

# Ortografi i norsk som andrespråk

*En undersøkelse av ortografien i norsk hos voksne innlærere med vietnamesisk og bosnisk/kroatisk/serbisk som morsmål*

Tone Vigrestad



Masteroppgave i norsk som andrespråk  
Institutt for lingvistiske og nordiske studier

UNIVERSITETET I OSLO

Våren 2016



© Tone Vigrestad

2016

Ortografi i norsk som andrespråk. En undersøkelse av ortografien i norsk hos voksne innlærere med vietnamesisk og bosnisk/kroatisk/serbisk som morsmål.

Tone Vigrestad

<http://www.duo.uio.no/>

Trykk: Representeren, Universitetet i Oslo

# Sammendrag

Temaet for oppgaven er ortografi i norsk som andrespråk. Jeg har undersøkt ortografien til 99 personer med vietnamesisk som morsmål og 99 personer med bosnisk/kroatisk/serbisk som morsmål – informantene er voksne. Det overordnede målet med undersøkelsen har vært å forsøke å finne ut om det er forskjeller i ortografien til de to informantgruppene, og om eventuelle forskjeller kan relateres til morsmålet og/eller morsmålsskriftspråket deres. Jeg utarbeidet, med utgangspunkt i dette, seks underordnede forskningsspørsmål. Disse forsøker å belyse trekk ved norsk ortografi jeg ville forventet at er vanskelig å tilegne seg for innlærere. Materialet i undersøkelsen har jeg hentet fra innlærerkorpuset ASK (Norsk andrespråkskorpus). For å komme fram til resultatene kategoriserte jeg alle de ortografiske feilene til informantene. Feilene ble plassert i beskrivende kategorier.

Noen av resultatene viste klare forskjeller mellom informantgruppene, mens andre viste ubetydelige forskjeller. Tre funn viste tydelige forskjeller: 1) Informantgruppa med bosnisk/kroatisk/serbisk som morsmål gjør flere feil per løpeord enn informantgruppa med vietnamesisk som morsmål. 2) Informantgruppa med vietnamesisk som morsmål gjør flere feil enn informantgruppa med bosnisk/kroatisk/serbisk som morsmål ved grupper av konsonantgrafemer. 3) Informantgruppa med bosnisk/kroatisk/serbisk som morsmål har flere tilfeller av enkelt for dobbelt konsonantgrafem enn informantgruppa med vietnamesisk som morsmål. Disse forskjellene er statistisk signifikante. En annen tydelig forskjell mellom gruppene er at de med bosnisk/kroatisk/serbisk som morsmål gjør flere substitusjonsfeil ved <y> og <ø> enn de som har vietnamesisk som morsmål.

Forskjellene tolker jeg som påvirkning fra informantenes morsmål og/eller morsmålsskriftspråk. Dette gjelder trekk som oppbygging av stavelser, grafemrestriksjoner og skriftkonvensjoner.



# Forord

Jeg vil rette en stor takk til hovedveilederen min, Else Ryen. Tusen takk for mye og god hjelp og for lærerike tilbakemeldinger på arbeidet! Jeg setter veldig stor pris på dette. Takk også til biveileder Liv Andlem Harnæs.

Den nesten takken går til Silje Ragnhildstveit for hjelp med ASK-korpuset, og til Anne Golden for spennende samtaler om emnet for oppgaven.

Victoria Rosén, Olaf Husby og Nguyen Ngoc Khang har hjulpet meg med spørsmål jeg har hatt om vietnamesisk – takk!

Bård Uri Jensen har kommet med nyttige råd om metode, og har hjulpet meg med viktige utregninger (?). Kenneth Utheim har også hjulpet meg med spørsmål om utregninger – og om Excel. Tusen takk til dere begge.

Kjersti Wictorsen Kola og Oda Røste Odden har gitt meg nyttige tilbakemeldinger på språk og innhold i oppgaven. Tusen takk! Jeg vil også takke dere – samt mange andre studieveinner – for mange hyggelige stunder på Blindern de siste årene.

Den siste takken går til familie og venner for mye og god støtte.

Oslo, juni 2016

*Tone Vigrestad*



# Innholdsfortegnelse

1 Innledning.....	1
2 Teoretisk ramme.....	5
2.1 Innledning.....	5
2.2 Ortografi.....	5
2.2.1 Ortografi og skriftspråk.....	5
2.2.2 Skriftkonvensjoner.....	6
2.2.3 Ortofoni.....	7
2.2.4 Grafotaks.....	7
2.3 Andrespråksortografi.....	8
2.3.1 Hva er andrespråksortografi?.....	8
2.3.2 Påvirkning fra fonologien og ortografien i morsmålet.....	8
2.3.3 Ortografisk kompetanse.....	11
2.4 Ortografisk feil.....	12
2.5 Tidligere norske undersøkelser i andrespråksortografi.....	13
2.6 Forskningsspørsmål.....	14
2.7 Oppsummering av kapitlet.....	17
3 Ortografien i norsk, vietnamesisk og BKS.....	18
3.1 Innledning.....	18
3.2 Hvorfor språkene BKS og vietnamesisk i undersøkelsen?.....	18
3.3 Ortografien i norsk.....	19
3.4 Ortografien i vietnamesisk.....	22
3.5 Ortografien i BKS.....	23
3.6 Oppsummering av kapitlet.....	25
4 Materiale og metode.....	27
4.1 Innledning.....	27



4.2 Bruk av korpus i undersøkelsen .....	27
4.2.1 Hvorfor korpus? .....	27
4.2.2 ASK- korpuset.....	28
4.2.3 Utvalg av materiale .....	28
4.2.4 Variasjon i tekstlengde .....	28
4.2.5 Søk i ASK og sortering av materiale.....	29
4.3 Valg av metode.....	30
4.4 Kategoriinndelingen .....	32
4.4.1 Formål .....	32
4.4.2 Forklaring til kategoriseringen .....	33
4.4.3 Kriterier for kategorisering av feilene .....	33
4.5 Metode for utregning.....	38
4.6 Oppsummering av kapitlet .....	39
5 Resultater.....	40
5.1 Innledning.....	40
5.2 Framstilling av resultatene .....	40
5.3 Hovedfunn.....	41
5.4 Vokalgrafemfeil .....	44
5.4.1 Utelatte og innskutte vokalgrafemer .....	44
5.4.2 Substitusjon av vokalgrafemer .....	46
5.5 Konsonantgrafemfeil.....	48
5.5.1 Feil ved gruppe av konsonantgrafemer .....	48
5.5.2 Feil ved grafemer/polygrafemer som gjengir fonemene /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/ .....	50
5.5.2 Feil ved enkelt og dobbelt konsonantgrafem .....	52
5.6 Variasjon innad i informantgruppene.....	54
5.7 Oppsummering av kapitlet .....	54

6 Drøfting .....	55
6.1 Innledning.....	55
6.2 Drøfting av funn som viser forskjeller mellom VIET og BKS .....	55
6.2.1 Oversikt over funn som viser forskjeller mellom VIET og BKS.....	55
6.2.2 Feil ved grupper av konsonantgrafemer.....	57
6.2.3 Feil ved enkeltskriving og dobbeltskriving av konsonantgrafem .....	59
6.2.4 Substitusjonsfeil ved <u>, <y> og <ø> .....	61
6.3 Drøfting av funn som ikke viser forskjeller mellom VIET og BKS .....	61
6.3.1 Oversikt over funn som ikke viser forskjeller mellom VIET og BKS.....	62
6.3.2 Feil ved grafemer/polygrafemer som gjengir /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ .....	63
6.3.3 Feil ved innskudd av vokalgrafemer i grupper av konsonantgrafemer .....	64
6.3.4 Feil ved diftongering av monoftong og monoftongering av diftong.....	65
6.4 Drøfting av noen andre funn .....	65
6.4.1 Substitusjon av vokalgrafemer .....	66
6.4.2 Ombytting av grafemer .....	67
6.4.3 Overgeneralisering, system, variasjon og påvirkning fra engelsk .....	67
6.4.4 Andre feil.....	69
6.5 Drøfting av undersøkelsen og videre forskning .....	70
6.6 Oppsummering av kapitlet og oppgaven.....	71
Litteraturliste.....	73
Vedlegg 1: Innlærertekster .....	76
Vedlegg 2: Søk i ASK.....	77
Vedlegg 3: Kategorisering av feilene.....	82
Vedlegg 4: Utrekning av statistisk signifikans.....	87
Vedlegg 5: Fordeling av antall feil.....	90
Vedlegg 6: Vokalgrafemfeil:.....	91
Vedlegg 7: Konsonantgrafemfeil.....	93
Vedlegg 8: Resultater i feilfrekvens per samlet antall feil.....	98

# Figurer

Figur 1: Konkordans.....	29
Figur 2: Hovedkategorier i prosent. ....	43
Figur 3: Utelatte, innskutte og substituerte konsonant- og vokalgrafemer i prosent. ....	43
Figur 4: Konsonantgrafemfeil i prosent. ....	44
Figur 5: Vokalgrafemfeil i prosent.....	44
Figur 6: Feilforekomster i prosenttall – hovedkategorier og underkategorier. ....	45
Figur 7: Feil ved gruppe av konsonantgrafemer i prosent. ....	49
Figur 8: Feil ved gruppe av konsonantgrafemer fordelt på underkategorier – tallene er i prosent. ....	49
Figur 9: Utelatte konsonantgrafemer i gruppe av konsonantgrafemer – tallene er i faktiske forekomster.....	50
Figur 10: Innskutte konsonantgrafemer ved siden av annet konsonantgrafem i faktiske forekomster.....	50
Figur 11: Feil ved fonemene /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ i prosent.....	51
Figur 12: Utelatte grafemer ved fonemene /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ i faktiske forekomster. ....	51
Figur 13: Enkelt for dobbelt og dobbelt for enkelt konsonantgrafem i prosent. ....	52
Figur 14: Enkelt for dobbelt konsonantgrafem i faktiske forekomster. ....	53
Figur 15: Dobbelt for enkelt konsonantgrafem i faktiske forekomster. ....	53
Figur 16: Søk i ASK.....	77
Figur 17: Hovedkategorier i prosent av det totale antallet feil.....	94
Figur 18: Feil i prosent av det totale antallet feil ved utelating, innskudd, substitusjon og ombytting. ....	94
Figur 19: Konsonantgrafemfeil i prosent av det totale antallet feil.....	95
Figur 20: Vokalgrafemfeil i prosent av det totale antallet feil. ....	95
Figur 21: Resultater i prosent av det totale antallet feil – hovedkategorier og underkategorier. ....	96

# Tabeller

Tabell 1: En sammenligning av fonologiske og ortografiske trekk, samt av alfabeter, i norsk, vietnamesisk og BKS.....	26
Tabell 2: Variasjon i tekstlengde i VIET og BKS.....	29
Tabell 3: Kategoriinndeling.....	34
Tabell 4: Eksemplifisering av plassering av feil i hovedkategorier og underkategorier.....	35
Tabell 5: Hvordan skille ut ortografiske feil fra andre typer feil.....	38
Tabell 6: Innskutt og utelatt ⟨e⟩ i faktisk forekomst og prosent.....	46
Tabell 7: Innskudd av vokalgrafemer i gruppe av vokalgrafemer – tallene er i faktiske forekomster.....	46
Tabell 6: Substitusjon av vokalgrafemer i faktiske forekomster.....	47
Tabell 7: Forskjell i feilfrekvens av løpeord mellom VIET og BKS – spredning av verdier.....	87
Tabell 8: Antall personer med feil ved enkelt for dobbelt konsonantgrafem hos VIET og BKS.....	87
Tabell 9: Feilfrekvens ved grupper av konsonantgrafemer – spredning av verdier.....	87
Tabell 10: Feilfordeling i de to morsmålsgruppene.....	90
Tabell 11: Utelatte vokalgrafemer.....	91
Tabell 12: Innskutte vokalgrafemer.....	91
Tabell 13: Innskutt ⟨e⟩ medialt og finalt.....	91
Tabell 14: Utelatt ⟨e⟩ medialt og finalt.....	91
Tabell 15: Substitusjon av vokalgrafemer hos BKS.....	92
Tabell 16: Substitusjon av vokalgrafemer hos VIET.....	92
Tabell 17: Oversikt over hvor mange personer som gjør hvor mange feil ved gruppe av konsonantgrafemer.....	93
Tabell 18: Antall personer som gjør x antall feil ved utelatt konsonantgrafem i gruppe av konsonantgrafemer.....	94
Tabell 19: Substitusjon av konsonantgrafemer i gruppe av konsonantgrafemer hos BKS...	94
Tabell 20: Substitusjon av konsonantgrafemer i gruppe av konsonantgrafemer hos VIET...	94

Tabell 21: To forskjellige konsonantgrafemer for dobbelt konsonantgrafemer og dobbelt konsonantgrafem for to forskjellige konsonantgrafemer.....	95
Tabell 22: Utelatte konsonantgrafemer ved grafemer som gjengir /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/.....	95
Tabell 23: Innskutte grafemer ved /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/.....	96
Tabell 24: Substitusjon av grafemer ved /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/.....	96
Tabell 25: Substitusjon av konsonantgrafemer hos BKS.....	96
Tabell 26: Substitusjon av konsonantgrafemer hos VIET.....	97



# 1 Innledning

I oppgaven tar jeg for meg ortografi i norsk som andrespråk. I undersøkelsen ser jeg på ortografiske feil av innlærere, og forsøker å forklare mulige feilkilder. Med *ortografiske feil*<sup>1</sup> mener jeg avvik fra norsk<sup>2</sup> rettskriving. Jeg har valgt å se på ortografiske feil av innlærere med vietnamesisk og bosnisk/kroatisk/serbisk som morsmål (heretter henholdsvis VIET og BKS<sup>3</sup>). Materialet er hentet fra Språkprøven i ASK-korpuset – en samling tekster skrevet av innlærere med ulike morsmål. Undersøkelsen er kvantitativ; jeg ser på 99 tekster fra hver morsmålsgruppe. Det overordnede målet med undersøkelsen er å sammenligne ortografien i VIET og BKS, og forsøke å finne ut om eventuelle forskjeller mellom informantgruppene kan forklares med påvirkning fra morsmålet eller morsmålskriptspråket til innlærerne.

Emnet for oppgaven springer ut av en semesteroppgave jeg skrev i emnet *NOAS4180 – Veiledet lesing i masterstudieretningen i norsk som andrespråk*. Oppgaven kan ses som et pilotprosjekt for undersøkelsen, på den måten at jeg begynte å utforske emnet andrespråksortografi. I likhet med i denne undersøkelsen brukte jeg også i oppgaven ASK-korpuset som materiale. Til forskjell fra denne undersøkelsen så jeg på de ortografiske feilene til innlærerne i alle de ti morsmålsgruppene som er representert i ASK – uten å legge vekt på å skille mellom innlærernes ortografiske feil ut fra morsmål. Jeg la heller vekt på å få et innblikk i hva slags typer feil som gjøres, og å forsøke å se sammenhenger eller mønstre. Oppgaven inneholdt også en beskrivelse av ortografien i norsk, med vekt på hva som kan være vanskelig å tilegne seg for innlærere. Arbeidet med semesteroppgaven gjorde meg interessert i å finne ut mer om andrespråksortografi. Jeg ble nysgjerrig på spørsmål som: Er det slik at innlærere med samme morsmål gjør samme typer ortografiske feil, og vil innlærere med forskjellige morsmål dermed gjøre forskjellige typer ortografiske feil i norsk – og hvordan kan dette eventuelt forklares? På hvilken måte kan fonologien og skriftkonvensjoner i innlærernes morsmål/morsmålskriptspråk spille inn ved læring av norsk ortografi? Dette var spørsmål jeg satt igjen med etter at jeg hadde levert inn semesteroppgaven. Jeg bestemte meg

---

<sup>1</sup> Man kan hevde at innlærerspråk er egne system, og at begrepet feil derfor ikke passer inn. Noen bruker begreper som *overgangsform* eller *avvik*. Jeg velger å bruke feil fordi det er målspråket som er sammenligningsgrunnlaget, og former som ikke tilhører målspråket synes jeg det er hensiktsmessig å omtale som feil.

<sup>2</sup> *Norsk* – når det gjelder skriptspråket – vil bli brukt synonymt med målformen bokmål.

<sup>3</sup> BKS refererer enten til språket *bosnisk/kroatisk/serbisk* eller *informantene med bosnisk/kroatisk/serbisk som morsmål*. For å vise til *informantene med vietnamesisk som morsmål* bruker jeg VIET. Jeg bruker ikke VIET om språket vietnamesisk – her bruker jeg *vietnamesisk*.

for at dette var noe jeg hadde lyst til å utforske nærmere, og slik ble altså temaet for denne masteroppgaven ortografi i norsk som andrespråk.

I arbeidet med undersøkelsen ble det etter hvert tydelig at det både er forskjeller og likheter mellom ortografien til VIET og BKS. Men det er også forskjeller innad i hver informantgruppe. Jeg vil nå peke på noen trekk ved norsk ortografi som kan være vanskelige for innlærere – og jeg vil gjøre dette ved å illustrere med eksempler fra en innlærertekst fra materialet<sup>4</sup>. Teksten er skrevet av en kvinne på 28 år med BKS som morsmål. Hun oppgir å ha vært i Norge i ett til to år, og svarer på en oppgave som lyder: Skriv en tekst om å flytte.

Jeg har flyttet flere ganger og kan si hvor vanskelig det er. For første gang flyttet jeg da jeg var 18 år. Det var i en krig og jeg måtte flytte fra min landsby. Den første gang, da jeg flyttet, hadde jeg bare en kofert med meg. I slike situasjoner tenker man bare på livet sitt, men jeg kunne hennte koferten også. Jeg flyttet hos kosina mi og bodde der bare par måneder. Det var vanskelig for meg og jeg visste at jeg kunne ikke bo lenger hos henne. Jeg og kjaresten min fant et forlatet hus for å bo. Vi måtte reparere den huse og det tok tid og penger. Men, etter par år ville eieren av den huse tilbake og leve i sitt hus. Han la oss å bo i sitt hus bare i en kort periode til, til vi finner en løsning. Til slut hadde vi hellet med oss, da vi fikk lov å flytte til Norge. Vi eide ikke mye da, men vi måtte selge noe av tingene som vi kunne ikke hente med oss. I hvert fall hadde vi fått alt som trenges for liv da vi kom i Norge. Nå bar vi et sted for å bo uten stress. Vi bor i et lite sted og det er vanskelig å finne noen jobb der, men vi skal prøve alt for å finne jobb akurat i det stedet. Vi er ikke villig til å flytte igjen og både vi og vår datter fant nye venner og for føste gang etter krig føller oss som hjemme.

Vi ser at størstedelen av ordene skrives riktig, men at det også gjøres feil. Kan feilene si noe om hva som er vanskelig å tilegne seg i norsk ortografi? For eksempel er det flere tilfeller i teksten av enkelt for dobbelt konsonantgrafem, som i "kofert" for *koffert*, "slut" for *slutt* og "akurat" for *akkurat*. Men det er også tilfeller av det omvendte, altså dobbelt for enkelt konsonantgrafem, som i "føller" for *føler* og "hennte" for *hente*. Skriftkonvensjonen med enkelt og dobbelt konsonantgrafem virker å være vanskelig for denne personen.

Konvensjonen kan være vanskelig av flere grunner – blant annet på grunn av dens mange unntak.

Vi har flere konsonantfonemer i norsk som har ulike stavemåter i ulike ord. Noen av dem skrives med polygrafemer – gruppe av konsonantgrafemer som uttales som ett fonem<sup>5</sup>. Det som kan vanskeliggjøre tilegningen av skrivemåten til noen av konsonantfonemene, er flere ulike gjengivelser av samme fonem, som for eksempel at /s/-fonemet gjengis <skj> i *skjorte* og

<sup>4</sup> For flere innlærertekster, se vedlegg 1.

<sup>5</sup> Jeg vil også bruke digraf og trigraf i oppgaven om polygrafemer som består av henholdsvis to og tre grafemer.



⟨sj⟩ i *sjal*<sup>6</sup>. Feilstavingen "føste" for *første* illustrerer vanskeligheter med stavemåten av fonemet /ʃ/ – her skal det gjengis ⟨rs⟩. I første linje i teksten er ordet likevel stavet riktig, men så er jo også variasjon et vanlig kjennetegn ved innlærertekster. Det finnes også en del feil ved grupper av konsonantgrafemer der grafemene står for hvert sitt konsonantfonem. Slike feil ser vi ikke i teksten over.

Av feil ved vokalgrafemer i teksten kan nevnes ⟨o⟩ for ⟨u⟩ i "kosina" for *kusina*, ⟨a⟩ for ⟨æ⟩ i "kjaresten" for *kjæresten* og ⟨ø⟩ for ⟨o⟩ i "situasjonjer" for *situasjoner*. Vokalgrafemer kan være vanskelig å tilegne seg i norsk av flere grunner. For det første har vi flere fremre, rundede vokaler i norsk – /u/ og /u:/, /y/ og /y:/ og /ø/ og /ø:/. Disse er merkede, det vil si sjeldne i verdens språk, og kan derfor være vanskelige å oppfatte og uttale for mange innlærere. Om dette betyr at også skrivemåten for vokalene – som er henholdsvis ⟨u⟩, ⟨y⟩ og ⟨ø⟩ – vil være vanskelig å få på plass, vet jeg ikke. I tillegg har mange vokaler en lite ortofon stavemåte, som at /ʊ/ staves ⟨u⟩ i *tung*, /ɔ/ staves ⟨o⟩ i *sokker* og /æ/ staves ⟨e⟩ i *verre*.

Ved utvelgelsen av morsmålsspråkene til innlærerne i undersøkelsen var det et poeng at språkene/skriftspråkene skulle være betydelig annerledes enn norsk/skriftspråket i norsk – og at (skrift)språkene seg imellom også skulle være forskjellige. At (skrift)språkene skulle være forskjellige seg imellom, er viktig for å gjøre det mulig å påvise eventuell påvirkning fra morsmålet/morsmålsskriftspråket. Språkene jeg falt ned på, var BKS og vietnamesisk. Vietnamesisk skiller seg fra norsk og BKS ved at det er monosyllabisk – hvert ord består av bare én stavelse. Stavelsesstrukturen i vietnamesisk er (K)V(K). Dette betyr at konsonantgrupper ikke finnes i vietnamesisk uttale, men noen av fonemene gjengis med polygrafemer i skrift. Som norsk har også vietnamesisk mange vokaler. Ortografien i BKS er mindre forskjellig fra norsk ortografi enn vietnamesisk ortografi, blant annet fordi det i BKS er vanlig med grupper av konsonantgrafemer. Når det gjelder vokalgrafemer, er forskjellen mellom norsk ortografi og ortografien i BKS derimot større. I BKS finnes kun fem vokaler, og disse representeres av hvert sitt vokalgrafem. Ortografien i BKS er langt mer ortofon enn ortografien i norsk.

---

<sup>6</sup> Ved beskrivelse av uttale i oppgaven er det uttalen på det sentrale Østlandet (*standard østnorsk*) jeg har som norm. Dette er ikke et vedtatt standardtalemål, men den varietet mange tenker på som et standardtalemål (Lie 2010:6).

Ortografi i norsk som andrespråk er et lite utforsket fagfelt – ja, forskningen er nærmest ikke-eksisterende. Dette har vært spesielt i arbeidet med undersøkelsen, på den måten at jeg har manglet rettesnor – noe som har gjort arbeidet med undersøkelsen både utfordrende og spennende.

Oppgaven består av 6 kapitler. I dette kapitlet – kapittel 1 – har jeg innledet med å presentere temaet for undersøkelsen. I kapittel 2 beskriver jeg teorien som ligger bak drøftingen av resultatene. Jeg presenterer også forskningsspørsmålene. Kapittel 3 er en beskrivelse av ortografien i norsk, BKS og vietnamesisk. Metoden jeg har brukt for å komme fram til resultatene, vil bli beskrevet i kapittel 4, før jeg presenterer funnene i kapittel 5. Deretter drøfter jeg disse funnene – i kapittel 6 – med vekt på forskningsspørsmålene. Kapitlet og oppgaven avsluttes med en drøfting av undersøkelsen og videre forskning.

## 2 Teoretisk ramme

### 2.1 Innledning

I dette kapitlet vil jeg presentere teori som undersøkelsen hviler på. Jeg vil også definere og avklare sentrale begreper, trekke inn relevant forskning, samt presentere og drøfte forskningsspørsmålene mine. Som nevnt i kapittel 1, finnes det svært få undersøkelser av ortografi i norsk som andrespråk. I engelsk som andrespråk finnes det derimot undersøkelser i ortografi som det er relevant å vise til.

Jeg starter kapitlet med å beskrive begrepet ortografi, samt noen begreper som er sentrale i oppgaven (2.2). Deretter beskriver jeg hvordan jeg forstår begrepet andrespråksortografi og hva som ligger i det (2.3), før jeg definerer begrepet ortografisk feil (2.4). En beskrivelse av tidligere norske undersøkelser i ortografi i norsk som andrespråk følger deretter (2.5). Videre følger en presentasjon av forskningsspørsmålene (2.6). Jeg avslutter kapitlet med en oppsummering (2.7).

### 2.2 Ortografi

#### 2.2.1 Ortografi og skriftspråk

Skriftspråk *er* ikke språk; det er en gjengivelse av språk – skriftspråk er med andre ord sekundært i forhold til talespråket. Denne gjengivelsen består av ortografiske tegn, som skal representere talespråket. Begrepet ortografi har jeg valgt å definere som Coulmas (2003): "[...] the standardized variety of a given, language-specific writing system" (Coulmas 2003:35). Coulmas bruker begrepet *staving* (*spelling*) med samme betydning som ortografi (ibid.), noe jeg også vil gjøre.

Gjengivelse av muntlig språk kan ta mange ulike former. Det finnes litt ulike måter å dele inn skriftspråk på (jf. f.eks. Borgwaldt & Joyce 2013). Det vanligste er likevel en tredeling i alfabetiske skriftspråk, logografiske skriftspråk og syllabariske skriftspråk. Skriftsymbolene i

et logografisk system (som f.eks. kinesisk) representerer hele ord med semantisk innhold, mens i syllabariske skriftsystem (som f.eks. japansk, tamil og koreansk) representerer hvert tegn en stavelse (Husby & Kløve 1998:120–124). I alfabetiske skriftspråk er grafemene (skriftegnene) basert på foner (språklyder). Noen skriftspråk har et mer direkte forhold mellom fonem og grafem enn andre (jf. 2.2.3). Vanligvis har et alfabetisk skriftspråk mellom 11 og 70 tegn, noe som er lite sett i forhold til logografiske systemer. Alfabetisk skrift er derfor økonomisk og lett å lære, men har den begrensningen at skriftlig tekst bare forstås av mennesker som snakker samme språk. Dette i motsetning til for eksempel kinesisk, som kan leses av alle som forstår skriftspråket – og uavhengig av om de forstår hverandres muntlige språk eller ikke (Husby & Kløve 1998:120–125).

Alfabetisk skrift er ikke et entydig begrep. Det kan kategoriseres på tre måter: 1) etter skrifttype (f.eks. latinsk, gresk, kyrillisk), 2) etter skriveretning (høyre mot venstre eller venstre mot høyre) og/eller 3) om det er konsonantspråk eller konsonant- og vokalspråk. (Husby & Kløve 1998:125). I de latinske skrifttypene representeres fonemer av grafemer, men hvilke fonemer grafemene representerer, vil variere fra språk til språk. Det å skulle snakke om fonem/grafem-forholdet i det latinske skriftspråket vil derfor ikke gi mening; man må oppgi hvilket språk som omtales. For eksempel uttales <j>-grafemet /j/ i tysk, nederlandsk, dansk, svensk og norsk, /ʒ/ i fransk og portugisisk, /x/ i spansk og /dʒ/ i engelsk. Digrafer og trigrafer har ofte mer forskjellig uttale enn enkle konsonantgrafemer, både innenfor et språk og mellom språk. <th> har for eksempel tre forskjellige uttaler i engelsk; /θ/ i *think*, /ð/ i *brother* og /t/ i *Thailand* (Coulmas 2003:100). Det latinske alfabetet er også tilpasset hvert enkelt språk – i mange tilfeller betyr denne tilpasningen en utvidelse. Ekstra grafemer eller diakritiske tegn er lagt til i mange skriftspråk.

## 2.2.2 Skriftkonvensjoner

Som beskrevet i 2.2.1 er muntlig språk det primære, og skriftspråket en gjengivelse av dette. I gjengivelsen har man regler for hvordan det talte språket skal uttrykkes. Disse reglene kalles *skriftkonvensjoner*, og er altså språkspesifikke. Sebba (2007) sier om skriftkonvensjoner at de er produktet av sosial praksis (Sebba 2007:27), noe som understreker at skriftkonvensjonene er konstruerte. Forskjellige skriftspråk har altså sine gjeldende konvensjoner for hvilke språktrekk som uttrykkes, og hvordan disse uttrykkes. Dette kan blant annet medføre at det

samme uttaletrekket ikke blir representert på samme måte i ulike skriftspråk – eventuelt at det ikke gjengis i et skriftspråk. For eksempel har man både i norsk og BKS betydningsforskjell mellom kort og lang vokal i uttale, men bare i skriftspråket i norsk uttrykkes dette ortografisk. I BKS kommer denne distinksjonen normalt ikke til uttrykk i skrift (Mønnesland 2002:28, jf. 3.5). I norsk uttrykker man dette *indirekte*, ved at kort vokal etterfølges av dobbelt konsonantgrafem og lang vokal etterfølges av enkelt konsonantgrafem – det er altså konsonantgrafemet/-ene som viser vokalens lengde, ikke vokalgrafemet. Distinktiv vokallengde har man også i finsk. Dette uttrykkes ortografisk, men her markeres kort og lang vokal *direkte* ved at kort vokal markeres med enkelt vokalgrafem og lang vokal med dobbelt vokalgrafem (Kulbrandstad 2005:64).

### 2.2.3 Ortofoni

Ortofoni kan defineres som graden av samsvar mellom skriftbilde og uttale. Fenomenet kalles også *gjennomsiktighet* og *dybde*. Selv om alle alfabetiske skriftspråk bygger på prinsippet om at grafemer skal representere uttalen, gjør de det bare i varierende grad. For eksempel representerer grafemet ⟨o⟩ i norsk fonemene /ɔ/ (i *ost*), /u:/ (i *mot*), /ɔ/ (i *som*) og /o:/ (i *sov*). Motsatt vei ser vi at for eksempel fonemet /f/ har flere stavemåter, som ⟨skj⟩ i *skjorte*, ⟨sk⟩ i *ski* og ⟨sj⟩ i *tusj* (mer om ortofoni i norsk i 3.3). Skriftspråk befinner seg et sted på skalaen mellom ortofont og ikke-ortofont; det er aldri et enten/eller. Skriftspråkene i dansk og engelsk er eksempler på mindre ortofone skriftspråk, mens skriftspråkene i finsk og spansk er mer ortofone. Norsk ortografi befinner seg et sted mellom disse (Kulbrandstad 2005:77).

### 2.2.4 Grafotaks

Grafotaks (jf. Wiggen 1992) betegner mulige grafemsammensetninger i et skriftspråk. Cook & Bassetti (2005) kaller fenomenet *orthographic constraints*, og beskriver det slik:

Writing systems constrain the position and co-occurrence of symbols: not all graphemes can occur in every position; many are limited in how they may be combined with other graphemes. Alphabetic writing systems may restrict where letters can occur in the word or syllable and what combinations they may take, unrelated to the occurrence of phonemes in the spoken language (Cook & Bassetti 2005:11).

At mulig grafemsammensetning og tilsvarende fonemsammensetning ikke alltid samsvarer, har vi flere tilfeller av i norsk. For eksempel er ⟨hv⟩ en vanlig grafemkombinasjon i opptakt i

mange spørreord, som i *hva*, *hvem* og *hvorfor*. Men /hv/ er derimot ingen mulig fonemkombinasjon i samme posisjon i stavelsen. Det samme er tilfelle med <gj> i ord som *gjest*, *gjemme* og *gjøre*. Og /ʃ/-fonemet opptrer både i opptakt og koda – men polygrafemene <sj>, <skj> og <skj>, som er blant grafemene/polygrafemene som gjengir fonemet, finnes bare i opptakt (Husby & Kløve 1998:128). I koda gjengis fonemet blant annet av <rs>. Jo mer ortofont et skriftspråk er, jo bedre er samsvaret mellom fonotaksen og grafotaksen.

## 2.3 Andrespråksortografi

### 2.3.1 Hva er andrespråksortografi?

En forutsetning for å bruke et begrep som andrespråksortografi må være at det er noe som kjennetegner ortografien i et andrespråk, uansett språk som omtales. Jeg vil beskrive andrespråksortografi slik: Ortografi i en innlærers andrespråk der hun/han bruker kompetanse fra tidligere lærte (skrift)språk og kompetanse om skriftspråket og fonologien i andrespråket.

Ved læring av et andrespråk, vil altså kompetansen i tidligere lærte språk påvirke. Om fenomenet brukes betegnelser som påvirkning, tverrspråklig innflytelse, transfer og overføring. Jarvis & Pavlenko (2008) definerer fenomenet som "[...] the influence of a person's knowledge of one language on that person's knowledge or use of another language" (Jarvis & Pavlenko 2008:1). Det handler altså om at tidligere erfaring tas med inn i ny erfaring. I dette tilfellet handler det mer spesifikt om at morsmålet og morsmålsskriftspråket kan ha påvirkning på ortografien i andrespråket. Dette vil jeg beskrive nærmere i 2.3.2, før jeg drøfter ortografisk kompetanse i et andrespråk i 2.3.3.

### 2.3.2 Påvirkning fra fonologien og ortografien i morsmålet

Forskning viser at morsmålet har størst påvirkning på andrespråket i en begynnerfase. Ifølge Figueredo (2006) gjelder dette også i ortografien i andrespråket. Figueredo har i en metastudie av 27 undersøkelser i ortografi i engelsk som andrespråk sett tendenser på flere områder. Blant annet viser hun til flere studier som viser at påvirkning fra skriftkonvensjoner i morsmålsortografien skjer på et tidlig nivå i ortografiinnlæringen i andrespråket, og at påvirkningen minker med økt kompetanse i engelsk ortografi (Berkel 1987; Ferroli &

Shanahan 1993; Wang & Geva 2003, referert etter Figueredo 2006:888, jf. 6.5). I den samme studien viser Figueredo til at i femten av studiene hun undersøkte, ble det påvist positiv og negativ overføring, åtte av dem fant bare positiv overføring, tre fant bare negativ overføring og én studie fant ikke overføring (op.cit.:881). Påvirkning fra morsmålet og morsmålsskriftspråket kan altså resultere i feil, men påvirkningen kan også være positiv. Videre er det også slik at mange innlærere kan andre språk og andre skriftspråk enn morsmålet og morsmålsskriftspråket. Mange kan for eksempel engelsk i tillegg til morsmålet og andrespråket. De forskjellige språkene og skriftspråkene en innlærer kan, vil kunne påvirke hverandre (jf. 6.4.3 og 6.4.4).

Både morsmålsskriftspråket og fonologien i morsmålet kan påvirke andrespråksortografien. Cook & Bassetti (2005) sier følgende om det de kaller *writing system transfer* og *language transfer*:

Because of the distinction between language and writing system, it is not so much aspects of the language itself that may be carried over as the attributes of a particular writing system [...] it's the move from one type of linguistic system to another that is important, not from one language to another. In principle, it is important to separate *language transfer* from *writing system transfer* (Cook & Bassetti 2005:29).

Cook & Bassetti gjør altså et klart skille mellom overføring fra språket (*language transfer*) og skriftspråket (*writing system transfer*). Begge deler påvirker andrespråksortografien, men det ene kommer fra morsmålet (språket), og det andre fra skriftspråket i morsmålet. For å eksemplifisere dette viser de til skrivemåten "termometer" for *thermometer* i engelsk av en person med italiensk som morsmål. Dette mener de er påvirkning fra morsmålet, og forklarer det med at fonemet /θ/ ikke finnes i italiensk. Skrivemåten "ingiury" for *injury* av en person med samme morsmål, mener de derimot er påvirkning fra morsmålsskriftspråket fordi ⟨gi⟩ tilsvarende /dʒ/ foran ⟨u⟩ i italiensk. Jeg vil trekke paralleller til to informanter i min egen undersøkelse, da disse har feil jeg mener demonstrerer begge disse feiltypene: Det feilstavede ordet "vare" for *være* mener jeg muligens kan være påvirkning fra fonologien i BKS. I BKS har man en /a/ som ifølge Mønnesland (1990) uttales lenger framme enn østnorsk /a/, og en med BKS som morsmål kan dermed oppfatte norsk /æ/ som /a/ (Mønnesland 1990:23), noe som kan ha resultert i denne feilstavingen. En annen informant har den ortografiske feilen "myje" for *mye*. Dette mener jeg illustrerer Cook & Bassettis begrep *writing system transfer* da man i BKS får ⟨j⟩ mellom en ⟨i⟩ og et vokalfonem, som for eksempel at ⟨j⟩ blir satt inn mellom ⟨i⟩ og ⟨e⟩. I dette tilfellet dreier det seg riktignok om grafemet ⟨y⟩, og ikke ⟨i⟩. Men siden BKS ikke har ⟨y⟩, og norskinnlærere med BKS som morsmål kan uttale /y/ som /i/

(Mønnesland 2002:28), tolker jeg likevel dette som mulig ortografisk påvirkning fra ortografien i BKS.

Slik jeg ser det, er skillet mellom språkoverføring og ortografisk overføring et viktig skille når det gjelder bevisstgjøringen om at andrespråksortografien kan påvirkes både fra en innlærers morsmålsfonologi og morsmålsskriftspråk. Jeg tror likevel ikke at det i alle tilfeller vil være mulig å påvise om en ortografisk feil er overføring fra en innlærers morsmål eller morsmålsskriftspråk, eller om feilen har helt andre årsaker. Dette kommer jeg tilbake til i drøftingskapitlet.

Den fonologiske påvirkningen fra italiensk til engelsk i eksemplet over, der *thermometer* ble feilstavet "termometer", var på fonem/grafem-nivå. Men fonologisk påvirkning kan også dreie seg om å gjøre stavelsesoppbyggingen mer lik den i morsmålet/morsmålsskriftspråket. Jeg vil trekke en parallell til templathypotesen – en hypotese som i utgangspunktet handler om uttale. Men i og med at ortografien til noen av innlærerne kan være påvirket av fonologi, mener jeg den kan være relevant her. Templathypotesen er et mentalt mønster som forteller oss hva som er mulig i en stavelse. Dersom en innlærer bruker sitt morsmålstemplat i uttalen i andrespråket, kan det oppstå uttalefeil (jf. Husby & Kløve 1998:70). Når dette igjen kan påvirke andrespråksortografien, blir altså påvirkningsmønstret slik: morsmålstemplatet påvirker andrespråkstemplatet, som igjen påvirker andrespråksortografien. Utfall av å tilpasse morsmålstemplatet til andrespråksfonologien, kan være som substitusjon, konsonantbortfall og/eller epentese (op.cit.:71). I min egen undersøkelse finner jeg eksempler på feil i alle disse tre kategoriene, og som alle kan være et utslag av tilpassing til morsmålstemplat. Feilen "kuntur" for *kultur* av en informant med vietnamesisk som morsmål er en substitusjonsfeil (<n> for <l>). I koda i vietnamesisk har man bare nasaler og plosiver (Husby & Kløve 1998:61) – /n/ er altså mulig i koda, men ikke /l/. Av konsonantbortfall gjør både VIET og BKS feil, som henholdsvis "spennede" for *spennende* og "taker" for *takler*. I begge informantgruppene har man kortere grupper av konsonantgrafemer enn i norsk (jf. 3.4 og 3.5) – og vietnamesisk har ikke konsonantgrupper i uttale i det hele tatt. Både VIET og BKS har også epentesefeil, som i "voksene" for *voksne* og "egene" for *egne*. I hvor stor grad dette handler om tilpassing til morsmålstemplat, eller om feilene har andre årsaker, vet jeg ikke. Dette er drøftet i 6.3.3. I Hvenekildes undersøkelse (1995, jf. 2.5) gjøres det også feil i disse tre kategoriene – substitusjon, konsonantbortfall og epentese.



I en undersøkelse av Cook (1997) ble det også satt inn det Cook definerer som epentesevokaler – i feilstavinger som "adovocated" for *advocated*, "cousese" for *course*, "Engilish" for *English* og "subejcts" for *subjects*. Informantene har japansk som morsmål (Cook 1997:484).

Husby & Kløve nevner også et annet utfall av tilpassing av morsmålstemplatet til andrespråket – metatese (ombytting) (Husby & Kløve 1998:72). I en undersøkelse av Kløve (1985) gjorde innlærerne med kantonesisk som morsmål uttalefeilen [bli:] for *bil* (Kløve 1985, referert etter Husby & Kløve 1998:73). I mitt eget materiale gjøres ombytting av grafemer av både VIET og BKS, men kun i mindre grad (jf. 6.4.2). Hvorvidt denne ombyggingen kan relateres til informantenes morsmålstemplater eller ikke, synes jeg det er vanskelig å vite.

### 2.3.3 Ortografisk kompetanse

Å tilegne seg et nytt skriftspråk innebærer å tilegne seg ortografisk kompetanse i språket. Ortografisk kompetanse kan beskrives som "[...] the information that is stored in memory that tells us how to represent spoken language in written form" (Apel 2011:592). Hva slags ortografisk kompetanse man som innlærer trenger, vil avhenge av skriftspråket som skal læres, fordi skriftspråk er bygd opp på forskjellige måter. Det som likevel i alle tilfeller må læres i det nye skriftspråket, er skriftkonvensjoner (jf. 2.2.2) og grafotaks (jf. 2.2.4). Å ha fonologisk bevissthet regnes også som en viktig ferdighet ved tilegning av andrespråksortografi (jf. f.eks. Stuart 1999:587; Wade-Woolley & Siegel 1997:387). Fonologisk bevissthet kan defineres som evnen til å forstå lydstrukturen i et språk. Dette innbefatter evnen til å segmentere tale til fonemer og evnen til å påvise og behandle fonemer (Jongejan, Verhoeven & Siegel 2007:836). Riktig persepsjon av fonemer er en ferdighet som gradvis opparbeides ved læring av et nytt språk. Spesielt ved tilegning av ortofone skriftspråk vil fonologisk bevissthet være en viktig kompetanse. Dich & Pedersen (2013) mener at ikke-ortofone skriftspråk er vanskeligere å lære enn ortofone. De begrunner dette med at tilegnere av ikke-ortofone skriftspråk må tilegne seg flere ferdigheter for å lære skriftspråket enn tilegnere av ortofone skriftspråk (Dich & Pedersen 2013:52).

Hagtvet, Helland & Lyster (2006) sier at man trenger grammatisk og lingvistisk metakunnskap for å kunne lese og skrive i norsk (Hagtvet, Helland & Lyster 2006:19). At grammatisk kunnskap skal ha noe med ortografisk kunnskap å gjøre, er kanskje ikke åpenbart. Cook (2010) beskriver at i læring av engelsk ortografi kan det være nyttig å ha noe morfologisk kompetanse. For eksempel har man alltid <s> finalt i tredje person presens, som i *fits*, *annoys* og *snatches*. Dette selv om grafemet representerer tre forskjellige uttaler – /s/, /z/ og /ɪz/ (Cook 2010:475). I materialet mitt finner jeg eksempler som "nivåe". Det er usikkert om det menes *nivå* eller *nivået* (eller *nivåer*), men med mer morfologisk kompetanse ville man uansett visst at dette ikke er en mulig endelse. En innlærer som vil opparbeide seg ortografisk kompetanse i norsk, vil også kunne ha nytte av å kjenne til avledningsaffikser. Hvis man for eksempel kjenner prefikset *be-*, vil man sannsynligvis kunne gjette seg til at det heter *besvare* og ikke *bisvare*. Og tilsvarende om man kjenner suffikset *-het*; man vil vite at *leilighet* ender på *-het*, og ikke for eksempel *-hit* (jf. 6.4.4).

## 2.4 Ortografisk feil

Jeg slutter meg til Hvenekildes (1995) beskrivelse av galt ordbilde: "[...] ordformer som ikke er medlemmer av det aktuelle rotlemmet, det vil si ikke er avlednings- og/eller bøyingsform av den samme rota [...]" (Hvenekilde 1995:176). Hvenekilde omtaler altså ord med ortografisk feil som *galt ordbilde* – tilsvarende bruker hun *riktig ordbilde* om riktig stavede ord. Hun eksemplifiserer beskrivelsen over med at "dammen" for *damen* og "yey" for *jeg* følgelig vil være ord med galt ordbilde, mens "reiste" for *reisen* vil ha riktig ordbilde. Om ordet er brukt feil semantisk eller grammatisk, har altså ikke betydning. For eksempel vil feil verbtid eller feil i bestemthet eller verb for substantiv (som i eksemplet over) ikke ha betydning. Bare ordets *ortografiske* form har betydning (ibid.). Til Hvenekildes beskrivelse vil jeg legge til at hvis det går klart fram av konteksten at det er et annet ord som menes (selv om ordet er riktig stavet), har det likevel ortografisk feil – som for eksempel ved homofoner (*godt/gått*).

Av en informants ortografiske feil kan man skille mellom systematiske og ikke-systematiske. Systematiske feil (*error*) er feil som gjøres flere ganger, mens ikke-systematiske (*mistake*) – glipp – er feil som bare gjøres én gang (Ellis & Barkhuizen 2005:64, jf. Corder 1974). Når en innlærer gjør systematiske feil, kan det tolkes som at den aktuelle målspråksformen ikke er

tilegnet. En ikke-systematisk feil trenger derimot ikke å være knyttet til innlærerens kompetanse. I mitt eget materiale finner jeg flere tilfeller av systematiske feil, som fem tilfeller av "ulike/uliker/brantuliker" for *ulykke/ulykker/brannulykker*. Også ikke-systematiske feil som forekommer én gang hver, finner jeg flere tilfeller av materialet – "jer" for *jeg*, "på" for *på*, "eg" for *er* og "ot" for *at* (jf. 6.4.3). I oppgaven vil jeg ikke skille mellom *feil* og *glipp*, men bruke feil som overordnet begrep for dem begge.

## 2.5 Tidligere norske undersøkelser i andrespråksortografi

Etter hva jeg vet, er det ingen andre norske undersøkelser i andrespråksortografi enn Anne Hvenekildes to undersøkelser som kom ut på 1980-tallet (Hvenekilde 1985 og 1987). Den siste undersøkelsen tok for seg ortografien til elever i andre til niende klasse med arabisk og berbersk som morsmål. I undersøkelsen gjorde elevene med berbersk som morsmål flere feil enn elevene med arabisk som morsmål. Begge elevgruppene hadde flere feil ved vokalgrafemer enn ved konsonantgrafemer (jf. Hvenekilde 1987:1, 20, jf. Vigrestad 2015:56). Av Hvenekildes to undersøkelser, er det likevel den første som har mest relevans for min egen undersøkelse, fordi en av informantgruppene i undersøkelsen har vietnamesisk som morsmål. Etter at artikkelen ble publisert første gang i 1985, ble den revidert og utgitt på nytt i 1995 i Bjørkavåg, Hvenekilde & Ryen (red.) – det er denne utgaven jeg vil vise til.

Elevene i undersøkelsen gikk i andre til niende klasse, og hadde vært i Norge fra ett til seks år. De hadde vietnamesisk og tyrkisk som morsmål. Det viste seg i undersøkelsen at elevene med tyrkisk som morsmål hadde flest feil. De gjorde blant annet en del feil ved grupper av konsonantgrafemer, noe Hvenekilde mener kan være relatert til ulik stavellesstruktur i norsk og tyrkisk (jf. Hvenekilde 1995:175, 179-181, 187, jf. Vigrestad 2015:58–59). Mens norsk har en kompleks stavellesstruktur (jf. 3.3), er stavellesstrukturen i tyrkisk enklere – (K)V(K)(K). Hvenekilde så spesielt på grupper av konsonantgrafemer initialt i ord. Elevene med tyrkisk som morsmål gjorde følgende feil her: 1. Innskudd av én eller flere vokalgrafemer – "sipille" for *spille*, 2. Ombytting i grafemrekkefølgen slik at det kommer et vokalgrafem mellom konsonantgrafemene – "sipse" for *spise*, 3. Redusering av antall konsonantgrafemer – "sømt" for *svømt*. Elevene med vietnamesisk som morsmål gjorde derimot få feil ved grupper av konsonantgrafemer initialt i ord: Bare 3 av 93 elever hadde slike feil (jf. Hvenekilde

1995:179-180, 187, jf. Vigrestad 2015:58–61). I mitt eget materiale finner jeg også en del feil ved grupper av konsonantgrafemer (jf. 6.2.2 og 6.3.3). Den mest frekvente feiltypen her tilsvarer Hvenekildes kategori 3.

## 2.6 Forskningsspørsmål

Som nevnt i innledningen, er det primære målet med undersøkelsen å forsøke å svare på om eventuelle forskjeller i ortografien til VIET og BKS kan relateres til informantenes morsmål og/eller morsmålskriftspråk. Det overordnede forskningsspørsmålet lyder:

**Er det forskjeller i ortografien til innlærerne med vietnamesisk og BKS som morsmål, og kan eventuelle forskjeller relateres til morsmålet og/eller morsmålskriftspråket deres?**

For å kunne svare på dette, har jeg formulert noen underordnede forskningsspørsmål. Disse er relatert til ortografiske forskjeller mellom de to informantgruppene, som jeg mener kan belyse om det er morsmålspåvirkning.

### **Forskingsspørsmål 1: Har VIET en høyere andel feil ved gruppe av konsonantgrafemer enn BKS?**

Verken ortografien i vietnamesisk eller BKS har like store grupper av konsonantgrafemer eller like mange mulige kombinasjoner av konsonantgrafemer som norsk. Jeg ville derfor se på om dette kan være vanskelig å tilegne seg for VIET og BKS. Det er også kjent at tilegning av konsonantgrupper i uttalen i norsk kan være vanskelig (jf. f.eks. Husby & Kløve 1998:185). Jeg ville undersøke om det samme gjaldt for grupper av konsonantgrafemer i skrift. Det muntlige og det skriftlige fenomenet vil av flere grunner ikke være helt sammenlignbart, blant annet fordi jeg bruker en videre definisjon av det jeg kaller gruppe av konsonantgrafemer enn for det muntlige fenomenet konsonantgruppe (jf. 4.4.3.1).

I BKS har man grupper av konsonantgrafemer, men disse er ikke like lange som i norsk, og man har færre mulige kombinasjoner av konsonantgrafemer (jf. 3.5). I vietnamesisk kan man ha to konsonantgrafemer i gruppe – men disse uttales som ett fonem. Man har altså ikke konsonantgrupper i uttalen i vietnamesisk (jf. 3.4). Om det er slik at informantenes

konsonantstruktur i morsmålet og/eller morsmålsskriftspråket vil ha noen påvirkning på tilegning av grupper av konsonantgrafemer i norsk, vil jeg derfor anta at VIET vil ha en høyere andel feil ved gruppe av konsonantgrafemer enn BKS.

**Forskningsspørsmål 2: Er det spesielt mange feil ved grafemer/polygrafemer som gjengir /t/, /d/, /l/ <sup>7</sup> /n/, /s/, /ç/ og /ŋ/ – og er det forskjeller mellom VIET og BKS?**

Fonemene /t/, /d/, /l/, /n/, /s/, /ç/ og /ŋ/ har svært inkonsekvente stavemåter i norsk – dette er hovedgrunnen til at jeg ville undersøke grafemer/polygrafemer som gjengir dem. /ç/-fonemet skrives for eksempel <k> i *kino*, <kj> i *kjole* og <tj> i *tjue*, og /ŋ/-fonemet skrives <ng> i *sang* og <n> i *tank*. Det fonemet som har flest forskjellige skrivemåter, er /s/, som blant annet har skrivemåten <skj> i *skjorte*, <rs> i *norsk* og <j> i *journalist*. I tillegg er det begrensninger på hvor i stavelsen en bestemt stavemåte kan stå, som for eksempel at <rs> for /s/ bare kan stå i koda – som i *norsk* – og aldri i opptakt (jf. 2.2.4 og 3.3). Den andre grunnen til at jeg ville undersøke om skrivemåten for fonemene /t/, /d/, /l/, /n/, /s/, /ç/ og /ŋ/ er vanskelig å tilegne seg, er at få av fonemene finnes i BKS og vietnamesisk (jf. henholdsvis 3.5 og 3.4). I og med at fonemene i de fleste tilfeller gjengis med grupper av konsonantgrafemer, vil jeg anta at VIET vil ha flere feil her enn BKS (jf. forskningsspørsmål 1).

Fonemene /t/, /d/, /l/, /n/, /s/, /ç/ og /ŋ/ representeres av et polygrafem, og ikke med ett grafem, i de fleste ord; <rn> i *stjerne*, <skj> i *skjerf*, <kj> i *kjelke*, <rt> i *ert* og så videre. Dette forskningsspørsmålet kan derfor ses i sammenheng med forskningsspørsmål 1 og 3, som begge ser på feil ved grupper av konsonantgrafemer. En mulighet kunne derfor ha vært å slå dette forskningsspørsmålet sammen med forskningsspørsmål 1. Likevel er det slik at selv om flesteparten av fonemene i de fleste tilfeller skrives med et polygrafem, er det flere av fonemene som i noen ord gjengis med bare ett grafem. Eksempler på ord der fonemene gjengis med ett grafem, er: *tank*, *punkt*, *adjunkt*, *senke*, *kompetanse* (der /ŋ/ skrives med <n>), *kino*, *kirke*, *kikke*, *kile*, *kylling* (der /ç/ skrives med <k>) og *journalist*, *journal* og *giro* (der /s/ skrives med <j> og <g>). Jeg mener derfor det er mest riktig at de to fenomenene undersøkes hver for seg.

**Forskningsspørsmål 3: Har VIET innskudd av vokalgrafemer i grupper av konsonantgrafemer i større grad enn BKS?**

---

<sup>7</sup> Jeg ser bare på polygrafisk gjengivelse av /l/, altså der den gjengis <rl>.

Jeg viser til problemstillingen i forskningsspørsmål 1. Om det nå er slik at VIET gjør feil ved grupper av konsonantgrafemer, går denne feilen ut på å skyte inn vokalgrafemer – jf. det muntlige fenomenet epentese (jf. 2.3.2). I så fall kan dette være en måte å gjøre strukturen i ortografien i norsk mer lik den i uttalen og ortografien i vietnamesisk.

**Forskningsspørsmål 4: Er enkelt for dobbelt konsonantgrafem en mer frekvent feil enn dobbelt for enkelt konsonantgrafem, og er det være forskjeller mellom VIET og BKS?**

Jeg vil anta at skriftkonvensjonen med enkeltskriving og dobbeltskriving av konsonantgrafemer vil være vanskelig å tilegne seg i norsk. Vietnamesisk har heller ikke dobbelt konsonantgrafem (jf. 3.4), mens BKS kun har det i liten grad (jf. 3.5). Jeg antar at VIET vil gjøre flere feil enn BKS fordi strukturen av konsonanter/konsonantgrafemer i vietnamesisk er mer forskjellig enn mellom konsonant(grafem)strukturen i norsk og BKS. I BKS har man også distinktiv vokallengde, selv om dette ikke uttrykkes i skrift. Vietnamesisk har derimot ikke distinktiv vokallengde.

**Forskningsspørsmål 5: Er diftongering av monoftong og monoftongering av diftong frekvente feil, og er det forskjeller mellom BKS og VIET?**

I skriftbildet i norsk har vi både monoftonger og diftonger. Dette varierer fra ord til ord (som i *teppe* og *skøyter*), men det kan også være flere mulige ordformer av det samme ordet – der den ene varianten har monoftong og den andre diftong, som i *ren/rein* og *brøt/brøyt*. Også i uttalen varierer vi mye mellom monoftong og diftong, både fra det ene ordet til det andre – og fra den ene dialekten eller sosiolekten til den andre. Denne vekslingen mellom monoftong og diftong i norsk var grunnen til at jeg ville undersøke om dette er vanskelig å lære seg. Jeg antar at VIET vil mestre vekslingen mellom monoftonger og diftonger bedre enn BKS fordi de i eget morsmål har mange vokaler og vokalgrafemer, og at de slik sett vil mestre vokalgrafemer – da også vekslingen monoftonger/diftonger – på en god måte. I vietnamesisk har man også diftonger (jf. 3.4), i motsetning til i BKS (jf. 3.5).

**Forskningsspørsmål 6: Er substitusjonsfeil ved ⟨u⟩, ⟨y⟩ og ⟨ø⟩ frekvente – og er det forskjeller mellom VIET og BKS?**

Det er kjent at de runde, fremre vokalene /u, u:/, /y, y:/ og /ø, ø:/, som gjengis henholdsvis ⟨u⟩, ⟨y⟩ og ⟨ø⟩, kan være vanskelige å uttale for innlærere. Vokalene finnes ikke i verken BKS eller vietnamesisk (jf. henholdsvis 3.5 og 3.4). Selv om jeg ikke vet i hvor stor eller liten grad eventuelle vanskeligheter med uttale er relevant for ortografitlegning, synes jeg det er

interessant å se på i hvor stor grad grafemene som gjengir disse fonemene, mestres. Jeg antar at VIET vil mestre ⟨u⟩, ⟨y⟩ og ⟨ø⟩ bedre enn BKS fordi de har et rikere vokalsystem enn BKS, og at denne variasjonen i morsmålet og morsmålsskriftspråket vil gjøre dem bedre i stand til å tilegne seg de norske vokalgrafemene. BKS har derimot få vokaler i morsmålet – fem – både i uttale og i ortografien.

## 2.7 Oppsummering av kapitlet

Det teoretiske utgangspunktet som drøftingen av resultatene hviler på, er presentert i dette kapitlet. Jeg har også behandlet sentrale begreper som ortografi, ortografisk feil, skriftspråk, skriftkonvensjoner, ortofoni og grafotaks. I tillegg har jeg forsøkt å avklare hva jeg forstår med andrespråksortografi – hva som ligger i begrepet og hvordan det kan forstås og tolkes. Påvirkning fra morsmålet og/eller morsmålsskriftspråket er sentralt, men andrespråksortografien farges også av skriftspråket som skal læres og hva slags egenskaper det har. I kapitlet har jeg også trukket fram tidligere norske undersøkelser i andrespråksortografi. Kapitlet ble avsluttet med en presentasjon av det overordnede forskningsspørsmålet – samt seks underordnede forskningsspørsmål.

# 3 Ortografien i norsk, vietnamesisk og BKS

## 3.1 Innledning

I dette kapitlet vil jeg beskrive ortografien i norsk, vietnamesisk og BKS. Det er viktig å belyse ortografien i vietnamesisk og BKS da det er grunn til å anta at denne kan ha betydning for ortografien i norsk hos disse innlærerne. De fonologiske systemene i vietnamesisk og BKS vil også være med i beskrivelsene siden disse også kan ha påvirkning på andrespråksortografien, slik jeg ser det. I kapitlet beskriver jeg også ortografien og fonologien i norsk, tilegnelsesspråket i undersøkelsen. Beskrivelsene av ortografien og fonologien i norsk, vietnamesisk og BKS er ikke fullstendige beskrivelser; jeg legger vekt på det jeg mener er relevant for undersøkelsen.

I første del av kapitlet forklarer jeg hvorfor jeg har valgt å undersøke ortografiske feil hos innlærere med BKS og vietnamesisk som morsmål (3.2). Deretter beskriver jeg ortografien og fonologien i henholdsvis norsk (3.3), vietnamesisk (3.4) og BKS (3.5). Til slutt oppsummeres kapitlet (3.6).

## 3.2 Hvorfor språkene BKS og vietnamesisk i undersøkelsen?

Da jeg skulle velge ut språk til undersøkelsen, var de viktigste kriteriene av språklig art. Det var i tillegg bare språk som informantene i ASK-korpuset hadde som morsmål som var aktuelle fordi det var dette korpuset jeg skulle bruke. Jeg ville også velge ut språk som har en relativt høy andel språkbrukere i Norge. De viktigste kriteriene var likevel, som nevnt, av språklig art. For siden jeg antar at ortografien jeg undersøker, vil være påvirket av ortografien og fonologien i morsmålet til innlærerne, er det et poeng at de innlærerne jeg undersøker ortografien til, har morsmål/morsmålsskriftspråk som er ulike hverandre. Med ulike mener jeg at de har forskjellige trekk hva gjelder for eksempel skriftkonvensjoner, fonotaks, grafotaks



og fonem- og grafeminventar. Hvis denne undersøkelsen skulle vise at BKS og VIET gjør forskjellige ortografiske feil, er det grunn til å anta at noe av årsaken til dette kan ligge i påvirkning fra innlærernes morsmålsortografi/fonologi. Det er derfor viktig at innlærerne har morsmål som er ulike hverandre. Hvis jeg hadde hatt innlærere med mer like morsmål, ville det vært vanskeligere å påvise om feilene kan være knyttet til morsmålene. I tillegg til at det er viktig at morsmålene til de to informantgruppene er forskjellige fra hverandre, ville jeg også velge innlærere med morsmål som er forskjellige fra norsk. Valget falt derfor på BKS og vietnamesisk.

### 3.3 Ortografien i norsk

I norsk bruker vi det latinske alfabetet. Det består av 29 grafemer – 20 konsonantgrafemer og 9 vokalgrafemer. Vokalgrafemene representerer 9 korte og 9 lange vokaler – altså har vi atten vokaler med betydningsskillende funksjon. Husby & Kløve (1998) sier om språk med mange vokaler at en følge er at man tar i bruk sekundærvokaler i tillegg til primærvokalene. Disse vokalene er sjeldne i verdens språk, og kalles også merkede. De fremre, rundede vokalene /ʌ/, ʊ:/, /y/, /y:/, /ø/ og /ø:/ i norsk er merkede. Vokalene gjengis alltid av henholdsvis <u>, <y> og <ø> (Husby & Kløve 1998:41, 44-45, 127).

I norsk har vi også diftonger. Disse er /ei/, /ai/, /ɔy/, /ʉy/, /œy/ og /æʌ/. Diftongene er fallende, det vil si at de begynner på en mer sonor vokal enn de slutter på (Husby & Kløve 1998:50). Kulbrandstad (2005) definerer en diftong som en sekvens av to ulike vokaler i samme stavelse (Kulbrandstad 2005:46). To etterfølgende vokaler som ikke opptrer i samme stavelse, er dermed ikke en diftong. Eksempler på dette finner man i fremmedordene *proteiner* og *naiv*, og i sammensatte ord som *babyelefant* og *sjokoladeis*. I skriftbildet har vi mye variasjon mellom enkelt og dobbelt vokalgrafem. Variasjonen finnes både mellom ord (enkelt vokalgrafem i *se* – gruppe av vokalgrafemer i *øy*), men vi har også variasjon på ordnivå i det at vi har valgfrie stavemåter som *sten/stein* og *skjøt/skøyt*.

Standard østnorsk har 20 konsonantfonemer (Husby & Kløve 1998:53). Sett i et andrespråkperspektiv er det enkelte forhold som peker seg ut, som for eksempel at vi har aspirasjon av ustemt /p t k/ (/p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>/). De fleste språk har en stemthetskontrast mellom plosivene /b d g/ og /p t k/, men ingen aspirasjonskontrast i tillegg. Vi har også stemthets- og aspirasjonskontrast mellom retrofleksene /t/ og /d/. Andre retroflekser i østnorsk er /ŋ/ og /ʃ/ og

/l/. Fonemet /t/ gjengis blant annet av <rt> (*ert*), /d/ gjengis av <rd> (*ferdig*), /l/ gjengis av blant annet <rl> (*herlig*) og /ŋ/ gjengis av <rn> (*stjerne*). /ʃ/ har en langt mer inkonsekvent stavemåte. Fonemet gjengis av flere forskjellige grafemkombinasjoner. De vanligste er <sk> (*skjorte*), <sj> (*sjonglere*) og <rs> (*norsk*)<sup>8</sup>. Det er også restriksjoner på hvor i ordet de forskjellige stavemåtene kan stå – for eksempel kan <rs> bare stå i koda, som i *vers*, *farse* og *person*. Nasalen /ŋ/ og frikativen /ç/ har også vanskelige stavemåter, hovedsakelig fordi de gjengis på flere måter. /ŋ/ representeres av <ng> i *sang*, <g> i *agn* og <n> i *tanke*, mens /ç/ representeres av <k> i *kino*, <kj> i *kjole* og <tj> i *tjue*.

Norske stavelser er bygd opp av en stavelsesbærende vokal i kjernen og konsonant(er) i opptakt og/eller i koda. Stavelseskjernen kan også være en diftong, og i noen tilfeller også en konsonant; /n/ og /l/ er stavelsesbærende i /hatn/ (*hatten*) og /tødl/ (*tøddel*) (Kulbrandstad 2005:54). Norsk har en kompleks stavelsesstruktur som tillater flere konsonanter både i opptakt og koda. Opptakter kan ha maksimalt tre konsonanter, mens koda kan ha opptil fem – som i ordet *skjelmskt*. Det er stor variasjon på hvor kompleks opptakt og koda er i stavelser i forskjellige språk. Norsk befinner seg i den ene enden av skalaen sammen med blant annet engelsk, russisk og BKS. I den andre enden av skalaen har man blant annet japansk og vietnamesisk. Forskjeller språk imellom på stavelsesnivå dreier seg likevel ikke bare om *antallet* konsonanter i stavelser, men like mye om *distribusjonen* (Husby & Kløve 1998:65–66). I norsk har vi for eksempel mange mulige konsonantkombinasjoner. Av konsonantgrupper på tre eller flere konsonanter er åtte kombinasjoner mulige i opptakt og trettifire i koda (op.cit.:83). Vietnamesisk har flere begrensninger (jf. 3.4). Videre er konsonanten /r/ den som forekommer i flest konsonantgrupper, og er en av de konsonantene som har hyppigst forekomst i ordinventaret i norsk. I tillegg til at den forekommer hyppig i leksikalske morfem, er den også presensmorfem og flertallsmorfem (op.cit.:87).

*Det ortofone prinsippet* er underliggende i norsk rettskriving på den måten at hovedregelen er at ord blir representert i skrift ved hjelp av tegn for hvert av fonemene i uttalen av ordet (Kulbrandstad 2005:63, jf. 2.1.3). Det finnes likevel mange unntak fra denne regelen – jeg vil nevne noen av dem. Vi har mange stumme konsonanter i norsk. Eksempler er det siste grafemet i *mulig*, *hard*, *land* og *kald*. Intetkjønns-substantiver i bestemt form entall har stum *-t* finalt, som i *huset*, *vannet* og *spannet*. I ord som *hjem* og *gjøre* uttales ikke det første

---

<sup>8</sup> Inkluderer man lånord, vil lista bli mye lengre. For eksempel har man <g> i *giro* og *gelatin* og <j> i *journalist* og *ajour*.

konsonantgrafemet. Det er også slik at grafemer kan få forskjellig uttale avhengig av konteksten. Hva er for eksempel uttalen av *kosten*, /kɔ̃sten/ eller /kɔsten/? Av ordet *lyst* finnes også to forskjellige uttaler – /lyst/ og /ly:st/ (Husby & Kløve 1998:126). Tilsvarende kan samme lyd ha forskjellige skrivemåter. Fonemet /ʊ/ gjengis med både ⟨o⟩ og ⟨u⟩, som i henholdsvis *sol* og *sukker*. Ellers er trykk og tonem er viktige uttaletrekk i norsk, men de blir ikke markert i ortografien.

Over nevnte jeg at vi har atten vokaler i norsk – ni korte og ni lange – og at vi altså har distinktiv vokallengde i norsk. Vokallengden blir indirekte gjengitt, på den måten at den leses ut ifra vokalgrafemets etterfølgende konsonantgrafem(er) – man kan altså ikke lese om vokalen er lang eller kort ut ifra selve vokalgrafemet (Kulbrandstad 2005:64). Ifølge Kulbrandstad har fordeling av enkeltskrevet og dobbeltskrevet konsonantgrafem først og fremst sammenheng med stavelsestrykk, vokallengde og antall konsonantlyder etter vokalen i stavelsen (op.cit.:70). Hovedregelen er at lang vokallyd etterfølges av enkelt konsonantgrafem og kort vokallyd etterfølges av dobbelt konsonantgrafem. Dette er likevel en regel med mange unntak. Grafemet ⟨v⟩ dobbeltskrives ikke, som i *lever* og *høvel*. Grafemet ⟨m⟩ dobbeltskrives ikke finalt i ord: *fram* - *framme*, *rom* - *rommet*. Disse kaller Kulbrandstad *generelle unntak*. Han sier at vi også har en del *enkeltunntak*. Disse kan deles i tre kategorier: 1) Mange vanlige småord har enkelt konsonantgrafem selv om vokalen er kort, som i *an*, *at*, *bør*, *din*, *hos*, *hvis*, *kan*, *man*, *men*, *nok*, *skal*, *til*, *vel* og *vil*. 2) Noen fremmedord skrives med enkelt konsonantgrafem selv om vokalen i trykkstavelsen er kort, som i *infinitiv*, *konditor* og *negativ*. 3) Noen fremmedord skrives med dobbelt konsonantgrafem selv om stavelsen ikke har trykk, som i *abonnement*, *attest* og *kollekt* (ibid.). Kulbrandstad sier videre om regler for bruk av enkelt/dobbelt konsonantgrafem at mange ord blir skrevet med dobbelt konsonantgrafem i noen ordformer og med enkelt konsonantgrafem i andre. Forenkling av dobbelt konsonantgrafem har man i disse tre tilfellene: 1) Når konsonantgrafemet står foran en bøyingsendelse som begynner på et konsonantgrafem: *eventuell-eventuelt*, *bygge-bygde*, *fylle-fylte* og *sann-sant*. 2) I sammendratte bøyingsformer: *artikkel-artikler*, *middel-midler* og *sikker-sikre*. 3) Foran noen avledningsendelser som begynner på et konsonantgrafem: *-ling* (*ætt-ætlings*), *-ne* (*tett-tetne*), *-ning* (*redd-redning*) og *-sk* (*sprell-sprelsk*)(op.cit.:71).

Avslutningsvis om ortografi i norsk, vil jeg trekke fram at det i mange tilfeller er valgfrihet mellom ortografiske former. Dette er ikke vanlig i andre skriftspråk. Vi kan for eksempel velge mellom former som *sen* og *sein*, *juice* og *jus*, *kapitlet* og *kapittelet*, *lete* og *leite*.

### 3.4 Ortografien i vietnamesisk

Vietnamesisk skrives med det latinske alfabetet. Skriftspråket ble laget midt på 1600-tallet, og gjenspeiler en datidig nordlig dialekt med distinksjoner som ikke finnes i språket i dag. Skriftspråket var antagelig nært et fonemisk system da det ble skrevet ned, men da språket har forandret seg siden den gang, er det ikke lenger fonemisk i dag (Rosén 2001:11). Alfabetet består av 29 grafemer – 17 konsonantgrafemer og 12 vokalgrafemer. I norsk har vi tre vokalgrafemer og fire konsonantgrafemer som ikke finnes i vietnamesisk. Disse er: <æ>, <ø>, <å>, <f>, <j>, <w> og <z>. Det at vietnamesisk har resten av de norske grafemene i sitt eget skriftspråk, betyr ikke nødvendigvis at de vil være lette å bruke i norsk. Dette har to grunner. For det første vil uttalen av grafemene gjerne ha en annen kvalitet i norsk enn i vietnamesisk. For eksempel uttales <g> i vietnamesisk med en lyd som ligner på en kort fremre skarre-r. For det andre har grafemene ofte ulik distribusjon i skriftspråkene i norsk og vietnamesisk (Husby 1991:33–39). Et annet trekk ved vietnamesisk ortografi er den store bruken av diakritiske tegn. Disse brukes for å angi fonemer og tonemer (Rosén 2001:12). Vietnamesisk har mange tonemer: 4–6 avhengig av hvilken dialekt man har (op.cit.:9).

Som norsk har også vietnamesisk mange vokaler. Språket har elleve vokalmonoftonger, og i tillegg mange diftonger – den sørlige dialekten av vietnamesisk har hele sytten. Norsk og vietnamesisk har begge merkede vokaler; norsk har rundede fremre vokaler, mens vietnamesisk har urundede bakre vokaler. I motsetning til i norsk har man ikke distinktiv vokallengde<sup>9</sup>. Vietnamesisk har videre 23 konsonantfonemer (Rosén 2001:3–6). Av norske fonemer som ofte skrives med en polygraf (jf. 3.3), har ikke vietnamesisk /d/, /l/, /ŋ/, /ʃ/<sup>10</sup> eller /ç/. Fonemet /t/ finnes bare i den sørlige dialekten (Husby 1991:49–55). I motsetning til norsk har ikke vietnamesisk konsonantgrupper. Dette gjelder likevel bare for uttalen. I skrift har man digrafer, men en digraf vil alltid bli uttalt som ett fonem. For eksempel er ordet for bie *ong*, og uttales [ɒŋ], og ordet for onkel er *chú* og uttales [cu]. Også trigrafer eksisterer i

<sup>9</sup> Av a-lyden finnes en kort og lang variant, men disse har litt ulik lyd kvalitet (Husby 1991:45).

<sup>10</sup> I den sørlige dialekten finnes en lyd som ligger mellom norsk /s/ og /ʃ/ (Husby 1991:56).

vietnamesisk, som <ng> (uttales [ŋ]) (Husby 1991:58-60). Dobbelte konsonantgrupper eksisterer ikke.

Rosén (2001) mener at de største forskjellene i norsk og vietnamesisk fonologi ikke handler om enkeltlyder, men om fonotaks. Måten lydene kombineres til stavelser, er veldig forskjellig i de to språkene (Rosén 2001:10). Og på samme måte som fonologien i vietnamesisk har fonotaktiske restriksjoner, har også skriftspråket grafotaktiske restriksjoner. Følgende konsonantgrupper/grupper av konsonantgrupper er de eneste som kan stå finalt i ord: <c>, <ch>, <m>, <n>, <ng>, <nh>, <p> og <t>. Det er derimot langt færre begrensninger initialt i ord (Husby 1991:60). Til sammenligning har norsk færre begrensninger på hvilke konsonantgrupper og grupper av konsonantgrupper som kan stå initialt og finalt (jf. 3.3).

Ord i vietnamesisk skiller seg fra ord i norsk ved at de er monosyllabiske. Ordnivå og stavelsesnivå blir derfor to sider av samme sak i vietnamesisk. Videre er kjernen den eneste obligatoriske delen av stavelsen. Vietnamesisk må ha en vokalisk kjerne – i motsetning til norsk som kan ha stavelsesbærende konsonanter (jf. 3.3). Kjernen er enten en monoftong eller en diftong. I tillegg til kjernen kan det være maksimalt én konsonant i opptakt og én i koda (Rosén 2001:10). To vokaler etter hverandre uttales da enten som en diftong – eller en glidelyd (først) og en monoftong (sist) (Husby 1991:47).

### 3.5 Ortografien i BKS

Bosnisk/kroatisk/serbisk er offisielle språk i Bosnia–Hercegovina, Kroatia, Serbia og Montenegro (Mønnesland 2002:11). Betegnelsen *serbokroatisk* (eventuelt *kroatoserbisk*) ble brukt fram til Jugoslavia ble oppløst i 1991–92. De største forskjellene mellom språkene er på ordforrådsnivå – grammatikken mellom språkene er veldig lik. Grunnen til at man i dag har et skille mellom bosnisk, kroatisk og serbisk, er fordi landområdet der språkene snakkes, tilhører forskjellige nasjoner. Distinksjonen mellom språkene har derfor et politisk snarere enn et lingvistisk grunnlag (op.cit.:25–26). I undersøkelsen behandler jeg språkene som tilhørende samme gruppe fordi språkene nesten er identiske, faktisk likere enn bokmål og nynorsk (op.cit.:11). I tillegg er språkene kategorisert sammen i ASK-korpuset, som jeg bruker i undersøkelsen. Dette igjen fordi språkene er veldig like.

BKS skrives med to alfabeter, det kyrilliske og det latinske. Det er nøyaktig samsvar mellom dem. Ord kan dermed overføres fra det ene alfabetet til det andre uten problemer (Mønnesland 1990:27). Det kyrilliske alfabetet har 30 grafemer, mens det latinske har 27. BKS er et ortofont skriftspråk – hvert fonem tilsvarer (som oftest) ett grafem. Best samsvar er det i det kyrilliske alfabetet. I det latinske alfabetet er det tre fonemer som skrives med en kombinasjon av to grafemer. Alfabetet har også diakritiske tegn ved noen av grafemene: ⟨č⟩, ⟨ć⟩, ⟨đ⟩, ⟨š⟩ og ⟨ž⟩. Disse grafemene finnes også uten diakritiske tegn; grafemene med og uten har ulik uttale (Mønnesland 2002:27).

Som nevnt over er skriftspråket i BKS svært ortofont. Dette er blant annet fordi assimilasjoner blir uttrykt i skrift. Mønnesland sier at man har *regressiv assimilasjon* i BKS, og forklarer det slik: "Dersom to konsonanter står ved siden av hverandre og den ene er stemt og den andre ustemt, blir den første påvirket av den som følger etter" (Mønnesland 2002:41). Regelen blir altså: stemt + ustemt = ustemt, og ustemt + stemt = stemt. Om man overfører dette mønsteret til norsk, ville *styggt* ha blitt skrevet *stykt*. Jeg vil trekke fram to eksempler fra BKS: *nizak* (lav) → *nisko* (lavt) og *svat* (bryllupsgjest) → *svadba* (bryllup)(ibid.).

I BKS har man et fem-vokalsystem. Vokallydene er det samme basissystemet som finnes i mange språk, nemlig /i e a o u/. De norske vokalfonemene /ʉ y æ ø/ i kort og lang variant finnes altså ikke i språket. Vokallengden er distinktiv i BKS (Mønnesland 2002:28), som i norsk. Husby og Kløve (2001) definerer systemet derfor som et ti-vokalsystem (Husby & Kløve 2001:48), mens Mønnesland (2002) definerer det som et fem-vokalsystem (Mønnesland 2002:28). I motsetning til i norsk gjengis ikke skillet mellom korte og lange vokallyder i skrift. Ved vokallydene i BKS har man heller ikke diftonger. Kombinasjonen av to etterfølgende vokalgrafemer vil derfor uttales som to separate lyder, som i ordet *sauna* (badstu): /sa-una/ (Mønnesland 2002:29).

BKS har 25 konsonantfonemer. Språket har flere frikativer og affrikater som ikke finnes i norsk. Andre veien har vi noen konsonanter i norsk som ikke finnes i BKS, blant annet /ç/, /ŋ/ og /h/ – i tillegg til retrofleksene /t/, /d/, /l/ og /ŋ/. Som norsk har også BKS /ʃ/-fonemet (Mønnesland 1990:24–25). BKS har også konsonantgrupper. Disse kan forekomme både i opptakt og koda i en stavelse. I BKS har man derimot flere begrensninger, sammenlignet med norsk, på hvilke konsonantforbindelser som kan stå i koda. De man finner i koda er /st/, /ʃt/,

/zd og /zd/ (Husby & Kløve 2001:85). Av enkeltkonsonantlyder er det 13 som kan forekomme i koda i stavelen i BKS mot 22 i norsk (Husby & Kløve 2001:78). Ved konsonanter som kan forekomme i opptakt, er det derimot et rikt utvalg – både når det gjelder enkeltkonsonanter og konsonantforbindelser (op.cit.:78, 85).

Grafemet ⟨r⟩ har to funksjoner i BKS. Når den står sammen med en vokal, markerer den konsonantfonemet /r/. Ellers fungerer den som en stavelsesbærende vokal. Den kalles i disse tilfellene for *vokalisk r* eller *syllabisk r*. Vokalisk r danner kjernen i en stavelse og forekommer mellom to konsonanter (som i *srpski* – serbisk) eller først i et ord foran en konsonant (som i *rt* – odde), og for det meste bare i betonte stavelser (Mønnesland 2002:29). Det som derfor med norske øyne kan se ut som en lang konsonantforbindelse, som for eksempel i ordet *čvrst* (fast), består derimot egentlig av en konsonantforbindelse på to fonemer ganger to med en vokalisk r i midten (Mønnesland 1990:25).

Fenomenet dobbelt konsonantgrafem er mye mindre brukt i BKS enn i norsk. Hovedregelen er at når to like konsonantgrafemer kommer side om side, faller det ene bort. Dette kan illustreres slik: Ordet for *skille* er satt sammen av *od* og *dijeliti*, og blir skrevet *odijeliti*. Og *betydningsløs* er satt sammen av *bez* og *značājan*, og blir skrevet *beznačājan*. I to tilfeller får ord likevel dobbelt konsonantgrafem: 1) etter forstavelsen *naj* i komparativ og superlativ av adjektiv når det begynner med ⟨j⟩, som i *najjačī* – sterkest, og 2) i noen sammensetninger, som for eksempel i *nuzzârada* – biintekt, og *poddijalekt* – underdialekt (Mønnesland 2002:43).

### 3.6 Oppsummering av kapitlet

Ortografien og fonologien i norsk, vietnamesisk og BKS har vært tema for kapitlet. At informantenes morsmål skulle være ulike hverandre og ulike norsk, var det viktigste kriteriet ved utvelgelsen av morsmålsspråk i undersøkelsen. Å påvise eventuell påvirkning fra morsmålet/morsmålsskriftspråket kunne vært vanskelig dersom informantene hadde morsmål som ikke var forskjellige nok fra norsk eller fra hverandre. Valget falt på vietnamesisk og BKS.

Jeg vil nevne noen av forskjellene mellom de ulike språkene for å summere opp kapitlet (jf. også tabell 1). Norsk har atten vokaler, vietnamesisk har elleve og BKS har ti. Av diftonger har vietnamesisk sytten, norsk har fem og BKS har ingen. Ved konsonanter er det mindre forskjell i antall: BKS har tjuefem konsonantfonemer, vietnamesisk har tjuetre og norsk har tjue. Oppbyggingen av stavelser er derimot et område med store forskjeller mellom språkene. I vietnamesisk kan en stavelse ha maksimalt én konsonantlyd i opptakt og én i koda – disse konsonantlydene skrives enten med ett konsonantgrafem eller med en gruppe på to konsonantgrafemer. Både BKS og norsk kan ha opptil flere konsonanter i opptakt og koda, men i norsk er det mulig med flere kombinasjoner av konsonantgrupper og med høyere antall konsonanter enn i BKS – både i tale og skrift. Videre er det i hovedsak bare norsk av de tre språkene som har dobbelt konsonantgrafem – selv om BKS også har det i noen tilfeller. BKS er det eneste av de tre språkene som er tilnærmet ortofont – i norsk og vietnamesisk er det større avstand mellom skriftbilde og uttale enn i BKS. Distinktiv vokallengde finnes i norsk og BKS, men bare norsk gjengir det i skrift. Vietnamesisk har ikke betydningsskilte mellom kort og lang vokallyd.

	<b>Norsk</b>	<b>Vietnamesisk</b>	<b>BKS</b>
Bruker det latinske alfabetet?	Ja.	Ja.	Det latinske og det kyrilliske alfabetet.
Konsonantforbindelse i uttale/skrift?	Ja/ja.	Nei/ja.	Ja/ja.
Polygrafemer	Ja.	Ja.	Noen polygrafemer i det latinske alfabetet, men ingen i det kyrilliske.
Har fonemene /t/, /d/, /l/, /ŋ/, /ʃ/, /ç/ og /ŋ/?	Ja.	Har /ŋ/ og /ŋ/. I sør finnes også /t/.	Bare /ʃ/.
Distinktiv vokallengde?	Ja.	Nei.	I uttale, men ikke i skrift.
Dobbelt konsonantgrafem?	Ja.	Nei.	I liten grad.
Diftong?	Ja. To vokalgrafemer ved siden av hverandre er enten en diftong eller de tilhører hver sin stavelse.	Ja. To vokalgrafemer ved siden av hverandre er enten en diftong eller en glidelyd pluss en monoftong.	Nei. To vokalgrafemer ved siden av hverandre tilhører hver sin stavelse.

Tabell 27: En sammenligning av fonologiske og ortografiske trekk, samt av alfabeter, i norsk, vietnamesisk og BKS.



# 4 Materiale og metode

## 4.1 Innledning

Som nevnt i kapittel 1 har det vært forsket lite på ortografi i norsk som andrespråk. Det betyr at det ikke er noen etablert praksis på området hva gjelder bruk av metode. I utføringen av undersøkelsen ser jeg først på det totale antallet ortografiske feil i VIET og BKS. Deretter kategoriserer jeg alle de ortografiske feilene etter en kategoriseringsinndeling som har som formål å belyse forskningsspørsmålene. Materialet har jeg hentet fra Språkprøven i ASK-korpuset. I kapitlet vil jeg i flere tilfeller eksemplifisere ved å vise til ord med ortografiske feil. Ordene er enten autentiske fra det utvalgte materialet eller konstruerte.

Jeg begynner kapitlet med å forklare hvorfor jeg har valgt å bruke korpus i undersøkelsen og å beskrive ASK-korpuset, samt beskrive hvordan jeg har brukt det (4.2). Deretter beskriver jeg henholdsvis valg av metode (4.3), kategoriinndelingen jeg plasserer feilene i (4.4) og metode for utregning (4.5). Til slutt oppsummerer jeg kapitlet (4.6).

## 4.2 Bruk av korpus i undersøkelsen

### 4.2.1 Hvorfor korpus?

I undersøkelsen ville jeg, som nevnt, finne ut hva slags feil som gjøres i ortografi i norsk som andrespråk av VIET og BKS, og om det er forskjeller mellom informantgruppene. For å kunne si noe om dette trengte jeg et relativt stort materiale. At materialet er stort, er viktig for å kunne se sammenhenger, og for å kunne sammenligne og trekke slutninger på et mer sikkert grunnlag enn om materialet var lite. Valget falt på å bruke korpus. Den største fordelen med korpus er at det har et stort materiale som er lett å behandle statistisk. Bruk av korpus gjør det enkelt å regne ut feilfrekvens, og å se variasjon og mønster i hva slags feil som gjøres. Det er også en fordel at man unngår å la egne hypoteser farge datainnsamlingen, og at nye og overraskende data kan dukke opp (jf. Johannessen 2003:149).

### 4.2.2 ASK- korpuset

Jeg har valgt å hente materialet i undersøkelsen min fra ASK-korpuset. Dette er et skriftlig innlærerkorpus som er utviklet ved Universitetet i Bergen. Det består av språkprøver på to nivåer, Språkprøven i norsk for voksne innvandrere og Test i norsk – høyere nivå. Jeg ser bare på førstnevnte. Språkprøvene består av tekster på 200–400 ord. Innlærerne som har tatt testene, har ti forskjellige morsmål: tysk, nederlandsk, engelsk, spansk, russisk, polsk, BKS<sup>11</sup>, albansk, vietnamesisk og somali. Fra hver morsmålsgruppe er det lagt inn 100 beståtte besvarelser fra hver av de to språkprøvene<sup>12</sup>. Korpuset inneholder også persondata fra alle informantene. Dette er informasjon om blant annet alder, kjønn, oppholdstid i Norge, antall timer med norskundervisning, utdanning og opplysninger om eventuelt andre språk personen kan. Det er også opplyst for hver tekst hvilken oppgave som er valgt (det er flere mulige), og hvilken tittel oppgaven har (Tenfjord, Hagen & Johansen 2009:52–57).

### 4.2.3 Utvalg av materiale

Materialet i undersøkelsen er altså tekster fra Språkprøven i ASK-korpuset. Fra tekstene har jeg hentet ut alle ord som er annotert med ortografisk feil fra de to morsmålsgruppene vietnamesisk og BKS. Som nevnt i 4.2.2, består hver morsmålsgruppe av hundre personer, som har skrevet én tekst hver. Ved gjennomgang av ord som var annotert med ortografiske feil i ASK, ble det klart at noen ord måtte tas ut av forskjellige grunner (jf. 4.2.5). I tillegg til enkeltord valgte jeg å ta ut alle de ortografiske feilene til én innlærer fra VIET. Denne innlæreren skriver på (tilnærmet) svensk, eventuelt en blanding av norsk og svensk – noe som ikke passer inn i denne sammenhengen. Siden jeg tok ut én person fra VIET, valgte jeg å ta ut én person fra BKS også. Utvelgelsen var tilfeldig – jeg fjernet den siste teksten i materialet.

### 4.2.4 Variasjon i tekstlengde

Materialet omfatter altså 99 personer i hver informantgruppe, der hver person har skrevet en tekst på omtrent 200–400 ord. BKS har litt lengre tekster enn VIET – de har i snitt 18 flere ord

---

<sup>11</sup> I ASK-korpuset brukes *serbokroatisk*.

<sup>12</sup> I Test i norsk – høyere nivå-materialet mangler det noen tekster. Dette er informanter med språkene albansk, vietnamesisk og somali som morsmål (Tenfjord, Hagen & Johansen 2009:56).

per tekst. Fordelingen i antall ord per tekst er likevel ikke bare ulik mellom VIET og BKS, men også innad i hver informantgruppe (jf. tabell 2).

Tekstlengde i antall ord	VIET (antall personer)	BKS (antall personer)
Under 150	1	1
151–200	15	10
201–250	38	30
251–300	25	28
301–350	12	16
351–400	7	8
401–450	1	2
451–500	0	4

Tabell 28: Variasjon i tekstlengde i VIET og BKS.

## 4.2.5 Søk i ASK og sortering av materiale

Ordene med ortografiske feil har jeg funnet fram til ved å søke etter ord med ortografiske feil i ASK. I søket har jeg også spesifisert morsmålsgruppe (vietnamesisk/BKS) og språkstest (Språkprøven i norsk for voksne innvandrere)(se figur 16 i vedlegg 2). Treffet kommer opp som en konkordans med ord med ortografiske feil (se figur 1). Når man stiller inn på KWIC (*key word in context*), kommer treffene opp i kontekst. I figur 1 er det stilt inn på minste kontekst av hensyn til plassen her, men jeg har stilt inn på største kontekst når jeg har gjort søket. At konteksten er der, er viktig for å kunne oppklare eventuelle uklårheter.

Treff 1 - 30 av 1734 | neste 30 treff | Gå til:  | Last ned (  Excel-modus) | Type:  | Vis:  |  Vis linjefilter  filter

count	cpos	Tekst-id	treff	Lemma	Ordklasse
1	617369	s0005	seg til nytt liv. Dessuten må man ta hensyn til informasjon	informasjon	subst
2	618363	s0008	i Norge snart i to år. I Norge er jo helt forskjell kulturen	kultur	subst
3	618546	s0008	rrre først for å få lov til å gjøre noe. Hva er bra tradisjoner	tradisjon	subst
4	618580	s0008	bare gjør noe som man føler det er ikke feil. For eksempel når vi spiser, må vi si værsågod til foreldre eller	for	prep
5	618581	s0008	re gjør noe som man føler det er ikke feil. For eksempel når vi spiser, må vi si værsågod til foreldre eller	eksempel	subst
6	618596	s0008	mpel når vi spiser, må vi si værsågod til foreldre eller de andre som sitter rundt. Vi ler ikke så høyt n	eller	konj
7	618638	s0008	e. Når vi spiser, kan ikke vi spise først noen de andre	annen	det
8	618698	s0008	å. Men uansett er vi menneske, må ikke vi leve uten oppdraelse fra foreldre eller noe gode vaner og	uten	prep
9	618701	s0008	en uansett er vi menneske, må ikke vi leve uten oppdraelse fra foreldre eller noe gode vaner og tradisjoner	oppdragelse	subst
10	618737	s0009	OPPGAVE 3 KONKURRANSE Det er konkurranser overalt. Det er i livet vårt i skoler, arbeid og id	konkurranse	subst
11	618931	s0009	igaver for eksempel å lære en dans. Det kan bli spennende å se på for eksempel " Vil du bli millionær? " Det	spennende	adj
12	619190	s0009	livet. Det er en sosial verdi hvor de som delta i konkurranser	konkurranse	subst
13	621910	s0016	alt de trenger slike som sofaen, vaskemaskin, t-v	tv	subst
14	621931	s0016	askin, t-v, cd spiller, bil... Det kan også være å nærhenten	nærhet	subst
15	621937	s0016	.. Det kan også være å nærhenten av butikker, restauranter	restaurant	subst
16	622163	s0016	sted et å bo og man kan puste et friskeluft et lyft	luft	subst
17	657593	s0091	Bolig og bosted Jeg har vært i Norge nesten 6 år og jeg forstår noe ting som man må ha når r	seks	det
18	657668	s0091	når barnskole og barnehage er det lang vei fra hjemme vårt. Mannen min må gå på jobb og jeg går på	hjem	subst
19	657774	s0091	ia nok tid som passer, omsorg, leker med barn. Dessuten boligen min måtte ha nok rom, boligen har bare	dessuten	adv
20	657798	s0091	måtte ha nok rom, boligen har bare to rom for 5 mennesker. Det er veldig vanskelig for familier	fem	det

Figur 1: Konkordans.

Jeg overførte deretter materialet til Excel, hvor jeg plasserte hver enkelt feil i kategori etter feiltype – dette vil jeg beskrive i ? (se også vedlegg 3). Før jeg plasserte ordene i kategorier, tok jeg ut ord/feil i ord tilhørende følgende kategorier:

- Egennavn.
- Feil der minuskel er brukt for majuskel og majuskel er brukt for minuskel.
- Feil der tallord er skrevet som sifre.
- Feil som omhandler feil bruk av diakritiske tegn og feil bruk av bindestrek og apostrof
- Ord som er stavet såpass feil at de (nesten) ikke er gjenkjennelige. For eksempel er "lobor" rettet til *laboratorium* i ASK. Ut fra konteksten ser denne rettingen riktig ut, men når såpass lite av ordet er på plass, er ikke dette lenger en ortografisk feil, slik jeg ser det – og følgelig er ikke feilen relevant for min undersøkelse.
- Feil som gjelder feil forkorting av ord, som "opp" for *oppg.* og "f.ex" for *f.eks.*

I tillegg var det en del tilfeller av like ord med samme feil – i både samme form av ordet og forskjellige former – gjort av samme person. Jeg valgte å beholde alle disse, det vil si at følgende kriterier gjelder for sorteringen av disse tilfellene:

- Flere tilfeller av et ord med samme feil i lik form av ordet, som for eksempel tre tilfeller av "løppe" for *løpe*: alle tre ordene beholdes (telles tre ganger).
- Ord i forskjellige former, men med samme feil, som for eksempel et verb i forskjellige bøyingsformer med samme feilen i alle formene, som i "sykkle", "sykkler" og "sykklet": alle tre ordene beholdes (telles tre ganger).

## 4.3 Valg av metode

På bakgrunn av forskningsspørsmålene har jeg kommet fram til en kategoriinndeling. Denne inndelingen består av hovedkategorier og underkategorier. Alle kategoriene er beskrivende; den videre behandlingen av funnene er forklarende. Det som ligger til grunn for bare å ha beskrivende kategorier, vil jeg forklare med det Wiggen (1994) sier om forholdet mellom beskrivelse og forklaring i kategoriseringen av språkfeil:

I de fleste eldre rettskrivingsstudiene i nordiske språksamfunn, ja i stor grad i andre språksamfunn òg, er beskrivelse og forklaring blitt blanda. I regelen er det slik at bestemte rettskrivingsavvik er blitt beskrevet og kategorisert ved hjelp av type-/klassenemninger og begrep som gir uttrykk for en bestemt oppfatning av hva de respektive avvikskategoriene/-klassene/-typene kommer av (forklaringer). Det kan f.eks. være tale om "Ranschburgske hemninger" og "dialektbetinga feil" (Wiggen 1994:10).

Wiggen sier videre at slike forklaringsbaserte kategorier vanskelig kan etterprøves og eventuelt omtolkes fordi undersøkelsene sjelden inneholder de primære språkdataene. Dette henger igjen sammen med at forklaringene som ligger til grunn for kategoriene, ifølge Wiggen, er dels språklig og dels psykologisk baserte – og styrt av de faglige forutsetningene til den som analyserer (ibid.). Wiggen sier videre at han har

[...] kommet til at det er best å avgrense avviks-*beskrivelsen* til strengt å gjelde bare *språklig karakteriserbare formelle karakteristika* (f.eks. tillegg, utelatelse eller ombytte av vokaler/konsonanter i gitte språklige kontekster), og så parallelt, men separat og uavhengig av den formelle beskrivelsen, argumentere for at det gitte avviket har den eller den *forklaringa* (ibid.).

Det å ha beskrivende kategorier med formelle kriterier for kategorisering av feilene har også den fordelen at det blir lettere å sammenligne resultater undersøkelser imellom. Har man derimot forklaringsbaserte kategorier, vil dette være vanskeligere fordi man da vil ha forskjellige kriterier som ligger til grunn for kategoriseringen. Jeg holder altså med Wiggen i at de ortografiske feilene først bør plasseres i beskrivende kategorier, og at forklaringen bør skje separat.

Den kategoriseringsinndelingen jeg har kommet fram til består, som nevnt over, av hovedkategorier og underkategorier. Hver feil plasseres først i en hovedkategori, og deretter i en eller flere underkategorier. Underkategoriene skal vise i hvilken forbindelse feilen befinner seg. Denne inndelingen vil jeg presentere i 4.4. Hovedkategoriene vil jeg likevel presentere først her. Disse er:

1. Utelatt konsonantgrafem
2. Utelatt vokalgrafem
3. Innskutt konsonantgrafem
4. Innskutt vokalgrafem
5. Substitusjon av konsonantgrafem
6. Substitusjon av vokalgrafem
7. Ombytting av grafemer

Hovedkategoriene er utformet slik at de skal fange opp alle ortografiske feil. Hensikten med hovedkategoriene er å bestemme om et grafem er utelatt, om et ekstra grafem er satt inn, om et grafem er byttet ut med et annet, eller om to grafemer har byttet plass. Kategoriinndelingen spesifiserer også om feilen gjelder et konsonantgrafem eller et vokalgrafem – unntatt i

kategori sju. Kategori sju er altså den eneste kategorien som både feil ved vokal- og konsonantgrafemer plasseres under. Dette fordi ombyttingen noen ganger skjer mellom vokal- og konsonantgrafemer. Det gir derfor mest mening å plassere disse i samme feilkategori. Det er også svært få feil som havner i kategorien, noe som gjør behovet for videre inndeling mindre. Jeg vil i tillegg presisere at jeg med *innskutt grafem* i kategori 3 og 4 mener grafem som er satt inn i ordet, uansett hvor i ordet dette måtte være.

Dette er en inndeling (med forskjellige varianter) som flere før meg har brukt – blant andre Brooks, Gorman & Kendal (1993, referert etter Cook & Bassetti 2005:34) og Elsness (2002). Begge ser på ortografiske feil hos morsmålsbrukere, henholdsvis hos 11- og 15-åringer med engelsk som morsmål og 7–8-åringer med norsk som morsmål. Av de som bruker en kategoriinndeling som ligner på den jeg bruker, er det den Cook (1997) bruker, som ligner mest. Det er likevel noen forskjeller mellom inndelingene. Dette er Cooks inndeling:

1. Insertion of single letters: *untill* for *until*
  2. Omission of a single letter: *occurring* for *occurring*
  3. Substitution of one letter by another: *definate* for *definite*
  4. Transposition of two consecutive letters: *freind* for *friend*
  5. Grapheme substitution, i.e. multiple related changes: *thort* for *thought*
  6. Other mistakes, such as local accent: *fought* for *thought*
- (Cook 1997:478–479)

Hovedforskjellen mellom min og Cooks inndeling er at Cook ikke skiller mellom vokal- og konsonantgrafemer. Hun har også én kategori som jeg ikke har: *Other mistakes, such as local accent*. Forskjellige kriterier for kategorisering av substitusjonsfeil gjør at hun har to kategorier (3 og 5), der disse i min inndeling ville blitt plassert i samme kategori.

## 4.4 Kategoriinndelingen

### 4.4.1 Formål

Det overordnede målet med kategoriinndelingen er å forsøke å få rede på hva som er vanskelig å tilegne seg i ortografi i norsk som andrespråk. Målet er at funnene skal kunne

belyse emnet godt nok til at jeg vil kunne svare på forskningsspørsmålene. Det er derfor viktig å spisse metoden riktig, slik at jeg får svar på det jeg lurer på. Jeg vil likevel kartlegge alle feilene for å få svar på hvilke feil som faktisk gjøres. Et alternativ kunne vært å bare rette seg inn på feil som besvarer forskningsspørsmålene, og utelatt andre feil, men da kunne man risikere å gå glipp av overraskende og interessante funn (jf. f.eks. 6.4.1 om substitusjon av vokalgrafemer). En kartlegging av alle feilene vil derfor kunne vise tendenser og mønstre på hva slags feil som faktisk gjøres.

## 4.4.2 Forklaring til kategoriseringen

Jeg har altså kommet fram til en kategoriinndeling som består av hovedkategorier og underkategorier (se tabell 3). Kategoriseringen går ut på at alle feilene plasseres i de aktuelle kategoriene for akkurat disse feiltypene. Om et ord består av flere feil, plasseres disse i kategorier uavhengig av hverandre. Hver feil plasseres først i en hovedkategori, og deretter i én eller flere underkategorier dersom det finnes en underkategori som passer. Plasseringen i underkategori(er) er altså ikke obligatorisk. Underkategoriene er heller ikke gjensidig utelukkende, slik at en feil kan bli plassert i flere underkategorier. Hovedkategoriene er derimot gjensidig utelukkende. For en eksemplifisering av hvordan feilene blir plassert i kategorier, se tabell 4.

## 4.4.3 Kriterier for kategorisering av feilene

### 4.4.3.1 Generelle kriterier

Noen av kriteriene for kategoriseringen bør forklares nærmere. Ved hovedkategoriene dreier de fleste kategoriene seg om enkeltgrafemer, altså for eksempel at ett grafem er utelatt i kategori 1 eller 2, eller at ett grafem er innskutt i kategori 3 eller 4. I kategori 7 – ombytting av grafemer – involverer derimot feilen to grafemer. I kategori 5 og 6 dreier det seg i utgangspunktet om at ett grafem er byttet ut med ett annet. I noen tilfeller er det likevel to grafemer som er byttet ut med ett, eller ett grafem som er byttet ut med to. For eksempel er det flere tilfeller av at <ks> er byttet ut med <x>, som i "exempel" for *eksempel*. Og det er feil som "hobbie" for *hobby*, der ett grafem (<y>) er byttet ut med to (<ie>), men der disse uttales

Tabell 29: Kategoriinndeling.

Hoved-kategorier	Underkategorier	Eksempel
1. Utelatt konsonantgrafem	a) Enkelt for dobbelt konsonantgrafem	Utelatt <ɫ> i "fjelet" for <i>ffjellet</i> , utelatt <ɸ> i "hat" for <i>hatt</i> og utelatt <d> i "vide" for <i>vidde</i> .
	b) Utelatt konsonantgrafem i gruppe av konsonantgrafemer	Utelatt <d> i "bor" for <i>bord</i> , utelatt <ɾ> i "skive" for <i>skrive</i> og utelatt <j> i "kærlighet" for <i>kjærlighet</i> .
	c) Utelatt konsonantgrafem som enten alene eller som del av et polygrafem gjengir /t/, /d/, /ʎ/, /ɲ/, /ʃ/, /ç/ eller /ŋ/	Utelatt <n> for /ŋ/ i "kokurransen" for <i>konkurransen</i> , utelatt <j> ved /ç/, som i "kole" for <i>kjole</i> og utelatt <ɾ> ved /t/, som i "et" for <i>ert</i> .
	d) Utelatt konsonantgrafem mellom to vokalgrafemer	Utelatt <g> i "lae" for <i>lage</i> , utelatt <ɾ> i "gjøe" for <i>gjøre</i> og utelatt <d> i "øelegge" for <i>ødelegge</i> .
2. Utelatt vokalgrafem	a) Monoftongering av diftong	Utelatt <y> i "løve" for <i>løyve</i> , utelatt <a> i "knus" for <i>knaus</i> og utelatt <i> i "rese" for <i>reise</i> .
	b) Utelatt vokalgrafem mellom to konsonantgrafemer	Utelatt <i> i "leilghet" for <i>leilighet</i> , utelatt <e> i "vidergående" for <i>videregående</i> og utelatt <e> i "burgr" for <i>burger</i> .
3. Innskutt konsonantgrafem	a) Dobbelt for enkelt konsonantgrafem	Innskutt <d> i "moddig" for <i>modig</i> , innskutt <p> i "proppaganda" for <i>propaganda</i> og innskutt <ɫ> i "seill" for <i>seil</i> .
	b) Innskutt konsonantgrafem i gruppe av konsonantgrafemer	Innskutt <ɸ> i "skrfive" for <i>skrive</i> , innskutt <h> i "balsham" for <i>balsam</i> og innskutt <p> i "gpryte" for <i>gryte</i> .
	c) Innskutt konsonantgrafem ved grafem/polygrafem som gjengir /t/, /d/, /ʎ/, /ɲ/, /ʃ/, /ç/ eller /ŋ/	Innskutt <s> i "skjærlighet" for <i>kjærlighet</i> , innskutt <g> i "tengke" for <i>tenke</i> og innskutt <j> i "kjylling" for <i>kylling</i> .
	d) Innskutt konsonantgrafem ved siden av enkelt konsonantgrafem	Innskutt <n> i "nærheten" for <i>nærheten</i> , innskutt <b> i "lebbe" for <i>leke</i> og innskutt <k> i "stolk" for <i>stol</i> .
4. Innskutt vokalgrafem	a) Diftongering av monoftong	Innskutt <i> i "breitt" for <i>brett</i> , innskutt <u> i "tråud" for <i>tråd</i> og innskutt <a> i "bainders" for <i>binders</i> .
	b) Innskutt vokalgrafem mellom to konsonantgrafemer	Innskutt <e> i "egene" for <i>egne</i> , innskutt <e> i "andere" for <i>andre</i> og innskutt <e> i "sient" for <i>sint</i> .
5. Substitusjon av konsonantgrafem	a) Substitusjon av enkelt konsonantgrafem	Substitusjon av <b> for <ɸ> i "filde" for <i>bilde</i> , substitusjon av <g> for <k> i "skog" for <i>skog</i> og substitusjon av <ɫ> for <ɸ> i "atarm" for <i>alarm</i> .
	b) Substitusjon av konsonantgrafem i gruppe av konsonantgrafemer	Substitusjon av <v> for <ɸ> i "ofn" for <i>ovn</i> , substitusjon av <p> for <ɾ> i "koffept" for <i>koffert</i> og substitusjon av <t> for <k> i "kjære" for <i>tjære</i> .
	c) Substitusjon av konsonantgrafem ved grafem/polygrafem som gjengir /t/, /d/, /ʎ/, /ɲ/, /ʃ/, /ç/ eller /ŋ/	Substitusjon av <n> for <m> i "lamg" for <i>lang</i> , substitusjon av <ɾ> for <p> i "bapn" for <i>barn</i> og substitusjon av <j> for <h> i "shonglere" for <i>sjonglere</i> .
6. Substitusjon av vokalgrafem	a) Substitusjon av enkelt vokalgrafem	Substitusjon av <i> for <e> i "togskenne" for <i>togskinne</i> , substitusjon av <o> for <a> i "kapp" for <i>kopp</i> og substitusjon av <o> for <u> i "blumst" for <i>blomst</i> .
	b) Substitusjon av vokalgrafem i diftong	Substitusjon av <ø> for <æ> i "æy" for <i>øy</i> , substitusjon av <u> for <o> i "sao" for <i>sau</i> og substitusjon av <e> for <a> i "faie" for <i>feie</i> .
7. Ombytting av grafemer		I ordet "ofentilge" for <i>offentlige</i> har <i> og <ɫ> byttet plass, i "guvl" for <i>gulv</i> har <ɫ> og <v> byttet plass og i "satf" for <i>saft</i> har <ɸ> og <ɸ> byttet plass.



Ord med ortografisk feil i kontekst	1. steg: plassering av feil i hovedkategori	2. steg: plassering av feil i underkategori(er)
I Norge er jo helt forskjell <b>kuntur</b> med Viet Nam, tror jeg.	Hovedkategori 5: Substitusjon av konsonantgrafem (<n> for <l>)	Underkategori: 5b: Substitusjon av grafem i gruppe av konsonantgrafemer. Her: <lt> er byttet ut med <nt>.
Hva er bra <b>tradisoner</b> ?	Hovedkategori 1: Utelatt konsonantgrafem (<j>)	Underkategorier: 1b: Utelatt grafem i gruppe av konsonantgrafemer. Her: <j> er utelatt i <sj>. 1c: Utelatt grafem som enten alene eller som del av et polygrafem gjengir /t/, /d/, /l/, /n/, /s/, /ç/ eller /η/. Her: utelatt grafem (<j>) i polygrafem (<sj>) som uttales /f/.
Neste dag ringte faren flyselskapet og <b>bestillte</b> 5 flybilleter.	Hovedkategori 3: Innskutt konsonantgrafem (<l>)	Underkategori: 3a: Dobbel for enkelt grafem. Her: <ll> for <l>.
I hoved sak må det <b>vare</b> service på veldig høy nivå.	Hovedkategori 6: Substitusjon av vokalgrafem (<a> for <æ>)	Underkategori: 6c: Substitusjon av enkelt vokalgrafem. Her: <æ>.
Jeg er ofte trøtt <b>fødi</b> jeg må jobbe hele tiden.	Hovedkategori 6: Substitusjon av vokalgrafem (<ø> for <o>) +	Underkategori: 6c: Substitusjon av enkelt vokalgrafem. Her: <ø>.
	Hovedkategori 1: Utelatt konsonantgrafem (<r>)	Underkategorier 1b: Utelatt konsonantgrafem i gruppe av konsonantgrafemer. Her: utelatt <r> i <rd>. 1c: Utelatt konsonantgrafem som enten alene eller som del av et polygrafem gjengir /t/, /d/, /l/, /n/, /s/, /ç/ eller /η/. Her: utelatt grafem (<r>) i polygrafem (<rd>) som uttales /d/.

Tabell 30: Eksemplifisering av plassering av feil i hovedkategorier og underkategorier.

med ett fonem. Uavhengig av om feilene involverer ett eller to grafemer, vil slike feil telle som én feil.

Kategori 5 og 6 (substitusjon av henholdsvis konsonant- og vokalgrafem) trenger også en utdyping om hvilke grafemer det korrekte grafemet har blitt erstattet med. I de fleste tilfeller gjelder substitusjon av konsonantgrafem at et konsonantgrafem har blitt byttet ut med et annet konsonantgrafem, og ved substitusjon av vokalgrafem har et vokalgrafem blitt byttet ut med et annet vokalgrafem – slik at man får feil som henholdsvis "sund" for *sunt* og "vare" for *være*. Det forekommer likevel tilfeller av vokalgrafem byttet ut med konsonantgrafem eller konsonantgrafem byttet ut med vokalgrafem – her mener jeg byttet ut i den betydningen at et vokalgrafem står på plassen til et konsonantgrafem eller at et konsonantgrafem står på plassen til et vokalgrafem. Det er likevel ikke sikkert at disse feilene bør karakteriseres som substitusjon, slik jeg ser det. For eksempel finnes ordet "kafee" for *kaffe* i materialet. Her står det første <e>-grafemet på plassen til det som skulle vært et <f>-grafem. Dette trenger likevel ikke bety at det har skjedd et bytte mellom de to grafemene. Jeg vurderer ikke dette som sannsynlig, og har derfor valgt å ikke kategorisere dette som substitusjonsfeil, men som

utelatt/innskutt grafem. Det nevnte eksemplet vil derfor kategoriseres i utelatt konsonantgrafem (<f>-grafemet) og innskutt vokalgrafem (<e>-grafemet) – henholdsvis i kategori 1 og 4. Hovedregelen er altså at substitusjon bare gjelder bytte av vokalgrafem for vokalgrafem og konsonantgrafem for konsonantgrafem. Denne regelen har likevel ett unntak; substitusjon av <j> for <i> eller <y> og substitusjon av <y> eller <i> for <j>. Her tenker jeg på feil som substitusjon av <i> for <j> i "miljtær" (*militær*) og <y> for <j> i "kjøretøj" (*kjøretøy*). I motsetning til tilfellet over med <e>-grafemet på plassen til <f>-grafemet i "kafee", som jeg ikke mener kan karakteriseres som substitusjon, mener jeg at byttet mellom grafemene i "miljtær" og "kjøretøj" kan karakteriseres som substitusjon. Dette fordi jeg tolker det som at det her har skjedd et bytte mellom to grafemer, og ikke en utelatelse av et grafem og innskudd av et annet.

Underkategoriene 1b, 2b, 3b og d, 4b og 5b – som dreier seg om feil som har oppstått i forbindelse med gruppe av konsonantgrafemer – trenger en utdyping med tanke på hva begrepet *gruppe av konsonantgrafemer* betyr. I denne undersøkelsen betegner jeg ikke begrepet på samme måte som konsonantgruppe vanligvis betegnes – som en gruppe konsonanter uttalt innenfor én stavelse. Jeg har valgt å definere gruppe av konsonantgrafemer som to eller flere konsonantgrafemer i gruppe, uansett om disse går over en stavelsesgrense eller ikke. For eksempel regner jeg <lb> i *stolbein* som en gruppe av konsonantgrafemer. Med den vanlige definisjonen ville ikke dette blitt karakterisert som en konsonantgruppe fordi den krysser en stavelsesgrense. Jeg har valgt denne definisjonen blant annet fordi jeg mener at stavelsesgrenser er noe mindre relevant i skrift enn i uttale, på den måten at det som tydelig oppleves som en stavelsesgrense i uttale, vil bli mindre tydelig i skrift. I tillegg kunne kategoriseringskriterier som skulle ta hensyn til stavelsesgrenser, vist seg vanskelige i praksis fordi det ikke alltid er enighet om hvor stavelsesgrensene går (jf. Endresen 2000:263–264).

Hvilken uttalenorm jeg bruker, er relevant ved kategorisering av kategoriene som har å gjøre med feil ved /t/, /d/, /l/, /ɲ/, /ʃ/, /ç/ og /ŋ/ (underkategoriene 1c, 3c og 5c). Som nevnt i kapittel 1 er det standard østnorsk jeg har som norm for uttale. Grunnen til at dette er relevant ved kategorisering av akkurat disse feiltypene, er fordi det er til dels store dialektforskjeller i Norge ved disse fonemene. Det er likevel ikke helt klart hva som menes med standard østnorsk, med mindre dette beskrives nærmere. Det eneste som likevel er relevant i denne sammenhengen er at feil som utelatt <l> i *gult* – "gut" for *gult* – vil havne i kategorien. Dette

fordi uttalen av <l>-grafemet blir tolket som [ɫ], og [ɫ] før [t] fører til at disse assimileres til [t] – slik at *gult* blir uttalt /gʉ:ɫ/. Jeg synes det er best å kategorisere feilen slik fordi hensikten jo er å finne mulige forklaringer på ortografiske feil. Det at jeg plasserer feilen i denne kategorien, betyr derfor at dette er én mulig forklaring på hvordan feilen kunne oppstå; andre forklaringer kan også være mulige. Og siden dette er én mulig forklaring, er det viktig å fange opp denne.

#### 4.4.3.2 Kriterier for å skille ortografiske feil fra andre typer feil

I ASK-materialet er ord med ortografiske feil annotert, og retting er foreslått. I all hovedsak har jeg fulgt rettingene til ASK, men i noen tilfeller har jeg vært uenig. I disse tilfellene har jeg erstattet med min egen retting. Hovedregelen er, i tråd med ASK, at morfologiske feil ikke rettes, bare ortografiske. I visse tilfeller kan det være vanskelig å skille ortografiske feil fra andre typer feil, da spesielt grammatiske. Jeg ser et mønster i at en del ord har feil finalt i ordet. I norsk foregår det meste av bøyningen her, og jeg tolker derfor ofte slike feil som tegn på usikkerhet ved bøyning. At det kan være vanskelig å skille ortografiske feil fra andre typer feil, har også vært en utfordring i annoteringsarbeidet. Dette er nevnt i *Kodebok for ASK*:

Ortografiske feil er i noen tilfeller vanskelige å skille fra andre typer koder. Derfor kan denne koden inneholde en del avvik som kunne ha vært kodet annerledes. Det er viktig å skille mellom feil av typen INFL og ORT [henholdsvis *feil form, men rett morfosyntaktisk kategori og ortografisk feil* – mitt tillegg]. Ortografiske feil gjelder kun for et ords lemma og ikke bøyningssendelser. Likevel vil for eksempel manglende forenkling av dobbeltkonsonant når skriftnormen krever det regnes med til denne feiltypen, og manglende stumme ende-konsonanter (<http://clu.uni.no/humfak/ask/brukerveiledning/feilkoder.htm#leksikalske>).

For å forsikre meg om at jeg er konsekvent i rettingen av ortografiske feil, har jeg utarbeidet kriterier som gjelder for denne rettingen. Hovedregelen er at morfologi ikke skal rettes – bare ortografi (jf. også beskrivelse av ortografisk feil i 2.4) – se eksempel 1, 2, 6, 7, 9, 10 og 12 i tabell 5. I eksempel 2 rettes "arbeidledigheter" til *arbeidsledighet*, uten at endelsen *-er* regnes som feil. Endelsen *-er* er den vanligste endelsen på substantiv i flertall. Jeg regner derfor feilen som morfologisk og ikke ortografisk. Det samme blir tilfelle i eksempel 6; "role" rettes til *rolle* (og ikke *roller*) fordi det er en utelatt <l> som er den ortografiske feilen i ordet.

I noen tilfeller har det likevel ikke vært mulig å rette et ord ortografisk uten samtidig å ta hensyn til hva som er morfologisk riktig, med andre ord når det er en morfologisk og ortografisk feil på samme sted i ordet. I disse tilfellene er ordet rettet til det som er morfologisk riktig – se eksempel 3–5, 8 og 11 i tabell 5. Eksempel 3 og 4 er ord som har feil i

bøyningsendelsen; endelsene er ikke gyldige innenfor ordklassen (i motsetning til tilfellet i eksempel 2). "Omkome" for *omkomne* i eksempel 5 er annerledes enn eksempel 3 og 4 fordi ordet har indre bøyning. I eksempel 7, 9, 10 og 12 er ordet rettet til en form som ikke er grammatisk korrekt i konteksten. Ordene står likevel i en form som er en eksisterende form innenfor den aktuelle ordklassen; de har derfor fått beholde den formen da den ortografiske feilen uansett er et annet sted i ordet enn i bøyningsendelsen. I eksempel 8 og 11 er ordet derimot rettet til den grammatisk korrekte formen i og med at den ortografiske feilen er på samme sted i ordet som den morfologiske feilen.

Sammensatte ord er ikke spesielt i fokus i undersøkelsen. Et sammensatt ord med ortografisk feil vil bli rettet på samme måte som andre ord – og om et sammensatt ord er særskrevet, vil ikke dette telle som en ekstra feil. "Fotbal kamp" vil dermed rettes til *footballkamp*, og registreres med én feil – at *football* er skrevet med én <l> i stedet for to.

Eks. nr.	Ord med ortografisk feil i kontekst	Rettes til
1	Selfølgelig er penger viktig delen av vår <b>sammfunnet</b> (...)	samfunnet
2	Her i Norge er veldig lavt <b>arbeidledigheter</b> .	arbeidsledighet
3	I dag kjører en hoved del av folks <b>utstyre</b> på elektrisitet.	utstyr
4	Uansett at vi har mange ofentile <b>konture</b> som vi kan får hjelp, det er ikke nok.	kontorer
5	Hver dag kan vi høre om bombe eksplosjon, med mange <b>omkome</b> mennesker.	omkomne
6	Mange politikere r misbrukte sine <b>role</b> i dag.	rolle
7	(...) stor forskjell mellom landene i verden om demokrati og <b>polytisk</b> systemer	politisk
8	Men <b>dene</b> ordet mister sine verdier i dag	dette
9	I dag har alle TV, data maskin er mer og mer <b>popular</b> og barn er altfor mye hjemme.	populær
10	Det er problemet med de unge som kjører biler og har sterke <b>høytoler</b> (...)	høytaler
11	<b>Spesiel</b> det kan være vanskelig for foreldrene som har små barna.	spesielt
12	Myndighetene prøver å redusere <b>trafik</b> på mange forskjellige måter men det er aldri nok.	trafikk

Tabell 31: Hvordan skille ut ortografiske feil fra andre typer feil.

## 4.5 Metode for utregning

Resultatene jeg har kommet fram til, viser i hvor stor grad en viss ortografisk feil gjøres av VIET og BKS. Tallene oppgir jeg i prosent av antall løpeord. Jeg har kommet fram til tallene på denne måten: Først har jeg regnet ut hvor stor prosentandel en viss type feil utgjør hos alle enkeltpersonene innenfor hver informantgruppe. Dette regnet jeg ut ved å dividere antallet av

en viss feil hos hver enkelt person på antall løpeord i denne personens tekst. Deretter regnet jeg ut gjennomsnittet av denne feiltypen innenfor hver informantgruppe. Disse resultatene vil jeg presentere i neste kapittel.

Noen av resultatene viser betydelige forskjeller mellom gruppene, noe som gjorde at jeg ville finne ut om disse forskjellene er statistisk signifikante. Disse utregningene fikk jeg en statistiker til å hjelpe meg med<sup>13</sup> (jf. vedlegg 4).

## 4.6 Oppsummering av kapitlet

I dette kapitlet har jeg beskrevet metoden i undersøkelsen. Jeg har også beskrevet utregningsmetoden som er brukt for å komme fram til resultatene. Materialet, som jeg har hentet fra innlærerkorpuset ASK, består av språkprøver fra to informantgrupper med hvert sitt morsmål – vietnamesisk og BKS. For å finne de ortografiske feilene søkte jeg på ord merket med ortografisk feil i ASK. Deretter kategoriserte jeg feilene etter en kategoriinndeling jeg mente var hensiktsmessig for å få svar på forskningsspørsmålene. Kategoriinndelingen består av hovedkategorier og underkategorier. Kategoriene er beskrivende – feil plasseres i kategorier ved hjelp av formelle kriterier. I kapitlet har jeg også beskrevet kriteriene som gjelder for kategoriseringen av feilene.

---

<sup>13</sup> Takk til Bård Uri Jensen.

# 5 Resultater

## 5.1 Innledning

I dette kapitlet vil jeg presentere resultatene av undersøkelsen. Resultatene har jeg kommet fram til ved å plassere de ortografiske feilene i en kategoriinndeling som har som formål å belyse forskningsspørsmålene (jf. 2.6). Resultatene viser at det er variasjon mellom VIET og BKS. Variasjonen handler om både antall og type feil som gjøres. Det er også variasjon innad i hver informantgruppe, men det er likevel forskjeller mellom VIET og BKS som peker seg ut. Disse forskjellene vil jeg beskrive i kapitlet, samt at jeg også vil beskrive funn der det er ubetydelige forskjeller mellom VIET og BKS.

Jeg starter kapitlet med å beskrive hvordan jeg har valgt å framstille resultatene (5.2), før jeg gjør rede for hovedfunn (5.3). Deretter vil jeg presentere henholdsvis vokalgrafemfeil (5.4) og konsonantgrafemfeil (5.5). Videre følger en beskrivelse av variasjonen innad i hver informantgruppe (5.6). Til slutt oppsummerer jeg kapitlet (5.7).

## 5.2 Framstilling av resultatene

Funnene framstilles i to typer tall: prosenttall og faktiske forekomster. Prosenttallet viser feilfrekvens per løpeord (jf. 4.5 for forklaring til utregningen). Faktiske forekomster er tallene slik de framkommer i materialet – altså tall som ikke er omregnet. Jeg bruker prosenttall når jeg framstiller resultatene til kategoriene i kategoriinndelingen. Tall i faktiske forekomster bruker jeg når jeg spesifiserer hvilke grafemer disse feiltyperne gjelder. Dette betyr for eksempel at jeg bruker prosenttall for å vise feilfrekvens av dobbelt for enkelt konsonantgrafem, og at jeg bruker tall i faktiske forekomster for å vise hvilke grafemer denne feiltypen gjelder for. Grunnen til at jeg bruker tall i faktiske forekomster i tillegg til prosenttall, er fordi at om jeg hadde omregnet disse tallene til prosenttall, så hadde de blitt meningsløst små. Jeg bruker likevel ikke tall i faktiske forekomster mer enn nødvendig fordi de ikke er like sammenlignbare informantgruppene imellom som prosenttallene. For i utregningen av prosenttallene – som er regnet ut ved å dividere tallet i faktisk forekomst på antall løpeord – blir ulik tekstlengde i de to informantgruppene utjevnet. Tallene i faktiske

forekomster tar derimot ikke hensyn til ulik tekstlengde, og disse tallene er dermed mindre sammenlignbare.

I oversikter der prosenttall brukes, har jeg tatt ut feil som er gjort av én eller to personer (uavhengig av hvor mange feil disse personene har gjort). Tallet hadde uansett vært så lite at det ville blitt rundet ned til 0 %. I oversikter der jeg presenterer faktiske forekomster, har jeg ikke tatt ut noen tall, uansett hvor få forekomster feilene har. Lave forekomster er dokumentert i like høy grad som høye fordi det også er viktig å belyse det som ikke er vanskelig å tilegne seg i norsk ortografi.

En følge av å vise prosenttall som her er feilfrekvens per antall løpeord, er at prosenttallene blir veldig små. Dette fordi mer enn 96 % av tekstmengden hos både VIET og BKS er riktig stavet. Det som i noen tilfeller kan virke som små forskjeller i prosenttall mellom informantgruppene, kan likevel tilsvare ganske mange forekomster i faktiske forekomster.

I framstillingene har jeg tatt vekk konsonantgrafemer som ikke er aktuelle, som <x>, <z>, <q> og <c>. Og i oversikter over forekomster av enkelt for dobbelt konsonantgrafem og dobbelt for enkelt konsonantgrafem har jeg ikke tatt med konsonantgrafemer som ikke eksisterer som doble. Videre vil hovedkategorier alltid skrives med store konsonantgrafemer i tabeller og figurer, mens underkategorier vil skrives med små.

Som nevnt over viser prosenttallet jeg viser til, feilfrekvens per antall løpeord. Det er disse prosenttallene jeg vil bruke i beskrivelsen av funnene i kapitlet. Jeg har også regnet ut prosenttall for feilfrekvensen av det totale antallet feil – da jeg mener disse tallene kan bidra til å nyansere bildet. Disse prosenttallene, tallene som viser feilfrekvensen av det samlede antall feil, finnes i figur 17–21 i vedlegg 8.

## 5.3 Hovedfunn

Det er forskjell i feilfrekvens hos VIET og BKS. Høyest feilfrekvens har BKS, med 3,8 % feil per løpeord. VIET har en feilfrekvens på 2,4 %. Forskjellen mellom informantgruppene er statistisk signifikant (jf. vedlegg 4). Før jeg går inn på underkategoriene og beskriver funnene i detalj, vil jeg gi en oversikt over hovedfunnene ved å beskrive feil ved hovedkategoriene

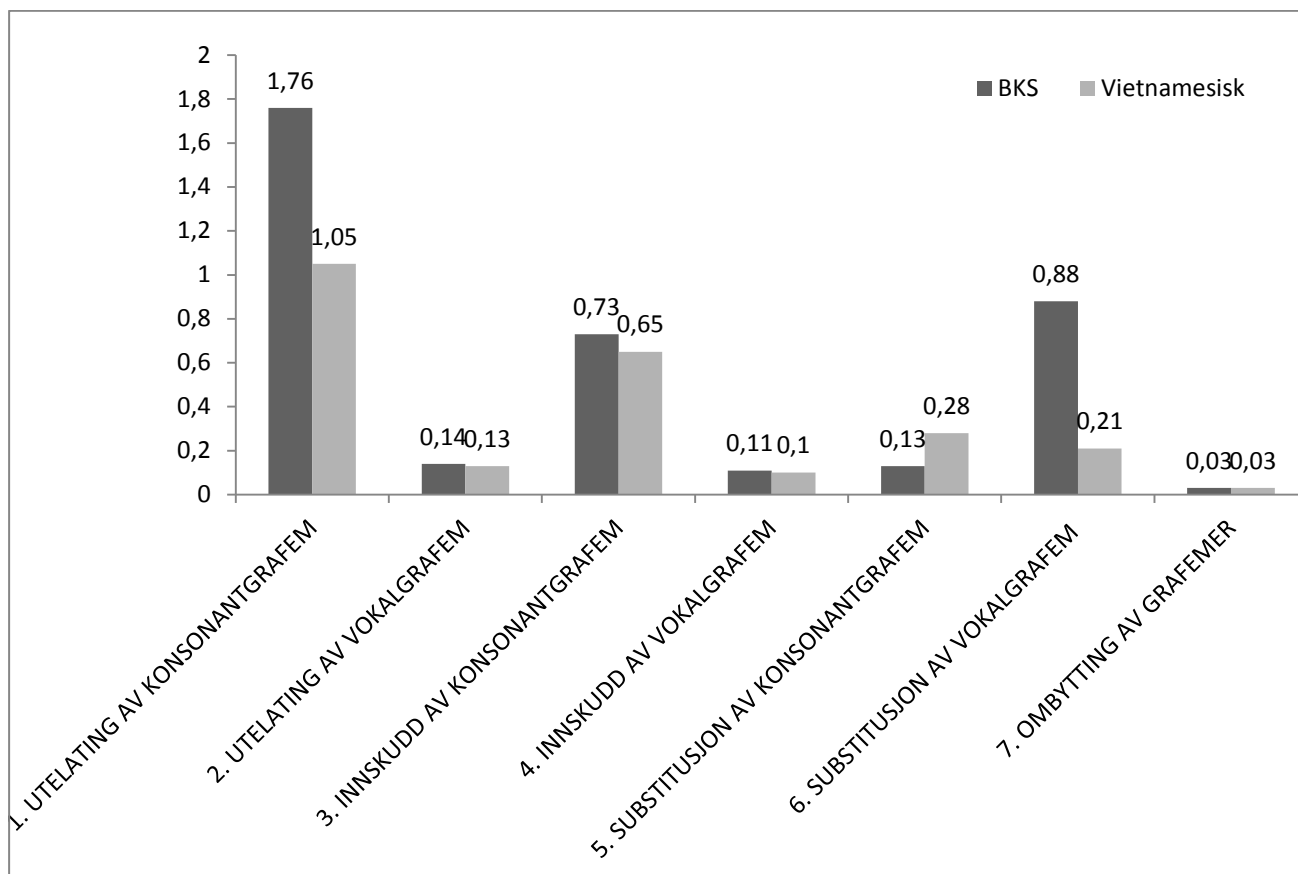
(for beskrivelse av hovedkategoriene, se 4.3). Dette for å gi et overordnet bilde av hva slags feil som gjøres.

Hovedkategorien med flest feil er den samme hos begge informantgruppene, utelating av konsonantgrafem. Kategorien har 1,05 % forekomster hos VIET og 1,76 % hos BKS (jf. figur 2). I faktiske forekomster tilsvarer det en differanse mellom gruppene på 208 forekomster. Kategorien med nest flest feil er derimot ikke den samme hos de to gruppene – hos BKS er det substitusjon av vokalgrafem som har nest flest feil (med 0,88 %), mens hos VIET er det innskudd av konsonantgrafem (med 0,65 %). Hos BKS er substitusjon av vokalgrafem tett etterfulgt av innskutt konsonantgrafem (0,73 %). Ombytting av grafemer er den av hovedkategoriene med færrest feil hos begge gruppene, med samme feilfrekvens (0,03 %) hos begge.

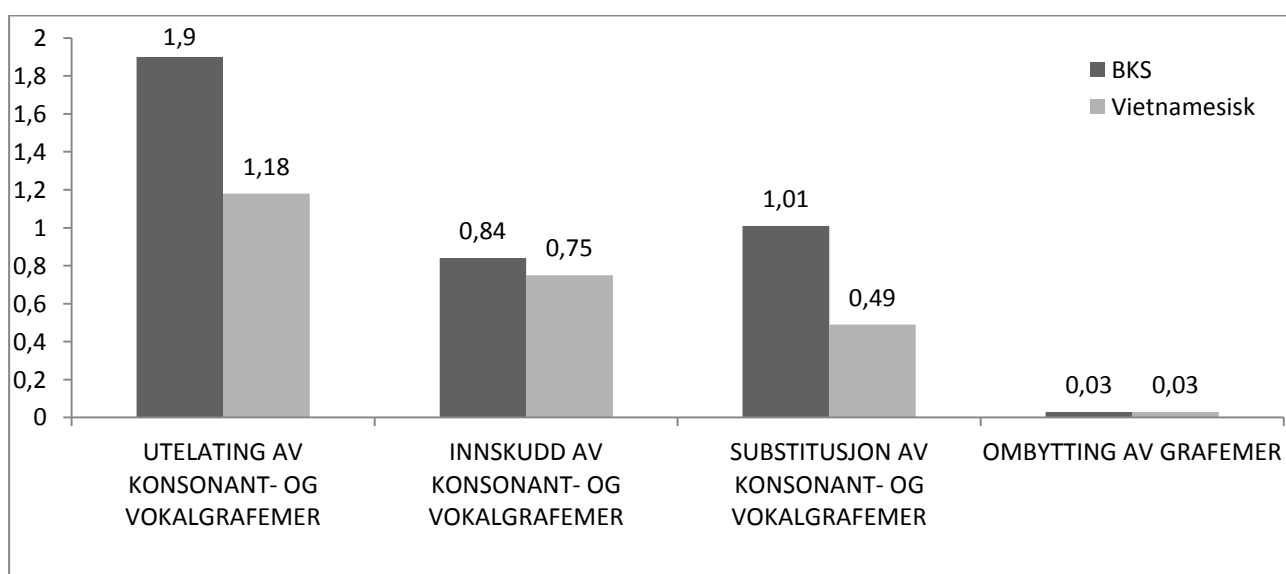
Ved en sammenlegging av konsonant- og vokalgrafemforekomstene i kategoriene utelating (kategori 1 pluss 2), innskudd (kategori 3 pluss 4) og substitusjon (kategori 5 og 6)(jf. figur 3) – kommer det fram at utelating av grafem er den vanligste feiltypen av disse kategoriene hos begge informantgruppene. BKS har 1,90 % feil per løpeord ved utelating av grafem, og VIET har 1,18 % feil. Deretter kommer substitusjon av grafem og innskudd av grafem for BKS, mens for VIET er det innskudd av grafemer som er den kategorien med nest flest feil etter denne inndelingen, og substitusjon av grafemer på tredje plass.

Hovedkategoriene kan også legges sammen på en annen måte, slik at forskjellen i feilfrekvens ved vokal- og konsonantgrafemfeil mellom informantgruppene blir tydeliggjort (jf. figur 4 og 5). Kategoriene utelating av vokalgrafem, innskudd av vokalgrafem og substitusjon av vokalgrafem er lagt sammen for å vise hvor mange vokalgrafemfeil som gjøres sammenlagt. Det er BKS som gjør flest vokalgrafemfeil, med 1,13 % per løpeord – VIET har 0,44 % vokalgrafemfeil per løpeord. Videre er kategorien som viser konsonantgrafemfeil en sammenlegging av kategoriene utelating av konsonantgrafem, innskudd av konsonantgrafem og substitusjon av konsonantgrafemfeil. Også her er det BKS som gjør flest feil – 210 flere feil enn VIET i faktiske forekomster. Alle resultatene er vist i figur 6. Inndelingen er den samme som ble beskrevet i metodekapitlet (jf. tabell 3). Med unntak av å spesifisere hvilke grafemer feilene gjelder, er figuren en fullstendig oversikt over alle feilene.

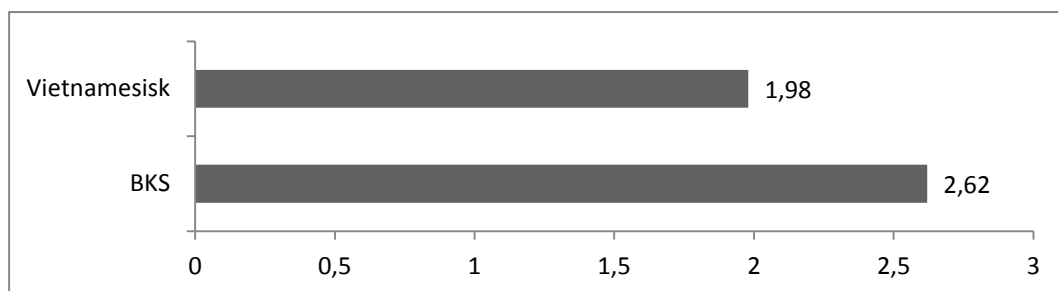




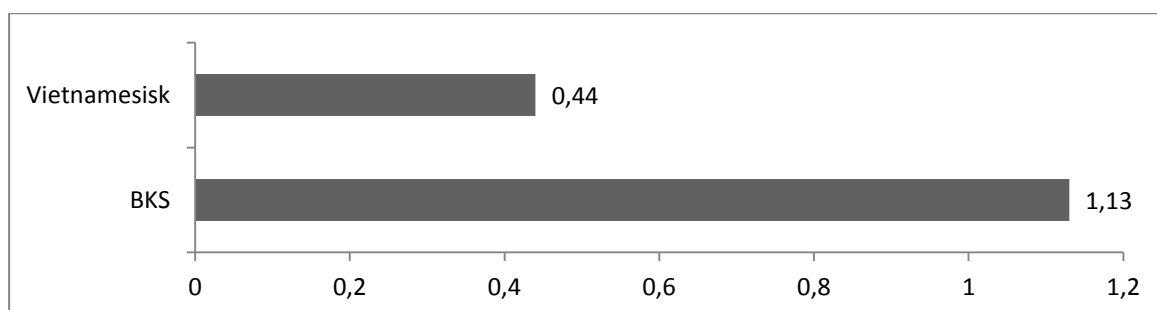
Figur 2: Hovedkategorier i prosent.



Figur 3: Utelatte, innskutte og substituerte konsonant- og vokalgrafemer i prosent.



Figur 4: Konsonantgrafemfeil i prosent.



Figur 5: Vokalgrafemfeil i prosent.

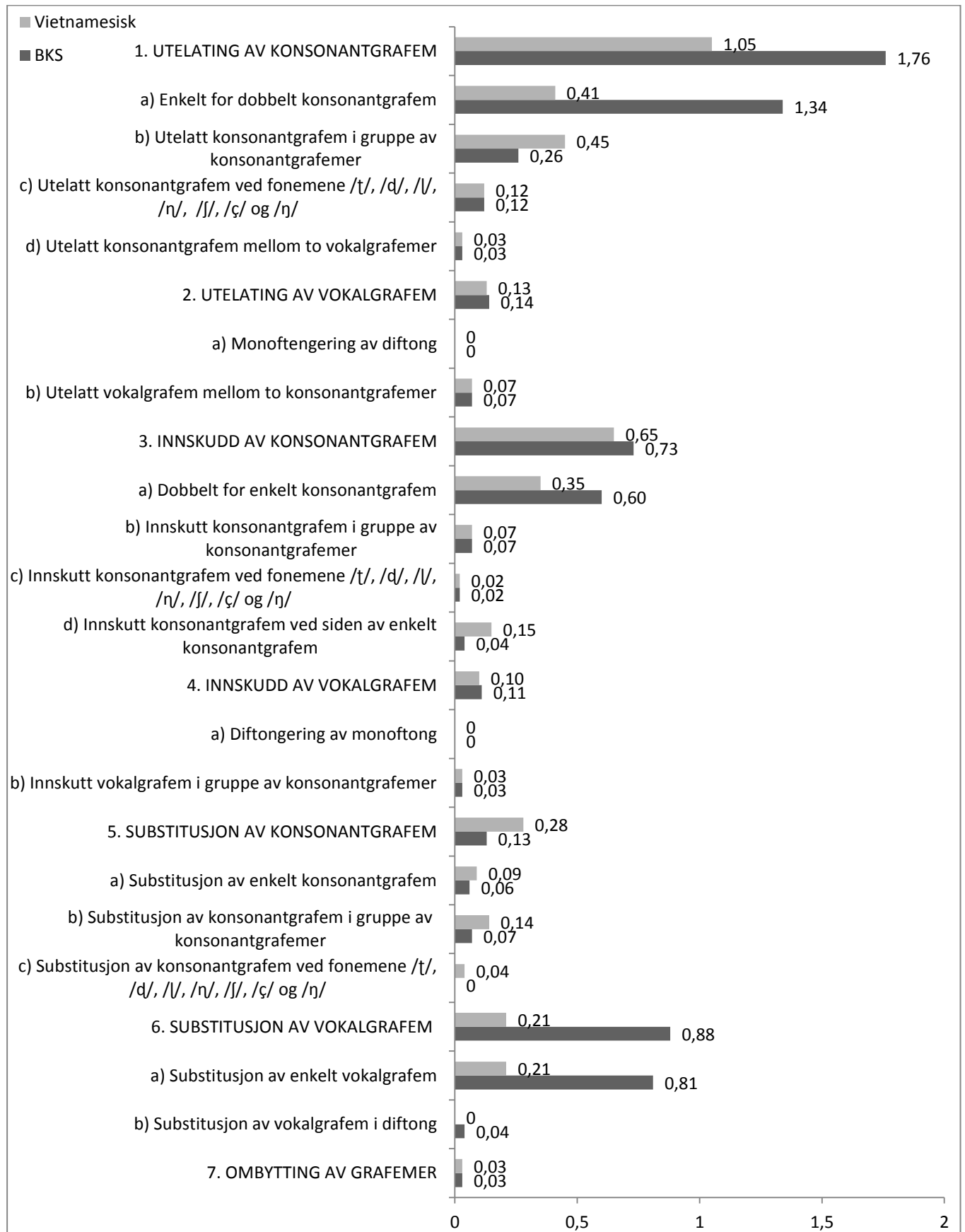
## 5.4 Vokalgrafemfeil

### 5.4.1 Utelatte og innskutte vokalgrafemer

Som vist (jf. figur 5) og beskrevet over (jf. 5.3) så gjør BKS flere vokalgrafemfeil enn VIET. Men selv om antallet er forskjellig hos de to informantgruppene, så er både antall og type feil ganske likt ved innskutte og utelatte vokalgrafemer (jf. figur 6). Av utelatte og innskutte vokalgrafemer er størstedelen av grafemene en «e» hos både VIET og BKS (jf. tabell 6).

Selv om den største delen av utelatte og innskutte vokalgrafemer er en «e», så er det også innskudd og utelating av noen få andre vokalgrafemer hos begge gruppene (se tabell 13 og 14

**Figur 6: Feilforekomster i prosenttall – hovedkategorier og underkategorier.**



i vedlegg 6). Hos begge grupper er <i> det av vokalgrafemene som blir innskutt og utelatt flest ganger etter <e>.

Det innskutte vokalgrafemet er en <e>		Det utelatte vokalgrafemet er en <e>	
Forekomst/det totale antallet innskutte vokalgrafemer. Tallet i parentes er forekomst av innskutt <e> i prosent.		Forekomst/det totale antallet utelatte vokalgrafemer. Tallet i parentes er forekomst av utelatt <e> i prosent.	
BKS	Vietnamesisk	BKS	Vietnamesisk
22/28 (79 %)	20/25 (76 %)	29/38 (81 %)	28/33 (85 %)

Tabell 32: Innskutt og utelatt <e> i faktisk forekomst og prosent.

Monoftongering av diftong og diftongering av monoftong er feil som verken gjøres av VIET eller BKS (jf. kategori 2a og 4a i figur 6). Jeg har også sett på om vokalgrafemer utelates når de står ved siden av et annet vokalgrafem, men der disse ikke er en diftong, men heller ikke i disse tilfellene utelates det vokalgrafemer.

I kategori 4b, "innskudd av vokalgrafemer i gruppe av konsonantgrafemer", gjøres det like mange feil hos VIET og BKS – 0,03 % per løpeord. Som ved innskutte vokalgrafemer generelt, som beskrevet over, er det <e> det gjøres feil ved også her (jf. tabell 7).

	VIET	BKS
Innskudd av <e>	7	8
Innskudd av <i>	2	1

Tabell 33: Innskudd av vokalgrafemer i gruppe av konsonantgrafemer – tallene er i faktiske forekomster.

## 5.4.2 Substitusjon av vokalgrafemer

Selv om informantgruppene ligner hverandre ved innskutte og utelatte vokalgrafemer, så er de svært ulike hverandre ved den siste kategorien av vokalgrafemfeil, "substitusjon av vokalgrafem". Dette er en av de underkategoriene jeg finner størst forskjell mellom de to gruppene. Forskjellen dreier seg både om antall feil, men også om hva slags feil som gjøres (se tabell 8). Mens VIET har en feilprosent på 0,21 % i kategorien, utgjør feilene i BKS 0,88 % (se figur 6).

Det er forskjeller i utbyttingene som gjøres ved substitusjon av vokalgrafemer hos VIET og BKS (se tabell 8). I tabellen viser kolonnen til venstre hvilken plass de forskjellige

substitusjonene kommer på, der substitusjonene som har høyest forekomst kommer øverst og substitusjonene med lavest forekomst nederst. Feil som bare har én forekomst, har jeg ikke tatt med i oversikten. Til tross for store forskjeller mellom morsmålsgruppene i hva slags substitusjoner som gjøres, er flere av substitusjonene like for de to gruppene. For eksempel ⟨a⟩ for ⟨e⟩ (men den gjøres over tre ganger så ofte av BKS), ⟨i⟩ for ⟨y⟩ (gjøres dobbelt så ofte av BKS som VIET) og ⟨ø⟩ for ⟨o⟩ (gjøres dobbelt så ofte av BKS som VIET). Videre er ⟨e⟩ det vokalgrafemet som oftest blir byttet ut med et annet vokalgrafem, og ⟨i⟩ og ⟨a⟩ de vokalgrafemene som oftest blir valgt til å bytte ut andre vokalgrafemer hos VIET. Hos BKS er ⟨o⟩ det vokalgrafemet som både blir byttet ut flest ganger med et annet vokalgrafem, og det vokalgrafemet som oftest velges til å bytte ut andre vokalgrafemer med (se tabell 17 og 18 i vedlegg 6).

Plass	BKS		VIET	
	Valgt grafem for hvilket riktig grafem	Antall forekomster	Valgt grafem for hvilket riktig grafem	Antall forekomster
1	o for ø	28	a for e	8
2	o for å	27	i for y/ø for o	7/7
3	a for e	26	e for a	5
4	a for æ	24	i for e/y for e/y for i/e for i/a for æ	2
5	ø for o	21		
6	å for o	17		
7	o for a	15		
8	i for y	14		
9	y for i	11		
10	a for o/a for å/e for æ	6/6/6		
11	i for e/u for y/ø for å	4/4/4		
12	æ for a/æ for e/u for o/y for u/å for ø	3/3/3/3		
13	ø for e/e for i/e for ø	2/2/2		

Tabell 34: Substitusjon av vokalgrafemer i faktiske forekomster.

## 5.5 Konsonantgrafemfeil

Andelen konsonantgrafemfeil er høyest hos BKS (jf. figur 4), som beskrevet i 5.3 – men både BKS og VIET har forholdsvis mange konsonantgrafemfeil. Forskjellen mellom de to informantgruppene her handler likevel ikke bare om antall feil, men også om type feil. VIET gjør betydelig flere feil ved grupper av konsonantgrafemer enn BKS, mens BKS gjør betydelig flere feil ved enkelt for dobbelt konsonantgrafem enn VIET.

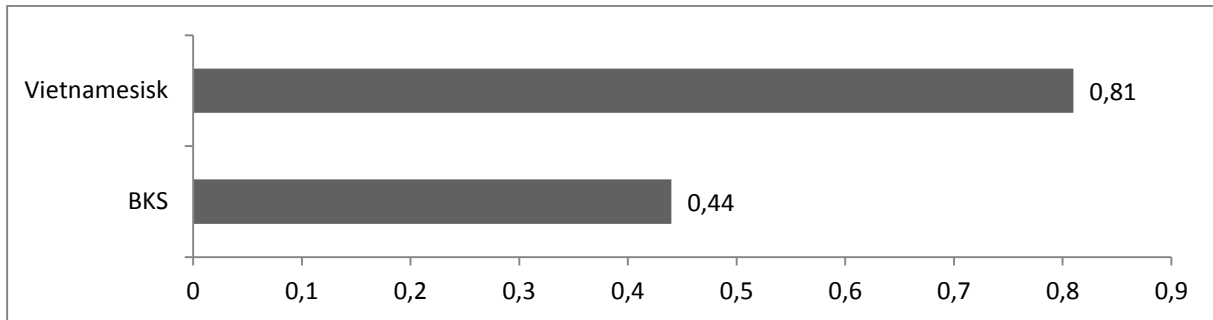
### 5.5.1 Feil ved gruppe av konsonantgrafemer

VIET gjør 0,81 % feil ved gruppe av konsonantgrafemer per løpeord, mens BKS som gjør 0,44 % feil (jf. figur 7). Forskjellen mellom gruppene er statistisk signifikant (jf. vedlegg 4). Samlekategorien feil ved gruppe av konsonantgrafemer har jeg kommet fram til ved å legge sammen tallene i kategoriene "Utelating av konsonantgrafem i gruppe av konsonantgrafemer" (kategori 1b), "Innskutt konsonantgrafem ved gruppe av konsonantgrafemer" (3b), "Innskutt konsonantgrafem ved siden av enkelt konsonantgrafem" (3d), "Innskutt vokalgrafem i gruppe av konsonantgrafemer" (4b) og "Substitusjon av konsonantgrafemer i gruppe av konsonantgrafemer" (5b). For å nyansere bildet kan man også se på funnene i de opprinnelige kategoriene (figur 8) – da kommer det fram at ved to av kategoriene gjør VIET og BKS like mange feil, mens ved tre av dem gjør VIET flest feil. Begge informantgruppene gjør flest feil ved utelatt konsonantgrafem i gruppe av konsonantgrafemer – BKS med 0,26 % feil per løpeord og VIET med 0,45 % feil.

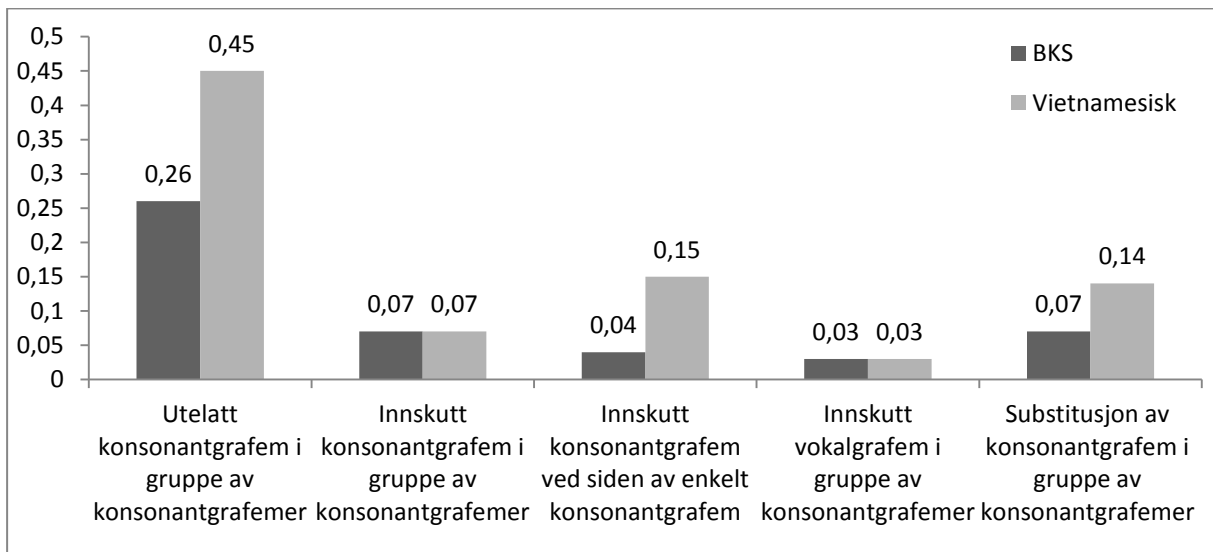
Ved utelatte konsonantgrafemer i gruppe av konsonantgrafemer utelater de to informantgruppene en del av de samme grafemene (jf. figur 9), som <d>, <g>, <j>, <k> og <n>. Men ved <r>, <s> og <t> er det store forskjeller mellom dem; VIET har henholdsvis 19, 27 og 17 forekomster av disse feilene, mens BKS har 8, 6 og 6 forekomster. Utelating av <v> har BKS 7 forekomster av, mens VIET har ingen. Av innskutte konsonantgrafemer ved siden av annet konsonantgrafem har VIET, som nevnt, mye flere feil enn BKS. VIET har 7 forekomster av innskutt <r> og <t>, og 5 forekomster av innskutt <d>, <g> og <k> (jf. figur 10).

Av innskutt konsonantgrafem i gruppe av konsonantgrafemer gjør begge informantgrupper få feil, og disse utgjør en like stor prosentandel hos begge gruppene – 0,07 % per antall løpeord.

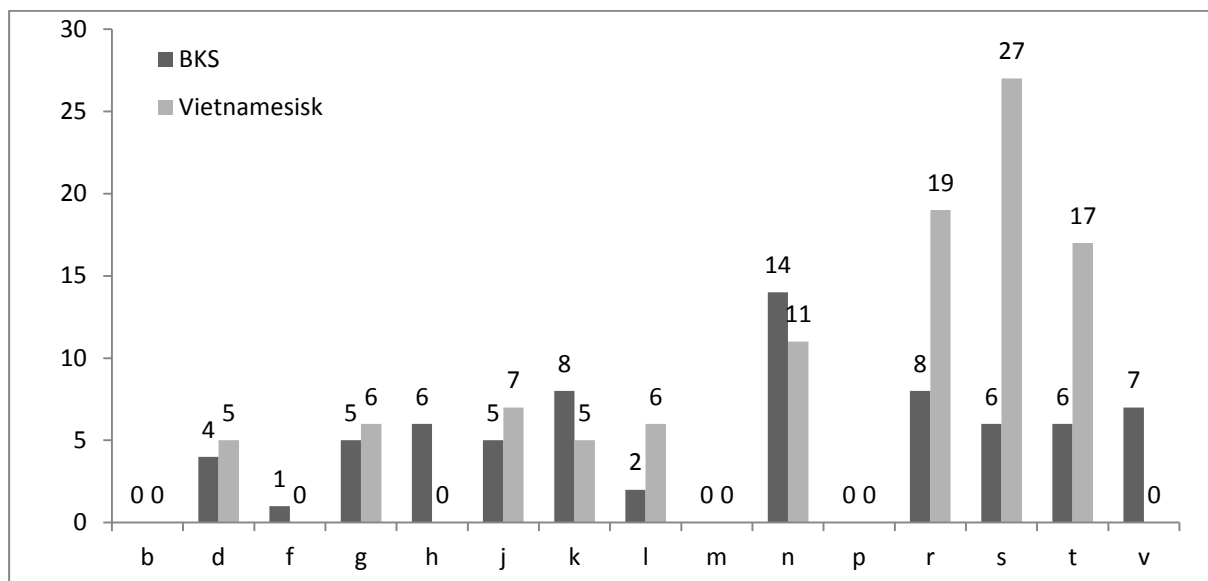
Ved substitusjon av konsonantgrafem i gruppe av konsonantgrafemer gjør VIET og BKS derimot ulik mengde feil – differansen er på 0,07; VIET gjør flest feil.



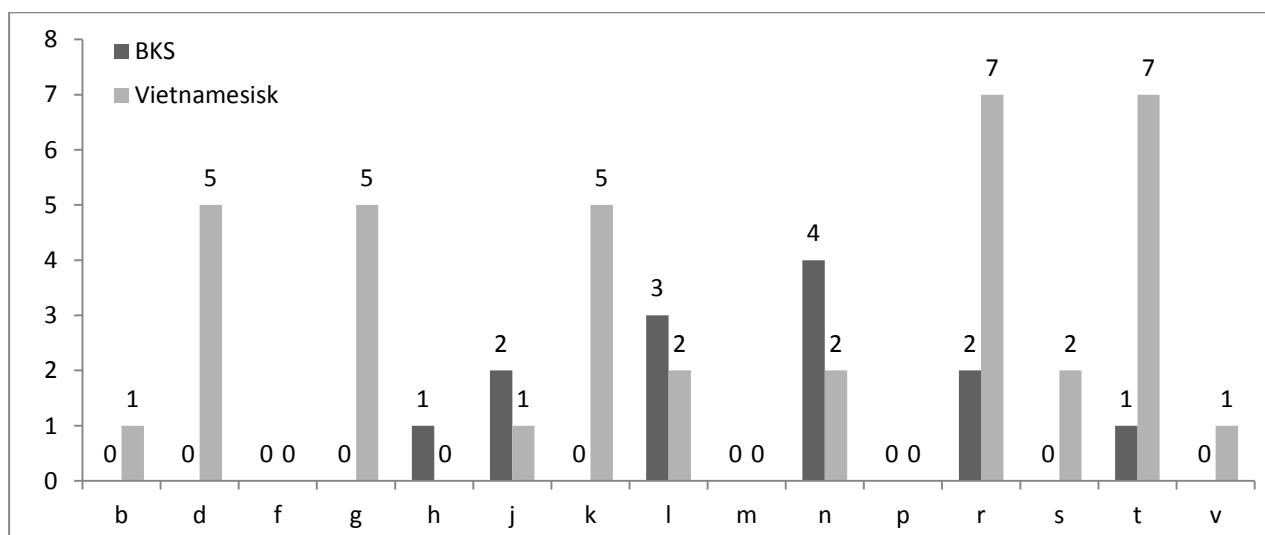
Figur 7: Feil ved gruppe av konsonantgrafemer i prosent.



Figur 8: Feil ved gruppe av konsonantgrafemer fordelt på underkategorier – tallene er i prosent.



Figur 9: Utelatte konsonantgrafemer i gruppe av konsonantgrafemer – tallene er i faktiske forekomster.



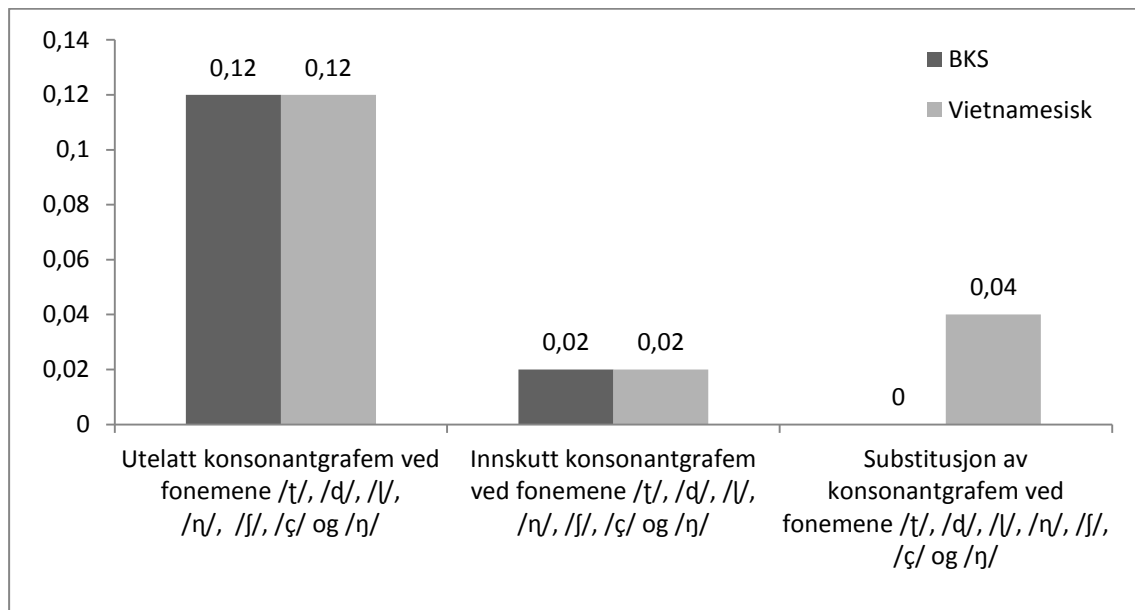
Figur 10: Innskutte konsonantgrafemer ved siden av annet konsonantgrafem i faktiske forekomster.

### 5.5.2 Feil ved grafemer/polygrafemer som gjengir fonemene /t/, /d/, /l/, /ŋ/, /ʃ/, /ç/ og /ŋ/

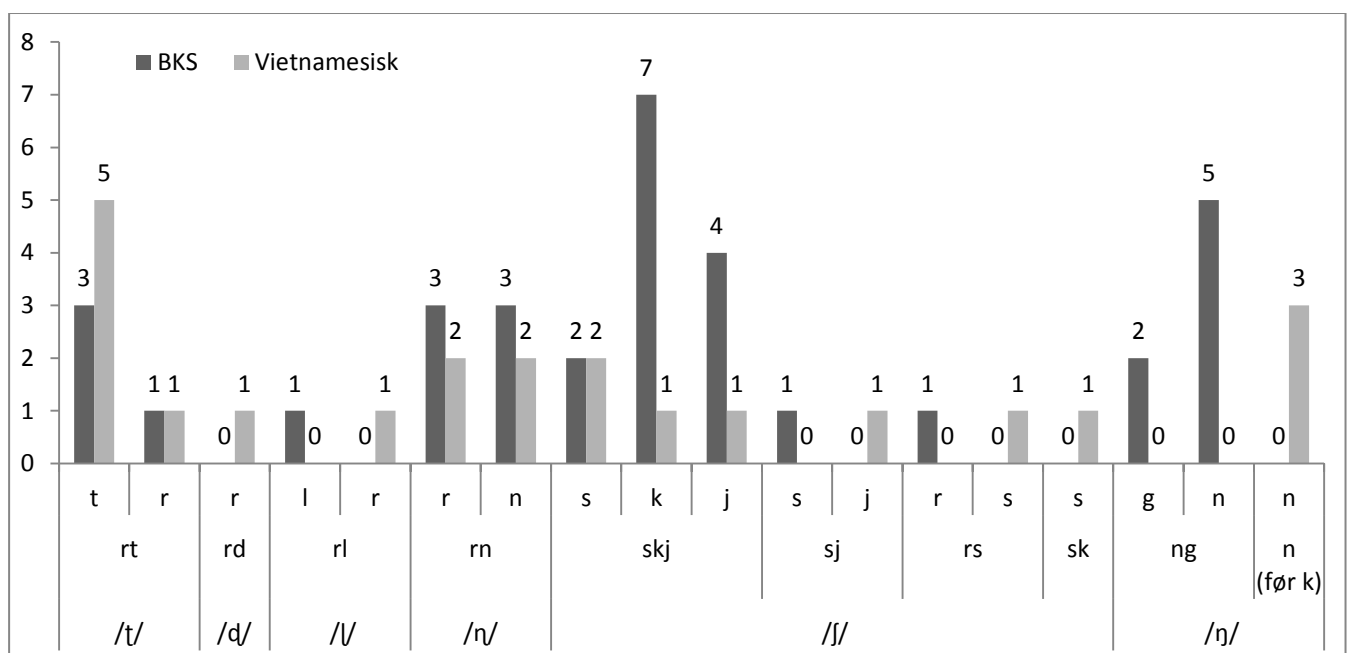
Feilprosenten ved grafemer/polygrafemer som gjengir fonemene /t/, /d/, /l/, /ŋ/, /ʃ/, /ç/ og /ŋ/, er den samme i VIET og BKS i to av tre feilkategorier (jf. figur 11). Dette er kategoriene "utelatt konsonantgrafem ved /t/, /d/, /l/, /ŋ/, /ʃ/, /ç/ og /ŋ/" og "innskutt konsonantgrafem ved



/t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ"/. Førstnevnte kategori har flest forekomster. I kategorien gjøres det forskjellige typer feil ved flere forskjellige grafemer – ingen spesifikk feiltype har mange forekomster (jf. figur x). I kategorien "substitusjon av konsonantgrafemer ved /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/" gjør VIET ingen feil, mens BKS gjør 0,04 % feil (se tabell 24–26 i vedlegg 7 for detaljer).



Figur 11: Feil ved fonemene /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ i prosent.

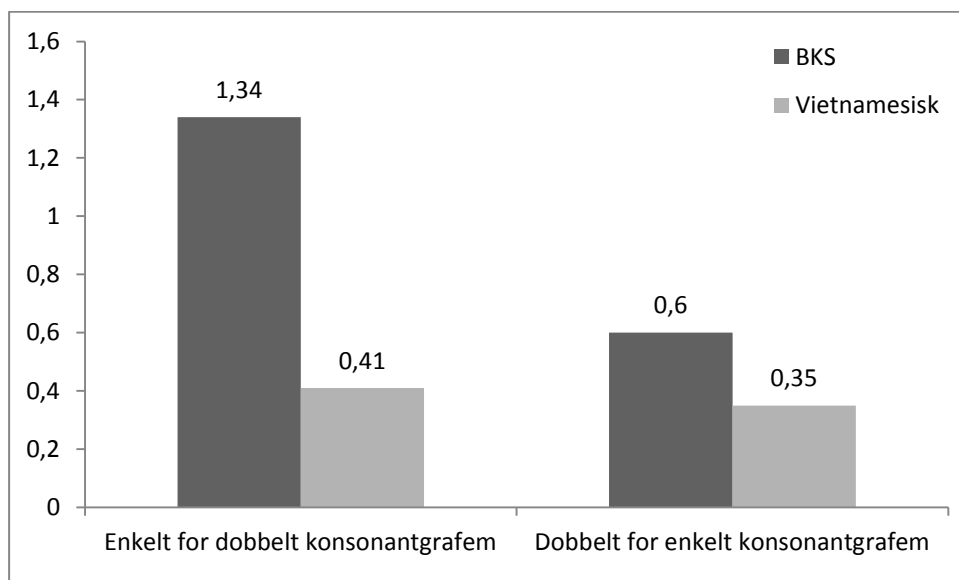


Figur 12: Utelatte grafemer ved fonemene /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ i faktiske forekomster.

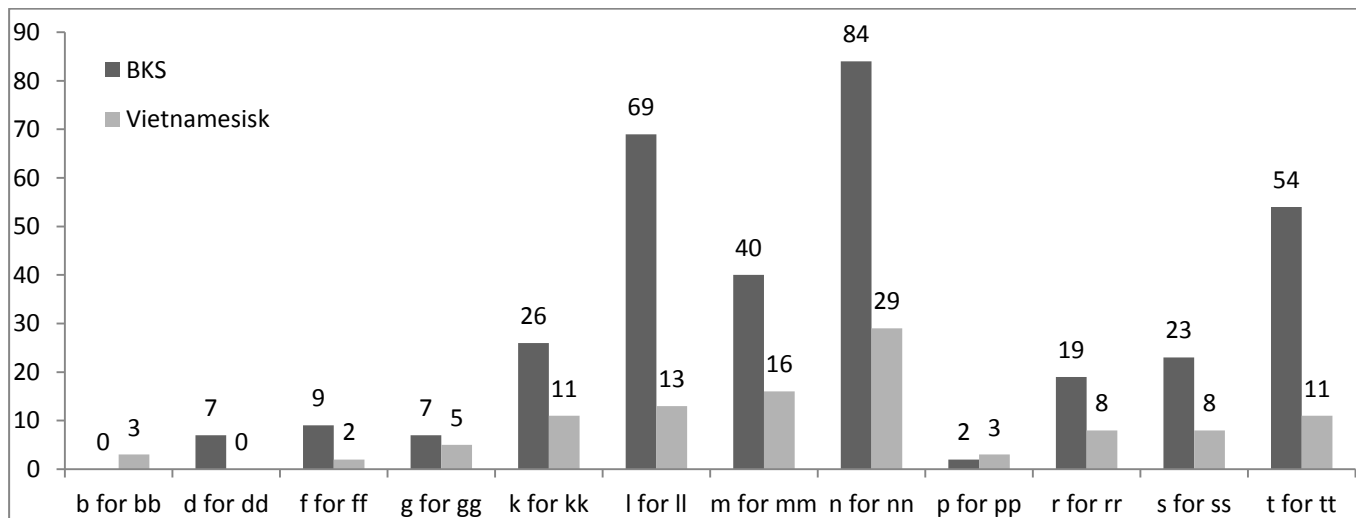
## 5.5.2 Feil ved enkelt og dobbelt konsonantgrafem

Både VIET og BKS har en del tilfeller av enkelt for dobbelt konsonantgrafem og dobbelt for enkelt konsonantgrafem (jf. figur 13). BKS gjør likevel flere feil enn VIET ved begge disse kategoriene. Av feil ved enkelt for dobbelt konsonantgrafem har BKS 1,34 % feil per løpeord, mens VIET har 0,41 % – forskjellen er statistisk signifikant (jf. vedlegg 4). Ved dobbelt for enkelt konsonantgrafem har BKS 0,6 % feil – her har VIET 0,35 % feil. I kategorien enkelt for dobbelt konsonantgrafem er det forskjeller innad i hver informantgruppe (jf. vedlegg 4).

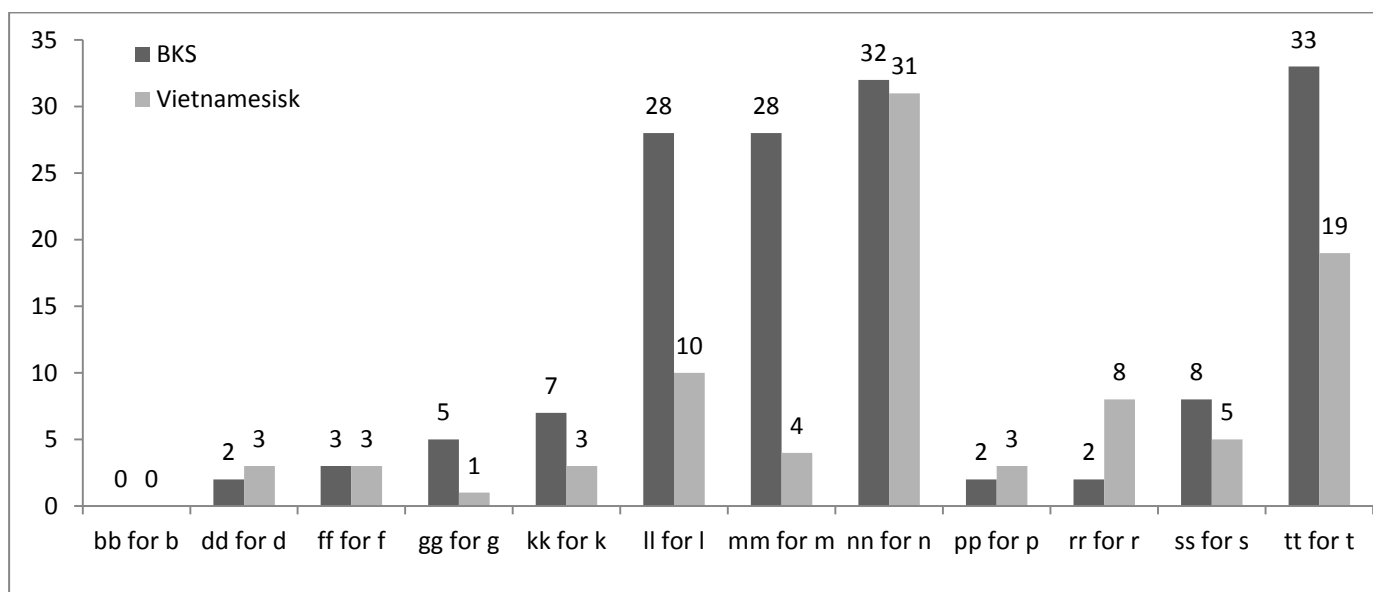
Hvilke grafemer det gjøres feil ved i kategoriene enkelt for dobbelt konsonantgrafem og dobbelt for enkelt konsonantgrafem, har store likheter gruppene imellom. Det gjøres flest feil ved <l>, <m>, <n> og <t>. Dette gjelder både enkelt for dobbelt konsonantgrafem og dobbelt for enkelt konsonantgrafem (jf. figur 14 og 15). Den største forskjellen ved feil ved enkeltskrivning og dobbeltskriving av konsonantgrafemer handler altså om antall feil som gjøres, ikke hvilke grafemer det gjøres feil ved. I tillegg er det en annen feiltype her som skiller de to informantgruppene: VIET gjør en del feil der dobbelt konsonantgrafem er byttet ut med to ulike konsonantgrafemer, eller der to ulike konsonantgrafemer er byttet ut med et dobbelt konsonantgrafem. BKS gjør derimot få slike feil (for detaljer, se tabell 23 i vedlegg 7).



Figur 13: Enkelt for dobbelt og dobbelt for enkelt konsonantgrafem i prosent.



Figur 14: Enkelt for dobbelt konsonantgrafem i faktiske forekomster.



Figur 15: Dobbelt for enkelt konsonantgrafem i faktiske forekomster.

## 5.6 Variasjon innad i informantgruppene

I noen tilfeller er det stor variasjon innad i VIET og BKS på hvor mange informanter som gjør en type feil, og hvor mange som ikke gjør den. Hvis få personer alene gjør alle forekomstene en feiltype har i en informantgruppe, betyr det at feiltypen er mindre representativ for gruppa som helhet enn om forekomstene er fordelt på flere personer. Når det er slik at én/noen få personer gjør stort utslag på et resultat, har jeg kommentert dette. Der variasjonen mellom personene ikke har blitt nevnt, betyr det at forekomstene fordeler seg nogenlunde jevnt mellom personene. Generelt er det viktigere å kommentere hvor mange informanter som gjør feilene ved lave forekomster enn ved høye, fordi en eventuell variasjon mellom personene i antall feil som gjøres, blir jevnet ut ved høyere forekomster.

## 5.7 Oppsummering av kapitlet

Kapittel 5 inneholder en presentasjon av resultatene i undersøkelsen. Resultatene viser at det er både forskjeller og likheter i ortografien til VIET og BKS. Det er også variasjon mellom informantenes ortografi innad i hver informantgruppe. Denne variasjonen – mellom informantgruppene og innad i hver informantgruppe – gjelder både feilfrekvens og type feil som gjøres. Resultatene har jeg presentert i prosenttall og faktiske forekomster. Prosenttallet viser feilfrekvens per løpeord, mens faktiske forekomster er tall som ikke er omregnet. Funnene er vist i tabeller og figurer. Dersom få personer i en informantgruppe alene står for et visst funn, er dette resultatet mindre representativt for gruppa som helhet enn hvis flere informanter gjør denne feilen. I slike tilfeller er dette kommentert.

Blant funnene som er presentert i dette kapitlet, vil særlig de følgende forskjellene mellom VIET og BKS bli sentrale i den videre drøftingen:

- VIET har flere feil enn BKS ved grupper av konsonantgrafemer.
- BKS har flere tilfeller av enkelt for dobbelt konsonantgrafem enn VIET.
- BKS har flere substitusjonsfeil enn VIET ved vokalgrafemer.

# 6 Drøfting

## 6.1 Innledning

I forrige kapittel presenterte jeg funnene i undersøkelsen. I dette kapitlet vil jeg sette feilene i kontekst, og drøfte hva som kan være årsakene til at de aktuelle feilene oppstår. Utgangspunkt for drøftingen vil være forskningsspørsmålene jeg presenterte i 2.6. I tråd med forskningsspørsmålene vil jeg sammenligne ortografiske trekk hos VIET og BKS og drøfte resultatene av undersøkelsen.

Kapitlet er inndelt som følger: Jeg starter med å drøfte funn som viser forskjeller mellom VIET og BKS (6.2), før jeg drøfter funn som ikke viser forskjeller mellom de to informantgruppene (6.3). Deretter ser jeg på andre funn (6.4). Til slutt vil jeg reflektere rundt undersøkelsen og beskrive mulig videre forskning (6.5), før jeg oppsummerer kapitlet (6.6).

## 6.2 Drøfting av funn som viser forskjeller mellom VIET og BKS

Det primære formålet med undersøkelsen har vært å forsøke å svare på om det er forskjeller i ortografien til VIET og BKS, og om eventuelle forskjeller kan relateres til morsmålet eller morsmålsskriftspråket deres. For å kunne svare på dette hovedspørsmålet utarbeidet jeg seks underordnede forskningsspørsmål (jf. 2.6). Av disse seks viste det seg, slik det går fram av resultatkapitlet, at VIET og BKS har relativt store forskjeller i resultater ved tre av dem. Ved de tre andre spørsmålene fant jeg ingen eller relativt små forskjeller.

### 6.2.1 Oversikt over funn som viser forskjeller mellom VIET og BKS

De tre forskningsspørsmålene som viste forskjeller mellom informantgruppene, var:

Forskningsspørsmål 1: Har VIET en høyere andel feil ved grupper av konsonantgrafemer enn BKS?

Forsknings spørsmål 4: Er enkelt for dobbelt konsonantgrafem en mer frekvent feil enn dobbelt for enkelt konsonantgrafem, og er det forskjeller mellom VIET og BKS?

Forsknings spørsmål 6: Er substitusjonsfeil ved ⟨u⟩, ⟨y⟩ og ⟨ø⟩ frekvente, og er det forskjeller mellom VIET og BKS?

Ved forsknings spørsmål 1 viste min forventning seg å stemme – altså forventningen om at VIET skulle ha flere feil enn BKS ved grupper av konsonantgrafemer (jf. figur 7). Differansen mellom informantgruppene viste seg å være statistisk signifikant (jf. vedlegg 4). Grunnen til at jeg mente VIET ville kunne ha flere feil ved grupper av konsonantgrafemer enn BKS, er at kontrasten mellom norsk og vietnamesisk ortografi er stor ved grupper av konsonantgrafemer – mens kontrasten mellom ortografien i norsk og BKS her er mindre. For mens vi i norsk har mange og lange grupper av konsonantgrafemer (jf. 3.3), har vietnamesisk ikke større grupper av konsonantgrafemer enn digrafer (men disse uttales alltid som ett fonem, jf. 3.4). I ortografien i BKS har man grupper av konsonantgrafemer, men disse er ikke like lange som i norsk, og de har større grafotaktiske begrensninger (jf. 3.5).

I forsknings spørsmål 4 forventet jeg at både VIET og BKS skulle ha flere tilfeller av enkelt for dobbelt konsonantgrafem enn dobbelt for enkelt konsonantgrafem, noe som viste seg å stemme (jf. figur 13). Differansen mellom feil ved enkelt for dobbelt konsonantgrafem og feil ved dobbelt for enkelt konsonantgrafem er likevel liten hos VIET. Jeg forventet også at VIET ville gjøre flere feil enn BKS. Denne forventningen ble avkreftet – det er BKS som gjør flest feil. BKS gjør flere feil enn VIET ved enkelt for dobbelt konsonantgrafem, og differansen mellom informantgruppene her er statistisk signifikant (jf. vedlegg 4). Jeg forventet at VIET skulle gjøre flere feil enn BKS fordi det ikke finnes dobbelt konsonantgrafem i vietnamesisk, og fordi konsonantgrafemer ikke brukes i like stor grad som i ortografien i BKS. Man har heller ikke distinktiv vokallengde i vietnamesisk (jf. 3.4). I ortografien i BKS brukes dobbelt konsonantgrafem kun i liten grad, men konsonantgrafemer brukes mye, og de forekommer ofte i grupper. I BKS har man distinktiv vokallengde, men dette gjengis ikke i skrift (jf. 3.5).

Av substitusjonsfeil ved ⟨u⟩, ⟨y⟩ og ⟨ø⟩, i forsknings spørsmål 6, viste det seg at BKS gjør svært få feil ved ⟨u⟩, men flere feil ved ⟨y⟩ og ⟨ø⟩. VIET gjør derimot få feil ved alle tre

vokalgrafemene (jf. tabell 8). I norsk har vi tre fremre, rundede vokaler: /u, u:/, /y, y:/ og /ø, ø:/. Disse gjengis alltid henholdsvis <u>, <y> og <ø> (jf. 3.3). Vokalene er sjeldne i verdens språk, og de finnes heller ikke i BKS eller vietnamesisk (jf. henholdsvis 3.5 og 3.4). Jeg ville derfor forsøke å finne ut av om grafemene for vokallydene er vanskelige å lære seg for innlærerne – selv om jeg ikke vet noe om i hvor stor eller liten grad eventuelle vanskeligheter med uttale/persepsjon av uttale er avgjørende for ortografien til innlærerne. Jeg antok at VIET ville mestre disse vokalgrafemene bedre enn BKS fordi VIET har mange vokaler/vokalgrafemer i morsmålet/morsmålsskriftspråket, og er dermed vant til stor variasjon. BKS har derimot bare fem vokaler både i uttale og skrift. Antagelsen om at BKS skulle gjøre flere feil ved disse vokalgrafemene enn VIET, viste seg altså å stemme for to av de tre vokalgrafemene – forskjellen mellom de to informantgruppene her er betydelig. Fordi jeg ser på forekomsten av feilene per antall løpeord, er tallene likevel for små til at det ville gitt noen mening å regne ut statistisk signifikans av denne forskjellen.

Slik jeg ser det, kan den store forskjellen mellom resultatene til VIET og BKS i forskningsspørsmål 1 og 6 tolkes som påvirkning fra innlærernes morsmål og morsmålsskriftspråk. I forskningsspørsmål 4 ser jeg på resultatene til BKS som påvirkning fra morsmålsskriftspråket – enkelt konsonantgrafem velges oftest fordi innlærerne er vant med å bruke enkelt konsonantgrafem i langt større grad enn dobbelt konsonantgrafem. At VIET bruker enkelt for dobbelt konsonantgrafem bare i litt større grad enn dobbelt for enkelt konsonantgrafem, synes jeg er vanskeligere å tolke. I deres morsmålsskriftspråk brukes bare enkelt konsonantgrafem, og aldri dobbelt konsonantgrafem. Forskjellen mellom informantgruppene her synes jeg derfor er vanskelig å forklare.

I det følgende vil jeg først gå nærmere inn på feil ved grupper av konsonantgrafemer (6.2.2), før jeg tar for meg feil ved enkeltskriving og dobbeltskriving av konsonantgrafemer (6.2.3). Jeg avslutter kapitlet med substitusjonsfeil ved <u>, <y> og <ø> (6.2.4).

## **6.2.2 Feil ved grupper av konsonantgrafemer**

Både VIET og BKS gjør en del feil ved grupper av konsonantgrafemer. VIET gjør likevel flere feil enn BKS i tre av de fem underkategoriene som beskriver feil ved grupper av konsonantgrafemer. Kategorien "utelating av konsonantgrafem i gruppe av

konsonantgrafemer" er en av kategoriene der VIET gjør flere feil enn BKS (jf. 8). Slik jeg ser det, kan forkortingen av konsonantgrafemgrupper ses på som påvirkning fra vietnamesisk ortografi og/eller fonologi (jf. 3.4) Påvirkningen fra fonologien kan knyttes til templathypotesen, som går ut på at stavelsesoppbyggingen i morsmålet kan påvirke andrespråket (jf. 2.3.2). Utelating kan også ses på som en form for forenkling, noe som er et vanlig kjennetegn ved innlærerspråk. De tre mest utelatte konsonantgrafemene i grupper av konsonantgrafemer for VIET er <ɾ>, <s> og <t> – som i "eke" for *ekte*, "næheten" for *nærheten* og "vankelig" for *vanskelig*. BKS har færre feil her.

Flere av konsonantgrafemene som utelates i gruppe av konsonantgrafemer, er stumme konsonanter. Dette vil jeg drøfte i 6.4.4. I noen tilfeller er det likevel ikke den stumme konsonanten som er utelatt, men den ved siden av. Dette finner jeg hos begge grupper, i tilfeller som "helpe" for *hjelp*, "genom" for *gjennom*, "muliget" for *mulighet* og "(i) hertfall" for *(i) hvert fall*. I andre tilfeller utelates et konsonantgrafem som er en del av et polygrafem som gjengis av /t/, /d/, /l/, /ɲ/, /ʃ/, /ç/ eller /ŋ/. Dette vil bli drøftet i 6.3.2.

VIET har også flere feil enn BKS ved substitusjon av konsonantgrafem i gruppe av konsonantgrafemer (jf. figur 8). Det er likevel vanskelig å se noe mønster i feilene fordi substitusjonene gjøres mellom mange forskjellige grafemer. Av forekomster med ett tilfelle hver har VIET 15 feil, mens BKS har 10 feil. Jeg synes det er interessant at det nesten ikke gjøres feil man kan beskrive som parvis plosivveksling (jf. Elsness 2002) – altså bytting begge veier mellom <p>, <t> og <k> og <b>, <d> og <g>. I materialet finner jeg blant annet feilstavingene "sund" for *sunt* og "oppmunndring" for *oppmuntring*. Informantgruppene har heller ikke tilfeller av parvis plosivveksling ved substitusjon av enkle konsonantgrafemer.

Selv om jeg forventet feil ved grupper av konsonantgrafemer av VIET, er jeg likevel overrasket over funnene i to av kategoriene, "innskudd av konsonantgrafem ved siden av annet konsonantgrafem" og "innskudd av konsonantgrafem i gruppe av konsonantgrafemer". Ved førstnevnte kategori gjør VIET flere feil enn BKS, mens ved sistnevnte kategori gjør de to informantgruppene like mange feil. Kategorien "innskudd av konsonantgrafem ved siden av annet konsonantgrafem" har en del forekomster av VIET – 38 (0,15 %). Dette er 25 flere enn gruppa med BKS som morsmål (som har 0,04 % slike feil). Feiltypen innebærer at der det tidligere var et enkelt konsonantgrafem, blir det nå to (for en nærmere beskrivelse av



feiltypen, se 4.4), som i ordene "nærhenten" for *nærheten*, "ordtakt" for *ordtak*, "fristid" for *fritid* og "aktivitert" for *aktivitet*. I den andre kategorien, "innskudd av konsonantgrafem i gruppe av konsonantgrafemer", gjøres det feil som "ulempler" for *ulemper*, "menst" for *mest* og "mistbruk" for *mistbruk*. Grunnen til at jeg er noe overrasket over feilene, er at jeg heller ville forventet at man lagde gruppene av konsonantgrafemer mindre enn større. En grunn til at innlærerne gjør slike feil, kan være at de generelt er usikre i skriving av konsonantgrupper.

I en undersøkelse av Hvenekilde (1995, jf. 2.6) gjorde informantene med vietnamesisk som morsmål svært få feil ved grupper av konsonantgrafemer. Av feil ved grupper av konsonantgrafemer ser hun likevel bare på forekomster initialt i ord. I hennes undersøkelse gjør bare 3 av 93 elever slike feil. Til sammenligning gjør 17 av 99 informanter med vietnamesisk som morsmål samme type feil i min undersøkelse. Tallene er likevel ikke helt sammenlignbare, blant annet fordi informantene i Hvenekildes undersøkelse er barn, mens i min er de voksne. I tillegg har tekstene i undersøkelsene forskjellig lengde. I BKS har 10 av 99 personer feil initialt i ordet ved grupper av konsonantgrafemer, men så gjør også BKS jevnt over færre feil enn VIET ved grupper av konsonantgrafemer.

### **6.2.3 Feil ved enkeltskriving og dobbeltskriving av konsonantgrafem**

BKS har altså flere tilfeller av enkelt for dobbelt konsonantgrafem enn VIET. Forskjellen er statistisk signifikant. BKS har også flere tilfeller av dobbelt for enkelt konsonantgrafem enn VIET, men forskjellen her er ikke stor (jf. figur 13). Den største forskjellen informantgruppene imellom handler om antall feil som gjøres ved enkelt for dobbelt og dobbelt for enkelt konsonantgrafem – ikke om hvilke grafemer det gjøres feil ved. De to informantgruppene gjør, jevnt over, feil ved de samme grafemene (for detaljer, se figur 14 og 15). For BKS er det interessant at de fire konsonantgrafemene det gjøres flest feil ved, er de samme for kategoriene enkelt for dobbelt konsonantgrafem og dobbelt for enkelt konsonantgrafem. Dette er <l>, <m>, <n> og <t>, og feil som "site" for *sitte*, "anerledes" for *annerledes*, "problemmer" for *problemer* og "bestillte" for *bestilte*.

BKS og VIET gjør altså for det meste samme type feil ved enkelt for dobbelt og dobbelt for enkelt konsonantgrafem. Det er likevel én feiltype som BKS har svært få av, men som VIET

har flere av. VIET gjør noen feil der to enkle konsonantgrafemer er byttet ut med et dobbelt konsonantgrafem, som i "kajakt" for *kajakk*, "arkuratt" for *akkurat* og "nygjerning" for *nysgjerrig*. De gjør også samme type feil, men motsatt vei – slik at et dobbelt konsonantgrafem er byttet ut med to enkle konsonantgrafemer, som i "utsikk" for *utsikt* og "detta" for *delta* (for oversikt over feilene, se tabell 23 i vedlegg 7).

Noen feil oppstår der regelen om enkeltskriving og dobbeltskriving av konsonantgrafemer har unntak (jf. 3.3). For eksempel har man enkelt konsonantgrafem etter kort vokal i mange småord. Er man ikke klar over dette, kan det resultere i feil som "denn" for *den* og "skall" for *skal*. Og gjør man stavefeilene "strømm" (*strøm*) og "hjemm" (*hjem*), er man ikke klar over at et ord aldri har <mm> finalt. Stavefeilene "fotballaget" og "busstasjon" er også interessante. Feilene tyder på at man ikke mestrer skriftkonvensjonen om at tre like konsonantgrafemer aldri kan stå sammen. Det er også slik at mange ord har dobbelt konsonantgrafem i én form av ordet, men ikke i en annen, som i infinitivsformen *brenne* og preteritumsformen *brant*. Eller på tvers av ordklasser, som i substantivet *sommel* og verbet *somle*. Dette kan skape forvirring, og resultere i stavefeil som "bestillte", "hjelpemidler", "sinnt" (to <n>-er i *sinne*), "sikkret", "ukjennt" (*kjenne* har to <n>-er), "gamle" og "størrste".

I materialet finner jeg også feil som "innformasjoner", "innroduksjonsprogram", "invandrere", "instans" (*innsats*) og "inbyggerne". Initialt har alle ordene /in/, men stavemåten er likevel forskjellig, enten *in* eller *inn*. *In-* er et latinsk prefiks, mens *inn* er et adverb som brukes mye i sammensetninger – derav to forskjellige stavemåter. Videre er feil som "kariere", "agresive" og "komunikasjoner" interessante fordi ordene skal ha dobbelt konsonantgrafem i en stavelse som ikke er betont. Slik er det i en del fremmedord. Vi trenger likevel ikke gå til fremmedordene for å finne dobbelt konsonantgrafem i ubetont stavelse. I materialet finner jeg feil som "landsman", "dareter" (*deretter*), "unttat", "samfunet" og "øyeblik" – der riktig stavemåte er med dobbelt konsonantgrafem selv om stavelsen det gjelder, er ubetont. Jeg vet likevel ikke om hvorvidt stavelsen er betont eller ikke, er relevant for innlærere ved innlæring av skriftkonvensjonen. Jeg vet heller ikke i hvor stor grad det spiller inn om man hører forskjell på kort og lang vokal i norsk eller ikke. Hvis dette har betydning, ville jeg mene at BKS ville lære skriftkonvensjonen raskere enn VIET i og med at de har distinktiv vokallengde i morsmålet. Resultatene viser likevel at VIET mestrer skriftkonvensjonen bedre enn BKS.

## 6.2.4 Substitusjonsfeil ved ⟨u⟩, ⟨y⟩ og ⟨ø⟩

Resultatene viste at BKS har en betydelig andel substitusjonsfeil ved vokalgrafemer. Mange av feilene gjelder ⟨y⟩ og ⟨ø⟩, men bare et lite antall gjelder ⟨u⟩ (jf. tabell 8). Feilene går ut på å erstatte ⟨y⟩ og ⟨ø⟩ med andre vokalgrafemer. Men substitusjonen går også andre veien, slik at ⟨y⟩ og ⟨ø⟩ er brukt for andre grafemer. VIET gjør derimot få feil ved substitusjon av vokalgrafemer, inkludert ved ⟨u⟩, ⟨y⟩ og ⟨ø⟩.

BKS gjør en del feil ved ⟨y⟩ – de aller fleste gjelder ⟨i⟩ for ⟨y⟩ (14 feil) og ⟨y⟩ for ⟨i⟩ (11 feil). I det latinske alfabetet man bruker i BKS, finnes ikke grafemet ⟨y⟩. I det kyrilliske alfabetet har man derimot grafemet ⟨y⟩, men dette tilsvarer /ʏ/ eller /u:/ – og /i, i:/ skrives ⟨и⟩. I tillegg uttales /y, y:/ og /i, i:/ på omtrent samme sted i munnen i norsk, noe som ytterligere kan vanskeliggjøre skillet mellom ⟨y⟩ og ⟨i⟩. Ifølge Mønnesland (2002) er å uttale /y/ som /i/ en vanlig uttalefeil i norsk av innlærere med BKS som morsmål (Mønnesland 2002:28). Dette kan være mulige forklaringer på substitusjonen av ⟨i⟩ for ⟨y⟩ og ⟨y⟩ for ⟨i⟩. VIET gjør derimot få feil ved ⟨y⟩. De har riktignok sju tilfeller av ⟨i⟩ for ⟨y⟩, men siden seks av tilfellene gjøres av samme person, er ikke dette representativt for informantgruppa som helhet.

Verken BKS eller vietnamesisk har /ø/ eller /ø:/, og de har heller ikke grafemet ⟨ø⟩ i skriftspråkene sine (jf. 3.4 og 3.5). At innlærerne gjør feil som involverer ⟨ø⟩, synes jeg derfor ikke er overraskende. Feilene går i all hovedsak ut på at ⟨ø⟩ er erstattet med ⟨o⟩, og at ⟨o⟩ er erstattet med ⟨ø⟩ – få feil gjelder substitusjon av ⟨ø⟩ med andre grafemer. BKS gjør flere feil enn VIET også ved ⟨ø⟩. ⟨o⟩ er valgt for ⟨ø⟩ i 28 tilfeller og ⟨ø⟩ er valgt for ⟨o⟩ i 21 tilfeller av BKS, mens VIET har 1 tilfelle av ⟨o⟩ for ⟨ø⟩ og 7 tilfeller av ⟨ø⟩ for ⟨o⟩. Begge informantgruppene har få substitusjonsfeil ved ⟨u⟩.

## 6.3 Drøfting av funn som ikke viser forskjeller mellom VIET og BKS

Som beskrevet i 6.2, viste resultatene på tre av de seks forskningsspørsmålene forskjeller mellom VIET og BKS, mens tre ikke viste forskjeller. Funnene som viste forskjeller, har jeg beskrevet i 6.2. De tre forskningsspørsmålene der funnene ikke viste forskjeller mellom VIET og BKS, presenterer jeg i det følgende.

### 6.3.1 Oversikt over funn som ikke viser forskjeller mellom VIET og BKS

De tre forskningsspørsmålene som ikke viste forskjeller mellom informantgruppene er:

Forskningsspørsmål 2: Er det spesielt mange feil ved grafemer/polygrafemer som gjengir /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/, og er det være forskjeller mellom VIET og BKS?

Forskningsspørsmål 3: Har VIET innskudd av vokalgrafemer i grupper av konsonantgrafemer i større grad enn BKS?

Forskningsspørsmål 5: Er diftongering av monoftong og monoftongering av diftong en frekvent feil, og er det forskjeller mellom BKS og VIET?

Ved forskningsspørsmål 2 viste det seg at VIET og BKS gjør omtrent like mange feil ved grafemer/polygrafemer som gjengir /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ (jf. figur 11). Jeg finner heller ingen forskjeller mellom informantgruppene hva gjelder typen feil som gjøres. I norsk har flere av fonemene /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/ ulike representasjoner i skrift. /ç/ staves for eksempel <k> i kinn og <kj> i kjeks, og /ŋ/ staves <ng> i sang, <g> i sagn og <n> i synke (for flere detaljer, se 3.3). Det som kan vanskeliggjøre tilegningen for VIET og BKS ytterligere, er at nesten ingen av fonemene finnes i morsmålene deres – men begge har polygrafemer (jf. 3.4 og 3.5). Jeg antok at VIET ville gjøre flere feil enn BKS fordi strukturen av konsonantgrafemer er likere i BKS og norsk enn i vietnamesisk og norsk. Denne antagelsen ble altså avkreftet.

I forskningsspørsmål 3 ville jeg undersøke om VIET har flere innskudd av vokalgrafemer i grupper av konsonantgrafemer enn BKS. Jeg fant omtrent like mange feil hos VIET og BKS (jf. tabell 7). Jeg hadde forventet at VIET skulle ha en høyere andel av innskudd av vokalgrafemer i grupper av konsonantgrafemer enn BKS fordi vietnamesisk ikke har lange grupper av konsonantgrafemer (jf. 3.4), som norsk. BKS har derimot grupper av konsonantgrafemer, men ikke like lange og med flere grafotaktiske restriksjoner enn i norsk ortografi. Innskudd av vokalgrafemer i grupper av konsonantgrafemer kunne, etter det jeg forventet, blitt brukt av VIET for å løse opp konsonantgrafemgruppene – likt det muntlige

fenomenet epentese (jf. 2.3.2). Dette er også blitt gjort, men i mindre grad enn jeg hadde forventet – og ikke i større grad enn BKS.

Monoftongering av diftong og diftongering av monoftong er feil verken BKS eller VIET gjør (figur 6). Min hypotese i forskningsspørsmål 5 om at BKS skulle gjøre flere feil enn VIET viste seg derfor å ikke stemme. Det som gjorde meg nysgjerrig på om det gjøres feil der monoftonger blir diftongert og diftonger monoftongert, er knyttet til den omfattende vekslingen i norsk mellom monoftonger og diftonger, både i skrift og uttale (jf. 3.3). I BKS har man derimot ikke diftonger – to vokalgrafemer etter hverandre blir uttalt som tilhørende hver sin stavelse (jf. 3.5). I vietnamesisk har man diftonger. Og i vietnamesisk, som i norsk, skrives disse med to vokalgrafemer (jf. 3.4). I utgangspunktet trodde jeg derfor at VIET ville mestre vekslingen mellom monoftonger og diftonger i norsk bedre enn BKS. At VIET ikke monoftongerer diftonger eller diftongerer monoftonger, viste seg altså å stemme. Men jeg antok ikke at heller ikke BKS ville gjøre slike feil, slik funnene altså viser.

Det er interessant at innlærere med såpass forskjellige morsmål som vietnamesisk og BKS viser store likheter på noen punkter – som ved funnene på forskningsspørsmål 2, 3 og 5. Jeg tolker det dit hen at ikke alle funn kan relateres til innlærernes morsmål/morsmålskriftsspråk. Funnene ved forskningsspørsmål 2, 3 og 5 sier heller mye om norsk skriftspråk. De sier noe om hva som er vanskelig – og ikke vanskelig – å tilegne seg i norsk ortografi.

I 6.3.2 vil jeg drøfte funnene ved feil ved grafemer som representerer fonemene /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/, før jeg drøfter feil ved innskudd av vokalgrafemer i grupper av konsonantgrafemer i 6.3.3. Deretter vil jeg drøfte feil ved monoftongering av diftonger og diftongering av monoftonger i 6.3.4.

### **6.3.2 Feil ved grafemer/polygrafemer som gjengir /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/**

Andelen feil som gjøres ved grafemer/polygrafemer som gjengir /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/ er omtrent lik i VIET og BKS (jf. figur 11). I underkategoriene "innskutt konsonantgrafem ved /t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/" og "substitusjon av konsonantgrafem ved t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/" gjøres det såpass få feil at det er vanskelig å se tendenser. I kategorien "utelating av konsonantgrafem ved t/, /d/, /l/, /n/, /f/, /ç/ og /ŋ/" gjøres det derimot flere feil. Her finner jeg

stavinger som utelatt ⟨g⟩ i "salonen" for *salongen*, utelatt ⟨r⟩ i "bana" for *barna* og utelatt ⟨r⟩ i "ansvalige" for *ansvarlige*. Disse feiltypene synes jeg ikke er overraskende fordi det grafemet som står igjen, er det som lydmessig ligner mest på fonemet som grafemene representerer. I for eksempel "bana" for *barna* er det ⟨rn⟩ som representerer fonemet /n/, og ⟨n⟩ (/n/) ligger nærmere /n/ enn ⟨r⟩ (/r/) gjør. Det finnes likevel eksempler på det motsatte, at grafemet som utelates, er det som er mest likt fonemet grafemene representerer, som i utelatt ⟨t⟩ for ⟨rt⟩ i "forsetter" for *fortsetter* og utelatt ⟨n⟩ for ⟨rn⟩ i "bara" for *barna*. Det finnes også tilfeller som utelatt ⟨n⟩ for /ŋ/ i "kokurranse" for *konkurranse*, der det utelatte grafemet alene representerer fonemet. Jeg hadde forventet flere feil ved /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/. Dette med tanke på den store inkonsekvensen i stavemåten til grafemene som representerer fonemene – spesielt ved /ʃ/, /ç/ og /ŋ/ (jf. 3.3).

### 6.3.3 Feil ved innskudd av vokalgrafemer i grupper av konsonantgrafemer

Innskudd av vokalgrafemer i grupper av konsonantgrafemer har relativt få forekomster hos både VIET og BKS – 0,03 % feil per løpeord hos begge informantgruppene (jf. figur 6). Vi ser likevel flere ulike eksempler, som i "foreldere" for *foreldre*, "personelig" for *personlig*, "voksene" for *voksne*, "inforemere" for *informere*, "fotballlekamp" for  *fotballkamp* og "Vietnameserene" for *vietnameserne*. Disse feilene er gjort av VIET. BKS gjør også denne typen feil – "egene" for *egne*, "andere" for *andre* og "viktigeste" for *viktigste*, "hjemeland" for *hjemland* og "baren" for *barn*. Det kan være flere årsaker til at feilene oppstår. Jeg vil illustrere det med å se det i sammenheng med utelatte vokalgrafemer mellom to konsonantgrafemer. Av slike feil er det interessant at mange av dem representerer en (vanlig) uttale, som i "fredlig" for *fredelig*, "vidergående" for *videregående* (det er det siste utelatte ⟨e⟩-grafemet det gjelder her), "hemmlig" for *hemmelig*, "plutselig" for *plutselig*, "vanskelige" for *vanskelige* og "genrasjon" for *generasjon* (se også 6.4.4 om feil som kan ha bakgrunn i en ortofon skrivemåte).

Slik jeg ser det, er det svært inkonsekvent når man bruker ⟨e⟩ og når man ikke gjør det i norsk ortografi – i slike kontekster som i eksemplene over. Felles for ordene over der ⟨e⟩-grafemet er utelatt, er at ⟨e⟩-en er ubetont, eller at den knapt eller ikke i det hele tatt blir uttalt. Og når ubetonte og ikke uttalte ⟨e⟩-grafemer skrives, så er det ikke uventet at innlærerne har skutt inn

vokalgrafemer i ord som ser ganske like ut som andre ord som har et ⟨e⟩-grafem. For eksempel skrives *fredelig* med ⟨e⟩, mens *personlig* skrives uten ⟨e⟩, enda begge ender på *-lig*.

Feilene "hjemeland" for *hjemland* og "fotballekamp" for *fotballkamp* kan muligens relateres til usikkerhet ved sammensetting av ord – innlærerne kan være usikre på om ordet skal ha fuge eller ikke (dette er drøftet i 6.4.4). Og feilene "Vietnameserene" for *vietnameserne* kan handle om at man er usikker på stavemåte ved substantivbøying, altså i dette tilfellet at man er usikker på hvordan man staver regelmessige hannkjønnssubstantiv i bestemt form flertall (se også 6.4.4).

Det er nesten bare ⟨e⟩-grafemer som blir innskutt i grupper av konsonantgrafemer (jf. tabell 7), noe som betyr at de andre åtte vokalgrafemene står for en veldig liten andel av feilene. Det samme er tilfelle ved utelatte vokalgrafemer. Dette er ikke overraskende da ⟨e⟩ er blant de aller mest frekvente vokalgrafemene i skriftspråket i norsk. Dette betyr ikke at det ikke gjøres feil som involverer andre vokalgrafemer enn ⟨e⟩, men disse feilene gjelder substitusjon, og ikke innskudd eller utelating.

#### **6.3.4 Feil ved diftongering av monoftong og monoftongering av diftong**

Resultatene viser altså at det ikke gjøres feil der monoftonger blir diftongert og diftonger monoftongert. Dette gjelder både VIET og BKS. En mulig forklaring på mestringsen kan være at innlærerne er vant til veksling mellom én og to vokalgrafemer også i sitt eget morsmålskriftspråk, og dermed er ikke denne vekslingen i norsk i seg selv noen stor utfordring. Kategoriene "monoftongering av diftong" og "diftongering av monoftong" utmerker seg derfor ved at det ikke gjøres feil. I norsk ortografi har vi også to vokalgrafemer etter hverandre som ikke er diftonger, men som uttales i hver sin stavelse. Men heller ikke her gjør innlærerne feil. Begge innlærergruppene gjør likevel feil ved vokalgrafemer, gruppa med BKS som morsmål med flest forekomster (jf. figur 5), men feilene er ikke knyttet til monoftongering eller diftongering.

### **6.4 Drøfting av noen andre funn**

I det foregående har jeg drøftet funn som kan belyse forskningsspørsmålene. I dette delkapitlet vil jeg drøfte noen andre funn.

### 6.4.1 Substitusjon av vokalgrafemer

Det er stor forskjell i antallet substitusjoner av vokalgrafemer mellom VIET og BKS – sistnevnte informantgruppe gjør flere feil enn førstnevnte (jf. tabell 8). Jeg har delt inn feilene i fire grupper etter hva jeg mener feilen kan relateres til:

1. Ortofon stavemåte av vokal der vokalen ikke er ortofon, som i "fåretrekker" for foretrekker, der ⟨o⟩ er byttet ut med ⟨å⟩.
2. Overføring av stavemåte fra ord der stavemåten er riktig til ord der stavemåten er feil – men der samme vokallyd uttrykkes i begge tilfellene. For eksempel skrives /ɔ/ og /o:/ ofte ⟨o⟩, som i henholdsvis *post* og *tog*. I feil som "motte" for *måtte* og "bode" for *både* blir denne staverregelen overgeneralisert til å gjelde ord der samme vokallyd uttrykkes, men der denne stavemåten ikke brukes.
3. Feil ved grafemer som representerer fremre, rundede vokaler – særlig ved ⟨y⟩ og ⟨ø⟩
4. Andre typer feil

I gruppe 1 har BKS en del tilfeller av ⟨å⟩ for ⟨o⟩ der vokallyden er en /ɔ/ eller /o:/, men der riktig vokalgrafem er ⟨o⟩. Informantgruppa har 17 feil der ⟨o⟩ er byttet ut med ⟨å⟩ – av disse er 9 tilfeller der vokalen uttales /ɔ/ eller /o:/. Dette er feil som "fåreldre" for *foreldre* og "får" for *for*. BKS gjør hovedsakelig denne typen feil bare ved grafemer som gjengir /ɔ/ og /o:/. VIET gjør derimot svært få feil i denne gruppa.

VIET gjør også få feil i den neste kategorien – gruppe 2. Jeg finner bare én feil. BKS gjør derimot en del feil jeg mener tilhører gruppe 2. Hele 27 av 27 feil der ⟨å⟩ er byttet ut med ⟨o⟩ er tilfeller der vokalen uttales med en /ɔ/ eller /o:/. Dette er feil som "uforstoelig" for *uforståelig*, "fo" for *få*, "motten" for *måten*, "stor" for *står* og "sonn" for *sånn*. Det som gjør substitusjonen ⟨o⟩ for ⟨å⟩ av BKS ekstra interessant er at ⟨o⟩ uttales /ɔ/ eller /o:/ i BKS. Denne informantgruppa gjør også flere feil enn VIET i gruppe 3. Dette har jeg drøftet i 6.2.4. Sammenlagt gjør altså BKS flere substitusjoner ved vokalgrafemer enn VIET – selv om tallene i faktiske forekomster ikke er helt sammenlignbare informantgruppene imellom (jf. 5.3).



## 6.4.2 Ombytting av grafemer

Det er to typer ombytting av grafemer som innlærerne gjør. Den ene er der grafemene står ved siden av hverandre, som i "frøst" for *først* – her har <r> og <ø> byttet plass. Denne feiltypen tilsvare kategori 7 ("Ombytting av grafemer") i kategoriinndelingen (jf. tabell 3). Det er flest ombyttinger der et vokalgrafem og et konsonantgrafem har byttet plass, som i "fra" (*far*), "strek" (*sterk*), "sikkerhet" (*sikkerhet*) og "elder" (*elder*). Det er også noen ombyttinger der to konsonantgrafemer har byttet plass. Det finner jeg i "large" (*lagre*), "evl" (*elv*), "søvl" (*sølv*) og "vedlig" (*vedlig*). Det er ingen tilfeller i materialet der to vokalgrafemer som står ved siden av hverandre har byttet plass.

Den andre typen ombytting av grafemer har ingen egen kategori i kategoriinndelingen. Dette er tilfeller der grafemer i ordet har byttet plass, men der grafemene ikke står ved siden av hverandre. Jeg vil illustrere med eksempler som "førnoyd" (*fornøyd*), "bølep" (*beløp*), "mårsmol" (*morsmål*), "lastebilsjøfår" (*lastebilsjåfør*) og "sølfelig" (*selvfølgelig*). I alle tilfellene er det vokalgrafemer som har byttet plass.

Det er ingen betydelige forskjeller mellom VIET og BKS i antall og type feil ved ombytting av grafemer. Dette gjelder begge typene av ombyttinger. Jeg synes det mest interessante og overraskende med denne feiltypen er hvor få feil det gjøres av dem. VIET og BKS demonstrerer likevel i min undersøkelse at de ikke finner det vanskelig å lære seg rekkefølgen på grafemer.

## 6.4.3 Overgeneralisering, system, variasjon og påvirkning fra engelsk

Både VIET og BKS gjør ortografiske feil jeg mener samsvarer med typiske kjennetegn for innlærerspråk generelt. Påvirkning fra morsmålet/morsmålsskriftsspråket og forenkling i form av utelatt grafem (jf. 6.2.2) har jeg drøftet i det foregående. Her vil jeg drøfte andre egenskaper ved ortografien til VIET og BKS som jeg mener kan knyttes opp mot kjennetegn for innlærerspråk. Dette er egenskaper som variasjon, system, overgeneralisering og påvirkning fra andre språk. Når det gjelder variasjon, som her betyr å ha flere stavemåter for det samme ordet i den samme teksten, finner jeg eksempler som "forskjelige" og "forsjelig"

for *forskjellig/e*, "bana" og "bara" for *barna* og "vidergående" og "vidergående" for *videregående*. Jeg finner mer variasjon hos BKS enn hos VIET. Variasjon i stavemåte kan også bety at et ord blir stavet både riktig og feil i samme tekst. Jeg har ikke oversikt over ord som staves riktig i og med at jeg bare har sett på ord som er annotert med ortografisk feil (se kapittel 4 for en beskrivelse av metoden i undersøkelsen). Det kan altså hende det er variasjon i form av både riktig og feil stavede ord, men dette har jeg ikke undersøkt.

At en innlærer har flere tilfeller av et feilstavet ord der ordet er stavet på den samme måten, har jeg derimot oversikt over. Slike tilfeller er det flere av, som én informant som har tre tilfeller av "kigarette" for *sigarett* og fem tilfeller av "ulike/uliker/brantuliker" for *ulykke/ulykker/brannulykker*. Jeg tolker dette som at innlæreren har funnet sitt system for en bestemt skrivemåte, og holder seg til den. Disse gjentakende feilene i samme tekst kan kalles systematiske feil, i motsetning til ikke-systematiske feil (glipp) (jf. 2.4). Jeg finner også eksempler i materialet på ikke-systematiske feil, som "jer" for *jeg*, "pa" for *på* og "ot" for *at*. De feilstavede ordene forekommer bare én gang i informantenes tekster, og jeg mener derfor det ikke er grunnlag for å si sikkert at informantene ikke egentlig kan stavemåten – dette også fordi ordene er høyfrekvente. Jeg tolker disse feilene som en glipp.

Innlærerne gjør også feil jeg vil beskrive som overgeneralisering av rettskrivingsregler. Dette gjelder begge grupper. Dette er stavefeil som "hvar" for *var*, som jeg vil tolke som en overgeneralisering av regelen med <h> foran <v> i spørreord. Innskudd av <d> i "lærde" for *lærte* og "marsipanjordbærd" for *marsipanjordbær* kan være en overgeneralisering av <d> etter <r>, som i *bord*, *verdi* og *verden*. Stavefeilen "gjente" for *jente* mener jeg også kan være en overgeneraliseringsfeil fordi vi i mange ord har <g> foran <j> – *gjenta*, *gjennom*, *gjemme* og *gjøre*.

Av påvirkning fra andre språk og skriftspråk er det ikke bare morsmålet eller morsmålsskriftspråket som er relevant å vurdere. Andre språk og skriftspråk innlæreren kan, kan også påvirke. I materialet finner jeg eksempler på det jeg mener er påvirkning fra engelsk ortografi. Ordene det gjelder, har en stavemåte som er enten helt eller delvis lik stavemåten til tilsvarende ord i engelsk. Dette er ord som "disco", "individer" (*individual*), "medicin", "patienter", "chokolade", "exporterer", "advansert" (*advanced*), "boxing" og siste del i ordet "flycrash". Ordet *eksempel* har spesielt mange feil i materialet, og disse mener jeg kommer av

påvirkning fra stavemåten i engelsk. Ordet staves da enten som på engelsk, eller det gjøres én eller flere av følgende feil: <x> for <ks> ("exempel"), <a> for det andre <e>-grafemet i ordet ("eksampel") og/eller ombytting av de to siste grafemene – <le> for <el> ("eksempel"). Jeg finner ingen forskjeller mellom VIET og BKS ved påvirkning fra engelsk stavemåte.

#### 6.4.4 Andre feil

I materialet finner jeg noen typer feil som er felles for begge informantgruppene, og som jeg mener ikke kan relateres til morsmålene/morsmålsskriftspråkene deres. Noen av disse har vært drøftet allerede, slik som innskudd av vokalgrafemer i grupper av konsonantgrafemer, som i "vanskelig" for *vanskelig* (jf. 6.3.3), samt alle feiltypene jeg drøftet i 6.4.3. Andre slike feil er utelating av fuge-s, som i ordene "oppholdtillatelse" og "opptakprøven". Det finnes også eksempler på innskudd av (det som kan være tenkt som) fuge-s. Her finner jeg eksemplene "regelsmessig" og "søksnadd". Jeg finner også eksempler på innskudd og utelating av det jeg tolker som en fuge-e i materialet. Dette har jeg drøftet i det foregående (jf. 6.3.3).

Mange ortografiske feil opptrer finalt i ord, og jeg mener det kan være grunn til å tro at feilene kan skyldes usikkerhet om ordform og bøyning (jf. 2.3.3). I ordene "skoge", "kroppe", "blomste", "meldinge" og "bedrifte" står ikke ordene i noen riktig bøyingsform. Det kan være at informantene er usikre på substantivbøyning. I en del tilfeller mangler et <e>-grafem finalt i ordet der siste stavelse er ubetont, som i "høyskol" for *høyskole*, "helsepersonal" for *helsepersonale*, "T-ban" for *t-bane* og "hjørn" for *hjørne*. Det er også en del tilfeller av <e> finalt i ord, som for eksempel i "vaere" for *hver*, "forstode" for *forstod* og "skifte" for (*et*) *skift*. Slik jeg ser det, kan noen av feilene skyldes påvirkning fra engelsk ortografi (se også 6.4.3). Skriftspråket i engelsk er svært lite ortofont, blant annet på den måten at mange ord slutter på <e>, uten at grafemet blir uttalt – som i *phone, shine, lane, bike*. Andre ord igjen har ikke <e> finalt – *soup, hat, chair, lamp*. Norsk har også både <e> finalt og ikke <e> finalt, men norsk er likevel mer ortofont her, på den måten at om et ord slutter på /e/, så skrives også <e>-grafemet.

Av feil finalt i ord dreier det seg altså ofte om innskudd eller utelatelse av <e>. Men i en del tilfeller er det et <r>-grafem som er utelatt eller innskutt. De ortografiske feilene "gamler" for

*gamle*, "senerer" for *senere* og "beger" for *begge* kan ha med usikkerhet om ordform å gjøre. I noen tilfeller mangler <r>-grafemet finalt i ord når stavelsen er ubetont, som i "ansva" for *ansvar*, "vakke" for *vakker* og "søste" for *søster*. Det er likevel også tilfeller av betonte <r>-grafemer som utelates; dette finner jeg i feil som "expeditø" for *ekspeditør* og "papi" for *papir*. Noen ganger er det andre grafemer enn <e> eller <r> som blir utelatt, som i utelatt <l> i "motoskykke" for *motorsykkkel* og utelatt <d> i "fornøy" for *fornøyd*.

Det gjøres også noen feil der jeg tolker det som at en ordform er blitt blandet med en annen, altså at deler av formen til et semantisk relatert ord i en annen ordklasse er blitt brukt i stavingen. Her tenker jeg på feilstavinger som "produksere" for *produsere*, "brant" for *brann*, "langtvarig" for *langvarig*, "organisasere" for *organisere*. I det førstnevnte ordet er stavemåten i substantivet *produksjon* blandet inn i stavemåten for verbet *produsere*. Det samme er tilfelle for *organisasere* for *organisere*. I "brant" for *brann* er preteritumsformen for verbet *brenne* valgt for substantivet *brann*. Nøytrumsformen av adjektivet *lang* er brukt i stedet for den ubøyde formen i "langtvarig" for *langvarig*. I de feilstavede ordene "tanke" for *tenke* og "fortalle" for *fortelle* er feil vokalgrafem valgt. I det førstnevnte ordet er substantivformen brukt for verbet, og i det sistnevnte er vokalgrafemet som blant annet finnes i preteritumsformen av verbet, brukt i infinitivsformen.

Noen feilstavinger kan være en mer ortofon skrivemåte av ordet. Dette kan være utelating av grafem som representerer trykklett vokal, som i "fredlig" for *fredelig* (jf. 6.3.3) eller substitusjon av vokalgrafem for et mer ortofont vokalgrafem, som i "fåretrekker" for *foretrekker* (jf. 6.4.1). Det kan også være utelating av stum konsonant. Dette er tilfeller som "beskje", "hyggelie", "veldi", "viktieri", "fårsto", "forskjelli", "vanskelriere" og "lane" for *landet*. Jeg finner også noen tilfeller av substitusjon av konsonantgrafem til en mer ortofon konsonant, som i "jikk" for *gikk*.

## 6.5 Drøfting av undersøkelsen og videre forskning

I arbeidet med undersøkelsen har jeg stått overfor en rekke avgjørelser. Noen av avgjørelsene har jeg måttet ta for å begrense undersøkelsens omfang. Andre avgjørelser gjaldt valg av metode. En bakgrunn for drøftingen av funnene har vært beskrivelser av BKS og VIET (jf.

3.5 og 3.4). Det har likevel vært en ulempe at jeg ikke kan disse språkene, fordi jeg kan ha risikert å overse interessante trekk ved informantenes ortografi.

Om jeg skulle ha utvidet undersøkelsen eller bygd videre på den, ville jeg ha sammenlignet funnene mine med:

1. Ortografien i språkprøver på lavere eller høyere nivå
2. Ortografien til innlærere med andre morsmål

I punkt 1 hadde det vært interessant og sammenlignet feilfrekvens og type feil som gjøres. Jeg ville ha forventet færre feil av informanter som er på et høyere nivå enn på et lavere nivå. Det ville også vært interessant å undersøke om feilene jeg har tolket som mulig påvirkning fra morsmålet/morsmålsskriftspråket, hadde vært lavere i språkprøver på et høyere nivå – og om ortografien der hadde vært mer målspråkslik. Og tilsvarende om påvirkningen fra morsmålet/morsmålsskriftspråket i språkprøver på et lavere nivå ville vært høyere. I så fall ville det ha stemt med det Figueredo (2006) sier om at ortografisk påvirkning fra morsmålet skjer på et tidlig nivå i skrive læringen, når man ikke har tilegnet seg andrespråkets skriftkonvensjoner. Riktige skriftkonvensjoner tas etter hvert i bruk, ettersom man får bedre kompetanse i ortografien i andrespråket (Figueredo 2006:888, jf. Vigrestad 2015:54). I punkt 2 ville det vært interessant og sett om innlærere med andre morsmål ville gjort feil som ligner på feilene til VIET og/eller BKS, eller om de hadde gjort andre typer feil.

## 6.6 Oppsummering av kapitlet og oppgaven

I dette kapitlet har jeg drøftet funnene i undersøkelsen, samt undersøkelsen som helhet. I tillegg har jeg foreslått videre forskning. I tråd med det overordnede forskningsspørsmålet har hovedmålet vært å drøfte om forskjellene i ortografien til VIET og BKS kan relateres til morsmålet og/eller morsmålsskriftspråket. De største forskjellene mellom informantgruppene er at VIET gjør flere feil enn BKS ved grupper av konsonantgrafemer (forskningsspørsmål 1), BKS har flere tilfeller av enkelt for dobbelt konsonantgrafem enn VIET (forskningsspørsmål 4) og at BKS gjør flere substitusjonsfeil ved <y> og <ø> enn VIET (forskningsspørsmål 6). Slik jeg ser det, kan forskjellene mellom VIET og BKS relateres til påvirkning fra informantenes morsmål og morsmålsskriftspråk. Dette gjelder trekk som oppbygging av stavelser, grafotaks og skriftkonvensjoner.

I denne undersøkelsen har jeg tatt for meg to informantgrupper, VIET og BKS. Jeg har undersøkt ortografien til informantene, og sammenlignet de to gruppene. Jeg mener andrespråksortografi er et område som bør undersøkes nærmere da det tidligere har vært forsket svært lite på andrespråksortografi i Norge og Norden. I dette kapitlet har jeg presentert noen temaer jeg mener kan egne seg for videre forskning.

# Litteraturliste

- Apel, Kenn 2011: "What Is Orthographic Knowledge?". I *Language, Speech, and Hearing Services in Schools*. Vol. 42.
- Borgwaldt, Susanne R. & Terry Joyce 2013: *Typology of Writing Systems*. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Cook, Vivian 1997: "L2 Users and English Spelling". I *Journal of Multilingual and Multicultural Development*.
- Cook, Vivian 2008: *Second Language Learning and Language Teaching*. Hodder Education. London.
- Cook, Vivian & Benedetta Bassetti 2005: "An Introduction to Researching Second Language Writing Systems". I Cook, Vivian & Benedetta Bassetti: *Second Language Writing Systems*. Clevedon. Multilingual Matters.
- Coulmas, Florian 2003: *Writing Systems. An introduction to their linguistic analysis*. Cambridge. Cambridge University Press.
- Dich, Nadia & Bo Pedersen 2013: "Native Language Effects on Spelling in English as a Foreign Language: A Time-Course Analysis". *The Canadian Journal of Applied Linguistics*.
- Ellis, Rod & Gary Barkhuizen 2005: *Analysing Learner Language*. Oxford. Oxford University Press.
- Elsness, Turid Fosby 2002: *Vi skriver. Stavestrategier hos barn i alderen 7-8 år*. Oslo. Unipub forlag.
- Endresen, Rolf Theil 2000: "Språklydlære: fonetikk og fonologi". I Endresen, Rolf Theil, Hanne Gram Simonsen og Andreas Sveen (red.): *Innføring i lingvistikk*. Oslo. Universitetsforlaget.
- Figueredo, Lauren 2006: "Using the known to chart the unknown: A review of first-language influence on the development of English as a second language spelling skill". I *Reading and Writing*. Vol. 19. Nr. 8.
- Husby, Olaf 1991: *Vietnamesisk grammatikk. En innføring i vietnamesisk grammatikk sett med norske øyne*. Oslo. Friundervisningens Forlag.
- Husby, Olaf & Helene Kløve 1998: *Andrespråksfonologi. Teori og metodikk*. Oslo. Gyldendal Norsk Forlag.

- Hvenekilde, Anne 1987: "Feil i vokalgrafemene i norske stiler skrevet av elever med arabisk og berberisk som morsmål". I *NOA. Norsk som andrespråk*. Nr. 5.
- Hvenekilde, Anne 1995: "Riktige og gale ordbilder i stiler skrevet av tyrkiske og vietnamesiske elever". I Bjørkavåg, Lise Iversen; Anne Hvenekilde & Else Ryen (red.): *Men hva betyr det, lærer? Norsk som andrespråk - Fagdidaktisk bidrag*. Cappelen Forlag.
- Jarvis, Scott & Aneta Pavlenko 2008: *Crosslinguistic Influence in Language and Cognition*. New York/London. Routledge.
- Johannesen, Janne Bondi 2003: "Innsamling av språklige data: Informanter, introspeksjon og korpus". I (Red) Johannesen, Janne Bondi (red.): *På språkjakt. Problemer og utfordringer i språkvitenskapelig datainnsamling*. Oslo. Unipub.
- Jongejan, Wilma; Ludo Verhoeven & Linda S. Siegel 2007: "Predictors of Reading and Spelling Abilities in First- and Second-Language Learners". I *Journal of Educational Psychology*. Vol. 99. Nr. 4.
- Kodebok for ASK: <http://clu.uni.no/humfak/ask/brukerveiledning/feilkoder.htm#leksikalske>  
<http://clu.uni.no/humfak/ask/brukerveiledning/feilkoder.htm#INFL>
- Kulbrandstad, Lars Anders 2005: *Språkets mønstre. Grammatiske begreper og metoder*. Oslo. Universitetsforlaget.
- Lie, Svein 2010: "Standardtalemålet – fins det?". I *Språklig samling*. 51. årgang.
- Mønnesland, Svein 1990: "Serbokroatisk-norsk kontrastiv grammatikk". I Hvenekilde, Anne (red.): *Med to språk. Fem kontrastive språkstudier for lærere*. Oslo. Cappelen Forlag.
- Mønnesland, Svein 2002: *Bosnisk, kroatisk, serbisk grammatikk*. Oslo. Sypress Forlag.
- Rosén, Victoria 2001: *Vietnamesisk. En kontrastiv og typologisk introduksjon*. Trondheim. Tapir Akademisk Forlag.
- Sebba, Mark 2007: *Spelling and Society. The Culture and Politics of Orthography around the World*. Cambridge. Cambridge University Press.
- Stuart, Morag 1999: "Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners". I *British Journal of Educational Psychology*. Vol. 69. Nr. 4.
- Tenfjord, Kari, Jon Erik Hagen & Hilde Johansen 2009: "Norsk andrespråkskorpus (ASK) – design og metodiske forutsetninger". I *NOA. Norsk som andrespråk*. Årgang 25. Nr. 1.



- Vigrestad, Tone 2015: "Andrespråkslæring og ortografi". I Golden, Anne & Elisabeth Selj (red.): *Skrijving på norsk som andrespråk. Vurdering, opplæring og elevens stemmer*. Oslo. Cappelen Damm.
- Wade-Wooley, Lesly & Linda S. Siegel 1997: "The spelling performance of ESL and native speakers of English as a function of reading skill". I *Reading and Writing: An Interdisciplinary Journal*. Nederland. Vol. 9. Nr. 5-6.
- Wiggen, Geirr 1992: *Rettskrivingsstudier 2. Kvalitativ og kvantitativ analyse av rettskrivingsavvik hos østnorske barneskoleelever*. Avhandling for den filosofiske doktorgraden. Universitetet i Oslo.
- Wiggen, Geirr 1994: *Innledning*. I Wiggen, Geirr og Carsten Elbro (red.): *Språkvitenskapelige beskrivelsesmodeller i studiet av lesing og skrijving. Nordiske bidrag*. Oslo. Universitetet i Oslo, Senter for lærerutdanning og skoletjeneste.

# Vedlegg 1: Innlærertekster

Under følger to tekster skrevet av innlærere med henholdsvis vietnamesisk og BKS som morsmål. Tekstene er hentet fra ASK-korpuset, og er tatt med for å illustrere hvordan tekster i materialet jeg har brukt, kan se ut.

Den første teksten er skrevet av en kvinne på 34 år med vietnamesisk som morsmål. Hun oppgir å ha vært 1–2 år i Norge. Oppgaven lyder: Skriv en tekst om en bok du har lest.

Jeg har lest mange bøker, men jeg liker best en bok som heter " Nye mål 2 ". Jeg skal fortelle hvorfor jeg liker den boka, hva som skjer i den boka... Til slutt skal jeg si min konklusjon .

Den boka snakker om José. Han er fra Chile å bo alene i Norge. Han har gått på jobb, der kjente han Kåre. De pratet mye hverandre i pause i kantinner på fabriken . En dag inviterte Kåre José reiste til Bergen for å besøke familien til Kåre. José har sett mange pene naturlig i Norge feks fjelet var bratt ned til sjøen... Kåres søster som heter Marit. Hun var veldig pen, synes José.

En dag var José ulyke på jobben, han var nødt til å legge på sykehus. Den tiden savnet han mye om familien sin i Chile . Han husket til Marit også. Han fikk et brev fra Marit imen han var på sykehus. Han var så glad i henne. Marit fikk jobb i Oslo, må hun flytte til Oslo fra Bergen .

Hennes leilighet var umøbler . Hun kjøpte noe nye møbler og noe har brukt fra en dame som avert av moren sin. Kåre og José har hjulpet Marit å flytte inn. Etter lenge kjente José og Marit . De ble kjæresten. De inviterte familien til José kom til Norge for å José og Marit bryllupet . Etter bryllupet reiste de til Danmark.

Jeg likte den boka fordi teksten er lett å forstå , den boka har mange ordtakt også. Jeg har noen som lignet på José første tiden kom jeg til Norge, derfor likte jeg den boka.

Den neste teksten er skrevet av en mann på 37 år med BKS som morsmål. Han har vært i Norge i 2–3 år. Oppgaveteksten lyder: Skriv en tekst om å flytte.

Jeg flyttet mange ganger i mit liv. Derfor jeg liker ikke så mye dette ord . Når det oppstår en krig i et land man må regne med tvunget flytting.

Det vær en dårlig opplevelse for meg og min familie, da huset vår ble direkt truffet med noen granater. Heldigvis har vi gjemt under taket av kjeleren så vi hørte når taket av huset ble falt sammen. Etter et kort tid var det så stille sånn at vi kunne komme ut av ruiner . Sånn er livet , alt ble borte i bare noen minutter .

Hele familien min var hjemmeløs i noen år i Bosnia og de måtte finne sted å bo. I dag bor de på forskjellige steder i mange land. Jeg har ikke sett min bror f.eks. på 8 år . Vi har forhold bare gjennom telefon og vi vet ikke når skal det samles hele fammilien i hjemmested vårt . Min nære fammilien bor sammen med meg i Norge nå. Vi bosette oss i en deilig landsby og vi tryves veldig godt her.

Men, hver flyktninger som ble tvunget å forlate sitt hjem lengter etter det. Følelsene er noen som ikke kan man kontrollere spesiell om morges når sola skinner føler man seg litt depriment.

Sølfelgelig man kan ikke ha alt hva forventer av livet. Heldigvis jeg og kona mi har to barn som tryves veldig godt her i Norge. De fikk mange norske venner til å leke på skolen og fritiden. De snakker flitende norsk nå og jeg håper at de skal ikke ha vanskelig som vi i framtiden deres .

## Vedlegg 2: Søk i ASK

For å finne de ortografiske feilene i ASK-korpuset søkte jeg på ord som er annotert med ortografiske feil (se figuren under). Jeg søkte på én morsmålsgruppe av gangen (i første linje, til høyre for "Morsmål ="), og spesifiserte at det var ortografiske feil fra *Språkprøven* jeg ville se på (i andre linje, til høyre for "Språktest="). *ORT* (=ortografiske feil) ble deretter valgt for hva slags feiltype jeg ville undersøke.

Her kan du lage et søk ved å bruke menyer.

Velg et **delkorpus**:

<b>Morsmål</b> = "vietnamesisk"
<b>Språktest</b> = "Språkprøven"
legg til:
attributt: -

Velg **posisjonelle betingelser**. |  Ignorer strukturelle posisjoner

<input type="checkbox"/> <i>target</i>
<b>Ord</b> = <input type="text"/> <input type="checkbox"/> %c
<b>Feiltype</b> = "ORT"
gjentakelse: 1
legg til:
attributt: -
struct: -

Søk

Figur 16: Søk i ASK

## Vedlegg 3: Kategorisering av feilene

For å komme fram til resultatene i undersøkelsen kategoriserte jeg de ortografiske feilene i en kategoriindeling, beskrevet i 4.4. For en fullstendig oversikt over alle kategoriseringene av de ortografiske feilene til VIET og BKS gå til denne nettsida:

<https://www.dropbox.com/sh/hv2e50mnrzdzojr/AABY0zHx1tP8ejdx8GiD3f1Za?dl=0>

Jeg vil på de neste fire sidene ta med et utdrag fra denne oversikten. På grunn av begrenset plass her, har jeg kuttet bort det meste av konteksten. Under arbeidet med kategoriseringen har jeg hatt mye større kontekst. Det har vært nyttig de gangene jeg har vært i tvil om hvilket ord som menes, eller hvilken form av ordet det skal rettes til. I konteksten rundt ordet med feil vil man noen ganger finne en @. Symbolet er satt inn i stedet for navn på personer, byer, steder eller lignende for å anonymisere innholdet.

I oversikten er feil rettet på den måten at i kolonnen som heter "Hvor i ordet er feilen(e)", er understrek brukt for å markere utelatt grafem – ved innskutt grafem, substitusjon og ombytting av grafemer er fet skrift brukt. Ved ord som har flere feil, markeres én av gangen. Ved retting av for eksempel "ulike" for *ulykke*, vises først *ulike*, og så forklares det i kolonnen "Grafem(er) feilen gjelder" at feilen er at <y> er byttet ut med <i>. Deretter, på linja under, vises den andre feilen i ordet, *ulik\_e* – og det presiseres at feilen er at en <k> er blitt utelatt.

Som nevnt i 4.2.5 har jeg tatt ut noen ord som er annotert med ortografiske feil i ASK-korpuset av forskjellige grunner. I kolonnen "Grunn til at ordet er tatt ut" har jeg tallkoder for grunner til at ordet er tatt ut. Her er forklaring på disse tallkodene:

- 1: Sifre for tallord – f.eks. "7" for *sju*, "13" for *tretten* og "2" for *to*.
- 2: Egennavn – f.eks. "Norger" for *Norge*, "Osebergg" for *Oseberg* og "Tvaita" for *Tveita*.
- 3: Minuskel for majuskel/majuskel for minuskel – f.eks. "Nordmann" for *nordmann*, "Juni" for *juni* og "Pappa" for *pappa*.
- 4: Ordet er forkortet på feil måte – f.eks. "opp." for *oppg.*, "osv" for *osv.* og "ad." for *adr.*
- 5: Feil i diakritiske tegn/bindestrek/apostrof – f.eks. "mannen's" for *mannens*, "idè" for *idé* og "bok-hylle" for *bokhylle*.
- 6: Feilannoterte ord og ord som ikke har ortografisk feil etter min definisjon (men som ikke innbefattes av de andre punktene). For eksempel er ordene "flere" og "mere" annotert som ord med ortografiske feil; jeg velger å ta disse ut da stavemåten finnes i dagens rettskriving. Også ord som ikke eksisterer i norsk, er tatt ut. For eksempel er "offise" rettet til *office* – dette ordet har jeg tatt ut. Ord som "lobor" for *laboratorium* er tatt ut da jeg ikke ser dette som en ortografisk feil.

Det er to informanter i BKS og fem informanter i VIET som ikke har noen ortografiske feil. Dette er informantene 0780 og 0789 fra BKS, og 0660, 0712, 0849, 0857 og 0896 fra VIET.

## Oversikt over ortografiske feil hos informanter med vietnamesisk som morsmål

Min nummerering	Personens nummer i ASK	Ord med ortografisk feil i kontekst	Grunn til at ordet er tatt ut	Rettet til	Hvor i ordet er feilen(e)?	Antall feil	Grafem(er) feilen gjelder	1 Utelatt konsonantgrafem	1a Enkelt for dobbelt	1b I gruppe av konsonantgrafemer	1c Ved /t/, /d/, /l/, /n/, /j/, /s/ eller /ŋ/	1d Mellom to vokalgrafemer	2 Utelatt vokalgrafem	2a Monoftongering av diftong	2b Mellom to konsonantgrafemer	3 Innskudd av konsonantgrafem	3a Dobbelt for enkelt	3b I gruppe av konsonantgrafemer	3c Ved /t/, /d/, /l/, /n/, /j/, /s/ eller /ŋ/	3d Ved siden av enkelt konsonantgrafem	4 Innskudd av vokal	4a Diftongering av monoftong	4b I gruppe av konsonantgrafemer	5 Substitusjon av konsonantgrafem	5a Enkelt konsonantgrafem	5b I gruppe av konsonantgrafemer	5c Ved /t/, /d/, /l/, /n/, /j/, /s/ eller /ŋ/	6 Substitusjon av vokalgrafem	6a Enkelt vokalgrafem	5b I diftong	7 Ombytting av grafemer								
	s0 00	hensyn til <u>informasjon</u>	informasjon	informasjon	imformasjon	1	m for																																
	1 5	På en <u>anne</u> måte	annen	anne_		1	n x																x		x														
	s0 00	<u>muligheter</u> til barn	muligheter	muli_heter		1	g x			x																													
	s0 00	helt forskjell <u>kultur</u>	kultur	kuntur		1	n for																																
	3 8	Hva er bra <u>tradisjoner</u> ?	tradisjoner	tradis_oner		1	j x			x	x													x		x													
		For <u>eksempel</u> når foreldre <u>elle</u> de andre	eksempel	exsempel		1	x for ks																	x		x													
		noen de <u>andre</u> ute oppdraelse fra foreldre	andre	endre		1	r e for a																																
		<u>ute oppdraelse</u> fra foreldre	uten oppdraelse	ute_oppdraelse		1	n x																																
		<u>allerede</u> s med andre land.	allerede	alledes		1	g x					x																											
	s0 00	Det er <u>kokurranser</u> overalt.	kokurranser	ko_kurranser		1	n x				x	x																											
		Det kan bli <u>spennede</u> hvor de som delta i <u>kokurranser</u>	spennede	spenne_de		1	n x				x																												
	s0 01	<u>t-v</u> , cd spiller, bil...	nærheten	nærheten		1	n e for a																																
	5 6	være å <u>nærheten</u> av butikker, <u>restauranter</u> , sentrum	restauranter	restauranter		1	y for a																																
		<u>lyft</u> av kjærligheten hjelpsomme <u>man</u> omkring oss.	luft	lyft		1	u for a																																
			menn	man		2	e for																																
	s0 09	nesten <u>6</u> år lang vei fra <u>hjemme</u> vårt. <u>Dessuten</u> boligen min måtte ha <u>to</u> rom for <u>5</u> mennesker. i gammelt hus i <u>6</u> år.	hjemmet	hjemmet		1	t x																																
			Dessuten	Dessuten		1	t x																																
	s0 09	Noen familier på <u>3</u> personer				1																																	

Min nummerering	Personens nummer i ASK	Ord med ortografisk feil i kontekst	Grunn til at ordet er tatt ut	Rettet til	Hvor i ordet er feilen(e)?	Antall feil	Grafem(er) feilen gjelder	1 Utelatt konsonantgrafem	1a Enkelt for dobbelt	1b I gruppe av konsonantgrafemer	1c Ved /t/, /d/, /l/, /n/, /j/, /ç/ eller /ŋ/	1d Mellom to vokalgrafemer	2 Utelatt vokalgrafem	2a Monoftongering av diftong	2b Mellom to konsonantgrafemer	3 Innskudd av konsonantgrafem	3a Dobbelt for enkelt	3b I gruppe av konsonantgrafemer	3c Ved /t/, /d/, /l/, /n/, /j/, /ç/ eller /ŋ/	3d Ved siden av enkelt konsonantgrafem	4 Innskudd av vokal	4a Diftongering av monoftong	4b I gruppe av konsonantgrafemer	5 Substitusjon av konsonantgrafem	5a Enkelt konsonantgrafem	5b I gruppe av konsonantgrafemer	5c Ved /t/, /d/, /l/, /n/, /j/, /ç/ eller /ŋ/	6 Substitusjon av vokalgrafem	6a Enkelt vokalgrafem	5b I diftong	7 Ombytting av grafemer					
2		De prøver å <u>lette</u> etter måtte jeg hverdag <u>lesse</u> @avis	lete	lette	1 t											x	x																			
		<u>sikke</u> på flere leiligheter	kikke	sikke	1 k																		x	x												
		<u>ganske mange unnbygere</u>	innbygere	unnbygere	2 i																									x	x					
		<u>for å for få informasjonen</u>	informasjon	informasjon	1 n											x	x																			
8	4	har masse <u>trær</u>	trær	trær	1 n											x																				
9	5	nær <u>skoge</u>	skog	skog	1 n	x																														
		<u>nær</u> barnehage	nær	nær	1 æ																											x	x			
		jeg bor <u>nær</u> jeg og min <u>man</u>	nær	man	1 n	x	x																									x	x			
		Forresten liker jeg å <u>pinte eke</u> trær og blomster i <u>ballkong</u>	pynte	pinte	1 y																															
		<u>ekte</u> blomster i <u>ballkong</u>	ekte	ekte	1 t	x	x																													
1	13	i kantinner på <u>fabriken</u>	fabrik	fabrik	1 k	x	x																													
		<u>feks</u> fjelet var bratt																																		
		En dag var José <u>ulyke</u> på jobben Hennes leilighet var <u>umøbler</u>	ulykke	ulykke	1 k	x	x																													
		boka har mange <u>ordtak</u> også	ordtak	ordtak	1 t											x																				
		pause i <u>kantinner</u> på fabriken	kantinner	kantinner	1 n											x	x																			
		<u>fjelet</u> var bratt ned til sjøen... <u>imen</u> han var på sykehus	fjellet	fjellet	1 l	x	x																													
		<u>avert</u> av moren sin	imens	imens	1 s	x	x																													
		sin	arvet	avert	2 r	x	x																													
1	14	Vi <u>leidde</u> en luke	leide	leidde	1 d											x	x																			
		Vi satte alle <u>sjeke</u>	sekke	sjeke	1 j											x																				
		da hadde en <u>kløv</u>	klovn	kløv	2 o																															
				kløv	2 n	x	x																													

Oversikt over ortografiske feil hos informanter med BKS som morsmål

Min nummerering	Personens nummer i ASK	Ord med ortografisk feil i kontekst	Grunn til at ordet er tatt ut	Rettet til	Hvor i ordet er feilen(e)	Antall feil	Grafem(er) feilen gjelder	1 Utelatt konsonantgrafem	1a Enkelt for dobbelt	1b I gruppe av konsonantgrafemer	1c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	1d Mellom to vokalfrafemer	2 Utelatt vokalfrafem	2a Monofongering av diftong	2b Mellom to konsonantgrafemer	3 Innskudd av konsonantgrafem	3a Dobbelt for enkelt	3b I gruppe av konsonantgrafemer	3c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	3d Ved siden av enkelt konsonantgrafem	4 Innskudd av vokalfrafem	4a Diftongering av monofong	4b I gruppe av konsonantgrafemer	5 Substitusjon av konsonantgrafem	5a Enkelt konsonantgrafem	5b I gruppe av konsonantgrafemer	5c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	6 Substitusjon av vokalfrafem	6a Enkelt vokalfrafem	b I diftong	7 Ombytting					
1	1	Jeg kunne <u>sit</u> e og se flere timer		sitte	sit_e	1	t	x	x																											
		Jeg skulle sitte i <u>salonen</u>		salonge	salon_e	1	g	x		x	x																									
		Nå er det blitt <u>forskjellige</u>		forskjellige	forskjilleg	2	e																					x	x							
					forskjilleg	1	l	x	x																											
2	2	han må gå på en <u>skol</u> .		skole	skol_	1	e						x																							
		Det var litt <u>letere</u> for meg		lettere	let_ere	1	t	x	x																											
		hadde <u>problemer</u> av og til.		problem	problem	1	m								x	x																				
		har vært helt <u>anerledes</u> enn å kjøre forbi andre store		annerledes	an_erledes	1	n	x	x																											
		<u>kjøretøy</u> .		kjøretøy	kjøretøy	1	y																						x	x						
		Den jobb kan være veldig <u>tøf</u> .		tøff	tøf_	1	f	x	x																											
		Den jobb kan være veldig <u>tøf</u> .		være	være	1	æ																													
		kanskje litt <u>stresende</u> .		stresende	stres_ende	1	s	x	x																											
		vanskeligst for <u>nybeggin</u> te sjåfør		nybeggin	nybeggin	2	g									x	x																			
					nybeggin	1	y																													
		Min første <u>sjåførs</u> jobb det er helt <u>spesiel</u>	6	spesiel	spesiel_	1	t	x		x																										
		i god tid tilbake om <u>morg</u> er tidlig		morgen	morg	1	r																													
		Jeg har hatt i <u>begyn</u> nelse som har sett i <u>begyn</u> nelse ut.		begyn	begyn_e	1	n	x	x																											
				begyn	begyn	2	g								x	x																				
					begyn_	1	n	x	x																											
3	3	<u>angasier</u> er en person fisisk		engasjer	angasjer	1	e																													
		<u>fisisk</u> og <u>psikisk</u> veldig mye.		fysisk	fisisk	1	y																													
		<u>fisisk</u> og <u>psikisk</u> veldig mye.		psykisk	psikisk	1	y																													
		arbeider i <u>skif</u> te, har også <u>familie</u> og barna.		skift	skifte	1	e																													
		må de ofte		familie	familie	1	o														x															
				glemme	glem_e	1	m	x	x																											

Min nummerering	Personens nummer i ASK	Ord med ortografisk feil i kontekst	Grunn til at ordet er tatt ut	Rettet til	Hvor i ordet er feilen(e)	Antall feil	Grafem(er) feilen gjelder	1 Utelatt konsonantgrafem	1a Enkelt for dobbelt	1b I gruppe av konsonantgrafermer	1c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	1d Mellom to vokalfgrafermer	2 Utelatt vokalfgrafem	2a Monoftongering av difftong	2b Mellom to konsonantgrafermer	3 Innskudd av konsonantgrafem	3a Dobbelt for enkelt	3b I gruppe av konsonantgrafermer	3c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	3d Ved siden av enkelt konsonantgrafem	4 Innskudd av vokalfgrafem	4a Diftongering av monoftong	4b I gruppe av konsonantgrafermer	5 Substitusjon av konsonantgrafem	5a Enkelt konsonantgrafem	5b I gruppe av konsonantgrafermer	5c Ved /t/, /d/, /l/, /n/, /r/, /s/ eller /n/	6 Substitusjon av vokalfgrafem	6a Enkelt vokalfgrafem	6b Idiftong	7 Ombyting					
		<u>gleme</u> egne problemer. de ofte <u>gleme</u> egne problemer. forskjellige <u>helsepersonale</u> som arbeider i <u>lobor</u> , <u>rentgen</u> lobor, <u>rentgen</u> o.s.v. lobor, <u>rentgen</u> o.s.v.																																		
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .	6		egne helsepersonale	1	e													x	x															
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		rentgen	rentgen	1	ø																					x	x							
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .	4	stress	stres_	1	s	x	x																											
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		lønn	lon	2	ø																					x	x							
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		den	denn	1	n								x	x																				
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .	4	forskjellige	forskjellige	1	l	x	x																											
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		veldig forskjellig	veldig forskjellig	1	g	x																												
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		veldig forskjellig	forsjellig	2	k	x	x	x																										
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		veldig forskjellig	forsjellig	1	l	x	x																											
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		andre forskjellige	andere	1	e														x	x														
		Det er også mange <u>stres</u> sykepleiere ikke god <u>lon</u> .		bra organisert egenes <u>levet</u> .	livet	1	e																					x	x							
4	03	flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.	4	utdanne	utdan_e	1	n	x	x																											
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.	1	medlem	medlem_er	1	m	x	x																											
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.		av det som far <u>kjente</u> og <u>bestillte</u> 5 flybilleter.	tjente	1	t															x			x											
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.	1	bestilte	bestilte	1	l								x	x																				
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.		flybillett	flybillett_er	1	t	x	x																											
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.		opphold	opphold_tillatelse	1	s	x	x																											
		flyttet jeg ca 8 mil fra mitt <u>utdanet</u> @yrke. stor familie på 6 medlemmer.		fredelig	fred_lig	1	e					x	x																							



# Vedlegg 4: Utregning av statistisk signifikans

Som beskrevet i kapittel 5 viser funnene på tre forskjellige områder betydelige forskjeller mellom VIET og BKS. For å finne ut om forskjellene er statistisk signifikante, har jeg fått hjelp av Bård Uri Jensen. Nedenfor følger en oversikt over hans utregninger for de tre områdene. På de to neste sidene følger programkoder for disse utregningene.

## 1. Forskjell i feilfrekvens mellom BKS og VIET.

VIET har en feilfrekvens på 2,4 % per løpeord, mens BKS har en feilfrekvens på 3,8 %. Utregningen tar sikte på å finne ut om forskjellen mellom de to informantgruppene er statistisk signifikant. Tallene er regnet ut ved å dividere antall feil på antall løpeord. Testen som er brukt for å regne ut dette, er den ikke-parametriske testen Wilcoxon. Testen viser at forskjellen mellom gruppene er statistisk signifikant:  $W=6542$ ,  $p<0.001$ . Effekttørrelse: Cohens  $d\approx 0.55$ . Tallene under viser spredning av verdiene innenfor hver informantgruppe:

	min.	median	middelverdi	maks.	standardavvik
<b>BKS</b>	0	0.030	0.038	0.115	0.027
<b>VIET</b>	0	0.020	0.024	0.113	0.022

Tabell 35: Forskjell i feilfrekvens av løpeord mellom VIET og BKS – spredning av verdier.

## 2. Forskjell i enkelt for dobbelt konsonantgrafem mellom BKS og VIET

BKS har 1,34% tilfeller av enkelt for dobbelt konsonantgrafem, mens VIET har 0,41%. Wilcoxon-testen viser at forskjellen mellom informantgruppene her er statistisk signifikant:  $W=7891$ ,  $p<0.001$  og Cohens  $d\approx 1.11$ . Det er likevel stor forskjell innad i hver informantgruppe på hvor mange slike feil som gjøres, som tabell x viser:

Antall feil	BKS	VIET
0	9	44
1	19	33
2	15	8
3	17	7
4	11	4
5	7	2
6	5	1
7	5	0
9	3	0
10	4	0
11	2	0
12	2	0

Tabell 36: Antall personer med feil ved enkelt for dobbelt konsonantgrafem hos VIET og BKS.

## 3. Forskjell i feilfrekvens ved grupper av konsonantgrafemer mellom BKS og VIET

VIET har en feilfrekvens på 0,81% per løpeord ved grupper av konsonantgrafemer, mens BKS har en feilfrekvens på 0,44%. Wilcoxon-testen gir et signifikant resultat på forskjellen mellom gruppene her:  $W=3379$ ,  $p<0.001$  og Cohens  $d\approx 0.49$

	Antall null feil	Min.	Median	Middelverdi	Maks.	Standard-avvik
<b>BKS</b>	36	0	0.00355	0.00476	0.0338	0.00593
<b>VIET</b>	22	0	0.00621	0.00842	0.0614	0.00877

Tabell 37: Feilfrekvens ved grupper av konsonantgrafemer – spredning av verdier.

## Programkoder for utregningene:

```
#####
## Noen utregninger for Tone
## 2016-04-20 BUJ
## Bruker to filer toneA.txt og toneB.txt

#####
## Noen parametre til diagrammer

fillBKS <- "red"
fillVIE <- "dark grey"
FILL <- c(fillBKS, fillVIE)
lineBKS <- "red"
lineVIE <- "black"

#####
## Funksjon som regner ut Cohens d for to utvalg.
## En opsjon (m=F) regner ut en median-basert verdi av d i stedet
## den vanlige middelverdi-baserte.

cohens.d <- function (s1, s2, m=T){
  n1 <- length(s1)
  n2 <- length(s2)
  cohens.s <- sqrt(((n1-1)*var(s1,na.rm=T)+(n2-1)*var(s2,na.rm=T)) / (n1+n2))
  cohens.d <- (mean(s1,na.rm=T) - mean(s2,na.rm=T))/cohens.s
  cohens.d.median <- (median(s1,na.rm=T) - median(s2,na.rm=T))/cohens.s
  ifelse(m, return(cohens.d), return(cohens.d.median))
}

#####
## Leser inn data
## Plasserer dem i tabeller toneA, toneB

toneA <- read.table("C:/Users/bardurij/Documents/0Veiledning/MultiLing/Tone/toneA.txt", header=T, sep="\t")
str(toneA) #
head(toneA)
TONEA <- toneA

toneB <- read.table("C:/Users/bardurij/Documents/0Veiledning/MultiLing/Tone/toneB.txt", header=T, sep="\t")
str(toneB) #
head(toneB)
TONEB <- toneB

toneA$feilF <- toneA$Feil / toneA$nOrd
toneA$feilF <- log(toneA$feilF) ## Gir feil. Må justere for 0-verdier

sum(toneA$plD==toneB$plD) == dim(toneA)[1] # Should be true
sum(toneA$nOrd==toneB$nOrd) == dim(toneA)[1] # Should be true
toneA$nEnk_Dob <- toneB$nEnk_Dob
toneA$nDob_Enk <- toneB$nDob_Enk
toneA$nDobfeil <- toneA$nEnk_Dob + toneA$nDob_Enk
toneA$DobfeilF <- toneA$nDobfeil / toneA$nOrd
toneA$Enk_DobF <- toneA$nEnk_Dob / toneA$nOrd
toneA$Dob_EnkF <- toneA$nDob_Enk / toneA$nOrd
toneA$DobfeilDiff <- toneA$Enk_DobF - toneA$Dob_EnkF

tBKS <- toneA$L1=="BKS"
tVIE <- toneA$L1=="VIE"

#####
## Kapittel 1: Forskjell i feilfrekvens mellom VIET og BKS

summary(toneA$feilF[tBKS]); sd(toneA$feilF[tBKS])
summary(toneA$feilF[tVIE]); sd(toneA$feilF[tVIE])
sum(toneA$nFeil[tBKS]==0)
sum(toneA$nFeil[tVIE]==0)

par(mfrow=c(1,3))
qqnorm(toneA$feilF, main="Total")
qqnorm(toneA$feilF[tBKS], main="BKS")
qqnorm(toneA$feilF[tVIE], main="VIE")
par(mfrow=c(1,1))

plot(density(toneA$feilF[tVIE]), main="Feilfrekvens")
lines(density(toneA$feilF[tBKS]), col="red")
legend("topright", fill=c("red", "black"), c("BKS", "VIET"))
```

```

boxplot(toneA$feilF~toneA$L1, col=c(fillBKS, fill))

wilcox.test(toneA$feilF[tBKS], toneA$feilF[tVIE])
cohens.d(toneA$feilF[tBKS], toneA$feilF[tVIE])

#####
## Kapittel 2: Flere enkelt_for_dobbelt enn motsatt? (Hele utvalget.)

plot(density(toneA$Dob_EnkF), col="red", main="")
lines(density(toneA$Enk_DobF))
legend("topright", fill=c("black", "red"), c("Enkelt for dobbelt", "Dobbelt for enkelt"))

plot(density(toneA$DobfeilDiff), col="grey", lwd=2, main="")
abline(v=0, lty=2, lwd=2, col="grey")

shapiro.test(toneA$DobfeilDiff)

wilcox.test(toneA$Enk_DobF, toneA$Dob_EnkF, paired=T)
wilcox.test(toneA$DobfeilDiff) # Skal gi akkurat samme resultat
cohens.d(toneA$Enk_DobF, toneA$Dob_EnkF)

wilcox.test(toneA$Enk_DobF[tBKS], toneA$Dob_EnkF[tBKS], paired=T)
cohens.d(toneA$Enk_DobF[tBKS], toneA$Dob_EnkF[tBKS])

wilcox.test(toneA$Enk_DobF[tVIE], toneA$Dob_EnkF[tVIE], paired=T)
cohens.d(toneA$Enk_DobF[tVIE], toneA$Dob_EnkF[tVIE])

par(mfrow=c(1,2))
boxplot(toneA$Enk_DobF[tBKS], toneA$Dob_EnkF[tBKS], main="BKS", col=fillBKS, names=c("Enk_Dob", "Dob_Enk"))
boxplot(toneA$Enk_DobF[tVIE], toneA$Dob_EnkF[tVIE], main="VIET", col=fillVIE, names=c("Enk_Dob", "Dob_Enk"))
par(mfrow=c(1,1))

#####
## Kapittel 3: Har BKS flere Enk_Dob-feil enn VIET?

plot(density(toneA$Enk_DobF[tVIE]), main="", col=lineVIE)
lines(density(toneA$Enk_DobF[tBKS]), main="", col=lineBKS)
legend("topright", fill=c(lineVIE, lineBKS), c("VIET", "BKS"))

wilcox.test(toneA$Enk_DobF~toneA$L1)
cohens.d(toneA$Enk_DobF[tBKS], toneA$Enk_DobF[tVIE])

boxplot(toneA$Enk_DobF~toneA$L1, col=c(fillBKS, fillVIE))

table(toneA$Enk_Dob, toneA$L1)

## Lagt til 2016-05-02
#####
## Kapittel 4
## Hypotese 1
##
toneC$ngrfeil <- rowSums(toneC[,c("grf1", "grf2", "grf3", "grf4", "grf5")])

## Rekkefølgen i pID er ikke den samme i toneA og toneC.
## Derfor går jeg gjennom tabellen linje for linje
for (i in toneA$pID){
  toneA$ngrfeil[toneA$pID==i] <- toneC$ngrfeil[toneC$pID==i]
  toneA$nOrdC[toneA$pID==i] <- toneC$nOrd[toneC$pID==i]
}
sum(toneA$nOrd==toneA$nOrdC) == length(toneA$pID) # Should be TRUE
toneA$ngrfeilF <- toneA$ngrfeil / toneA$nOrd
sum(toneA$ngrfeil==0) # Antall nullverdier er 58 av 198

sum(toneA$ngrfeil[tBKS]==0)
sum(toneA$ngrfeil[tVIE]==0)
summary(toneA$ngrfeilF[tBKS]); sd(toneA$ngrfeilF[tBKS])
summary(toneA$ngrfeilF[tVIE]); sd(toneA$ngrfeilF[tVIE])

plot(density(toneA$ngrfeilF)) # Tydelig høyreskjev
plot(density(toneA$ngrfeilF[toneA$L1=="BKS"]), col=lineBKS, main="")
lines(density(toneA$ngrfeilF[toneA$L1=="VIE"]), col=lineVIE)
legend("topright", fill=c("red", "black"), c("BKS", "VIET"))

boxplot(toneA$ngrfeilF~toneA$L1, col=FILL)

wilcox.test(toneA$ngrfeilF~toneA$L1)
cohens.d(toneA$ngrfeilF[tBKS], toneA$ngrfeilF[tVIE])

```

## Vedlegg 5: Fordeling av antall feil

Det er forskjeller i antall feil som gjøres, både innad i hver informantgruppe og mellom informantgruppene (jf. tabell 12). I VIET er det for eksempel fem informanter som ikke gjør feil; hos BKS er det to. Og mens det i VIET er sju personer som gjør 16 feil eller mer, er det atten personer som gjør 16 feil eller mer i BKS. Det er likevel viktig å ha i mente at informantene i BKS i snitt har 18 flere ord per tekst enn VIET, noe som ikke gjør tallene helt sammenlignbare.

Antall feil	Vietnamesisk (antall personer)	BKS (antall personer)
Ingen	5	2
1-5	52	31
6-10	23	27
11-15	12	21
16-20	3	6
21-25	2	5
26-30	2	6
31-35	0	1

Tabell 38: Feilfordeling i de to morsmålsgruppene.

## Vedlegg 6: Vokalgrafemfeil

Både antall og type feil av innskutte og utelatte vokalgrafemer er ganske likt hos VIET og BKS (jf. henholdsvis tabell 13 og 14). Begge informantgruppene gjør flest feil ved ⟨e⟩ – dette gjelder både innskutte og utelatte grafemer – det gjøres også noen feil ved ⟨i⟩.

Av innskutte og utelatte ⟨e⟩ gjøres disse feiltypene enten medialt eller finalt i ord (jf. tabell 15 for innskutte ⟨e⟩ og tabell 16 for utelatte ⟨e⟩). Oversiktene viser eksempler på slike feil.

Grafem	a	e	i	o	u	y	æ	ø	å
<b>BKS</b>	1	29	4	1	1	0	0	0	0
<b>VIET</b>	0	28	3	0	1	1	0	0	0

Tabell 39: Utelatte vokalgrafemer.

Grafem	a	e	i	o	u	y	æ	ø	å
<b>BKS</b>	2	22	3	1	1	0	0	0	0
<b>VIET</b>	0	20	3	0	1	1	0	0	0

Tabell 40: Innskutte vokalgrafemer.

Medialt		Finalt	
BKS	V	BKS	V
egene for egne	Vietnameserene for vietnameserne	skifte for skift (subst.)	sucsesse for suksess
andere for andre	voksene for voksne	forstode for forstod	kigarette for sigarett
hjemeland for hjemland	klærene for klærne	vaere for hver	samarbeide for samarbeid
viktigeste for viktigste	foreldere for foreldre	frekvensie for frekvens	presse for press
baren for barn	personelig for personlig		tide for tid
seer for ser	inforemere for informere		inntrykke for inntrykk
	fottballekamp for fotballkamp		live for liv
			folke for folk
			kafee for kaffe
			kontakte for kontakt
			mosjone for mosjon

Tabell 41: Innskutt ⟨e⟩ medialt og finalt.

Medialt		Finalt	
BKS	V	BKS	V
fredlig for fredelig	jerbarnstasjon for jernbanestasjon	skol for skol	høyskol for høyskole
kjentegner for kjennetegner	genrasjon for generasjon	helsepersonal for helsepersonale	yrkerskol for yrkesskole
vidergående for videregående	morn for morgen	plant for plante	hjørn for hjørne
likvel for likevel	inspirete for inspirerte	direkt for direkte	famili for familie
hemmlig for hemmelig		T-ban for t-bane	list for liste
plutslig for plutselig			direkt for direkte
vansklige for vanskelige			tank for tanke
menneskhet for menneskehet			grupp for gruppe
vanskligere for vanskeligere			
familine for familiene			

Tabell 42: Utelatt ⟨e⟩ medialt og finalt.

Her gis en fullstendig oversikt over substitusjonsfeil av vokalgrafemer hos VIET og BKS – jf. tabell 17 og 18. Tabellene leses på samme måte. Jeg vil forklare hvordan tabellene leses ved å illustrere med eksempler fra tabell 17, tabellen som viser forekomster av substitusjonsfeil av vokalgrafemer hos BKS. Øverste vannrette rekke av vokalgrafemer viser grafemet som skulle vært valgt. Den loddrette rekke over vokalgrafemer til venstre viser det valgte grafemet. Stavefeilen ⟨e⟩ for ⟨i⟩ i for eksempel "lett" for *litt* vil derfor bli plassert der den vannrette e-aksen møter den loddrette i-aksen. På den plassen står det et 2-tall, som betyr at feilen ⟨e⟩ for ⟨i⟩ har to forekomster. Videre har for eksempel ⟨a⟩ for ⟨o⟩ 6 forekomster, og ⟨æ⟩ for ⟨e⟩ har 3 forekomster.

I enden av den loddrette og vannrette rekke av grafemer står det "Sum". Tallene i hver rekke og kolonne er altså summert, slik at man kan se hvor mange forekomster det er av hvert tilfelle. De summerte tallene i rekke nederst viser hvor mange ganger de korrekte grafemene er blitt byttet ut. For eksempel har ⟨e⟩ blitt erstattet med et annet grafem 35 ganger, mens ⟨u⟩ bare har blitt erstattet 5 ganger av et annet grafem. Den summerte rekke loddrett til høyre viser hvilket grafem som er valgt i hvert tilfelle. Grafemet ⟨a⟩ er valgt for et riktig grafem 62 ganger, mens ⟨æ⟩ bare har erstattet et riktig grafem 6 ganger.

Vannrett: riktige grafemer. Loddrett: valgte grafemer	a	e	i	o	u	y	æ	ø	å	j	S u m
a		26		6			24		6		62
e	1		2				6	2			11
i		4				14		1			19
o	15				1			28	27		71
u				3		4					7
y			11		3						14
æ	3	3									6
ø		2		21	1				4		28
å	1			17				3			21
j						1					1
Sum	20	35	13	47	5	19	30	34	37	0	

Tabell 43: Substitusjon av vokalgrafemer hos BKS.

Vannrett: riktige grafemer. Loddrett: valgte grafemer.	a	e	i	o	u	y	æ	ø	å	j	Sum
a		8			1		2				11
e	5		2				1	1			9
i		2		1	1	7					11
o	1	1			1			1			4
u			1			1					2
y		2	2		1						5
æ		1									1
ø				7	1						8
å				1							1
j			1								1
Sum	6	14	6	9	5	8	3	2	0	0	

Tabell 44: Substitusjon av vokalgrafemer hos VIET.

## Vedlegg 7: Konsonantgrafemfeil

Som beskrevet i 5.5.1 gjør VIET flere feil enn BKS ved grupper av konsonantgrafemer. Det er likevel store forskjeller innad i både VIET og BKS på hvor mange informanter som gjør feil ved grupper av konsonantgrafemer; mange informanter gjør ikke feil i det hele tatt her – jf. tabell 19.

"Utelatt konsonantgrafem ved gruppe av konsonantgrafem" er den av underkategoriene av feil ved grupper av konsonantgrafemer som har flest forekomster (jf. figur 7). De fleste av informantene har ett tilfelle av et bestemt utelatt grafem, mens andre har to eller tre utelatelinger av det samme grafemet – for detaljer se tabell 20.

I kategorien "substitusjon av konsonantgrafem i gruppe av konsonantgrafemer" gjøres substitusjonene mellom mange ulike grafemer, få substitusjoner gjøres mellom de samme grafemene, jf. tabell 21 og 22. Unntaket er "x" for *ks* – BKS har 5 forekomster av denne ombyttingen, mens VIET har 6. I tabell 27 og 28 finnes en fullstendig oversikt over alle substitusjonene av konsonantgrafemer, ikke bare de som gjøres i gruppe av konsonantgrafemer (tabellene leses på samme måte som tabell 17 og 18).

Ved "enkelt for dobbelt konsonantgrafem" og "dobbelt for enkelt konsonantgrafem" har VIET en feiltipe som ikke BKS har, de har flere tilfeller av to ulike konsonantgrafemer for dobbelt konsonantgrafem og dobbelt konsonantgrafem for to ulike konsonantgrafemer – jf. tabell 23.

"Innskutte grafemer som gjengir /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/" og "substitusjon av grafemer som gjengir /t/, /d/, /l/, /n/, /ʃ/, /ç/ og /ŋ/" er av underkategoriene det gjøres færrest feil ved (jf. figur 6). Dette gjelder både VIET og BKS. Tabell 24–26 viser hvilke grafemer det gjøres feil ved her. Bare de fonemene det gjøres feil ved, er med i tabellen.

Antall feil	Utelatt konsonantgrafem i gruppe av konsonantgrafemer		Innskutt konsonantgrafem ved siden av annet konsonantgrafem		Innskutt konsonantgrafem i gruppe av konsonantgrafemer		Substitusjon av konsonantgrafem i gruppe av konsonantgrafemer	
	Antall personer		Antall personer		Antall personer		Antall personer	
	BKS	V	BKS	V	BKS	V	BKS	V
0	53	43	89	73	84	85	87	77
1	33	22	7	17	12	12	10	13
2	7	23	3	7	2	1		7
3	4	5					1	1
4		3		2		1		1
5	1	2			1		1	
6		1						
7								
8	1							

Tabell 45: Oversikt over hvor mange personer som gjør hvor mange feil ved gruppe av konsonantgrafemer.

Grafem	Antall personer som har gjort x antall feil ved disse grafemene:					
	1 feil		2 feil		3 feil	
	BKS	VIET	BKS	VIET	BKS	VIET
b						
d	4	7		2		
f	1					
g	5	5				
h	6					
j	2	6		1	1	
k	5	3		1	1	
l	2	4		1		
m						
n	9	8	1		1	1
p						
r	8	11		3		1
s	3	21		2	1	1
t	6	12		3		
v	7					

Tabell 46: Antall personer som gjør x antall feil ved utelatt konsonantgrafem i gruppe av konsonantgrafemer.

1 feil	3 feil	5 feil
k for t	d for t	x for ks
l for j		
c for k		
m for n		
t for k		
r for f		
n for l		
d for g		
n for d		
t for d		

Tabell 47: Substitusjon av konsonantgrafemer i gruppe av konsonantgrafemer hos BKS.

1 feil	2 feil	3 feil	6 feil
m for n	t for l	h for j	x for ks
n for l	t for s	c for k	
x for k			
n for m			
k for b			
s for r			
m for t			
n for r			
l for n			
k for t			
c for s			
t for k			
k for l			
g for p			
c for n			

Tabell 48: Substitusjon av konsonantgrafemer i gruppe av konsonantgrafemer hos VIET.



Gjelder ved grafem:	To forskjellige konsonantgrafemer for dobbelt konsonantgrafem		Dobbelt konsonantgrafem for to forskjellige konsonantgrafemer	
	Vietnamesisk	BKS	Vietnamesisk	BKS
a				
b				
c				
d				
e				
f	oftentlige for offentlige x 2			
g				
h				
j				
k	kajakt for kajakk x 2, arkurat for akkurat		utsikk for utsikt	
l		(i hvert) falt for fall	Fitter for filter, detta for delta	ville for vilje
m				
n	brantuliker for brannulykker, brant for brann		utlenning for utlending	utlannet for utlandet
p				
q				
r	nygjerning for nysgjerrig		mossommer for morsomt, strekk for sterk	
s	yrkerskole for yrkesskole			
t	støster for støtter			
v				
w				
x				
<b>Sum:</b>	10 feil	1 feil	6 feil	2 feil

Tabell 49: To forskjellige konsonantgrafemer for dobbelt konsonantgrafemer og dobbelt konsonantgrafem for to forskjellige konsonantgrafemer.

Fonem	Skrivemåte	Utelatt grafem	Antall personer som gjør x antall feil					
			1 feil		2 feil		3 feil	
			BKS	V	BKS	V	BKS	V
/t/	rt	t	3	5				
		r		1				
/d/	rd	r		1				
/l/	rl	r	1	2				
/ŋ/	rn	r	2			1		
		n	2	2	1			
/ʃ/	skj	s	2	3				
		k	4	1			1	
		j	1	1			1	
	sj	s	1					
		j		4				
	rs	s		2				
		r	2					
sk	s		1					
/ŋ/	ng	n	2		1			
		g	3					
	n (foran k)	n		1		1		

Tabell 50: Utelatte konsonantgrafemer ved grafemer som gjengir /t/, /d/, /l/, /ŋ/, /ʃ/, /ç/ og /ŋ/.

Fonem	/f/	/ç/	/t/
Stavemåte	sj	kj	rt
Innskutte grafemer	k	s	d
Vietnamesisk	0	3	0
BKS	1	2	1

Tabell 51: Innskutte grafemer ved /t/, /d/, /l/, /n/, /f/, /ç/ og /η/.

Fonem	/f/				/η/	
Stavemåte	sj		rs		ng	n (før k)
Substituerte grafemer	h for j	c for s	s for r	t for s	d for g	c for n
Vietnamesisk	3	1	1	2	0	0
BKS	0	0	0	0	1	0

Tabell 52: Substitusjon av grafemer ved /t/, /d/, /l/, /n/, /f/, /ç/ og /η/.

Vannrett: riktige grafemer Loddrett: valgte grafemer	b	c	d	f	g	h	j	k	l	m	n	p	q	r	s	t	v	w	x	z	Sum
b																					0
c								2													2
d					1											3					4
f									1												1
g			2											1							3
h								1													1
j					1																1
k					1											1					2
l							1							1							2
m											1										1
n			1		1			1								1					4
p																					0
q																					0
r					2			3		1											6
s																					0
t			1					1						1							3
v				1																	1
w																					0
x																					0
z															1						0
Sum	0	0	4	1	6	0	1	4	5	0	2	0	0	3	1	5	0	0	0	0	

Tabell 53: Substitusjon av konsonantgrafemer hos BKS.

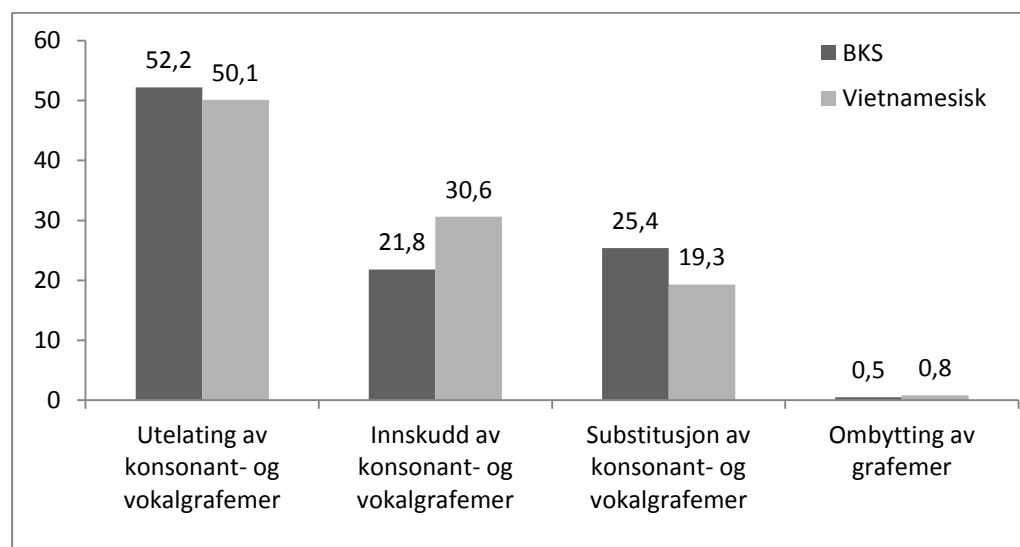
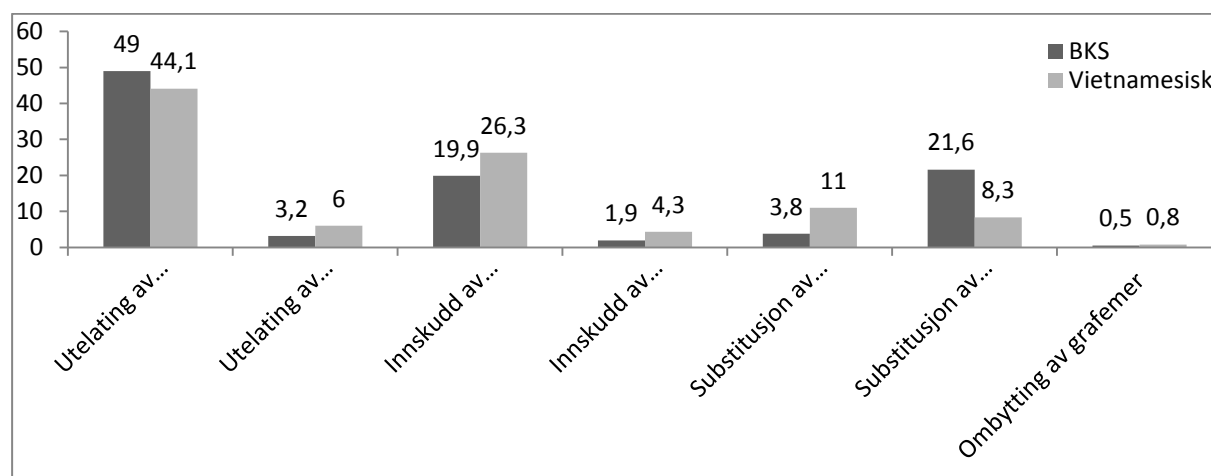
Vannrett: riktige grafemer Loddrett: valgte grafemer	b	c	d	f	g	h	j	k	l	m	n	p	q	r	s	t	v	w	x	z	Sum
b				1																	1
c								3			1				2						6
d									1							1					2
f					1																1
g												1									1
h							2														2
j																					0
k	1								1						3	1					6
l			2								2			2		1					7
m											2					1					3
n			2						1	1				3							7
p	1																				1
q																					0
r								1			1				2						4
s			2					1						1		1					5
t				2				4	4		2				4						16
v																					0
w																					0
x								1													1
z																					0
<b>Sum</b>	2	0	6	3	1	0	2	10	7	1	8	1	0	6	11	5	0	0	0	0	

Tabell 54: Substitusjon av konsonantgrafemer hos VIET.

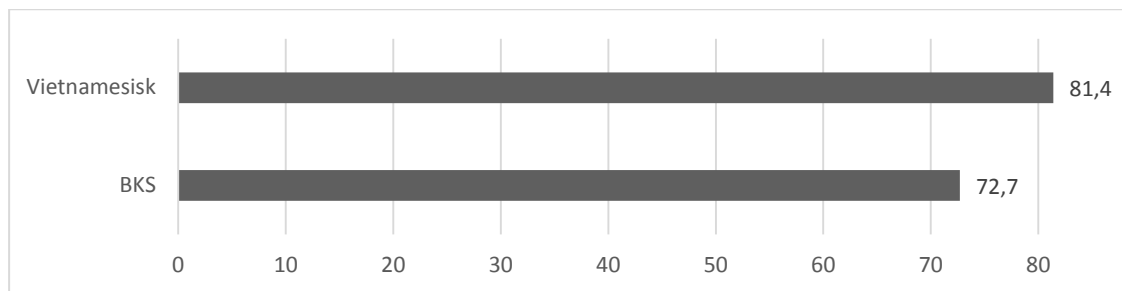
## Vedlegg 8: Resultater i feilfrekvens per samlet antall feil

Som beskrevet i 5.3, er den feilfrekvensen jeg hovedsakelig bruker i undersøkelsen feilfrekvens per løpeord. Jeg har også regnet ut feilfrekvens per samlet antall feil, jf. figur 17–21.

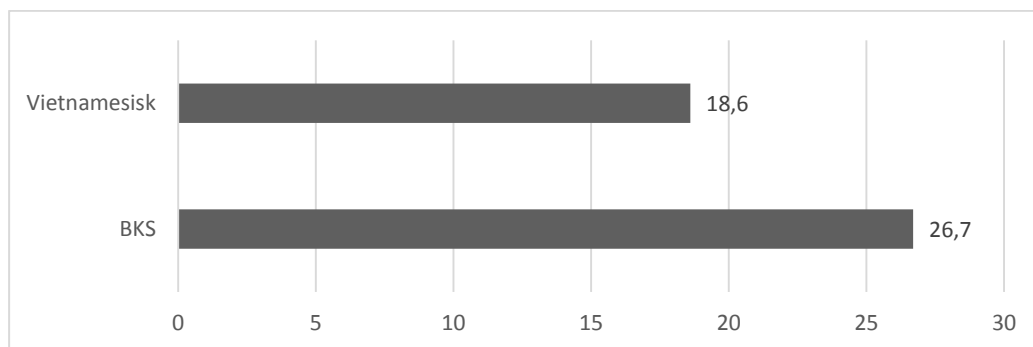
Figur 17: Hovedkategorier i prosent av det totale antallet feil.



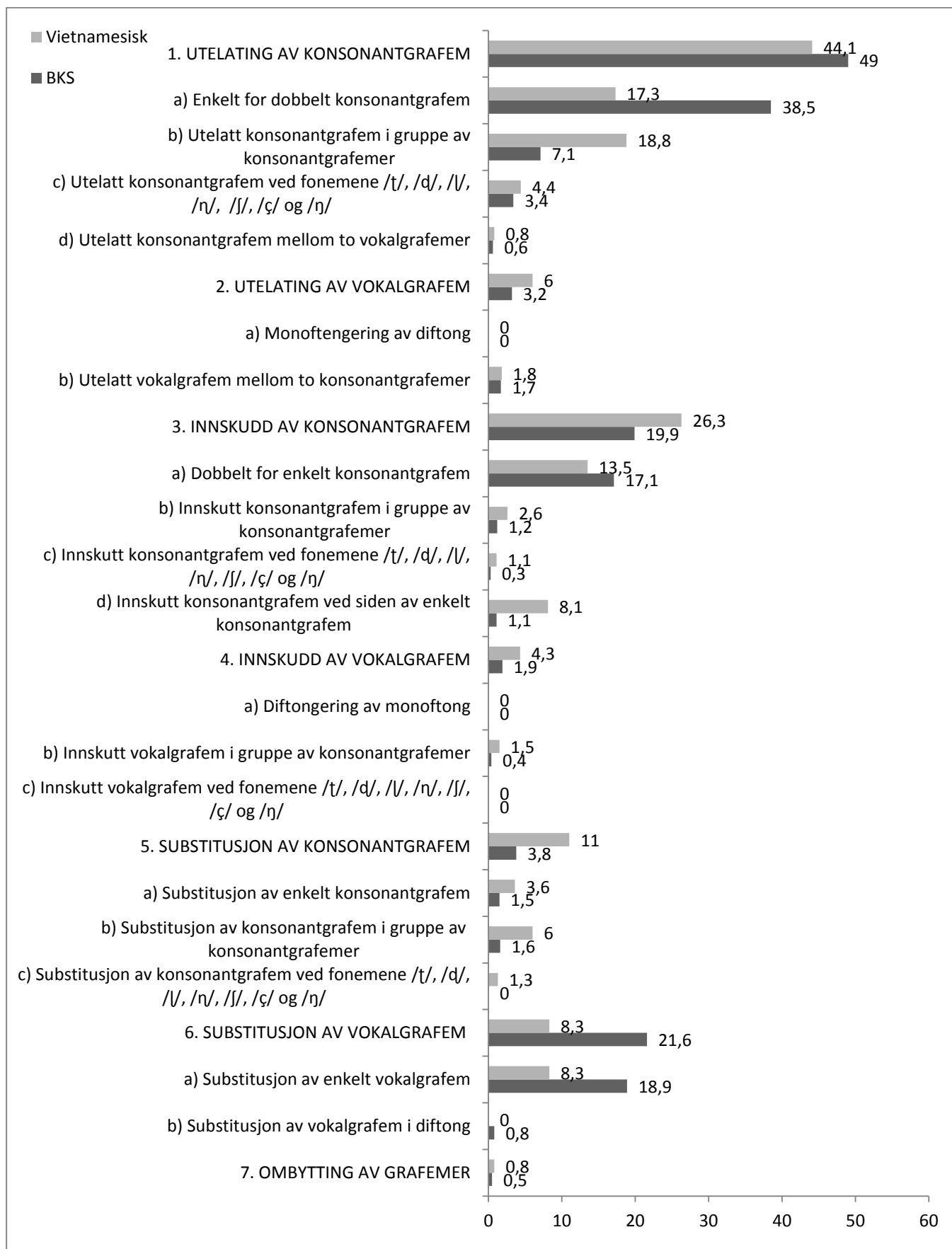
Figur 18: Feil i prosent av det totale antallet feil ved utrelating, innskudd, substitusjon og ombyting.



**Figur 19: Konsonantgrafemfeil i prosent av det totale antallet feil.**



**Figur 20: Vokalgrafemfeil i prosent av det totale antallet feil.**



Figur 21: Resultater i prosent av det totale antallet feil – hovedkategorier og underkategorier.

