



Adjektivlæring hos vaksne innlærarar av norsk – ein empirisk studie

Anna-Marie Kjøde Olsen

Universitetet i Bergen

Denne artikkelen utforskar empirisk nokre sider ved læring og bruk av adjektiv hos vaksne innlærarar av norsk som andrespråk. Han er basert på ein kvantitativ tverrsnittstudie av adjektivbruken i 300 tekstar frå Norsk andrespråskorpus. Eg har samanlikna innlærarmaterialet med tekstar skrivne av førstespråksbrukarar av norsk og i tillegg sett etter samanhengar mellom adjektivbruken og nokre bakgrunnsvariablar. Som mål på adjektivbruken har eg valt frekvens, noko som er motivert av eit bruksbasert teoretisk fundament. Eg ser ein skilnad mellom andrespråksbruken og førstespråksbruken ved at innlærartekstane har ein lågare adjektivfrekvens og mindre variasjon i adjektivbruken sin enn tekstane frå kontrollkorpuset. Vidare er det skilnader mellom innlærarar med tre ulike førstespråksbakgrunnar (somal, tysk og vietnamesisk) som kan tyde på tverrspråkleg påverknad mellom førstespråket og norsk. Det er ikkje sikkert grunnlag for å seie at adjektivfrekvensen varierer med tekstane sitt nivå etter Det felles europeiske rammeverket for språk eller utdanningsbakgrunnen til informantane i denne studien. Det overordna bildet som teiknar seg av adjektivbruken, er den omfattande individuelle variasjonen, noko som er eit venta trekk i innlærarspråk og som finn klangbotn i det teoretiske grunnlaget.

Nøkkelord: adjektiv, adjektivlæring, norsk som andrespråk, bruksbasert, frekvens

1 Innleiing¹

Denne artikkelen presenterer ein kvantitativ tverrsnittstudie av adjektivbruken i tekstar skrivne av vaksne med norsk som andrespråk (S2). Studien nyttar språkbruksdata frå 300 tekstar henta frå Norsk andrespråkskorpus (ASK) (Meurer 2012), og informantane har tre ulike førstespråksbakgrunnar (somali, tysk og vietnamesisk). Målet er å finne ut meir om forekomsten av adjektiv i innlærarnorsk, og kva faktorar adjektivforekomsten kan vise seg å ha ein samanheng med. Adjektiv har i relativt liten grad vore studert i nordisk andrespråkssamanheng ut ifrå eit frekvensperspektiv og leksikalsk perspektiv, og, etter det eg kjenner til, ikkje med basis i same type data eller teoretisk tilnærming som i denne studien. Slik kan studien vere eit empirisk bidrag til utforsking av norsk innlærarspråk basert på språkbruksdata og ut frå eit bruksbasert perspektiv. Studien bruker data frå ulike ferdigheitsnivå i norsk til å skildre utvikling, noko som er ein vanleg framgangsmåte. Vidare føreset studien at den bruken som viser seg i tekstane, gir eit blikk inn i tileigningsprosessen. Dette følgjer av at bruksbaserte tilnærmingar ikkje set noko klart skilje mellom kompetanse og performanse. Både kunnskap om språket og evna til å bruke det er ein del av innlæraren sin kompetanse, og begge komponentane blir etablerte gjennom språkbruk (Mitchell, Myles & Marsden 2013: 7–8).

Forskingsspørsmåla som blir stilte i artikkelen, er:

- 1: Kva slags frekvensmønster kjem til syne i innlærartekstane sin bruk av adjektiv?
- 2: Er det ein kvantitativ skilnad i bruken av adjektiv mellom innlærartekstane og kontrollkorpuset?
- 3: Er det ein kvantitativ skilnad i bruken av adjektiv mellom tekstar på ulike rammeverksnivå?
- 4: Er det ein kvantitativ skilnad i bruken av adjektiv etter kva utdanningsnivå innlæraren har?
- 5: Er det ein kvantitativ skilnad i bruken av adjektiv mellom dei ulike S1-gruppene?

Artikkelen er strukturert slik: Del to omtalar tidlegare forskning, og i del tre skriv eg om adjektiv som ordklasse i norsk og i kontrastivt perspektiv. Del fire tar for seg teoretiske sider ved rolla frekvens har innanfor det bruksbaserte ram-

1. Eg vil gjerne takke ein anonym fagfelle for konstruktive tilbakemeldingar på ein tidlegare versjon av manuset og Jonar Eikeland ved UiB for hjelp med statistikken.

meverket og andre variablar som kan spele inn i bruken av adjektiv. I del fem blir dataa og metoden presentert og i del seks resultatata av undersøkinga. Til slutt blir funna diskuterte i del sju før eg oppsummerer i del åtte.

2 Tidlegare forskning

Det har blitt gjort få studiar av adjektivfrekvens i norsk som andrespråk, men Axelsson (1989, 1994) og Kotsinas (1982) har begge undersøkt adjektiv i munnlege data i svensk S2. Studiane inkluderer adjektiv i både attributiv og predikativ stilling. Kotsinas gjer ei longitudinell kartlegging av ordforrådet til seks voksne, lågt utdanna informantar (op.cit.: 14). Generelt nyttar dei adjektiv (token) med om lag same frekvens (4,2 %) som førstespråksbrukarar (4,4 %), men den individuelle variasjonen er stor (op.cit.: 118). Adjektiva er distribuerte talmessig over få ulike adjektiv (typar) (op.cit.: 40–45). Ho finn at innlærarane med svakast utvikla S2-vokabular, lener seg på høgfrequente ord i større grad enn dei andre, men det kjem ikkje fram om det gjeld adjektiv (op.cit.: 211).

Axelsson sin studie er ein kombinert longitudinell og tverrsnittstudie av 60 voksne S2-innlærere med tre ulike språkbakgrunnar (finsk, polsk og spansk) og to ulike ferdigheitsnivå (Axelsson 1994: 7). Dei spanske og polske innlærarane med eit høgt ferdigheitsnivå i svensk brukte eit høgare tal adjektiv (token) og fleire ulike adjektiv (typar) i intervjuet enn dei på eit lågare ferdigheitsnivå, men hos dei finske informantane var det berre så vidt skilnad mellom ferdigheitsnivåa (op.cit.: 116). Oversikta Axelsson gir, er ikkje justert for talet på ord i intervjuet totalt. Dei finske informantane har samtidig ein lågare frekvens av det Axelsson definerer som «foreign adjectives» enn dei spanske og polske, noko ho tilskriv tverrspråkleg påverknad (op.cit.: 127 & 149). Ho konkluderer med at det ikkje er éin enkelt faktor som bestemmer korleis adjektiv blir lært, men at fleire faktorar, som kommunikative behov, utdanning og personlegdom, tverrspråkleg påverknad, formularlæring og ferdigheitsnivå, verkar saman (op.cit.: 147–150).

I hovudoppgåva si brukte Svendsen frekvens og semantikk for å sjå om det var mogleg å predikere ei læringsløype for adjektiv hos 42 voksne innlærere av norsk som andrespråk (Svendsen 1996: 1). Svendsen gjorde ein tverrsnittstudie med skriftlege elisiteringstestar. Adjektiva i testane var valte både på grunnlag av høg frekvens i ulike frekvensordlister og for å representere ulike semantiske typar. Konklusjonen var at frekvens aleine ikkje kunne predikere ei læringsløype for adjektiv, men at frekvens får høgare forklaringsverdi i kombinasjon med semantikk (op.cit.: 131–132).

Til slutt ser Nilsen (2018) i sin doktorgradstudie på utviklinga av substantivfrasen i tekstar skrivne til Norskprøven.² Tekstane spenner frå A1 til B2 i Det felles europeiske rammeverket for språk (CEFR/Rammeverket) (Europarådet 2001). Ho går ut frå at innlærarane prosentvis tar i bruk stadig fleire ledd i substantivfrasen samtidig i takt med at den språklege erfaringa deira aukar, og indirekte gjeld altså studien attributiv bruk av adjektiv. Resultata så langt kan tyde på at adjektivfrekvensen i substantivfrasar aukar svakt mellom kvart av ferdigheitsnivåa, og at auken ser ut til å vere tydelegast mellom A1 og A2.³

3 Ordklassen adjektiv

3.1 Adjektiv i norsk

Orda i studien er valte basert på at dei er adjektiv i norsk. Adjektiv i norsk er kjenneteikna av at dei lar seg gradbøye og samsvarbøye i kjønn, tal og bestemtheit (Faarlund, Lie & Vannebo 1997: 345). Semantisk uttrykker dei vanlegvis eigenskapar eller tilstandar (op.cit.: 346). Dei utgjer kjernen i ein adjektivfrase og kan fungere predikativt eller attributivt til eit substantiv op.cit.: 345.). I setninga «*Det vesle huset i skogkanten er raudt og kvitt.*» er *vesle* eit eksempel på attributiv bruk og *raudt* og *kvitt* eksempel på predikativ bruk.

Eg har operasjonalisert adjektivomgrepet etter Norsk referansegrammatikk (Faarlund, Lie & Vannebo 1997) sin definisjon fordi han korresponderer med definisjonen tagginga i ASK bygger på. Definisjonen der er morfologisk fundert og reknar gradbøying som eit tilstrekkeleg kriterium for medlemskap (op.cit.: 22). Det tyder mellom anna at tidsnemnarar (*ofte, fort* og *lenge*), mengdeord (*mange, få*), ordenstal og plasseringsord (*høgre, nedre*) er rekna med. I tillegg inneheld klassen presens partisipp, perfektum partisipp i attributiv stilling og perfektum partisipp i predikativ stilling når dei er kongruensbøyge.

Analyseeininga i den aktuelle studien er ord. Eg bruker leksemdefinisjonen av ord (Sveen, Simonsen & Theil 2000: 74). Medlemmar i eit leksem høyrer til i same ordklasse og har lik tyding, men kan ha ulike bøyingsformer. ASK nyttar *lemma* for leksem, mens eg i teksten held på termen *type* i tråd med Bybee (2008) sin token-/typedistinksjon.

2. Norskprøven – <https://www.kompetansenorge.no/Norsk-og-samfunnskunnskap/Norsk-prove> (05.12.19)

3. Gølin K. Nilsen, personleg kommunikasjon i e-post 14.november 2018.

3.2 Adjektiv i kontrastivt lys

Informantane i den studien har tysk (TY), somali (SO) og vietnamesisk (VI) som førstespråk. Valet av desse språkbakgrunnane er motivert av den ulike avstanden dei har til norsk typologisk og genetisk, og særleg når det gjeld korleis ord som i norsk høyrer til i adjektivklassen, blir klassifiserte. Både generelt og når det gjeld adjektivordklassen meir spesifikt, har tysk langt meir til felles med norsk enn dei to andre språka. Ei side ved slektskapen mellom tysk og norsk som er relevant, er den store mengda kognatar, noko som lettar læringsbyrden både reseptivt (Golden 2014: 112) og produktivt (Jarvis & Pavlenko 2008: 85, Llach 2011: 5). Om vi ser vekk frå kasussystemet, blir ordklassen adjektiv i det store og heile handsama på same måte som i norsk både syntaktisk, morfologisk og semantisk (Janik 2015: 72).

Somali blir rekna som eit syntetisk språk i større grad enn norsk (Husby 2017: 92). Adjektiv blir anerkjent som ein eigen ordklasse i somali av både Husby (2017: 99) og Saeed, i kontrast til grammatikkar som har rekna orda som ei undergruppe av verb (Saeed 1993: 4). Saeed inkluderer både det han kallar «basic adjectives» og «derived adjectives» (ibid.), og sistnemnde er ei stor og ope gruppe som kan avleiast både frå verb og substantiv (op.cit.: 185). Adjektiva i somali kan stå både etterstilt attributivt til substantiv og predikativt i setninga (Husby 2017: 99). Dei kan ha gradbøying og blir bøygd i samsvar med substantivet, unntatt når det gjeld bestemtheit (op.cit.: 100).

Vietnamesisk er på si side typologisk svært ulikt norsk (Rosén 1999: 1). Det kan klassifiserast som eit isolerende språk ettersom orda manglar bøyging. Ord med det vi assosierer med adjektivisk tyding, blir i vietnamesisk vanlegvis sett på som ei undergruppe av verb (Husby 2017, Rosén 1999: 28). Felles for dei er at dei kan stå etter det forsterkande ordet *rât* 'veldig', og sjølv om orda gjerne blir oversett med norske adjektiv, er altså tydinga verbal (Husby 2017: 250). Ordet *lón* tyder til dømes snarare 'å vere stor' enn berre 'stor'. Vietnamesisk har ikkje noko som tilsvare norsk samsvarsbøying, men kan gradbøye uttrykk ved hjelp av eigne funksjonsord (ibid.). Den typologiske avstanden relatert til det adjektiviske kan derfor seiast å vere noko mindre mellom norsk og somali enn mellom norsk og vietnamesisk.

4 Bruksbaserte perspektiv

Det er mange faktorar som kan vere med å påverke tileigninga av eit andrespråk hos voksne. I det følgjande skriv eg om frekvens, påverknad frå førstespråket,

utdanningsnivå og ferdigheitsnivå, målt langs CEFR-skalaen, som alle er inkluderte som variablar i denne studien.

4.1 Frekvens

Denne studien plasserer seg innanfor ramma av eit bruksbasert perspektiv. Bruksbaserte tilnærmingar legg stor vekt på innputten si rolle i språklæringa (Ellis & Wulff 2014: 86), og frekvens blir rekna som ein sentral faktor ved innputten som drivar for læringa (Nistov et al. 2018: 116). For det første har innputten innlæraren møter visse strukturelle eigenskapar som gjer han meir eller mindre «prosesserbar» (Berggreen & Tenfjord 1999: 94), og av dei faktorane blir frekvens rekna som den viktigaste. Enkelt sagt går det raskare å identifisere og hente fram ein frekvent språkleg konstruksjon enn ein som er lite frekvent (Ellis 2006: 102). For det andre baserer prosesseringa av innputt seg på umedvitne, statistiske utrekningar, og frekvens er ein heilt nødvendig komponent for å kunne rekne ut kva som er statistisk sannsynleg (Mitchell, Myles & Marsden 2013: 104).

Ulike typar frekvens kan ha ulike effektar i læringsprosessen, og her er skiljet Bybee (2008) har utvikla mellom tokenfrekvens og typefrekvens relevant. Av dei effektane Bybee (2008: 218) set fram, er *bevaringseffekten* den mest relevante i denne studien. Bevaringseffekten er relatert til tokenfrekvens. Høg tokenfrekvens styrkar den kognitive representasjonen av den aktuelle språklege eininga og gjer eininga lettare tilgjengeleg i minnet. Frekvente einingar blir lært raskare og bevart betre (op.cit.: 219). Adjektiv som har høg frekvens i målspråket, vil derfor bli lært raskare enn lågfrekvente. Basert på bevaringseffekten vil vi mellom anna kunne forvente at stor grad av eksponering for målspråket vil vise seg i meir leksikalsk variert adjektivbruk og bruk av meir lågfrekvente adjektiv.

Frekvensseffektar aleine kan likevel ikkje forklare språklæring fullt ut. Det blir mellom anna peikt på at frekvens ikkje ser ut til å ha ein kumulativ effekt på læringa, men følgjer ei bratt læringskurve i starten og flatar gradvis ut (Ellis & Wulff 2014: 78).

4.2 Påverknad frå førstespråket

I eit bruksbasert perspektiv er det òg tilhøve knytte til den assosiative læringsprosessen som er med på å dempe frekvensseffekten (Mitchell, Myles & Marsden 2013: 105), og eitt av desse tilhøva er *learned attention*. Innan innlæraren møter S2, er han allereie kognitivt optimalisert til førstespråket sitt (Ellis & Wulff 2014: 82). På område der S1 og S2 skil seg frå kvarandre, kan resultatet bli at innlæraren ikkje legg merke til trekk i S2 fordi S1 har lært han å oversjå

dei (op.cit.: 83). Slik forklarer bruksbaserte tilnærmingar at sjølv om innlæraren får store mengder innputt, er det likevel ikkje alt som blir til inntak. Førstespråket vil altså kunne påverke korleis effekten av frekvensen i andrespråket blir.

Internasjonal og norsk forskning har vidare vist empirisk at ein ikkje kan sjå vekk frå S1-bakgrunn i S2-studier der innlærarane har ulikt førstespråk (Golden, Jarvis & Tenfjord 2017).

Tverrspråkleg påverknad kan observerast langs ulike aksar (Jarvis & Pavlenko 2008: 20). Særleg relevante for den aktuelle studien er lingvistisk transfer, i motsetnad til transfer på eit konseptuelt nivå, og leksikalsk transfer (ibid.). Når ordkunnskapen i eitt språk påverkar bruken av ord i eit anna språk, kan det vere relatert til ulike sider ved denne ordkunnskapen, for eksempel ordklasse og frekvens (op.cit.: 72–73).

Ei rekke faktorar kan føre til tverrspråkleg påverknad, mellom anna forholdet mellom målspråket og språk innlæraren kan frå før, og kva grad av likskap det er mellom dei (Jarvis & Pavlenko 2008:176). Eit viktig skilje går mellom det som blir kalla objektiv og subjektiv likskap. Kjelda til tverrspråkleg påverknad er primært subjektiv likskap, altså at innlæraren har ei subjektiv oppfatning om likskap mellom S1 og S2, og derfor vel å lene seg på S1 når han lærer og/eller bruker S2 (op.cit.: 178–179). Kva grad av samanfall som faktisk finst mellom språka, den objektive likskapen, spelar inn på kor sannsynleg det er at det som blir overført er positivt eller negativt (op.cit.: 179). Mens positiv transfer kan gjere at eit trekk blir lært raskare, kan negativ transfer på si side mellom anna føre til unngåing (Gujord & Ragnhildstveit 2018: 138). Objektiv likskap fører ikkje i seg sjølv til transfer, og det treng ikkje vere samanheng mellom subjektiv og objektiv likskap, men ein finn likevel langt fleire tilfelle av openberr transfer når målspråket liknar på kjeldespråket (Jarvis & Pavlenko 2008: 177–179).

4.3 Andre kjelder til variasjon

To andre variablar som kan gi opphav til systematisk variasjon i adjektivbruken mellom innlærarane, er utdanningsnivå og ferdigheitsnivået i S2, her forstått som nivåplassering i CEFR. Både desse variablane og informasjon om førstespråk finst i ASK.

Valet av CEFR-variabelen baserer seg på den generelle forventninga om ei gradvis utvikling frå det enkle til det meir komplekse (Berggreen & Tenfjord 1999: 65) og at ferdigheitsnivåa kan spegle ei slik utvikling i adjektivbruk, slik også Nilsen (2018) føreset. Dette samsvarer med tanken om at innlæraren sitt

språklege nettverk gradvis veks fram og aukar i kompleksitet i takt med den språklege erfaringa, noko som harmonerer med fleire teoretiske synspunkt.

Vidare kan utdanningsnivået vere opphav til variasjon på ulike språklege område og kan påverke både kor raskt og kor vellukka ein lærer eit S2 (t.d. Bugge, Carlsen & Gujord, 2014 via Gujord 2017). Andrespråksforskinga har i stor grad konsentrert seg om høgt utdanna⁴ innlærarar, og i studiar som omfattar innlærarar med låg eller manglande utdanning, har ofte ikkje utdanningsvariabelen vore isolert (van de Craats, Kurvers & Young-Scholten 2005: 9). Som ein følge av det, veit vi i dag mindre om utviklinga av eit S2 hos den sistnemnde gruppa, noko som er uheldig når det er ting som tyder på at utdanningsnivået til innlæraren er ein sterkt medverkande faktor for kor vellukka S2-læringa blir (Mitchell, Myles & Marsden 2013: 21). Utdanningsnivå er tatt med i studiane til både Kotsinas (1982), Axelsson (1994) og Svendsen (1996).

5 Metode og data

5.1 Dataa frå ASK

Dataa i studien kjem frå 300 tekstar i det elektroniske innlærarkorpuset ASK (Meurer 2012). Informantanene har anten somali (SO), tysk (TY) eller vietnamesisk (VI) som førstespråk. Dei aktuelle tekstane er svar på Språkprøven i norsk for vaksne innvandrarar (heretter Språkprøven), som skulle måle språkferdigheiter på B1-nivå (Tenfjord, Hagen & Johansen 2009: 55). Alle prøve-svara blei vurderte til bestått. Tekstane har i ettertid blitt plasserte på rammeverksnivå (Carlsen 2012: 31–33), og hamna då på tre ulike nivå, A2, B1 og B2. Det er denne plasseringa eg nyttar.

Tekstane er på om lag 200–400 ord kvar (jf. Tabell 5). Dei omhandlar ulike opne, kvardagslege tema, og til saman svarer dei på 32 ulike oppgåvetitlar. Til kvar tekst er det knytt sjølvrapporterte personopplysningar om mellom anna førstespråksbakgrunn, alder og utdanningsnivå.

Som kontrollkorpus har eg brukt 100 tekstar frå ASK som er skrivne av vaksne med norsk som S1. Tekstane blei til under dei same rammevilkåra som resten av ASK-materialet når det gjeld lengde, tema, hjelpemiddel, tid⁵ osv., men dei er ikkje skrivne i ein autentisk testsituasjon. Tematisk er tekstane svar

4. van de Craats et.al (2005: 8-9) definerer ein lågt utdanna som «an adult who has at most ten years of education in the country of origin», og høg utdanning som meir enn 10 års skulegang.

5. Ingvild Nistov, personleg kommunikasjon juni 2020.

på fire ulike oppgavetitler. Så få tekstar med så smal tematikk kan sjølvstekt ikkje gi eit representativt bilde av skriftleg norsk, men dei gir likevel eit visst empirisk samanlikningsgrunnlag for adjektivbruken i same typen tekst.

5.2 Handtering av dataa

Først blei alle dei tagga førekomstane (token) av ordklassen adjektiv søkte opp⁶ og lasta ned som ei typebasert frekvensliste. Feilannotering kan vere ei utfordring i ASK fordi Oslo-Bergen-taggaren (Johannessen et al. 2012) ikkje er konstruert for å handtere innlærarspråk (Tenfjord, Hagen & Johansen 2009: 62). Derfor var det behov for ein nøyte manuell gjennomgang av typelista. Der ein type verka potensielt problematisk, blei enkeltførekomstane søkte opp i kontekst. Når det var grunnlag for det, blei førekomstane og typar fjerna eller slått saman. Døme på endringar eg har gjort i materialet er at⁷:

- adjektiv med to gyldige skrivemåtar er slått saman til eitt (f.eks *gammel* og *gammal*)
- samskrivingar er delte opp og plasserte i rett type (*altforflink*, *mestmulig*)
- feilstavingar er retta opp (*islamisk* og *islamsk*)
- samansette ord som er særskrivne og derfor feilannoterte, er fjerna (*syke* i *sykepleier*)
- typar som tilhøyrer ein annan ordklasse, er fjerna (*butikkpersonal* og *1800-talle*)
- ikkje-adjektiveiske førekomstane som har same form som ein adjektivisk type, er fjerna (substantivførekomstane av *dyr* og *mobil*)
- partisippførekomstane som etter definisjonen er adjektiv, er endra frå infinitivform til partisippform (*leke* → *lekende*, *drepe* → *drept*) og ikkje-adjektiveiske partisipp fjerna
- førekomstane som ser ut til å vere samanblanda ut frå konteksten, er retta opp (ein førekomst av *ferdig* er flytta til *farlig*: «Man bør ikke røyke mye for røyking kan være ferdig for kroppen, det vil ødelegge lunger og brenne tenner (tekst-id s0882).»)

6. Søkestreng i ASK: [pos = «adj»] \ <> \ \ :: language = «somali|vietnamesisk|tysk» & testlevel = «Språkprøven»

7. I ettertid av analysane har eg gått gjennom ei gruppe ord i datasettet på nytt. Det gjeld nokre få ord som ASK har tolka som språk-/nasjonalitetsadjektiv, men som også kan vere substantiv. Totalt 12 treff (7 'norsk') av dei totalt 7408 treffa er feilkategoriserte som adjektiv.

- førekomstar utan ein klar parallell i ordboka, er forsøkt tolka og retta opp etter målspråksnormen (førekomsten *tryggdelig* er tatt inn i typen *trygg*: «Jeg fikk panikk og bad Herren om vi kunne komme ned tryggdelig (tekst-id s0100).»

Utover den gjennomgangen som er skildra over, er det ikkje tatt omsyn til om bruken av adjektiva er grammatisk eller målspråkslik. Etter gjennomgangen sat eg igjen med 7405 førekomstar fordelt på 593 typar, mot opphavleg 7408 førekomstar fordelt på 692 typar.

5.3 Statistisk testing

I den statistiske testinga held eg meg til alfanivået $p < 0,05$, slik vanleg praksis er i andrespråksforskinga (Larson-Hall 2010: 389). Utvala er normalitetstesta med ein Shapiro Wilks-test (jf. Vedlegg 1) i tillegg til at eg har vurdert fordelingskurva visuelt i eit histogram (Figur 1, s. 160). Ved samanlikning av to grupper med normalfordelte data bruker eg ein t-test, og tilsvarande One-way ANOVA ved samanlikning av tre normalfordelte grupper, med Tukey HSD som post hoc-test. Der post hoc-testen finn ein statistisk⁸ skilnad mellom to grupper, blir Cohens *d* (Hedges 1981) brukt som effektmål. Når tre grupper med ikkje-normalfordelte data blir samanlikna, bruker eg Kruskal-Wallis og Mann Whitney U som post hoc-test. Alfanivået er Bonferroni-korrigert manuelt i desse tilfella (Field 2009: 69) og som effektmål bruker eg *r* (op.cit.: 227). Testane i studien er to-hala.

6 Resultat

I det følgande vil eg presentere resultatata av adjektivbruken i datasettet ved å gå gjennom forskingsspørsmåla suksessivt. Først går eg gjennom sentral- og spreingsmål for dei viktigaste variablane, med vekt på tokenfrekvens, og så viser eg resultatata av den statistiske testinga for variabelen tokenfrekvens.

6.1 Forskingsspørsmål 1

Det første forskingsspørsmålet dreier seg om kva frekvensmønster som kjem til syne i innlærartekstane sin bruk av adjektiv. Tabell 1 viser sentrale verdiar for tekstlengde som kan gi oss eit inntrykk av datasettet. Analysane her omfattar

8. I tråd med tilrådingane frå Larson-Hall (2010: 103–104) har eg valt å omtale resultat under alfanivået som *statistiske*, snarare enn *signifikante*.

alle dei 300 tekstane, og tala er på individnivå. Som vi kan sjå, er det stort spenn i tekstlengde. Skeivskapsmålet viser at fordelinga er lett positivt asymmetrisk ($N(300) = 1,26, p < 0.05$).

Tabell 1: Oversikt over tekstlengde i talet på ord i datasettet på individnivå

Gjennomsnitt	Median	St.avvik	Kortaste tekst	Lengste tekst	Skeivskap	Kurtose
258	246	70	110	595	1,3	3,1

For å få eit overblikk over mengda adjektiv som er brukt på gruppenivå i heile datasettet, går vi til Tabell 2. Han viser at det totalt er brukt 7405 adjektiv, og dei er fordelte på 593 ulike typar. Det høge talet på typar med berre éin førekomst, viser at datasettet er tungt i den eine enden.

Tabell 2: Oversikt over talet på adjektiv i datasettet på gruppenivå

	Løpeord	Adjektivtoken	Ulike typar	Typar med eitt treff
Totalt	77466	7405	593	244

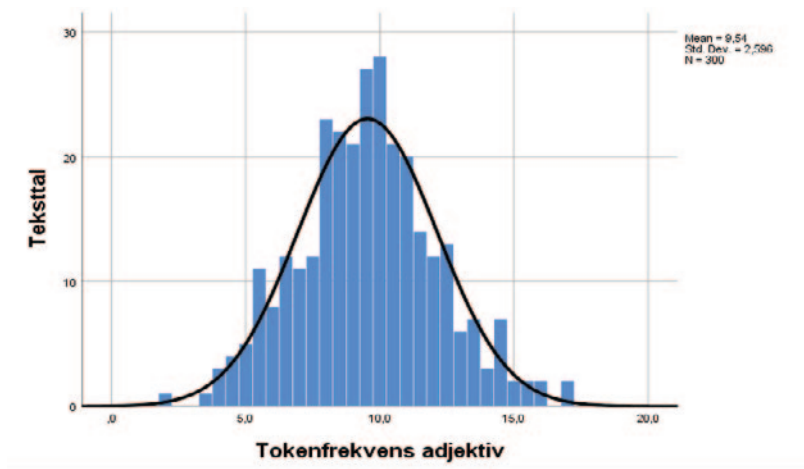
Tabell 3 viser til talet på og frekvensen av adjektiv (både token og type, rekna ut som talet på adjektiv dividert med talet på ord i teksten) i det samla materialet på individnivå.

Tabell 3: Oppsummering av frekvens og tal i datasettet på individnivå

Variabel	Gj.snitt	Median	St.avvik	Min	Maks	Skeivskap	Kurtose
Talet på adjektivtoken	24,7	23	9,7	5,0	68,0	1,0	1,8
Talet på adjektivtypar	16,3	15	6,2	5,0	42,0	1,0	1,6
Frekvensen av adjektivtoken	9,6	9,6	2,6	2,0	16,8	0,1	0,0

Oversikta syner at ingen av tekstane er utan adjektiv, og at spennet i talet på adjektivtoken er stort, mellom fem og 68. Gjennomsnittleg har dei 24,7 adjektiv, men standardavviket tilseier at variasjonen er relativt stor. Utvalet er igjen noko positivt assymmetrisk ($N(300) = 0,965, p < 0.05$). Adjektiva utgjer i snitt 9,6 % av løpeorda i tekstane. Vidare inneheld tekstane i snitt 16,3 ulike adjektivtypar kvar, med eit spenn på mellom fem og 42. Også når det gjeld mengda ad-

jektivtypar og frekvensen av adjektivtoken, er det ein del variasjon i datasettet. Av den grunn kan det vere nyttig å vise fordelinga av tekstane etter adjektivfrekvens med eit histogram (Figur 1). Det illustrerer korleis tekstane i stor grad følger kurva for normalfordeling når det gjeld tokensfrekvensen av adjektiv, med ein topp rundt 10,0. Den lågaste frekvensen er 2,0, og den med høgast frekvens inneheld 16,8 (jf. Tabell 3).



Figur 1: Histogram for fordelinga av tekstar etter tokenfrekvens

Samanfatta kan vi seie at tekstane i materialet er på i overkant av 250 ord gjennomsnittleg. Av desse orda er i snitt 25 adjektiv, og dei er fordelte på 16 ulike adjektivtypar. I frekvens utgjer adjektiva 9,6 % av den totale ordmengda. Spennet i verdiane på alle desse områda er likevel relativt stort.

6.2 Forskingsspørsmål 2

Det neste forskingsspørsmålet gjeld om innlærartekstane sin bruk av adjektiv skil seg frå kontrollkorpuset. I Tabell 4 ser vi ein oversikt over løpeord og adjektiv i dei to korpusa på gruppenivå. Sjølv om det er langt færre kontrolltekstar enn innlærartekstar, inneheld dei både fleire ulike adjektivtypar totalt og fleire typar med berre eitt treff enn innlærartekstane.

Tabell 4: Oversiktstal frå ASK på gruppenivå

	Løpeord totalt	Talet på adjektivtoken	Talet på adjektivtypar	Typar med eitt treff
ASK Språkprøven (N=300)	77466	7405	593	244
ASK kontrollkorpus (N=100)	27495	3068	614	345

I Tabell 5 kan vi sjå at tekstane i kontrollkorpuset i snitt er noko lengre enn innlærartekstane, men skilnaden er relativt liten. Om vi ser på førekomstane av adjektiv målt i tokenfrekvens, talet på token og typar, er skilnaden tydelegare – igjen med høgare verdiar i kontrollkorpuset. Ein t-test for uavhengige utval på variabelen tokenfrekvens slo ut statistisk mellom gruppene ($t(398) = -5,411$, $p < 0,0005$, $d = 0,60$).

Tabell 5: Oppsummering av datasettet på individnivå etter korpus

		Ord i teksten	Talet på adjektivtoken	Frekvensen av adjektivtoken	Talet på adjektivtypar
ASK Språkprøven (N=300)	Gj.snitt	258	24,7	9,6	16,3
	Median	246	23	9,6	15
	St.avvik	70	9,7	2,6	6,2
	Min	110	5	2	5
	Maks	595	68	16,8	42
ASK kontrollkorpus (N=100)	Gj.snitt	275	30,7	11,3	21,5
	Median	264	29,0	11,4	21,0
	St.avvik	71	11,1	3,5	7,5
	Min	78	5	2,7	6
	Maks	625	67	21,5	50

Oppsummert har altså førstespråksbrukarane ein noko høgare tokenfrekvens enn innlærarane i tekstane sine. Typetalet og talet på unike adjektiv er også høgare i S1-datasettet.

6.3 Forskingsspørsmål 3

Det tredje forskingsspørsmålet ser adjektivbruken i tekstane opp mot den plasseringa dei har fått på CEFR-skalaen. ASK opererer med mellom- og heilnivå, men for å få meir solide utval, har eg slått dei saman til heile nivå.⁹ Då står eg igjen med A2, B1 og B2. Dei to lågaste nivåa har omlag like mange tekstar kvar

9. A2 og A2/B1 blir rekna som A2. B1 og B1/B2 blir rekna som B1. B2 og B2/C1 blir rekna som B2.

(N=135 og N=144), mens B2-utvalet er lite (N=21). Tabell 6 viser sentral- og spreingsmål etter kva rammeverksnivå tekstane er tildelte.

Tabell 6: Oppsummering av datasettet på individnivå etter CEFR-nivå

	Ord i teksten	Talet på adjektivtoken	Frekvensen av adjektivtoken	Talet på adjektivtypar	
A2 (N=135)	Gj.snitt	225,7	21,0	9,4	13,6
	Median	219,0	21,0	9,4	13,0
	St.avvik	50,3	7,3	2,7	4,4
	Min	110	8,0	3,4	5,0
	Maks	393	44,0	15,8	26,0
B1 (N=144)	Gj.snitt	276,1	26,8	9,7	17,7
	Median	263,5	25,0	9,7	17,0
	St.avvik	63,2	9,9	2,5	6,1
	Min	177	5,0	2,0	5,0
	Maks	540	68,0	16,8	42,0
B2 (N=21)	Gj.snitt	345,1	33,8	9,8	23,4
	Median	332,0	34,0	9,9	23,0
	St.avvik	101,4	11,4	2,2	7,3
	Min	212	11,0	4,9	8,0
	Maks	595	58,0	16,1	36,0

Tekstlengda er kortast på A2-nivået og aukar tydeleg for kvart ferdigheitsnivå. Det same ser vi i talet på token og typar. Sentralmåla for tokenfrekvens viser den same auken mellom kvart av dei tre utvala, sjølv om skilnaden er liten mellom alle tre nivåa. Ein One-Way ANOVA-test viser heller ingen statistisk skilnad mellom gruppene for denne variabelen ($F_{2,297} = 0,516, p = 0,597$). Det ser altså ut til at innlærarane har omlag den same frekvensen av adjektiv i tekstane sine uavhengig av ferdigheitsnivå.

6.4 Forskingsspørsmål 4

Det neste forskingsspørsmålet dreier seg om utdanningsbakgrunnen til informantane. Av dei 300 informantane har om lag 1/3 rapportert at utdanninga deira ligg på vidaregåande nivå og om lag 1/3 har høgare utdanning. Den siste tredjedelen har anten gått grunnskule (46), har kryssa av for «Anna» (32) eller ikkje svart på spørsmålet (7). Dei informantane det manglar opplysningar om, er utelatne. Kategorien «Anna» er vanskeleg å tolke, og sidan kategorien kan påverke den statistiske analysen og dermed konklusjonen, vil heller denne ikkje bli tatt med vidare. Ein oversikt over sentral- og spreingsmål kjem fram i Tabell 7.

Tabell 7: Oppsummering av datasettet på individnivå etter utdanning

		Ord i teksten	Talet på adjektivtoken	Frekvensen av adjektivtoken	Talet på adjektivtyper
Grunnskule (N=46)	Gj.snitt	233,0	19,7	8,6	13,7
	Median	224,0	20,0	8,3	13,5
	St.avvik	63,2	7,9	3,0	5,6
	Min	110	5,0	2,0	5,0
	Maks	455	45,0	15,2	36,0
Vidaregåande (N=104)	Gj.snitt	263,8	25,3	9,6	16,3
	Median	248,0	24,0	9,7	16,0
	St.avvik	77,6	9,4	2,4	5,9
	Min	118	9,0	4,2	5,0
	Maks	595	58,0	15,8	34,0
Høgskule/ universitet (N=111)	Gj.snitt	269,6	25,9	9,5	16,8
	Median	253,0	24,0	9,3	16,0
	St.avvik	71,4	10,6	2,6	6,7
	Min	123	8,0	4,4	7,0
	Maks	540	68,0	16,8	42,0

Tekstlengda er stigande etter utdanningslengde, frå «Grunnskule» til «Vidaregåande» og så vidt frå «Vidaregåande» til «Høgskule/universitet». Det same er tilfellet for talet på token og typer. Likeeins er sentralmåla for frekvensen av adjektivtoken ganske like for gruppene med vidaregåande og høgare utdanning, og noko lågare i grunnskulegruppa. Ein One-Way ANOVA-test viser ikkje likevel nokon statistisk skilnad mellom utdanningsgruppene for tokenfrekvens ($F_{2,258} = 2,523, p = 0,082$). Oppsummert verkar det altså ikkje til å vere nokon samanheng i datasettet mellom utdanningsnivå og tokenfrekvensen av adjektiv.

6.5 Forskingsspørsmål 5

For å finne svar på om det er ein kvantitativ skilnad mellom S1-gruppene i bruken av adjektiv, deler eg datasettet i tre like store grupper (å N=100) etter kva førstespråksbakgrunn informantane har. I Tabell 8 blir sentral- og spreingsmål presentert på individnivå. Dei viser at innlærere med tysk som S1 (TY) skriv dei klart lengste tekstane og har høgast førekomst av adjektiv både i tal og frekvens. Tilsvarende har gruppa med somali (SO) som S1 dei lågaste verdiane, og dei vietnamesisktalande (VI) ligg i midten. Samtidig er spreinga i tekstlengde størst i gruppa med tysk som S1 og minst hos SO. Motsett er spreinga i tokenfrekvens minst i TY, mens VI og SO har litt større spreing.

Tabell 8: Oppsummering av datasettet etter S1 på individnivå

	Mål	Ord i teksten	Talet på adjektivtoken	Talet på adjektivtypar	Frekvensen av adjektivtoken
Somali (N=100)	Gj.snitt	228,5	20,0	12,9	8,8
	Median	220,5	19,0	13,0	8,4
	St.avvik	53,5	7,8	4,3	3,0
	Min	110	5	5	2,0
	Maks	410	49	25	15,8
Tysk (N=100)	Gj.snitt	291,3	30,5	20,7	10,5
	Median	270,5	29,0	20,0	10,1
	St.avvik	79,4	10,2	6,5	2,0
	Min	163	15	10	5,8
	Maks	595	68	42	16,8
Vietnamesisk (N=100)	Gj.snitt	254,9	23,6	15,2	9,3
	Median	240,0	23,0	15,0	9,3
	St.avvik	60,0	7,7	4,7	2,5
	Min	133	8	5	4,2
	Maks	420	44	26	16,8

Vidare vil eg sjå om skilnaden mellom S1-gruppene gir utslag statistisk. Ein Kruskal-Wallis H-test viser at skilnaden i tokenfrekvens mellom gruppene gir statistiske resultat ($H = 22,718$, $df=2$, $p < 0,0005^{10}$). Gruppene blir så samanlikna parvis med Mann Whitney U som post hoc-test ($\alpha = 0,017$). Post hoc-testane viser at skilnadene er statistiske både mellom SO ($mdn = 8,4$) og TY ($mdn = 10,1$) ($U = 3150,0$, $p < 0,0005$, $r = 0,32$) og mellom TY ($mdn = 10,1$) og VI ($mdn = 9,3$) ($U = 3603,0$, $p = 0,001$, $r = 0,24$), men ikkje mellom VI ($mdn = 9,3$) og SO ($mdn = 8,4$) ($U = 4427,5$, $p = 0,162$). Det er altså lite truleg at skilnaden i tokenfrekvens mellom TY på den eine sida og SO og VI på den andre, er tilfeldig. Effekten er middels mellom SO og TY og liten til middels mellom TY og VI.

Resultata frå forskings spørsmålet peikar i retning av ein skilnad mellom førstespråksgruppene, men det er likevel ikkje gitt at skilnaden ligg i S1-bakgrunn. Andre forskjellar som heng saman med språkbakgrunnen, kan ligge bak. Innanfor Jarvis sitt metodiske rammeverk for identifisering av transfer, er det tilrådd å kontrollere andre faktorar (Jarvis 2000: 260–261).¹¹ Eg vel derfor å

10. Verdiane i SPSS viser $p=0,000$, men eg rapporterer det som $p < 0,0005$ både her og seinare ettersom skilnaden i realiteten ikkje er null (Larson-Hall 2010: 279).

11. Studien kvalifiserer til krava i Jarvis (2000) sitt rammeverk når det gjeld intragruppe-homogenitet og intergruppe-heterogenitet i performansen ettersom eg bruker statistikk som kan vise at åtterden til ei gruppe er ulik nok ei anna gruppe – og lik nok internt – til å hevde at det observerte kan koplast til S1-bakgrunnen. Vidare gir eg under overskrifta «Andre påverknadsfaktorar» eit oversyn over nokre kontrastive forhold mellom dei aktuelle språka (Gujord & Ragnhildstveit 2018: 141).

sjå nærare på distribusjonen av språkbakgrunnar saman med høvesvis rammeverksnivå og utdanningsnivå.

6.6 Førstespråk og ferdigheitsnivå

Det blir generelt rekna som viktig å halde kontroll på ferdigheitsnivå i andre-språksstudiar. Om ferdigheitsnivået varierer mellom S1-gruppene, kan det vise seg i adjektivbruken deira. Om ferdigheitsnivået derimot er jamt, kan ein skilnad i adjektivbruken tyde på ein transfereffekt. Tabell 9 viser korleis tekstane i materialet fordeler seg på ulike nivå etter kva S1 informantane har. Ein langt høgare del av dei tysktalande innlærarane er vurderte til B-nivået enn i dei to andre språkgruppene. Hovudvekta av TY-tekstane ligg på B1-nivå, og fleire er på B2-nivå enn A2. Til samanlikning er ingen av tekstane frå VI og berre éin frå SO på B2-nivå, og fleirtalet av tekstane i desse gruppene ligg på A2-nivå.

Tabell 9: Krysstabell for talet på tekstar på gruppenivå etter S1 og CEFR-nivå

	Somali (N=100)	Tysk (N=100)	Vietnamesisk (N=100)	Total (N=300)
A2	69	11	55	135
B1	30	69	45	144
B2	1	20	0	21

Den skeive fordelinga av gruppene er ei potensiell feilkjelde i materialet. For å finne ut meir om nivåplasseringa i forhold til S1 og adjektivbruk, gjer eg derfor ein analyse med CEFR-nivå som kontrollvariabel. Frå Tabell 10 kan vi sjå at TY-utvalet skil seg ut med dei høgaste middelverdiene uavhengig av CEFR-nivå, mens SO-utvalet har dei lågaste. Variasjonen er noko større i SO og VI enn i TY.

Tabell 10: Frekvensen for adjektivtoken etter S1 og CEFR-nivå

		CEFR-nivå			
		A2	B1	B2	
Førstespråk	Somali	N=	69	30	1
		Gj.snitt	9,1	8,3	4,9
		St.avvik	3,0	2,7	
	Tysk	N=	11	69	20
		Gj.snitt	10,5	10,6	10,1
		St.avvik	1,1	2,1	2,0
	Vietnamesisk	N=	55	45	0
		Gj.snitt	9,5	9,2	
		St.avvik	2,5	2,5	

Den statistiske testinga vidare blir gjort utan B2-nivået. Når CEFR-nivåa blir haldne konstante, viser ein Kruskal-Wallis-test ikkje statistiske resultat mellom S1-gruppene på A2-nivå ($H = 3,855$, $df = 2$, $p = 0,146$), men derimot på B1-nivå ($H = 19,823$, $df = 2$, $p < 0,0005$). Ein Mann Whitney U-test viser at det er statistiske skilnader ($\alpha = 0,017$) på B1-nivå mellom både SO (mdn = 8,1) og TY (mdn = 10,5) ($U = 523,0$, $p < 0,0005$, $r = 0,39$) og TY (mdn = 10,5) – VI (mdn = 9,0) ($U = 987,5$, $p = 0,001$, $r = 0,31$), men ikkje mellom SO (mdn = 8,1) – VI (mdn = 9,0) ($U = 557,0$, $p = 0,202$). Resultata tyder altså på at skilnaden på B1-nivå mellom TY på den eine sida og SO og VI på den andre ikkje er tilfeldig, og i begge tilfella slår effekten ut som middels.

Eit usikkerheitsmoment ved CEFR-nivået er at heile nivå og mellomnivå er slått saman (jf. forsk.spm. 3). Som vi ser av Tabell 11, er fordelinga mellom S1-bakgrunnar ujamn, særleg innanfor det som blir rekna som B1-nivået (B1 og B1/B2), der testen viste statistiske skilnader. TY har totalt flest tekstar på B1-nivå, og ein stor del av dei er vurdert til mellomnivået B1/B2, i motsetnad til SO- og VI-tekstane, som for det meste er vurderte til heilnivået.

Tabell 11: Fordeling av CEFR-nivå med mellomnivå på førstespråksbakgrunn

	A2	A2/B1	B1	B1/B2	B2	B2/C1
Somali	14	55	27	3	1	0
Tysk	3	8	42	27	19	1
Vietnamesisk	3	52	44	1	0	0

For å sjekke om skilnaden på B1-nivå kunne skrive seg frå ei skeivfordeling av S1-bakgrunnane på B1 og B1/B2, gjorde eg ein analyse der eg fjerna mellomnivået og heldt det utanfor analysen av B1-nivået. Utvala var normalfordelte og relativt like i storleik: SO ($N=27$, $m=8,6$, $w=0,983$, $p=0,929$), TY ($N=42$, $m=10,6$, $w=0,985$, $p=0,863$) og VI ($N=44$, $m=9,1$, $w=0,961$, $p=0,143$). Ein One-way ANOVA viste statistiske resultat ($F_{2,110} = 6,478$, $p = 0,002$). Skilnaden er framleis statistisk mellom TY (mdn = 10,6) og SO (mdn = 8,3) ($p = 0,004$, $d = -0,88$) og TY (mdn = 10,6) og VI (mdn = 8,95) ($p = 0,019$, $d = -0,65$), men ikkje mellom SO (mdn = 8,3) og VI (mdn = 8,95) ($p = 0,640$). Effekten er stor mellom TY og SO og middels mellom SO og VI.

Utifrå dette kan vi konkludere med at skilnadene mellom den tyskspråklege gruppa og dei to andre mest truleg ikkje er eit tilfeldig resultat, og at det eksisterer ein reell skilnad mellom dei på B1-nivået. Vi kan ikkje seie det same om skilnaden mellom SO og VI.

6.7 Førstespråk og utdanningsnivå

Slik vi har sett på S1-gruppene i lys av CEFR-nivået til tekstane, vil eg òg undersøke utdanningsbakgrunnen deira. Tabell 12 viser ei fordeling av tekstane etter kva S1 og utdanningsbakgrunn informantane har. Vietnamesisktalende dominerer i gruppa for høgare utdanning, mens TY har flest på vgs-nivå og SO klart flest med grunnskulebakgrunn.

Tabell 12: Krystabell for talet på tekstar på gruppenivå etter utdanning og S1

		Utdanning		
		Grunnskule	Vidaregåande	Høgskule/universitet
Førstespråk	Somali	31	35	28
	Tysk	3	40	33
	Vietnamesisk	12	29	50
Total		46	104	111

I Tabell 13 finn vi den høgaste middelværdien for adjektivfrekvens i gruppa som har TY som S1 og utdanning rapportert som «Grunnskule» (11,4). Den lågaste middelværdien har SO-utvalet i same utdanningskategori (8,2). Samtidig er det ein del variasjon i utvala, særleg i SO.

Tabell 13: Sentral- og spreingsmål for adjektivfrekvens etter S1 og utdanning på individnivå

			Utdanning		
			Grunnskule	Vidaregåande	Høgskule/universitet
Førstespråk	Somali	N=	31	35	28
		Gj.snitt	8,2	9,2	9,0
		St.avvik	3,1	3	2,8
	Tysk	N=	3	40	33
		Gj.snitt	11,4	9,7	10,5
		St.avvik	2	1,6	2
	Vietnamesisk	N=	12	29	50
		Gj.snitt	8,9	9,9	9,2
		St.avvik	2,7	2,4	2,6

VI-utvalet med grunnskulebakgrunn er ikkje normalfordelt ($w=0,0844$, $p=0,031$). Ein Kruskal-Wallis-test der utdanningsbakgrunnen blir halden konstant gir statistiske resultat mellom S1-gruppene i «Høgskule/universitet» ($H = 7,297$, $df = 2$, $p = 0,026$), men ikkje i «Grunnskule» ($H = 3,561$, $df = 2$, $p =$

0,169) og «Vidaregåande» ($H = 1,906$, $df = 2$, $p = 0,386$). Ein Mann Whitney U viser ikkje statistiske resultat i gruppa «Høgskule/universitet» ($\alpha = 0,017$) verken mellom SO (mdn = 8,55) og TY (mdn = 10,1) ($U = 304,5$, $p = 0,023$), mellom SO (mdn = 8,55) og VI (mdn = 9,2) ($U = 671,0$, $p = 0,763$), berre mellom TY (mdn = 10,1) og VI (mdn = 9,2) ($U = 566,0$, $p = 0,016$). Det er altså berre grunnlag for å seie at det er skilnad mellom høgt utdanna tysk- og vietnamesisktalande.

7 Drøfting

7.1 Frekvens og språkleg utviklingsnivå

Utforskinga av forskingsspørsmål 1 har gitt oss eit inntrykk av datasettet, og i forskingsspørsmål 3 såg vi etter samanheng mellom tekstane sitt CEFR-nivå og adjektivbruken. Sjølv om dei deskriptive tala viste ein knapp auke i tokenfrekvens for kvart nivå frå A2 til B2, slo ikkje desse ut ved statistisk testing. Mine data gir altså ikkje nokon tydeleg indikasjon på at det eksisterer ein samanheng mellom tokenfrekvensen av adjektiv og det språklege utviklingsnivået til innlærarar av norsk på dei aktuelle CEFR-nivåa. Nilsen (2018) ser ein eventuell auke i adjektivfrekvens frå A1 til B2 som eit teikn på auka kompleksitet i innlærarspråket. Ei slik utvikling kan vere i tråd med dei forventingane vi har til læringsgang generelt om aukande kompleksitet (Berggreen & Tenfjord 1999: 65). Kompleksiteten kan ligge i at nominalfrasen blir bygd ut ved at substantivkjernen blir moderert eller presisert (Faarlund, Lie og Vannebo 1997: 346) ved hjelp av adjektiv. Det kan òg vere kognitive krav knytte til bruken av adjektiv som gjer dei meir komplekse å lære og ta i bruk enn til dømes substantiv og verb, noko som tilseier at innlærarane får grep om adjektivbruken på eit seinare tidspunkt i læringsprosessen (Llach 2011: 5). I Axelsson (1994: 116) sin studie brukte informantar på eit høgare utviklingsnivå talmessig fleire token og typar av adjektiv enn dei på eit lågare nivå, men det har låg samanlikningsverdi så lenge vi ikkje kjenner frekvensen. Kotsinas (1982: 45) rangerer ikkje informantane etter ferdigheitsnivå, men oppgir storleiken på ordforrådet dei bruker i materialet og adjektivfrekvensen. Eg kan ikkje sjå at det er nokon klar samanheng mellom dei to. Nilsen (2018) finn ein svak auke i frekvensen frå A1–B2 og at auken ser ut til å vere tydelegast mellom A1 og A2. Det same kan ikkje seiast med bakgrunn i mitt materiale, men her er heller ikkje tekstar på det lågaste rammeverksnivået representerte. Ein mogleg konklusjon kan altså vere

at innlærerne i materialet mitt allereie har tileigna seg ein adjektivtokenfrekvens som liknar førstespråksbruken.

På same tid som vi ikkje kan sjå nokon klar samheng mellom frekvens og utviklingsnivå (målt i CEFR-nivå) i studien min, viser dataa stor grad av variasjon mellom innlærerne seg i mellom. Det same viser seg både i Kotsinas (1982) og Axelsson (1994) sine studiar. Systematikk og variasjon i andrespråklæringa er tilbakevendande tema innanfor andrespråksforskninga. Ei rekke faktorar kan skape variasjon (Berggreen & Tenfjord 1999: 26–27), og bakgrunnsvariablane i denne studien kan representere nokre av dei. Men i eit bruksbasert perspektiv knyt ein òg variasjon til faktorar ved innputt:

Such variability is a natural consequence of input factors such as exemplar type and token frequency, recency, context, salience, contingency, regularity, and reliability, along with the various other associative learning factors that affect the emergence of attractors in the problem space (Ellis & Wulff 2014: 87).

Den framtrekande variasjonen i materialet mitt stemmer med at bruksbaserte tilnærmingar ser variasjon som sjølve standarden ved språklæring og -bruk (Nistov et al. 2018: 120). I sin natur er språkbrukaren sin kompetanse eit dynamisk system som blir forma av erfaringane og dei rammene språkbruken til ei kvar tid skjer innanfor (ibid.), anten det er i S1 eller S2. Dette synet representerer ei viss relativisering samanlikna med meir tradisjonelle tilnærmingar.

I forskings spørsmål 2 jamførte eg innlærartekstane med kontrollkorpuset i ASK. Frekvensen i dei to datasetta skil seg statistisk frå kvarandre ved at innlærartekstane inneheld relativt færre adjektiv enn kontrolltekstane. Kotsinas (1982: 233) konkluderte i studien sin av munnleg språkbruk med at det var lite som skilte adjektivfrekvensen til informantane hennar frå S1-bruk, men det blei ikkje testa statistisk. Det er mogleg at modusen, altså at Kotsinas sitt materiale er munnleg og mitt skriftleg, kan ha noko å seie for at Kotsinas får eit anna resultat, for i mitt materiale er det altså ein skilnad.

Ei forklaring på ulikt frekvensmønster i S1-bruk og S2-bruk kan ha med den ulike mengda norsk innputt kvar av desse gruppene har møtt. Vaksne S1-brukarar har opplevd betydeleg meir eksponering for norske adjektiv og kan dra veksling på sterkare og fleire kognitive representasjonar enn innlærerne i korpuset venteleg kan. I oppbygginga av desse mentale representasjonane er frekvens ein sentral faktor. Skilnader i læringsprosessane som ligg bak S1-læring og S2-læring kan òg gjere at innlærerne i mindre grad har nyttegjort seg

av innputten dei har møtt på norsk. Til saman kan det føre til eit bruksmønster av adjektiv som skil seg frå det personar med norsk som S1 har.

Det kan vere verdt å påpeike at ein skilnad i rammevilkåra kan ha verka inn på tekstproduksjonen. S2-brukarane skreiv som del av ein språkprøve, mens S1-brukarane berre var med i eit forskingsprosjekt. Presset på S2-informantane var derfor større, og skrivinga hadde eit anna siktemål. Vilkåra var likevel dei same når det gjeld tidsramme, oppgåvetype og tilgang på hjelpemiddel.

7.2 Leksikalsk variasjon

Materialet mitt gir òg ei viss innsikt i den leksikalske variasjonen i informantane sine tekstar jamført med kontrollkorpuset. Ein måte å måle leksikalsk variasjon på, er å sjå på forholdet mellom typar og token. Slike mål er sårbare for tekstlengde, ettersom lengre tekstar normalt inneheld færre typar i forhold til tekstlengda enn kortare tekstar. Som vi kan sjå av Tabell 5, er tekstane noko lengre i kontrollkorpuset enn i innlærartekstane. På gruppenivå i Tabell 4 ser vi òg at trass i at innlærarutvalet (N=300) er tre gongar større enn kontrollutvalet (N=100), finn vi flest ulike typar adjektiv og flest typar med berre éin førekomst i kontrolltekstane. For det tredje ser vi i Tabell 5 at kvar av tekstane i kontrollkorpuset i snitt inneheld langt fleire adjektivtypar enn innlærarkorpuset. Dette peikar i retning av større leksikalsk variasjon i tekstane til S1-brukarane.

Eit forhold som kan ha verka på variasjonen, er tematikken i tekstane, sidan informantane har svart på tematisk ulike oppgåver. Mens tekstane i kontrollkorpuset berre spenner over fire ulike oppgåver, spenner S2-tekstane over 32. Det kunne ut ifrå det vore rimeleg å vente større grad av leksikalsk spreiding i S2-gruppa, men det er altså ikkje tilfellet.

Axelsson (1994: 117) finn at informantane hennar i all hovudsak nyttar adjektiv som er relativt frekvente i svensk, og Kotsinas (1982: 211) observerer det same om ord generelt. Det kan stemme med at innlærartekstane samla ser ut til å ha mindre leksikalsk breidde enn S1-tekstane. Teoretisk kan ein innanfor det bruksbaserte perspektivet sjå leksikalsk variasjon som eit teikn på kor godt utbygd og konsolidert nettverket av mentale representasjonar er. Nettverket blir bygd opp gjennom eksponering for innputt. I tråd med bevaringseffekten treng innlærarane store mengder autentisk innputt for å tileigne seg mindre frekvente konstruksjonar. Den avgrensa språklege erfaringa gjer at dei derfor ikkje har dei same føresetnadene for leksikalsk breidde som vaksne S1-brukarar. I studien min tyder altså den leksikalske variasjonen i tekstane på skilnader mellom S1- og S2-bruk.

7.3 Tverrspråkleg påverknad

Eg har samanlikna korleis S2-innlærere med tre ulike S1-bakgrunnar bruker adjektiv i tekstane sine. Den tyskspråklege gruppa har høgast frekvens av adjektivtoken. Skilnadene er statistiske mellom TY og kvar av dei to andre gruppene, men ikkje mellom SO og VI. Også når CEFR-nivået blir halde konstant, viser det seg at TY bruker fleire adjektiv enn dei to andre gruppene på B1-nivå. Likeeins bruker TY fleire adjektiv enn VI på utdanningsnivået «Høgskule/universitet». Dette er den einaste skilnaden som slår ut med utdanning som kontrollvariabel.

Om skilnadene mellom gruppene kjem av tverrspråkleg påverknad, kan det ha samband med den typologiske og genetiske avstanden mellom S1 og norsk. Oppfatta nærleik mellom S1 og S2 gir større sjanse for transfer og motsett mindre sjanse ved oppfatta avstand. Både leksikalsk og morfologisk har adjektivkategorien i norsk og tysk mykje til felles. Det svært nære sambandet mellom norsk og tysk, og dei mange kognata, gir grunn til å forvente at TY oppfatar likskapar mellom språka, noko som igjen seinkar terskelen for å overføre S1-kunnskapen til S2 (Berggreen & Tenfjord 1999: 210). Tilsvarende kan den typologiske avstanden i handsaminga av adjektivkategorien og ein liten grad av felles ordforråd gjere terskelen høg for SO og VI for å overføre sin leksikalske kunnskap frå S1. Skilnadene i adjektivfrekvens mellom TY på den eine sida og SO og VI på den andre kan i lys av dette tolkast som utslag av høvesvis positiv og negativ lingvistisk transfer. Sjølv om rammeverksnivå og utdanning kan vere faktorar som verkar parallelt med S1 og fører til ulik adjektivbruk i datasettet, er det mogleg at skilnadene kan tilskrivast tverrspråkleg påverknad.

7.4 Utdanning og adjektivbruk

Analysen avdekkjer ikkje statistiske samband mellom utdanningsbakgrunn og adjektivbruk, trass i at det ut frå dei deskriptive tala ser ut til å vere eit skilje mellom grunnskulegruppa på den eine sida og dei med lengre skulegang på den andre.

Den tidlegare forskinga konkluderer ulikt om samanhengen mellom skulegang og adjektivlæring i skandinaviske språk. Axelsson (1994: 147) konstaterer kort at utdanning ser ut til å vere ein av faktorene som verkar inn på læringa. Kotsinas (1982: 225) skriv at den lange skulegangen til den eine informanten kanskje «varit henne til hjelp» i språklæringa generelt, men ikkje spesifikt kopla til adjektiv. Svendsen (1996: 129) meiner på si side at utdanningsbakgrunn i liten grad har nokon forklaringsverdi for resultatane. Studiane gir altså sprikande signal om effekten av utdanning på adjektivlæring, og min eigen studie gir hel-

ler ikkje grunnlag for å hevde nokon samanheng. Det er heller ikkje gitt at ein eventuell utdanningsrelatert skilnad skulle ligge i variabelen tokenfrekvens. Han kunne vere relatert til andre aspekt ved adjektivbruken, som leksikalsk variasjon.

8 Sluttord

Målet med undersøkinga har vore å skaffe eit større empirisk grunnlag for korleis adjektiv blir brukt og lært i norsk som andrespråk hos vaksne. Funna kan skaffe innsikter som betrar utgangspunktet for å stille meir spissa hypotesar i framtida. Drøftinga har vist at trass i stor individuell variasjon, tilsvarer ikkje andrespråksbruken av adjektiv i det aktuelle materialet førstespråksbruken i kontrollkorpuset verken når det gjeld tokenfrekvens eller leksikalsk variasjon. Vidare ser førstespråket til innlærarane ut til å henge saman med tokenfrekvensen av adjektiv, mens bruken ikkje varierer nemneverdig etter rammeverksnivå (A2–B2) og utdanningsbakgrunn. Funna gir meining å tolke i lys av bruksbaserte perspektiv.

Kontrollkorpuset er lite og tematisk avgrensa i tillegg til at vi veit lite om tekstforfattarane. Som kjelde til innputt kunne det vere aktuelt å bruke større norskspråkelege korpus, eventuelt justert i frekvens for adjektiv som er mykje brukt i læremateriell eller relevante kontekstar. Å kunne sjå fleire faktorar i samanheng er ein fordel når ein undersøker komplekse fenomen som språklæring. Derfor bør ein ha eit datagrunnlag som i større grad gjer det metodisk mogleg. Vidare ville det vore ønskeleg å ha meir kontroll på temaa i tekstane og utdanningsbakgrunnen til informantane enn det ASK har. Til slutt vil longitudinelle data kunne gi eit utfyllande bilde av variasjon og systematikk over tid hos enkeltindivid eller grupper. Eit design som legg vekt på individuelle erfaringar og rammevilkår for språklæring over tid, vil vere særleg aktuelt ut frå eit bruksbasert perspektiv.

Vedlegg 1: Resultat av normalitetstesting

Testa med Shapiro-Wilk. Alle testresultata gjeld variabelen tokenfrekvens.

Gruppe	Test-statistikk	P-verdi	Konklusjon	
FORSKINGSPØRSMÅL 2				
SAMANLIKNING MED S1-BRUK				
Innlærartekstane	w = 0,996	p = 0,668	Dataa er normalfordelte	
Kontrolltekstane	w = 0,993	p = 0,858	Dataa er normalfordelte	
FORSKINGSPØRSMÅL 3				
CEFR-NIVÅ				
A2	w = 0,990	p = 0,412	Dataa er normalfordelte	
B1	w = 0,995	p = 0,913	Dataa er normalfordelte	
B2	w = 0,929	p = 0,130	Dataa er normalfordelte	
FORSKINGSPØRSMÅL 4				
UTDANNINGSNIVÅ				
Grunnskule	w = 0,979	p = 0,569	Dataa er normalfordelte	
Vidaregåande	w = 0,991	p = 0,748	Dataa er normalfordelte	
Høgskule /universitet	w = 0,988	p = 0,411	Dataa er normalfordelte	
FORSKINGSPØRSMÅL 5				
PÅVERKNAD FRÅ S1				
SO	w = 0,985	p = 0,317	Dataa er normalfordelte	
TY	w = 0,971	p = 0,027	Dataa er ikkje normalfordelte	
VI	w = 0,988	p = 0,521	Dataa er normalfordelte	
PÅVERKNAD FRÅ S1 – CEFR-NIVÅ SOM KONTROLLVARIABEL				
SO	A2	w = 0,976	p = 0,218	Dataa er normalfordelte
SO	B1	w = 0,992	p = 0,997	Dataa er normalfordelte
TY	A2	w = 0,767	p = 0,003	Dataa er ikkje normalfordelte
TY	B1	w = 0,984	p = 0,504	Dataa er normalfordelte
VI	A2	w = 0,987	p = 0,797	Dataa er normalfordelte
VI	B1	w = 0,962	p = 0,141	Dataa er normalfordelte
PÅVERKNAD FRÅ S1 – UTDANNING SOM KONTROLLVARIABEL				
SO	Grunnskule	w = 0,981	p = 0,852	Dataa er normalfordelte
	Vidaregåande	w = 0,968	p = 0,397	Dataa er normalfordelte
	Høgskule /universitet	w = 0,970	p = 0,591	Dataa er normalfordelte
TY	Grunnskule	w = 0,900	p = 0,384	Dataa er normalfordelte
	Vidaregåande	w = 0,981	p = 0,719	Dataa er normalfordelte
	Høgskule /universitet	w = 0,952	p = 0,149	Dataa er normalfordelte
VI	Grunnskule	w = 0,844	p = 0,031	Dataa er ikkje normalfordelte
	Vidaregåande	w = 0,973	p = 0,642	Dataa er normalfordelte
	Høgskule /universitet	w = 0,974	p = 0,322	Dataa er normalfordelte

Litteraturliste

- Axelsson, Monica. 1989. «Vuxna SV2-inlärare lär sig adjektiv.» *Andra symposiet om svenska som andraspråk*, Göteborg.
- Axelsson, Monica. 1994. Noun phrase development in Swedish as a second language : a study of adult learners acquiring definiteness and the semantics and morphology of adjectives. Doktorgradsavhandling, Stockholm: Centrum för tvåspråkighet, Stockholms universitet.

- Berggreen, Harald & Kari Tenfjord. 1999. *Andrespråklæring*. Oslo: Ad notam Gyldendal.
- Bybee, Joan. 2008. Usage-based grammar and second language acquisition. I Robinson, Peter J. & Nick C. Ellis (red.): *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge, s. 216–236.
- Carlsen, Cecilie (red.). 2012. *Norsk profil: det felles europeiske rammeverket spesifisert for norsk: et første steg*. Oslo: Novus forlag.
- Ellis, Nick C. 2006. Cognitive perspectives on SLA: The associative-cognitive CREED. Themes in SLA Research: *AILA Review* 19, 100–121.
- Ellis, Nick C. & Stefanie Wulff. 2014. «Usage-based approaches to SLA.» I VanPatten, Bill & Jessica Williams: *Theories in second language acquisition: an introduction*. Florence, KY: Routledge, 75–93.
- Europarådet 2001. *Common European Framework of Reference for Languages: Learning, teaching and assessment*. Cambridge: Cambridge University Press.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Field, Andy. 2009. *Discovering Statistics using SPSS (and Sex and Drugs and Rock 'n' roll)* (3. utg.). Los Angeles: SAGE.
- Golden, Anne. 2014. *Ordforråd, ordbruk og ordlæring*. 4. utg. utg. Oslo: Gyldendal akademisk.
- Golden, Anne, Scott Jarvis & Kari Tenfjord (red.). 2017. *Crosslinguistic influence and distinctive patterns of language learning : findings and insights from a learner corpus*. Bristol: Multilingual Matters.
- Gujord, Ann-Kristin Helland & Silje Ragnhildstveit. 2018. Tverrspråkleg påverknad – Status og trendar i norsk andrespråksforskning. I Gujord, Ann-Kristin H. & Gunhild Randen (red.): *Norsk som andrespråk – perspektiver på læring og utvikling*. Oslo: Cappelen Damm Akademisk, 133–164.
- Gujord, Ann-Kristin Helland. 2017. «Hos en voksen er tanken helt full av masse ting». Personlege narrativar i Norsk andrespråkskorpus (ASK) – ei kjelde til innsikt i læringsprosessen? *Acta Didactica Norge* 11-1, 7–22.
- Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6, 107–128
- Husby, Olaf. 2017. *Innvandrerens morsmål: En ressursbok for lærere*. Bergen: Fagbokforlaget.
- Janik, Marta Olga. 2015. *Adjektivbøyning i norsk – kongruenstilegnelse hos polskspråklige norskinnlærere*. Doktorgradsavhandling, Institutt for skandinaviske studier, Adam Mickiewicz-universitetet i Poznan.

- Jarvis, Scott. 2000. Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language learning* 50-2, 245–309.
- Jarvis, Scott & Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. New York og London: Routledge.
- Johannessen, Janne Bondi; Hagen, Kristin, Lynum André & Nøklestad, Anders. 2012. OBT+stat. A combined rule-based and statistical tagger. I Andersen, Gisle (red.): *Exploring Newspaper Language. Corpus compilation and research based on the Norwegian Newspaper Corpus*. John Benjamins Publishing Company, 51–65.
- Kotsinas, Ulla-Britt. 1982. *Svenska svårt: Några invandrades svenska talspråk*. MINS, Stockholms universitet.
- Larson-Hall, Jenifer. 2010. *A guide to doing statistics in second language research using SPSS*. Second language acquisition research series: theoretical and methodological issues. New York/London: Routledge.
- Llach, María Pilar Agustín. 2011. *Lexical Errors and Accuracy in Foreign Language Writing*. Bristol/Buffalo/Toronto: Multilingual Matters.
- Meurer, Paul. 2012. *ASK – Norsk andrespråkskorpus*. Bergen: Uni Computing. Lokalisert 2019 på <http://clarino.uib.no/ask/ask>
- Mitchell, Rosamond, Florence Myles & Emma Marsden. 2013. *Second language learning theories*. 3. utg. London/New York: Routledge.
- Nilsen, Gølin Christine Kaurin. 2018. Substantivfrasen i norsk som andrespråk: Fra A1 til B2. ILN, UiO; 2018-10-11 – 2018-10-12.
- Nistov, Ingvild, Hana Gustafsson, Teresa Cadierno. 2018. Bruksbaserte tilnærminger til andrespråksspråklæring. I Gujord, Ann-Kristin H. & Gunhild Randen (red): *Norsk som andrespråk – perspektiver på læring og utvikling*. Oslo: Cappelen Damm Akademisk, 107–132.
- Rosén, Victoria. 1999. *Vietnamesisk: en kontrastiv og typologisk introduksjon*. Trondheim: Tapir.
- Saeed, John Ibrahim. 1993. *Somali reference grammar*. 2. utg. Kensington, Md: Dunwoody Press.
- Sveen, Andreas, Hanne Gram Simonsen & Rolf Theil. 2000. *Innføring i lingvistikk*. 2. utg. Oslo: Universitetsforlaget.
- Svendsen, Bente Ailin. 1996. *Adjektiv i norsk som andrespråk: En frekvensbasert leksikalsk-semantic tværnittstudie av adjektivtilegnelse hos voksne innlærere*. Hovedoppgåve, Institutt for nordistikk og litteraturvitenskap, Universitetet i Oslo.

- Tekstlaboratoriet. 2010. *Oslo-Bergen-taggeren*. Universitetet i Oslo. Lokalisert 2019 på <http://www.tekstlab.uio.no/obt-ny/index.html>
- Tenfjord, Kari, Jon Erik Hagen & Hilde Johansen. 2009. Norsk andrespråkskorpus (ASK) – design og metodiske forutsetninger. NOA. Norsk som andrespråk 25-1, 52-81.
- van de Craats, Ineke, Jeanne Kurvers & Martha Young-Scholten. 2005. Research on low-educated second language and literacy acquisition. *Proceedings of the Inaugural Symposium*. Tilburg, Utrecht.

Summary

This article aims to explore empirically some aspects of the acquisition and use of adjectives in adult learners of Norwegian as a second language. It is a quantitative cross-sectional study of the adjective use of texts written by 300 learners in a learner corpus of Norwegian (Norsk andrespråkskorpus, ASK). I have compared the learner data to first language texts. In addition, I have examined the connection between the use of adjectives and some background variables. Guided by a usage-based approach, frequency was chosen as the measure of adjective use. Firstly, adjectives are used less frequently and less varied in the learner data than in the first language data. Secondly, there are group differences based on three different first language backgrounds (German, Somali, and Vietnamese) that indicate cross-linguistic influence. Proficiency level and educational background, on the other hand, do not seem affect the frequency of adjectives. Individual variation is widespread throughout the data, a finding that is common in second language research and corresponds well with the theoretical framework of the current study.

Anna-Marie Kjøde Olsen
Universitetslektor ved Norskurska
Institutt for lingvistiske, litterære og estetiske studier
Universitetet i Bergen
Postboks 7805
N-5020 Bergen
anna-marie.olsen@uib.no