter, og vi vil rette en spesiell takk til Terje Lohndal for inngående gjennomlesing av manuskriptet og mange gode forslag. Klassifiseringsstudien som er brukt i denne artikkelen er en del av Rachel Klassen's MSCA-prosjekt (Marie Skłodowska Curie Action ved UiT 2017-2020, 748966 GenBiLex). Forskningen til denne artikkelen er utført i forbindelse med prosjektet Multi-Gender ved Senter for grunnforskning (CAS) 2019-2020 og NFR-prosjektene MiMS (250857) og GenVAC (301094).

1. Innledning

I løpet av de siste årene har flere studier vist at det grammatiske genussystemet er i forandring i flere norske dialekter (Lødrup 2011; Rodina & Westergaard 2015, 2021; Lundquist mfl. 2016, Busterud mfl. 2019, Klassen mfl. 2022). Denne forskningen har vist at forandringen har skjedd veldig raskt enkelte steder: bare på én generasjon har et stabilt tre-genussystem blitt utviklet til et system der hunkjønnsmarkører brukes overraskende sjelden. Språk endres dog ikke over natten, og vi forventer å se betydelig variasjon over lengre tid før et språk går fra et stabilt system til et annet, både innen en og samme språkbruker og i et språksamfunn. Ved å studere den pågående prosessen av språkforandring kan vi få bedre forståelse for hvorfor systemer forsvinner, og hvordan nye, stabile systemer utvikles. Utviklingen forventes å stabilisere seg over minst tre ulike nivåer:

1. Sosiolingvistisk nivå: Forandringen forventes ikke å påvirke alle talere av et språk på samme tid, men spre seg fra språkbruker til språkbruker, og til nye sosiolingvistiske kontekster og dialekter i henhold til mer eller mindre velkjente mønster.
2. Morfologisk nivå: Det gamle bøyningssystemet forsvinner ikke nødvendigvis i alle nominale kategorier samtidig; det nye systemet spres seg snarere til stadig flere av de nominale bøyingskategoriene, inntil det gamle systemet ikke lenger oppfattes som en distinkt kategori av språkinnlærere og språkbrukere.
3. Leksikalsk nivå: Forandringen når ikke nødvendigvis hele leksikonet på samme tid, men kan spre seg på ulike tidspunkter til ulike semantiske eller morfologiske klasser av leksikonet, eller muligens spesifikke substantiver.

Vi vet fortsatt ganske lite om hvor raskt utviklingen og stabiliseringen skjer, og heller ikke i hvilken grad utviklingen på de tre nivåene samvirker. Gjennom å studere språkforandringen i nåtid håper vi å kunne forstå mekanismene for språkforandring bedre, både på individnivå og generelt i språksamfunnet. I dagens Norge, som i de fleste andre land, er det likevel åpenbart at de fleste talere tilhører mer enn ett språksamfunn: de tilhører både et lokalt språksamfunn (lokal dialekt) og et større nasjonalt språksamfunn, det vil si Norge, som enhver taler kjenner fra skriftspråk og media. I tillegg kan det finnes andre nivåer, for eksempel et regionalt nivå (se f.eks. Sandøy 2013 for diskusjon). Innenfor det større nasjonale språksamfunnet finner vi åpenbart mye variasjon; enkelte

dialekter har et tre-genussystem og andre et to-genussystem, og de morfofonologiske formene for de tre genusene kan være forskjellige i ulike dialekter. Det er fortsatt noe uklart hvordan variasjon innenfor det større nasjonale språksamfunnet påvirker «stabile» dialekter og stabile dialekttalere, og også i hvilken grad enkelte talere er bevisst på om språklig variasjon er isolert til visse språklige kontekster.

I denne artikkelen undersøker vi hvordan spenningen som finnes mellom genusbruk i skriftlig bokmål og muntlig dialekt påvirker genussystemet på individnivå. Vi vet fra tidligere studier at talere som har vokst opp i Norge i løpet av de siste 30 årene med stor sannsynlighet har fått et tre-genussystem i inputen fra foreldrene sine, med mindre de har vokst opp i Bergen, (deler av) Oslo eller språkkontaktområder i Nord-Norge. Til tross for dette har mange utviklet et ustabil tre-genussystem, eller til og med et to-genussystem. Disse talerne har også blitt eksponert for et to-genussystem, eller et variabelt tre-genussystem i skriftlig bokmål. I denne artikkelen sammenlikner vi korpusdata fra skriftlig bokmål (NoWaC) og muntlig dialekt (NDC) med data fra en genusklassifiseringsstudie (Klassen mfl. 2022). Klassifiseringsstudien viser at deltakerne i studien grovt sett kan deles inn i tre grupper: en stabil to-genus-gruppe, en stabil tre-genus-gruppe, og en ustabil gruppe. Vi beskriver kort hvordan gruppeinndelingen henger sammen med demografiske faktorer, og deretter sammenlikner vi mønster fra klassifiseringsstudien med ulike typer av korpusfrekvenser, hentet både fra skriftspråkkorpus og talespråkkorpus: lemmafrekvens, frekvensen av ulike typer synlige genusmarkører og forhold mellom hun- og hankjønnmarkører. Når vi har identifisert de frekvenser som best korrelerer med genusklassifiseringen fra Klassen mfl. (2022), vil vi se om disse tallene også korrelerer med reaksjonstider fra våre «stabile» grupper. Dette vil gjøre oss i stand til å kunne svare på spørsmålet om såkalte stabile språkbrukere også påvirkes av den variasjonen i språket som finnes på nasjonalt nivå.

2. Genus og genusforandring i norsk

Mens noen germanske språk, særlig nederlandsk, svensk og dansk, har gjennomgått en historisk forandring fra et tre-genussystem til et to-genussystem (felleskjønn og intetkjønn, se f.eks. Trudgill 2013), har norske dialekter generelt holdt på alle tre genus: hankjønn, hunkjønn og intetkjønn. Det eneste unntaket fra dette har vært bergensdialekten, som mistet hunkjønn for flere hundre år siden, sannsynligvis på grunn av språkkontakt med nedertysk i Hansa-perioden (f.eks. Jahr 2001, men se også Heide 2003, Perridon 2003 og Enger 2013 for

fonologiske og morfologiske forklaringer). Imidlertid er hankjønns- og hunkjønnsparadigmene svært like, selv i dialekter med tre genus, og bare i noen få former finner vi unike eksponenter i hunkjønnsparadigmet. Synkretismen mellom hunkjønn og hankjønn har økt de siste 400-500 årene, mens intetkjønnsparadigmet har beholdt eller utviklet unike eksponenter. Hunkjønnsbøyning på de fleste adjektiver forsvant fra de siste dialektene for over 100 år siden (Sandøy 1988), og siden har de ubestemte hunkjønnsartiklene blitt mer sjeldne først i skriftspråk (bokmål), men også i talemål. I skriftspråket (bokmål) er også tidligere hunkjønnsuffixer (-*a*) på bestemte substantiv ofte erstattet med den typiske hankjønnsendelsen -*en*.

Det tradisjonelle tre-genussystemet for mange muntlige varianter av norsk er illustrert i Tabell 1.

Tabell 1. Det tradisjonelle tre-genussystemet i mange varianter av norsk.

	Hankjønn	Hunkjønn	Intetkjønn
Ubestemt artikkel	<i>en stol</i>	<i>ei bok</i>	<i>et hus</i>
Bestemt artikkel	<i>stolen</i>	<i>boka</i>	<i>huset</i>
Dobbelt bestemt form	<i>den røde stolen</i>	<i>den røde boka</i>	<i>det røde huset</i>
Adjektiv	<i>en fin stol</i>	<i>ei fin bok</i>	<i>et fint hus</i>
Possessivformer	<i>min stol/stolen min</i>	<i>mi bok/boka mi</i>	<i>mitt hus/huset mitt</i>

Den omfattende synkretismen mellom hankjønn og hunkjønn betyr at forskjellen mellom disse kun er synlig i den ubestemte artikkelen og de possessive formene, hvis man ser bort fra formen på bestemthetsuffixet. I vår studie står den ubestemte hunkjønnsartikkelen *ei* i sentrum. Vi vil anta at erstatning av *ei* med *en* er et sikkert tegn på bortfall av hunkjønns-kategorien. Vi vil også følge Lødrup (2011) og Rodina & Westergaard (2015) i antakelsen at bestemthets-suffixet -*a* ikke i seg selv uttrykker genus. Dette betyr likevel ikke at språkinnlæreren og språkbrukeren ikke kan bruke bestemthetsuffixet og andre mer eller mindre transparente morfofonologiske trekk på substantiver for å klassifisere substantiver i ulike genus-kategorier. Det er altså viktig å skille mellom grammatiske genusmarkører, som f.eks. form på frittstående artikler og adjektiver, og semantiske, fonologiske og til og med morfologiske trekk som en innlærer kan bruke for å kategorisere substantiver. I sistnevnte kategori kan vi for hunkjønn gi eksempel på substantiver som refererer til naturlig hunkjønn,

som *ku* og *dame* (semantisk trekk), substantiver som ender på trykksvak *-e*, som *veske* og *flaske*, og substantiver som tar *-a* som bestemthetsuffix (morfologisk trekk).

De siste årene har det blitt rapportert om endringer i genussystemet i en rekke dialekter. Basert på et muntlig korpus med 142 talere i Oslo viser Lødrup (2011) at den ubestemte artikkelen *ei* sjelden brukes blant eldre talere, og at den er nærmest ikke-eksisterende hos den yngre generasjonen. Flere studier har også vist at hankjønnsartikkelen *en* ofte brukes ved feminine ord i Finnmark og i andre nordnorske språkkontaktområder (Jahr 1984, Bull mfl. 1986, Conzett, Johansen & Sollid 2011). Et lignende mønster bekreftes blant yngre talere i to større byer, Tromsø og Trondheim (Rodina & Westergaard 2015, 2021; Busterud mfl. 2019). I alle disse tilfellene blir den ubestemte artikkelen i hunkjønn, *ei*, erstattet av hankjønnsformen *en*, mens bestemthetsuffixet *-a* beholdes i sin opprinnelige form. Som et resultat av denne utviklingen har det nye to-genussystemet mønsteret *en bok-boka* for tidligere hunkjønnsst substantiver. I en produksjonsstudie viser Lundquist mfl. (2018) likevel at en stor del ungdommer i Oslo (16-18 år, fra Lambertseter VGS) bruker *-en* som bestemthetsuffix (*en bok – boken*), i likhet med systemet i Bergen og Sverige. Samme studie viser også at yngre talere fra Sogndal konsekvent bruker hunkjønnsmarkører både på artikler og suffiks, hvilket kan tyde på at talere i nynorskområder har et mer stabilt tre-genussystem. En annen studie, Lundquist mfl. (2016), viser at yngre deltakere fra to mindre byer i bokmålsområder i Nord-Norge (Sortland og Finnsnes) også konsekvent bruker hunkjønnsmarkører på både artikler og suffiks, i kontrast til talere i samme aldersgruppe i Tromsø (Rodina & Westergaard 2015).

Til tross for at genustilordningen i språkproduksjonen var den samme på Sortland og Finnsnes som i Sogndal, fantes det forskjeller i prosessering av grammatisk hunkjønns-genus på disse stedene, noe som ble undersøkt ved et blikksporingseksperiment. I motsetning til deltakerne i Nord-Norge kunne deltakerne i Sogndal forutsi at et hunkjønnsord *ikke* kunne følge etter artikkelen *en*. De kunne også forutsi at kun et hunkjønnsord kunne følge etter *ei*. Ut fra disse resultatene ser det altså ut som de yngre språkbrukerne fra Nord-Norge har en genusbruk der de følger tre-genussystemet i sin produksjon, men de har ingen forventninger om at andre talere, fra samme eller ulik dialekt, skal bruke et tre-genussystem. Mer spesifikt: de har ikke noen forventninger om at andre talere bruker artikkelen *ei* framfor hunkjønnsord. På grunnlag av disse resultatene antar Lundquist mfl. (2016) og Lundquist & Vangsnes (2018) at to-genussystemet sprer seg ved å først påvirke språkforståelsen: talere slutter å legge merke til “feilaktig” bruk av hankjønnsartikkelen *en* (dvs., sammen med

hunkjønnsord), men fortsetter å bruke tre-genussystemet selv. Disse talere har altså allerede et delvis forandret system, men det manifesteres kun i reseptiv prosessering av språk (muligens både i skrift og tale). Eksakt hvordan og hvorfor språkproduksjonen påvirkes i en senere fase er ikke kjent, og vi vet heller ikke om steget fra persepsjon til produksjon skjer innen ett og samme individ, eller kun mellom generasjoner. Vi kan imidlertid være nokså sikre på at vi kan se tegn på språkforandringen i språkforståelse, og kanskje mer generelt i språkprosessering, før vi kan se konkrete tegn i språkproduksjonen. For å kunne forstå fullt ut stabiliteten i et genussystem i et språk, en dialekt eller en individuell språkbruker, bør man derfor bruke forskjellige forskningsmetoder: språkprosessering i diverse *online*-eksperiment (der for eksempel reaksjonstid måles), ulike typer av språklig produksjon, samt bedømminger av språkstimuli der man måler en kategorisk variabel.

I forbindelse med denne studien er det relevant å merke seg at genustilordning i norsk generelt ikke er gjennomsiktig, hvilket vil si at det ikke finnes klare og konsekvente regler som forteller en som lærer språket hvilket genus et substantiv har. Ifølge Trosterud (2001), som har basert sine beregninger på 31.500 substantiver i *Nynorskordboka*, utgjør hankjønnsst substantiver 52%, hunkjønns 32% og intetkjønn bare 16%. Rodina & Westergaard (2015) har foretatt en telling basert på et korpus av voksentale i et barnespråkkorpus (språkbrukere i Tromsø med et tre-genussystem) og funnet at frekvensen (*token*-frekvens) av hankjønnsst substantiver i input er enda høyere, 62,6%, mens hunkjønns- og intetkjønnsst substantiver er representert henholdsvis 18,9% og 18,5%. Dette betyr at i et to-genussystem vil felleskjønn forekomme betydelig hyppigere enn intetkjønn, rundt 80% (og dermed være lik den fordelingen man finner i f.eks. nederlandsk).

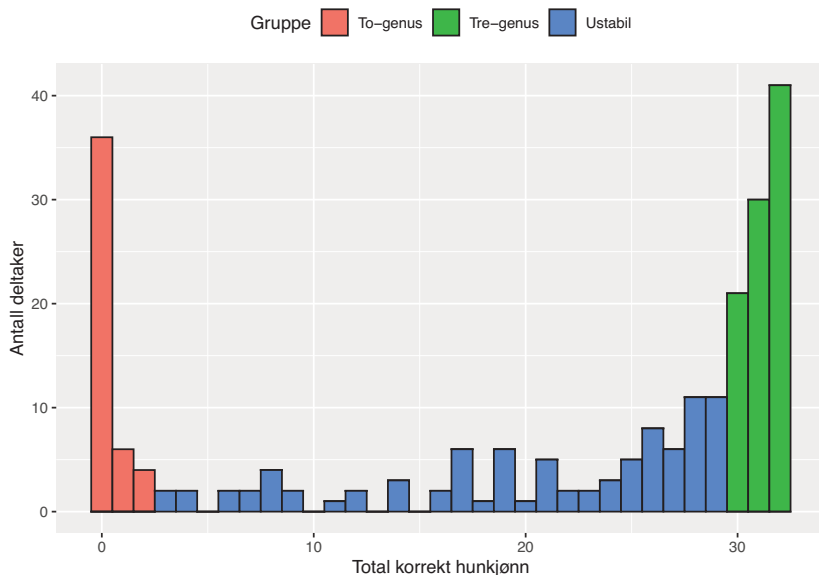
Studier av innlæring av norsk genus har også vist at barn iblant bruker hankjønnsmarkører ved hunkjønns- og intetkjønnsord, f.eks. *en bok* eller *en fly*, se Rodina & Westergaard (2013). Vi kan derfor anta at innlærere ser hankjønn som umarkert grammatisk kjønn i norsk og dermed må lære seg, ord for ord, hvilke substantiver som tilhører kategoriene hunkjønn og intetkjønn. Støtte for en slik hypotese kommer også fra Urek mfl. (dette nummeret), som studerer genustilordning av «pseudo-ord» (ord som ikke er faktiske norske ord, men som kunne ha vært det). Urek mfl. (dette nummeret) viser også at norsktalende til en viss grad kan bruke morfofonologiske trekk på stammer for å klassifisere substantiver i ulike genuskategorier. Det mest pålitelige trekket de finner, er stamme-final *-e*, som i *jent-e* og *vesk-e*, som språkbrukere ser som et tegn på hunkjønnsgenus (se også Trosterud 2001).

3. Bakgrunn: internett-basert genus-klassifiseringstudie (Klassen mfl. 2022)

Størstedelen av dataene som diskuteres i denne artikkelen kommer fra et eksperiment som opprinnelig var designet for å undersøke effekten av variasjon i morsmålet på den abstrakte representasjonen av genus i det mentale leksikonet (Klassen mfl. 2022). I studien fikk deltakerne presentert en rekke substantiver (ett om gangen) og ble bedt om å velge den tilsvarende ubestemte artikkelen (*en*, *ei* eller *et*) så raskt som mulig. I instruksjonen ble det understreket at det ikke fantes kun ett riktig svar, men at deltakerne skulle svare i henhold til sin egen dialekt. For å trene deltakerne på oppgaven og bruk av de tre svartastene (N for *en*, I for *ei* og T for *et*), var det en liten øvingsøkt før selve eksperimentet begynte. Gjennom hele eksperimentet ble hvert substantiv vist på en dataskjerm, enten til deltakeren svarte eller i maksimalt 3500 millisekunder. Hele eksperimentet (inkludert et kort spørreskjema om språkbakgrunn) tok omtrent fem minutter å fullføre. Oppgaven besto av 108 substantiver, 12 i øvingsøkten og 96 i selve eksperimentet (se vedlegg A), og presentasjonen av disse ble randomisert av programvaren. Disse var alle ikke-animate tellelige substantiver, som ble jevnt fordelt på alle tre genus basert på tradisjonell genustilordning (36 hankjønn, 36 hunkjønn og 36 intetkjønn, der 4 fra hver gruppe var en del av øvingsøkten). Listen med substantiver ble nøye valgt ut av to morsmålstalere av norsk, en fra Nord-Norge og en fra Vestlandet, for å unngå spesifikke substantiver som varierte i enten form eller genus i forskjellige dialekter. Til tross for disse kontrolltiltakene, måtte ett substantiv av hvert genus (totalt tre substantiver) ekskluderes på grunn av ustabil genustildeling eller avvikende frekvenser i de påfølgende korpusanalysene. Vi ekskluderer også de fire substantivene fra øvingsøkten, og har da igjen 31 substantiver i hver gruppe. Det var ingen signifikant forskjell i gjennomsnittlig frekvens på tvers av de tre genustilordningene ved nettkorpus (*TenTen*, Jakubiček mfl. 2013) og undertittelkorpus (*Opus2*, Tiedemann 2016). Datainnsamlingen ble utført via IbeX Farm (spellout.net/ibexfarm), en internettplattform for eksperimenter som gjennomføres i en nettleser. For hvert substantiv registrerte IbeX Farm tasten som deltakeren brukte, samt reaksjonstiden.

Det er framfor alt tre funn fra Klassen mfl. som er viktige for vår studie. Vi går gjennom disse funnene nedenfor, og vil etterpå utvikle og finjustere disse. Det første funnet handler om klassifiseringsmønsteret i svarene. Totalt ble data fra 227 deltakere analysert i studien, og ut fra svarmønstrene ble deltakerne delt inn i tre ulike grupper. Deltakere som klassifiserte alle hunkjønnord som «en»-ord, ble kategorisert som «stabile to-genustalere». Totalt har vi 46 deltakere i denne gruppen. Vi inkluderte også talere som klassifiserte 1-2 hunkjønnord

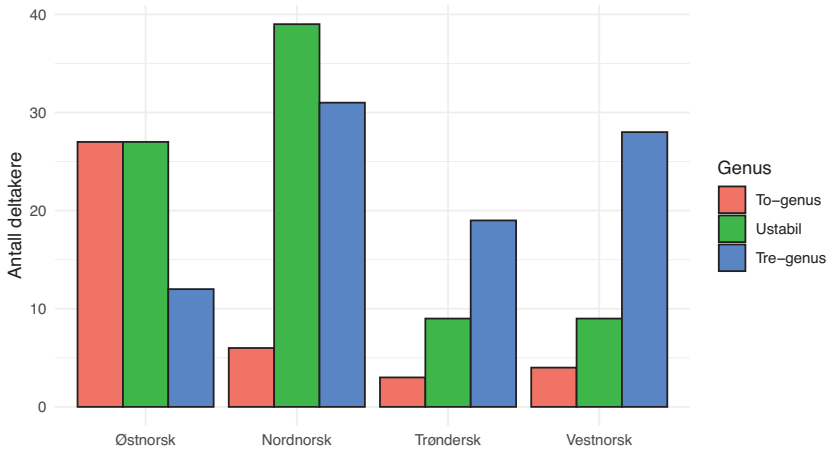
med «ei» i denne gruppen. Deltakere som klassifiserte alle hunkjønnsord som «ei» ble kategorisert som «stabile tre-genustalere». Deltakere som manglet hunkjønnsklassifisering på 1-2 ord i denne genuskategorien, havnet også i denne gruppen, til sammen 92 deltakere. De som havnet mellom disse gruppene ble regnet inn i gruppen «ustabil», totalt 89 talere. Vi presenterer fordelinga av de tre gruppene i figur 1.



Figur 1: Gruppeinndeling av deltakerne basert på kategorisering av hunkjønnsord.

Her bør man legge merke til at det er en relativt liten del av deltakerne som ikke brukte hunkjønn i det hele tatt, kun cirka 20%. De fleste har enten et stabilt eller ustabil tre-genussystem (40% i hver gruppe). Totalt ble 63,7% av hunkjønnsordene “korrekt” klassifisert.

I studien ble det også samlet inn informasjon om deltakernes dialektbakgrunn samt hvor de bodde på det tidspunktet de deltok i eksperimentet. Utfra dialektbakgrunn ble deltakerne delt inn i fire dialektgrupper, basert på den klassiske firedeelingen østnorsk, vestnorsk, trøndersk og nordnorsk. I figur 2 viser vi hvordan de tre genusgruppene er fordelt på de fire dialektområdene.



Figur 2: Størrelsen av de tre gruppene i ulike dialektområder.

Kun i det østnorske området ser vi en stor andel to-genustalere (27 av 66 deltakere, 41%). Vi kan se fra deltakernes svar at de fleste i to-genus-gruppen snakker «Oslo»-dialekt, og at de er bosatte enten i Oslo eller Akershus. Fra det vestnorske området ser vi at de fire deltakerne i to-genus-gruppen alle kommer fra Bergen og oppgir at de snakker bergensk. Ingen andre deltakerne oppgir at de snakker denne dialekten. I trøndersk-gruppen ser vi bare tre talere som har et to-genussystem (bosatt i Trondheim eller Oslo). I den østnorske tre-genus-gruppen finner vi først og fremst folk fra Hedmark, Oppland og Telemark, mens den østnorske ustabile gruppen er mer blandet. I de tre andre dialektgruppene finner vi kun 8-10% to-genustalere. Den nordnorske gruppen skiller seg fra trøndersk og vestnorsk ved at det ustabile systemet er det hyppigst forekommende. Dette beror ikke nødvendigvis på at genussystemet er mindre stabilt i Nord-Norge enn i andre deler av landet, men sannsynligvis heller på at vi hadde en større andel yngre deltakere i Tromsø. Majoriteten av deltakere med språktilknytning til Trøndelag eller Vest-Norge hadde et stabilt tre-genussystem. De ustabile deltakerne fra Vest-Norge er en gruppe der de fleste har flyttet til større byer (fire til Oslo, en til hhv. Bergen, Trondheim og Halden).

Utvalget av deltakere kan dessverre ikke sies å være representativt for den norske befolkningen: yngre deltakere fra Nord-Norge er klart overrepresentert, og vi har svært få deltakere fra mange fylker i Vest- og Øst-Norge. Til tross for at vi mangler systematiske data om alder og at deltakergruppen ikke er helt re-

presentativ for det norske språklandskapet, tør vi likevel trekke noen konklusjoner: (1) et fullt utviklet to-genussystem er en sjeldenhet utenfor Oslo og Bergen, (2) et ustabil system er vanlig rundt og i Trondheim og Tromsø blant yngre talere, samt på Østlandet (framfor alt rundt Oslo), og (3) et stabilt tre-genussystem er fortsatt det vanligste systemet på Vestlandet og Trøndelag, men sannsynligvis også blant eldre talere i Nord-Norge (med unntak av Finnmark, der vi ikke har mye data). Vi vil i denne studien bruke gruppeinndelingen fra Klassen mfl. og mer nøyaktig utforske variasjonen vi finner innenfor og mellom gruppene.

Det andre funnet fra Klassen mfl. (2022) av betydning for denne artikkelen er korrelasjonen mellom klassifiseringsmønster og skriftspråkmønster. Det viser seg at de hunkjønnsordene som oftest ble kategorisert som *en*-ord i eksperimentet, er de samme ordene som oftest brukes med bestemthetssuffikset *-en* i bokmålskorpuset NoWaC. Det finnes altså en korrelasjon mellom genusklassifiseringen og en form brukt i korpuset. Det er interessant at dette gjelder bestemthetssuffikset og ikke den utvetydige hunkjønnsmarkøren *ei* (i motsetning til *en*). Det tredje funnet gjelder en korrelasjon mellom korpusfrekvenser og reaksjonstider. Klassen m.fl. (2022) fant at den stabile tre-genusgruppen brukte lengre tid for å klassifisere hunkjønnsord som ofte realiseres med typiske hankjønnsmarkører i skriftlig bokmål. I dette tilfelle handlet korrelasjonen igjen om bestemthetssuffikset *-en* og ikke den frittstående artikkelen *ei/en*. Klassen mfl. tolket dette funnet som et tegn på at genussystemet til den stabile gruppen også var påvirket av variasjonen som finnes innenfor språksamfunnet. I denne studien ser vi nærmere på disse tre funnene, og bringer inn nye tall fra korpus for å få bedre forståelse av gyldigheten av resultatene.⁶

4. Denne studien: forskningsspørsmål og hypoteser

Klassen mfl. (2022) fant at det ikke var tilfeldig om et spesifikt hunkjønnsord klassifiseres som han- eller hunkjønnsord: responsmønstrene korrelerte med

6. Klassen mfl. (2022) fokuserer mer på forskjell i reaksjonstid mellom de to gruppene. I korthet var funnene følgende: Språkbrukere som konsekvent brukte alle tre kjønn (hankjønn, hunkjønn og intetkjønn) hadde betydelig raskere reaksjonstider med hunkjønn og intetkjønn enn med hankjønn. De som bare brukte hankjønn og intetkjønn, viste ingen forskjell i reaksjonstid med hensyn til genus, mens de som brukte alle tre genus, men som ikke var konsekvente i bruken av hunkjønn, hadde betydelig raskere reaksjonstider med intetkjønn enn med både hankjønn og hunkjønn. Vi henviser til Klassen mfl. (2022) når det gjelder hva disse forskjellige resultatmønstrene kan si om hvordan variasjon i morsmålet reflekteres i representasjonsnivået i det mentale leksikonet.

korpusfrekvenser fra skriftlig bokmål. Klassen mfl. fant også en korrelasjon mellom reaksjonstider og korpusfrekvenser. I denne studien ser vi mer detaljert på disse korrelasjonene og utvider studien ved å sammenligne resultatene fra klassifiseringsstudien ikke bare med skriftspråkmønster, men også med talemålsmønster.

Denne artikkelen har som sitt fremste mål å belyse språkforandringens dynamikk, men vi vil samtidig forsøke å gi et bilde av to-genussystemets framvekst i det norske språksamfunnet, i leksikonet og i det morfosyntaktiske systemet. Når et språk er i endring, er det ikke uvanlig at høyfrekvente ord beholder det gamle mønsteret lenger (se diskusjon om *entrenchment*/forankring i f.eks. Bybee 2006), men i den pågående genusendringen i norsk er det mulig at vi vil se et motsatt mønster: høyfrekvente ord kan i prinsippet endre seg tidligere, på grunn av at språkbrukerne ofte møter dem i f.eks. skriftlig tekst og media i den «nye» maskuline formen. I stedet for å ha en beskyttende effekt, kan da høy frekvens ha en akselererende effekt. Ved å sammenligne effekten av generell høy frekvens og effekten av høy frekvens av «feil» genusmarkør (hankjønn i stedet for hunkjønn), kan vi få bedre forståelse av spredningen til språkendringen. Vi vil i denne artikkelen se direkte på hvordan ulike typer av genusmarkører (artikkel og suffiks), og ulike typer av input (talespråk og skriftspråk) påvirker klassifisering og prosessering av substantiver hos norske språkbrukere.

Basert på tidligere litteratur (Trosterud 2001), behandler vi hankjønn som umarkert kjønn: språkinnlæreren antar at hvert nytt ord hun møter, tilhører hankjønnsklassen, og bare gjennom eksplisitte markører vil språkinnlæreren kunne klassifisere substantiver som intetkjønn eller hunkjønn. Det er imidlertid ikke klart hva disse eksplisitte markørene er. Vi vet at ubestemte artikler gir pålitelige bevis for skillet mellom maskulin og feminin, men vi er mindre sikre på hvilken rolle bestemthetssuffikset har, gitt at det ikke er en direkte korrelasjon mellom suffikset og ubestemt artikkel, hverken i språktilegnelsesprosessen (Rodina & Westergaard, 2013, 2015) eller ved språkendringer (f.eks. Lødrup, 2011, Conzett mfl. 2011, Lohndal & Westergaard, 2016). Språkinnlæreren kan også i prinsipp bruke formen på stammen for å tilordne genus på ordet; som påpekt i Trosterud (2001) og Urek mfl. (dette nummeret) er stammefinalt *-e* en relativt pålitelig feminin markør (dette til tross for at de fleste *e*-finale substantiver faktisk er hankjønnsord).

Som læringskurve kan vi anta at alle substantiv i starten er klassifisert som umarkerte hankjønnsord. Høyfrekvente ikke-maskuline ord bør tilegnes tidlig, så lenge genusmarkørene er transparente (og hyppige). Når innlæreren får

tilgang til skriftspråket vil likevel frekvensen av genusmarkører endres: de synlige hunkjønnsmarkørene blir nå mye sjeldnere. Innlæreren får også motstridende input som vil påvirke innlæringsprosessen av genussystemet. Stabiliteten til genussystemet kan nå bli mer påvirket av mengden åpenbare genusmarkører i skriftspråket i stedet for talespråket. Ved å sammenligne muntlige og skriftlige språkfrekvenser mot genusklassifisering, vil vi utforske den potensielle rollen til de to modusene for språkendringsprosessen, og ved å sammenligne ubestemt artikkelfrekvens og frekvensen av bestemthetsuffixer, håper vi å forstå hva språkbrukere faktisk tolker som pålitelige markører for genusklassifisering av substantiver. Som vi skrev i avsnitt 2, behøver dette ikke være det samme som morfologiske eksponenter av grammatisk genus: en språkinnlærer kan bruke fonologiske og semantiske trekk i input for å lære seg et genussystem, uten at disse trekkene i seg selv er uttrykk for genus.

Vi kan i prinsippet forvente oss at genustilordning i et ustabil system styres først og fremst av semantiske faktorer; f.eks. at prototypisk *feminine* ord (*jente*, *dame*) brukes oftere med hunkjønnsmarkører. Tidligere studier har likevel ikke funnet at semantiske trekk spiller noen stor rolle (Rodina & Westergaard 2015). Materialet til studien i Klassen mfl. inneholder kun ikke-animate substantiver, og vi kan derfor ikke undersøke disse semantiske trekkene. Fra et kvantitativt perspektiv kan vi imidlertid forvente oss at en eller flere av følgende faktorer vil korrelere med den variasjonen i genustilordning vi finner i eksperimentet:

1. Total frekvens, alle former uavhengig av genusmarkør: jo vanligere et ord er, desto oftere aktiveres ordet i det mentale leksikonet. Vi kan regne med at genus aktiveres hver gang et ord aktiveres, uavhengig av om genus markeres eksplisitt på substantivet eller andre ord i substantivfrasen; det vil si at genus aktiveres også når et substantiv står i flertall, i entall uten artikkel, eller i bestemt form uten artikkel. Om dette stemmer, vil total frekvens korrelere sterkt med genustilordning i et variabelt system.
2. Frekvens av ord som forekommer med eksplisitt hunkjønnsmarkering på ubestemt artikkel (*ei*), eller hunkjønnsrelatert bestemthetsuffix (-a): Vi kan hypotetisk anta at genus bare aktiveres av eksplisitte genusmarkører. Jo oftere disse aktiveres, desto mer stabil genustilhørighet bør ordet ha. Vi kan her sammenligne effekten av den utvetydige hunkjønnsartikkelen *ei* med bestemthetsuffixet *-a*. Et potensielt problem med denne mulige korrelasjonen er at det er stor forskjell mellom ulike substantiver når det gjelder fordelingen av enkelte ordformer (f.eks. ordet

grense opptrer oftest i bestemt form entall, *skulder* opptrer ofte i possessivkonstruksjoner, mens ord som *bok* og *tavle* fordeler seg mer jevnt over ulike former).

3. Frekvens av hunkjønnsmarkører i relasjon til hankjønnsmarkører: vi kan anta at det er frekvensen av hankjønnsmarkører på tradisjonelle hunkjønnsord i forhold til frekvensen av hunkjønnsmarkører som best korrelerer med resultatene i eksperimentet vårt, det vil si, andelen av hankjønnsmarkører av alle genusmarkører per ord.

For å undersøke en mulig effekt av de ulike faktorene i 1-3 har vi hentet ut frekvenser for de 31 hunkjønnsordene i eksperimentet fra Klassen mfl. fra NoWaC-korpuset (Guevara 2010) og Nordisk dialektkorpus (Johannessen mfl. 2009). Fem ulike mål ble hentet ut: (1) Lemma-frekvens, (2) *ei* fulgt av substantiv, (3) *en* fulgt av substantiv, (4) bestemt form med suffikset *-a* og (5) bestemt form med suffikset *-en*. I tillegg har vi søkt på strenger med et eller to adjektiver mellom artikkel og substantiv. Ut fra målene 2-4 har vi kalkulert andelen hankjønnsmarkører (artikkel og suffiks) for hvert substantiv i hunkjønnsklassen. Vi har sammenlignet korpusresultatene med eksperimentresultatene ved hjelp av lineære modeller i lme4-pakken i R (Bates mfl. 2015). Vi tester følgende variabler:

1. Lemma-frekvens i (a) skriftlig korpus og (b) muntlig korpus.
2. Frekvens ubestemt artikkel *ei* i (a) skriftlig og (b) muntlig korpus.
3. Frekvens ubestemt artikkel *en* i (a) skriftlig og (b) muntlig korpus.
4. Frekvens bestemthetssuffiks *-a* i (a) skriftlig og (b) muntlig korpus.
5. Frekvens bestemthetssuffiks *-en* i (a) skriftlig og (b) muntlig korpus.
6. Andel hankjønnsmarkering av alle ubestemte former (*en* vs. *ei*) i skriftspråkskorpus.
7. Andel hankjønnsmarkering av alle bestemte former (*-en* vs. *-a*) i skriftspråkskorpus.
8. Totalt antall hankjønnsmarkering (*ei* og *-a*) i talespråkskorpus.
9. Form på stammen: (stamme som ender på *-e*).

I neste avsnitt sammenligner vi genusbruk i de korpusene vi bruker. Deretter undersøker vi hvilke av variablene i listen over som korrelerer med genustilordning i den ustabile gruppen i klassifiseringsstudien (Klassen mfl.), og deretter ser vi hvilke variabler som korrelerer med reaksjonstid.

5. Korpusstudie

I dette avsnittet beskriver vi en korpusstudie som vi skal sammenligne med resultatene fra Klassen mfl. i avsnitt 6. I korpusstudien har vi undersøkt de samme ordene som i Klassen mfl. og sett på variasjonen i genusmarkering, med fokus på hunkjønnsordene (se liste over disse substantivene i figur 4). Vi har tatt utgangspunkt i to korpus: (1) NoWaC-korpuset (Guevara 2010), et web-basert korpus med 700 millioner ord fra diverse web-kanaler inklusive blogg-tekster, innsamlet i 2009-2010, og (2) Nordisk dialektkorpus (NDC, Johannessen mfl. 2009), et korpus med talemålsmateriale fra de nordiske landene som inneholder innspillinger fra 110 steder i Norge. Størrelsen på den norske delen er cirka to millioner ord fra yngre og eldre talere (cirka 900.000 ord fra yngre talere, 18-30 år). Dialektkorpuset har mange fordeler: materialet består av talemål, og alder og dialekt for hver enkelt taler kan lett finnes fram. Ulempen er at materialet ikke alltid er stort nok for å finne mønster for enkelte ord. Som vi vil se, kan en kombinasjon av talespråkskorpuset og web-korpuset fange viktige mønster i språkforandringen. Fra begge korpusene har vi sammenstilt statistikk over lemma-frekvens og bruk av han- og hunkjønnsformer for de 31 ordene i studien.

Lemma-frekvensen for de 31 ordene til sammen i NoWaC-korpuset er over en million, men i denne studien vil vi først og fremst diskutere entallsformer i bestemt eller ubestemt form (sammen med ubestemt artikkel). I NDC finner vi sammenlagt 2790 treff på lemma-nivå. I tabell 2 viser vi frekvensen for hunkjønns- og hankjønnsformer i entall fra de to korpusene. For de ubestemte formene har vi inkludert treff der ikke mer enn to adjektiver står mellom artikkelen og substantivet.

Tabell 2: Antall og prosent av ubestemt artikkel ei/en og bestemt suffiks -a/-en for de 31 ordene i testen, i skriftspråkskorpus (NoWaC) og talespråkskorpus (NDC).

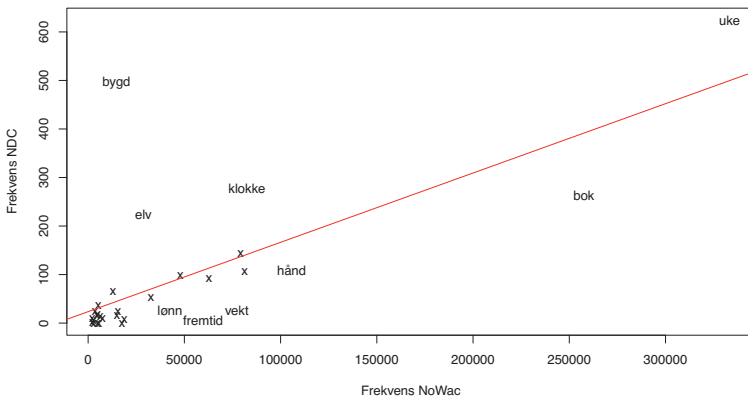
	#ubest. ei	#ubest. en	%ubest. ei	#best. -a	#best. -en	%best -a
NoWaC	5134	57047	8,25%	191935	342082	35,9%
NDC	221	44	83,4%	1074	157	87,2%

Fra tabell 2 kan vi notere oss at hunkjønnsmarkører er forholdsvis sjeldne i skriftspråket (bokmål), mens de brukes nesten kategorisk i dialektalt talemål. Forskjellen er størst for den ubestemte artikkelen, der *ei* kun brukes i 8,25% av tilfellene i NoWaC, sammenliknet med 83,4% i NDC. I klassifiseringsstudien

ble 63,7% av disse 31 ordene klassifisert som hunkjønn, og tallene derfra likner derfor mest på det muntlige korpuset (NDC). Forskjellen mellom tallene i NDC og Klassen mfl. kan nok best forklares som forskjeller mellom deltakergrupper: de fleste opptakene i NDC kommer fra dialekttalere i små bygder, mens mange av deltakerne i klassifiseringsstudien kommer fra større byer, med utviklede to-genussystem eller ustabile tre-genussystem. Fra resultatene i tabell 2 kan vi også notere oss at bestemte substantiver i entall er mye hyppigere enn ubestemte substantiver (cirka fem ganger vanligere i talemål og åtte ganger vanligere i skriftspråk for ordene i studien). Vi kommer tilbake til dette senere i artikkelen.

Til tross for at de to korpusene er forskjellige med hensyn til bruk av hunkjønnsmarkører, finner vi likheter mellom dem. Først og fremst er det en korrelasjon i lemma-frekvens mellom de to korpusene ($r = 0,69$, $p < 0,001$); det vil si at de ordene som er hyppige i talemål, også er hyppige i skriftspråk. Merk likevel at det finnes en del avvik her; for eksempel er lemmafrekvensen for *bygd*, *elv* og *klokke* høyere enn forventet i NDC, gitt frekvensene i NoWaC. Vi viser korrelasjonen mellom NDC-frekvenser og NoWaC-frekvenser i figur 3, der vi tar ut avvikende lemma (det vil si, lemma som havner langt fra regresjonslinja).

Figur 3: Korrelasjon mellom talespråkkorpus (NDC) og skriftspråkkorpus (NoWaC) i lemma-frekvens for de 31 ordene i studien (Klassen mfl.).

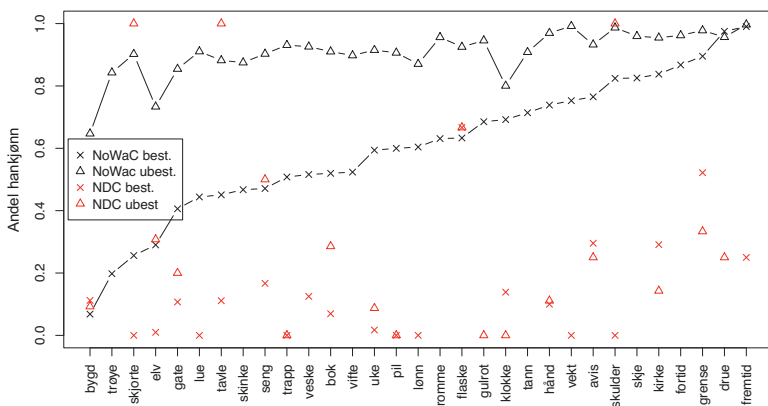


Det er vanskelig å korrelere bruk av hunkjønnsmarkører for enkelte ord mellom de to korpusene. Dette skyldes først og fremst at datamengden fra NDC ikke er tilstrekkelig stor. Flere av ordene i studien nevnes ikke en eneste gang i bestemt

form entall eller ubestemt form entall med artikkel, for eksempel nokså vanlige ord som *tromme* og *gulrot* (som forekommer i flertallsformer og som nakne nominaler i NDC). I NoWaC-korpuset finner vi stor variasjon i andelen hunkjønnsmarkører, framfor alt for bestemthetsuffixet (*-a/-en*): et ord som *bygd* brukes ofte med bestemthetsuffixet *-a*, mens ord som *drue* og *fremtid* nesten alltid forekommer med *-en*.

Siden ubestemt artikkel *ei* forekommer sjeldnere i NoWaC, ser vi også mindre variasjon i valg av *ei/en*. I figur 4 viser vi variasjonen i valg av bestemt artikkel sammen med valg av ubestemt artikkel fra NoWaC, her i form av andelen hankjønnsmarkører (*en/-en*) for hvert ord. I tillegg viser vi andelen hankjønnsmarkører for bestemte og ubestemte substantiver fra NDC. Legg merke til at det mangler data for enkelte ord, og at andelene iblant bygger på kun en eller to observasjoner, hvilket gjør talespråksandelene mindre pålitelige.⁷

Figur 4: Andel hankjønnsmarkører i bestemte og ubestemte substantivfraser for hunkjønnordene i studien (Klassen mfl.), fra NoWaC og NDC.



Det er uklart hva som ligger bak forskjellene mellom ordene når det gjelder valg av suffiks i skriftspråkskorpuset. Vi ser ingen effekt av formen på ordet. Eksempelvis har ord som slutter på *-e*, noe som er et relativt pålitelig trekk for hunkjønn, og ord som slutter på konsonant like stor sannsynlighet for å realiseres med *-a* i bestemt form entall. Det er mulig at forskjellen snarere er

7. Følgende ord mangler enten bestemte eller ubestemte entallsformer i NDC: *fortid*, *fremtid*, *gulrot*, *lønn*, *lue*, *skinke*, *skje*, *tann*, *tromme*, *trøye*, *vekt*, *veske*, *vifte*.

en effekt av stil og register: ord som *grense*, *drue*, *kirke* og *fremtid* brukes sannsynligvis i mer formelle situasjoner sammenliknet med *bygd*, *trøye*, *lue* og *elv*.

Til slutt vil vi si noe om effekten av dialekt og alder ut fra NDC-dataene. De yngre talerne i NDC-korpuset bruker *en* i 23% av forekomstene av hunkjønnordene i studien, mens de eldre kun bruker *en* i 10% av tilfellene. Disse 10% er heller ikke jevnt distribuert over Norge, men det er hovedsakelig i Nord-Troms og Finnmark vi finner *en* med hunkjønnord, samt i og rundt Oslo og Bergen. Blant de yngre talerne er spredningen av *en* noe større, men de fleste belegg er fortsatt fra Nord-Troms og Finnmark.

6. Sammenligning av klassifiseringsstudien og korpus

Vi så i avsnitt 4 at det finnes stor variasjon mellom talere og grupper av talere i Norge, og at denne variasjonen i stor grad er styrt av geografi og alder. Videre så vi i avsnitt 5 at det også finnes stor variasjon mellom ord i skriftspråket: noen hunkjønnord brukes veldig sjelden med hunkjønnmarkører, mens andre ord nokså ofte har hunkjønnmarkører, i det minste når det gjelder bestemthets-suffikset. Vi fant også et stort avvik mellom bruken av feminine markører i det talte og skriftlige språket. I denne delen vil vi sammenligne variasjonen i klassifiseringsstudien og resultatene i korpuset, noe vi mener kan gi oss en bedre forståelse av endringsprosessen generelt. Som nevnt innledningsvis, er hyppige leksikalske elementer ofte mindre påvirket av endring enn lavfrekvente ord. Ut fra dette, kunne vi altså forvente at hyppige feminine substantiver mer pålitelig ville bli klassifisert som hunkjønnord. Som også nevnt, vil imidlertid hyppige ord også kunne endres tidlig i prosessen, fordi språkbrukere ofte eksponeres for disse både i skrift og tale (se korrelasjon mellom talte og skriftlige lemmafrekvenser i tabell 2). Siden skriftspråket (bokmål) av historiske årsaker hovedsakelig benytter maskuline genusmarkører for hunkjønnord, vil språkbrukere oftere bli eksponert for høyfrekvente hunkjønnord med hankjønnformer enn substantiver med lavere frekvens i input. Nedenfor undersøker vi hvordan både talespråkfrekvenser og skriftspråkfrekvenser korrelerer med både genusklassifisering og reaksjonstidene fra Klassen mfl.

6.1 Faktorer som styrer genustilordning i den ustabile gruppen

I dette avsnittet vil vi bare se på variasjonen hos de såkalte ustabile talerne, det vil si de som veksler mellom hankjønn og hunkjønn i valg av artikkel for (tidligere) hunkjønnord. Gjør disse talere tilfeldige valg av genus, eller styres

valget av leksikalske og/eller morfologiske faktorer? Enkelt sagt: er alle hunkjønnord rammet i like stor grad av språkendringen, altså skjer spredningen i jevn takt gjennom hele leksikonet, slik som antydnet i tidligere studier, f.eks. Rodina & Westergaard (2015), som riktignok behandlet et svært begrenset antall hunkjønnord? Når vi ser nærmere på de individuelle hunkjønnssubstantivene i studien, ser vi en nokså stor variasjon i genustilordningen. Et ord som *bygd* blir klassifisert som hankjønn i så lite som 16% av tilfellene, mens ordet *skulder* klassifiseres som hankjønn i 60%. De andre ordene fordeles jevnt mellom disse to ytterpunktene.

Et interessant spørsmål er om svarmønstrene kan forklares ut fra ulike korpusfrekvenser og morfologiske egenskaper ved de ulike ordene. Vi undersøker dette gjennom statistiske tester, der vi ser på om ulike korpusfrekvenser korrelerer med valg av genus (*ei* eller *en*) i klassifiseringstesten. Vår avhengige variabel i den statistiske modellen er genusvalg, her kodet som 1 (hunkjønn, «korrekt») og 0 (hankjønn), og prediktorene er de som er gitt i 1-9 i avsnitt 4. Legg merke til at vi ikke tar med andelen av hankjønnsmarkører fra NDC som en variabel, ettersom datamaterialet ikke er stort nok. Vi velger også å slå sammen NDC-tallene for *ei* og *-a* (variabel 8) – dette for å få et større og mer pålitelig materiale. Som vi har notert tidligere, mangler vi tall på bestemte eller ubestemte former for noen av hunkjønnordene.

Korrelasjon er naturligvis ikke det samme som kausalitet, og frekvensene vi finner i korpusene må nødvendigvis være et resultat av en endring som er forårsaket av faktorer som er antydnet i tidligere forskning (f.eks. sosiolingvistiske faktorer, synkretisme, språktilegnelse; se Lødrup 2011, Rodina & Westergaard 2015, Busterud mfl. 2019). Vi vil imidlertid hevde at dynamikken i endringsprosessen kan påvirkes av frekvenser i talt og skriftlig input, slik at variasjonen i genusvalg i Klassen mfl. kan sies å være et resultat av disse frekvensene.

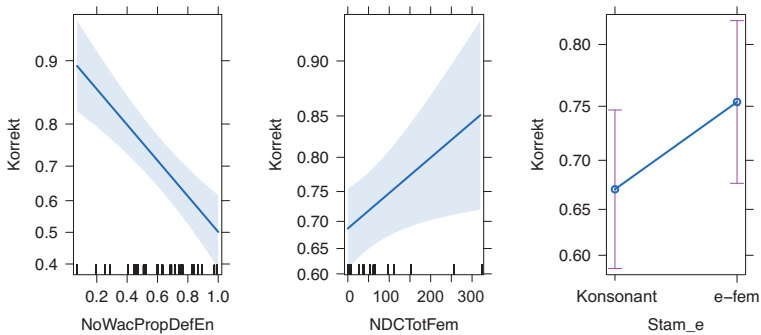
Vi tester først om de ulike variablene hver for seg korrelerer med valg av artikkel, og deretter setter vi opp en modell med samtlige variabler som har en effekt på resultatet og eliminerer stegvis de variabler hvis effekt dekkes bedre av andre variabler (det vil si, om to variabler *x* og *y* korrelerer med hverandre, kan vi eliminere *y* hvis den ikke forklarer noen varians som ikke *x* forklarer). I modellene har vi deltakere og ord («item»/ord) som tilfeldige variabler. Når vi tester variablene hver for seg, finner vi totalt fem variabler som korrelerer signifikant med genusvalget i studien: (1) andel bestemthetssuffiks tilsvarende hankjønn (*-en*) i NoWaC ($p < 0,001$), (2) andel hankjønnformer for ubestemt artikkel (*en*) i NoWaC ($p < 0,001$), (3) totalt antall hunkjønnformer (*ei*, *-a*) i

NDC ($p < 0,05$), (4) totalt antall ubestemt artikkel *ei* i NDC ($p < 0,05$) og (5) form på stammen. Verdiene 1 og 2 korrelerer sterkt ($r = 0,81$, $p < 0,001$), og likeså verdiene 3 og 4 ($r = 0,79$, $p < 0,001$). I en modell med samtlige fem prediktorer viser det seg at kun variabel 1, 3 og 5 har en effekt på genusvalget. Den endelige modellen gis nedenfor, fulgt av koeffisienter, standardfeil og *p*-verdier for de tre variablene.

Tabell 3: Koeffisienter, standardfeil og *p*-verdier fra modellen $glmer(\text{correct} \sim \text{AndelBestEnNoWaC} + \text{Stam_e} + \text{TotFemNDC} + (1|\text{deltaker}) + (1|\text{Stimulus}))$.

	Beta	Std.feil	z-verdi	Pr(> z)
(Intercept)	1,9729	0,3559	5,54	3,0e-08 ***
AndelBestEnNoWaC	-2,2852	0,4516	-5,06	4,2e-07 ***
Stam_e	0,4059	0,1941	2,09	0,037 *
TotFemNDC	0,0298	0,0132	2,26	0,024 *

Vi illustrerer effekten av de ulike variablene i figur 5.



Figur 5: Den estimerte korrelasjonen mellom (a) andel hunkjønnssuffiks (-en) i skriftspråk (NoWaC), (b) totalt antall hunkjønnsmarkører i talespråk (NDC), og (c) form på stammen (-e-final vs. konsonantfinal stamme) og resultatet i klassifiseringsstudien hos den ustabile gruppen.

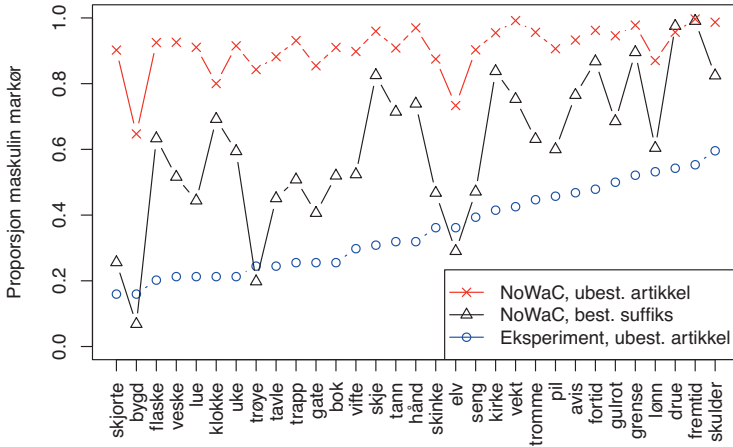
Her ser vi en klar negativ korrelasjon mellom høy andel *-en* som bestemthets-suffiks i skriftspråkkorpuset og klassifisering av (tidligere) hunkjønnssord som hunkjønn. Vi ser en motsatt effekt for talemålsfrekvensen av hunkjønnssord med synlige hunkjønnsmarkører (ubestemt *-ei* eller bestemt *-a*). Videre ser vi at hun-

kjønnsord som slutter på *-e* oftere blir klassifisert som hunkjønnsord. Vi ser derimot ingen direkte effekt av lemma-frekvens, hverken fra NDC eller NoWaC.

Disse resultatene tyder på at det som styrker eller svekker hunkjønnsklassifisering er input av eksplisitte genusmarkører, og ikke bare aktivering av hunkjønnsord. Dette kan tyde på at genusinformasjon ikke automatisk aktiveres når substantiver og substantivfraser prosesseres av språkbrukeren, men hovedsakelig når genusmarkører er synlige (men se Cubelli mfl. 2005 og Paolieri mfl. 2011 for resultater som viser at substantiver med og uten synlige markører aktiverer genus i italiensk, dog i større grad for ord med synlige genusmarkører). En annen mulig tolkning er at i denne allerede ustabile gruppen vil både hankjønn og hunkjønn aktiveres når en taler blir eksponert for et (tidligere) hunkjønnsord. Resultatene antyder også at bestemthetssuffikset (*-a*, *-en*) aktiverer genustrekk: det er spesielt ord med høy andel av bestemthetssuffikset *-en* i NoWaC som deltakerne i testen ofte klassifiserer som hankjønnsord. Likeså er det ord som hyppig forekommer i talespråk med bestemthetsmarkering *-a* og/eller ubestemt artikkel *ei* som deltakerne oftere klassifiserer som hunkjønn. Merk også at disse to målene, NoWaC *-en* og talespråk *-a/ei*, ikke korrelerer med hverandre ($r < 0,2$, n.s.).

Disse resultatene viser at de substantivene som hyppigst forekommer med bestemthetssuffikset *-en*, er kommet noe lenger i endringsprosessen fra hunkjønn til hankjønn enn andre substantiver. Men om man aksepterer at frekvenser i input påvirker dynamikken i endringsprosessen, kan man også hevde at resultatene antyder at frekvens av bestemthetssuffiks faktisk har en større effekt enn frekvensen av ubestemte artikler. Dette skyldes sannsynligvis to faktorer. For det første er bestemthetssuffikset betraktelig mer frekvent enn den ubestemte artikkelen, hvilket gjør de statistiske beregningene mindre sikre for den ubestemte artikkelen, først og fremst i NDC-materialet. Videre er det mindre variasjon ved de ubestemte artiklene i NoWaC: hunkjønnsordene i eksperimentet realiseres i over 90% av tilfellene med artikkelen *en* (kontra *ei*) i ubestemt form, mens de realiseres med *-en* (kontra *-a*) i bestemte kontekster i cirka 65%.

I figur 6 illustrerer vi variasjonen for de enkelte (tidligere) hunkjønnsordene i klassifiseringsstudien for den ustabile gruppen. Den blåprikkete linjen illustrerer i hvor stor andel av hvert hunkjønnsord deltakerne valgte hankjønnsartikkel *en* i stedet for *ei*. Figuren viser også valg av ubestemt artikkel (*en/ei*, linje med røde kryss) og form av suffiks (*-en/-a*, linje med svarte trekkanter), som tilsvarer mål 3 ovenfor.



Figur 6: Klassifisering av hunkjønnordene i eksperimentet hos den ustabile gruppen (blå linje), sammenliknet med korpustall.

For å sammenfatte dette avsnittet kan vi si at suffikset *-en* på typiske hunkjønnord kan skape usikkerhet hos språkbrukerne, noe som leder til variasjon. Hunkjønnord som er høyfrekvente i en talemålsdiskurs er mindre utsatte, men dette er tilfelle kun dersom synlig genusmarkering (*ei*, *-a*) er til stede. Resultatene viser også at talere av norsk til en viss grad tolker stammefinal *-e* som tegn på hunkjønnstilhørighet.

6.2 Faktorer som påvirker reaksjonstiden ved klassifisering av hunkjønnord

I avsnittet ovenfor så vi korrelasjoner mellom klassifisering i eksperimentet og korpusfrekvenser av genusmarkører i norsk (uformelt) skriftspråk og talemål. Dette er ikke i seg selv veldig overraskende: de deltakerne vi har i eksperimentet, tilhører samme populasjon som de som skriver og leser blogger, men det bør også påpekes at språket som disse talerne støter på i skriftspråket (bokmål) nok skiller seg fra deres talemål-input og talemål-produksjon. Vi kan være nokså sikre på at de ustabile talerne bruker bestemthetsmarkøren *-a* mer eller mindre konsekvent i talespråket sitt. Vi er nå interessert i å finne ut om de to stabile gruppene også kunne være påvirket av de samme faktorene. Talere i disse to gruppene er også eksponert for variasjon, fra skriftspråk, TV og radio. I de to stabile gruppene ser vi likevel kategoriske mønstre i svarene, så her må vi i stedet se på reaksjonstid (se Klassen mfl. for en mer nøyaktig diskusjon av

reaksjonstider fra dette eksperimentet). Vi kan i prinsippet forvente oss at de to stabile gruppene er helt upåvirket av skriftspråksfrekvenser: de har stabile systemer, som for to-genus-gruppen også likner på skriftspråket (bokmål). På den annen side kan man forvente at den variasjonen som åpenbart finnes på nasjonalt nivå og som alle talere er eksponert for (i varierende grad) vil skape en viss ustabilitet i individets språkssystem. Denne ustabiliteten skulle da kunne vise seg i lengre reaksjonstider. Tre-genus-gruppen forventes da å ha lengre reaksjonstider for hunkjønnsord som ofte realiseres med typiske hankjønnsmarkører (*en/-en*), mens to-genus-gruppen muligens ville ha problemer med ord som ofte realiseres med hunkjønnsmarkører (*ei/-a*).

For å undersøke mulig innflytelse av korpusfrekvenser og substantivform, satte vi opp statistiske modeller med variablene 1-9 (se seksjon 6) som avhengige variabler, på samme måte som i modellen for genusklassifisering i den ustabile gruppen, men nå med reaksjonstid i stedet for klassifisering som uavhengig variabel.⁸ Vi laget separate modeller for tre-genus- og to-genus-gruppen. For to-genus-gruppen finner vi kun en signifikant effekt av talespråksfrekvens, og denne er sterkest for entallsformene (bestemt eller ubestemt). Det er generelt lite variasjonen i to-genusgruppen i reaksjonstid for hunkjønnsordene, og den variasjonen som finnes er først og fremst variasjon mellom deltakerne. For tre-genus-gruppen finner vi at de samme korpusmålene som påvirker klassifisering i den ustabile gruppen, også har en effekt på reaksjonstiden. Formen på stammen (final *-e*) er kun marginalt signifikant for reaksjonstiden for tre-genus-gruppen. Vi gir koeffisienter, standardfeil og p-verdier for de to stabile gruppene i tabell 4. Her inkluderer vi også verdiene for effekten av stammeformen, for å gjøre sammenligningen med den ustabile gruppen (tabell 3) tydeligere (merk at frekvenser og reaksjonstid er log-transformerte, se fotnote 3).

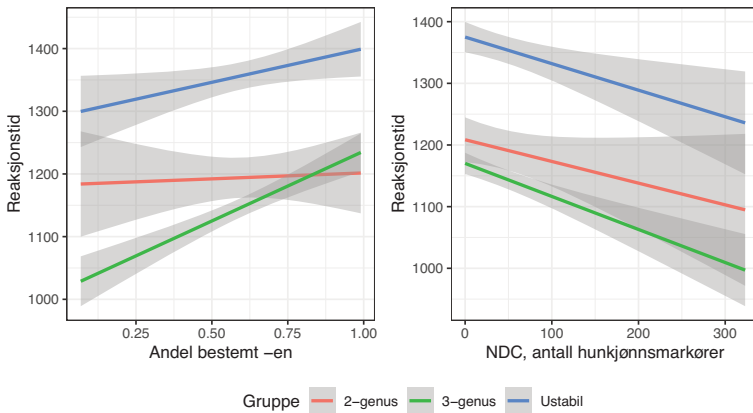
I figur 7 illustrerer vi effekten av andelen *-en* som bestemthetsuffix og NDC-frekvens av substantiv(fraser) med hunkjønnsmarkører på reaksjonstidene for de tre gruppene. Vi ser at det første målet, *-en* som bestemthetsuffix, korrelerer klart med tre-genus-gruppen, men ikke med to-genusgruppen (venstre graf), og at talemålsfrekvensen av genusmarkører (NDC) viser en svak korrelasjon i samtlige grupper (kun signifikant for to-genus-gruppen).

To-genus-gruppen viser derimot ikke noen tegn til å være påvirket av frekvensen av *-a/ei* i skriftspråket. Det er rimelig å anta at de stabile to-genus-

8. Den avhengige variabelen reaksjonstid samt samtlige korpusfrekvenser er log-transformert i modellen. Vi har også tatt bort reaksjonstider på under 300 ms og over 3000 ms. Dette gjøres for å oppfylle antagelser om normalitet og homogenitet for den statistiske modellen.

Tabell 4: Koeffisienter, standardfeil og p -verdier fra modellen $lmer(korrekt \sim AndelBestEnNoWaC + Stam_e + TotFemNDC + (1|deltaker) + (1|Stimulus))$, for tre-genus-gruppen og to-genus-gruppen.

	Tre-genus-gruppen (#2829 obs.)			To-genus-gruppen (#1438 obs.)		
	Beta	Stand.feil	p	Beta	Stand.feil	p
Intercept	6,96	0,04	<0,001	6,97	0,05	<0,001
AndelBestEnNoWaC	0,12	0,05	0,019	0,008	0,06	0,9
log(TotFemNDC)	-0,015	0,01	0,001	-0,018	0,007	0,013
Stam_e	-0,04	0,02	0,05	0,02	0,026	0,43



Figur 7: Effekten av proporsjonen av suffikset *-en* (i forhold til *-a*) i skriftspråk og frekvensen av hunkjønnsmarkører i talespråk på reaksjonstid for de tre gruppene.

talerne ikke på noen måte i dag assosierer den bestemte *-a*-endelsen med hunkjønn (se Lødrup 2011, Rodina & Westergaard 2015). Det er ikke særlig overraskende at vi ikke ser noen effekt av frekvensen av den ubestemte artikkelen *ei* hos disse talere heller: andelen *ei* er så liten i skriftspråk (bokmål) at den neppe vil påvirke talernes grammatikk eller prosessering (i hvert fall ikke om *ei* samtidig sjelden forekommer i talemålet).

Det bør påpekes at andelen av suffikset *-en* (i forhold til *-a*) i korpusdataene korrelerer sterkere med reaksjonstidene enn artikkelen *en* (i forhold til *ei*). Dette tyder på at tre-genustalere faktisk tolker bestemthetssuffikset *-a* som en markør

for genustilhørighet. Vi antar altså at økt variasjon i input – enkelte hunkjønnsord dukker opp med bestemthetssuffikset *-en* i høy grad – leder til usikkerhet hos de «stabile» tre-genustalerne, noe som manifesteres i lengre reaksjonstider og muligens variabel genusproduksjon i en senere fase.

7. Diskusjon og konklusjon

I denne artikkelen har vi vist at vi kan få et relativt godt bilde av variasjon både mellom talere og innen individuelle talere med henhold til grammatisk genus gjennom å sammenligne korpusdata og resultater fra psykolingvistiske tester. Om vi forstår denne variasjonen, kan vi forhåpentligvis lage bedre modeller av språkforandringen. Vi har gjort en grov inndeling av språkbrukere av norsk i tre grupper, og gjennom å sammenlikne resultater fra en klassifiseringsstudie, korpusundersøkelser og tidligere kartlegginger, håper vi nå å være i stand til å gjøre nokså kvalifiserte gjetninger om geografisk spredning samt språklig input, output og grammatikk til disse tre gruppene.

Vi finner en stabil to-genus-gruppe, som geografisk hovedsakelig er lokalisert til Oslo og Bergen, men også til språkkontaktområder i Nord-Norge. Disse talerne får både skriftlig og muntlig input som veldig sjelden inneholder den ubestemte artikkelen *ei*. Derimot vet vi at bestemthetssuffikset *-a* er hyppig forekommende både i skriftlig og muntlig input og output for disse talerne (med unntak for talere fra Bergen). Resultat fra sammenligningen mellom korpusdata og reaksjonstider tyder på at disse talerne ikke på noe vis assosierer vekslingen mellom *-a* og *-en* med grammatisk genus. Disse talerne tolker sannsynligvis vekslingen mellom artiklene *en* og *ei* som stilistisk og vekslingen mellom suffiksene *-en* og *-a* som koblet til bøyningsklasse (se Lødrup 2011; Westergaard & Rodina 2016, 2021; Svenonius 2017). Det samme kan ikke sies om de andre to gruppene, der vi finner en sterk effekt av andelen *-en* på klassifisering av hunkjønnsord. Den stabile tre-genus-gruppen inneholder talere fra stort sett hele Norge, med unntak av de byer og områder som allerede er nevnt. Disse talerne møter et stabilt tre-genus-system i sin daglige muntlige diskurs, men møter et helt annet system i skriftlig input, i hvert fall fra bokmål (som utgjør den største kilden for skriftspråk for alle nordmenn). Utfra NoWaC-resultatene estimerer vi at den ubestemte artikkelen *ei* brukes så lite som i 10% av tilfellene for (tidligere) hunkjønnsord, mens bestemthetssuffikset *-a* brukes i cirka 40%. Skriftlig input påvirker ikke genusklassifiseringen direkte for tre-genus-gruppen, siden disse deltakere er konsekvente i klassifiseringstesten, men det påvirker reaksjonstidene. Her er det viktig at det framfor alt er variasjonen

i bruk av *-en/-a* som suffiks som korrelerer med reaksjonstid: hunkjønnsord som ofte brukes med suffikset *-en* i skriftspråk, tar lengre tid å klassifisere, noe vi tolker som et tegn på usikkerhet. At variasjonen i valg av suffiks gir opphev til usikkerhet, kan da ses som et tegn på at disse talerne ser *-en* som en hankjønnsmarkør og *-a* som en hunkjønnsmarkør.

I vår tredje gruppe, den ustabile gruppen, viser denne usikkerheten seg også i klassifiseringen: ord som disse talerne ofte møter i skriftspråk med bestemt-suffikset *-en* blir ofte klassifisert som hankjønn. Vi ser likevel at ord med en høy talemålsfrekvens av former med genusmarkørene *-a/ei* er mer motstandsdyktige. At variasjonen i valg av bestemt-suffikset bedre forklarer variasjon i klassifiseringstid enn variasjon av ubestemt artikkel (*ei/en*), beror nok først og fremst på at bestemte former er hyppigere enn ubestemte former (med artikkel), og at det er mindre variasjon i valg av artikkel sammenlignet med suffiks i skriftspråk.

Resultatene fra klassifiseringsstudien og korpusstudien, i kombinasjon med resultatene og innsikten fra tidligere studier, kan nå ligge til grunn for en ganske plausibel modell for hvordan hankjønn gradvis går tapt i norske dialekter. Språkinnlæringsprosessen begynner med en antakelse om at substantiv er hankjønn som standardverdi. Bare gjennom input av substantiver og substantivfraser som bærer synlige morfologiske trekk som er forskjellige fra de maskuline markørene, vil språkinnlæreren kunne klassifisere substantiver som hunkjønnsord eller intetkjønnsord. Vi ser dermed de forventede frekvens-effektene i tilegnelse og bruk av spesielt hunkjønnsord: hunkjønnsord som er høyfrekvente i talespråk i former med synlige genusmarkører klassifiseres kjapt som hunkjønnsord. Resultatene våre antyder at språkinnlæreren ser både den ubestemte artikkelen *ei* og bestemt-suffikset *-a* som markører for feminint genus, og sannsynligvis også stammevokal *-e*. Innlæringsprosessen av hunkjønnsord endrer likevel retning når en økende del av inputen kommer fra skriftspråk, der de synlige hankjønnsmarkørene ofte mangler. Ord som er mindre hyppige i en talespråkskontekst sammenliknet med en skriftspråkskontekst blir da hardest rammet, mens ord som allerede er *dypt innlærte* på grunn av høy talespråksfrekvens, klarer seg bedre.

Det store spørsmålet som gjenstår, er da hvorfor visse talere, eller talere i visse områder, utvikler et variabelt to/tre-genussystem, mens andre fortsatt har et stort sett kategorisk tre-genus-system. Som Busterud mfl. (2019) viser, er det ikke helt tilfeldig hvor vi finner de ustabile talerne, nemlig i større byer som Trondheim og Tromsø, og ikke på landsbygda. Dette spredningsmønsteret har også blitt funnet ved andre språkforandringer og har blitt kalt byhopping (*city*

jumping, Taeldeman 2005, Vanderkerckhove 2009). Vi avslutter med å spekulere om hvorfor språkinnlærere i byer har større sannsynlighet for å utvikle et ustabil system enn et kategorisk tre-genus-system (som til tross for at det er «stabil») fortsatt er preget av en viss usikkerhet koblet til ord som hyppig brukes med *-en* i skriftspråk), og også hvorfor det skjer nå. En faktor er at større byer har en betydelig andel talere som ikke snakker den lokale dialekten: språkinnlærere i både Trondheim og Tromsø møter med stor sannsynlighet et inkonsekvent tre-genus-system fra både voksne (barnehageansatte, foreldre til innflyttede venner) og barn. Motstridende former finnes altså med stor sannsynlighet allerede i tidlig muntlig input (se også Trudgill 2013). I denne artikkelen har vi imidlertid argumentert for at variasjonen i skriftlig input kan spille en rolle i genussystemets instabilitet, og vi forventer oss da at både språkinnlærerens første skriftspråk (bokmål eller nynorsk), og hvor mye ethvert individ faktisk leser, påvirker genus-systemet. Slike spørsmål må vi overlate til fremtidig forskning.

Referanser

- Bates, Douglas, Martin Mächler, Ben Bolker og Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using *lm4*. *Journal of Statistical Software* 6: 1–48.
- Bull, Tove, Jorid Hjulstad Junttila og Aud-Kirsti Pedersen. 1986. Nominalfrasen i skibotnålet i Troms. *Norsk Lingvistisk Tidsskrift* 1 (2): 60–71.
- Busterud, Guro, Terje Lohndal, Yulia Rodina og Marit Westergaard. 2019. The loss of feminine gender in Norwegian: A dialect comparison. *Journal of Comparative Germanic Linguistics* 22: 141–167.
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82: 711–733.
- Conzett, Philipp, Åse Mette Johansen og Hilde Sollid. 2011. Genus og substantivbøying i nordnorske språkkontaktområder. *Nordand: nordisk tidsskrift for andrespråksforskning* 6: 35–71.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- Cubelli, Roberto, Lorella Lotto, Daniela Paolieri, Massimo Girelli og Remo Job. 2005. Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language* 53: 42–59.
- Enger, Hans-Olav. 2013. Inflectional change, ‘sound laws’, and the autonomy of morphology. *Diachronica* 30: 1–26.

- Guevara, Emiliano R. 2010. NoWaC: a large web-based corpus for Norwegian. I Kilgarriff og Lin (red.) *Proceedings of the NAACL HLT 2010 Sixth Web as Corpus Workshop*, Los Angeles: Association for Computational Linguistics. 1-7.
- Heide, Eldar. 2003. Felleskjønnet i bergensk. Resultat av mellomnedertysk kvantitetspåverknad? *Maal og Minne* 1: 23-35.
- Hockett, Charles F. 1958. *A course in modern linguistics*. New York: MacMillan.
- Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý og Vít Suchomel. 2013. The TenTen corpus family. *7th International Corpus Linguistics Conference CL*: 125-127.
- Jahr, Ernst Håkon. 1984. Language Contact in Northern Norway. Adstratum and Substratum in the Norwegian, Sami and Finnish of Northern Norway. *Acta Borealia* 1: 103–112.
- Jahr, Ernst Håkon. 2001. Historical sociolinguistics: The role of Low German language contact in the Scandinavian typological split of the late Middle Ages. *Lingua Posnaniensis* 43: 95–104.
- Klassen, Rachel. 2016. *Asymmetric Grammatical Gender Systems in the Bilingual Mental Lexicon*. Ottawa: University of Ottawa.
- Klassen, Rachel; Lundquist, Björn; Westergaard, Marit. L1 grammatical gender variation through the representation in the lexicon. 2022. *Journal of Psycholinguistic Research*. ISSN 0090-6905.
- Lohndal, Terje og Marit Westergaard. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology*. [https://doi:10.3389/fpsyg.2016.00344](https://doi.org/10.3389/fpsyg.2016.00344), 16 mars.
- Lundquist, Björn og Øystein A. Vangsnes. 2018. Language separation in bidialectal speakers: evidence from eye tracking. *Frontiers in Psychology* 9: 1394.
- Lundquist, Björn, Yulia Rodina, Irina Sekerina og Marit Westergaard. 2016. Gender change in Norwegian dialects: Comprehension precedes production. *Linguistics Vanguard* 2: 69–83.
- Lødrup, Helge. 2011. Hvor mange genus er det i Oslo-dialekten? *Maal og Minne* 2: 120–136.
- Paolieri, Daniela, Lorella Lotto, Debora Leoncini, Roberto Cubelli og Remo Job. 2011. Differential effects of grammatical gender and gender inflection in bare noun production. *British Journal of Psychology* 102: 19–36.
- Perridon, Harry. 2003. Language Contact and Grammatical Change: The case of Bergen. *Tijdschrift voor Skandinavistiek* 24: 235–256.

- Rodina, Yulia og Marit Westergaard. 2013. The acquisition of gender and declension class in a non-transparent system: Monolinguals and bilinguals. *Studia Linguistica* 67: 47–67.
- Rodina, Yulia og Marit Westergaard. 2015. Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics* 27: 145–187.
- Rodina, Yulia og Marit Westergaard. 2021. Grammatical gender and declension class in language change: A study of the loss of feminine gender in Norwegian. *Journal of Germanic Linguistics* 33: 235–263.
- Sandøy, Helge. 1988. Samsvarbøying av adjektiv og perfektum partisipp i norske dialektar. I Bjørkum og Borg (red.): *Nordiske studiar*, Oslo. 85–117.
- Sandøy, Helge 2013: Driving forces in language change – in the Norwegian perspective. I Tore Kristiansen og Stefan Grondelaers (red.): *Language (De)standardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus. 125–151.
- Svenonius, Peter. 2017. Declension class and the Norwegian definite suffix. I Vera Gribova og Stephanie S. Shih (red.): *The morphosyntax-phonology connection: Locality and directionality at the interface*. Oxford: Oxford University Press. 325–359.
- Tældeman, Johan. 2005. The influence of urban centres on the spatial diffusion of dialect phenomena. I Peter Auer, Frans Hinskens og Paul Kerswill (red.): *Dialect change: Convergence and divergence in European Languages*. Cambridge: Cambridge University Press. 263–284.
- Tiedemann, Johan. 2012. Parallel data, tools and Interfaces in OPUS. *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'2012)*.
- Trosterud, Trond. 2001. Genustilordning i norsk er regelstyrt. *Norsk lingvistisk tidsskrift* 19: 29–58.
- Trudgill, Peter. 2013. Gender maintenance and loss in Totenmålet, English, and other major Germanic varieties. I Terje Lohndal (red.): *In search of Universal Grammar: From Old Norse to Zoque*. Amsterdam: John Benjamins. 77–107.
- Urek, Olga, Terje Lohndal og Marit Westergaard. 2021. *En splyv eller et splyv?* Tilordning av grammatisk genus til pseudosubstantiv i norsk. *Norsk Lingvistisk Tidsskrift* (dette heftet).
- Vanderkerckhove, Reinhild. 2009. Urban and rural languages. I Peter Auer og Jürgen Erich Smith (red.): *Language and Space: An International Handbook of Linguistic Variation*. Berlin: Mouton de Gruyter. 315–332.

Vangsnes, Øysten A., Gören Söderlund og Morten Blekesaune. 2017. The effect of bidialectal literacy on school achievement. *International Journal of Bilingual Education and Bilingualism* 20: 346–361.

Westergaard, Marit og Yulia Rodina. 2016. Hvor mange genus er det i Tromsø-dialekten? *Maal og Minne* 2: 159-189.

Summary

Across Norwegian dialects and sociolects, we currently find considerable variation in the gender system, mainly regarding the feminine. Many speakers have a robust three-gender system (masculine, feminine, neuter), while others have a two-gender system (common, neuter), where all or most feminine forms are replaced by masculine forms. Independent of dialect, all speakers are exposed to both the two-gender and three-gender systems through written and spoken media, i.e., there is variation and instability in the input. In this article, we hypothesize that input and output affect each other; more specifically that variation in the input may influence Norwegian speakers' output and over time contribute to language change. We compare spoken and written corpus frequencies for feminine markers with response patterns and reaction times from a gender classification task. We find a strong correlation between classification patterns and corpus frequencies in the written corpus: feminine nouns that often appear with masculine markers in the corpus, most notably the definite suffix *-en*, are often classified as masculine and have longer reaction times.

Björn Lundquist
Førsteamanuensis i lingvistikk
Institutt for språk og kultur
UiT Norges arktiske universitet
Postboks 6050 Langnes
9037 Tromsø
bjorn.lundquist@uit.no

Rachel Klassen
Forsker i lingvistikk
Institutt for språk og kultur
UiT Norges arktiske universitet
Postboks 6050 Langnes
9037 Tromsø
r.hm.klassen@gmail.com

Marit Westergaard
Professor i engelsk lingvistikk
Institutt for språk og kultur
UiT Norges arktiske universitet
Postboks 6050 Langnes
9037 Tromsø
marit.westergaard@uit.no