

Masteroppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Det humanistiske fakultet
Institutt for språk og litteratur

Malin Bakke Frøystadvåg

Ekstra, ekstra eller ekestara?

Når somaliske stavelsesstrukturer møter
norsk ortografi

Masteroppgave i nordisk språk

Veileder: Kaja Borthen og Cecilie Slinning Knudsen

Juni 2020

Malin Bakke Frøystadvåg

Ekstra, ekstra eller ekestara?

Når somaliske stavelsesstrukturer møter
norsk ortografi

Masteroppgave i nordisk språk
Veileder: Kaja Borthen og Cecilie Slinning Knudsen
Juni 2020

Norges teknisk-naturvitenskapelige universitet
Det humanistiske fakultet
Institutt for språk og litteratur



Kunnskap for en bedre verden

Sammendrag

Dette er en av få studier av ortografi i norsk som andrespråk, og den første som tar for seg ortografien til innlærere av norsk med somali som morsmål. Jeg undersøker et utvalg ord som avviker fra bokmålsnormen av 2005, og diskuterer om disse avvika kan skyldes transfer av stavelsesstrukturer fra somali. Sentrale tema er mellomspråk, transfer, reparasjonsstrategier og forholdet mellom tale og skrift.

Datamaterialet mitt er henta fra Norsk andrespråkskorpus (ASK). Jeg undersøker 107 tekster skrevet av voksne andrespråksinnlærere med somali som morsmål. Disse tekstene inneholder til sammen 3048 ortografiske avvik, og av disse klassifiserer jeg 409 utvalgte ord etter hvor i stavelsen og hvor i ordet det er avvik, samt hvilket tegn som er involvert i avviket. Med de klassifiserte avvika til grunn stiller jeg to spørsmål basert på Jarvis' (2000) rammeverk for transferstudier: (1) Er det noen typer avvik som går igjen blant informantene mine? og (2) Er det likheter mellom stavelsesstrukturen i mellomspråket til informantene mine og stavelsesstrukturen i morsmålet deres?

I studien min finner jeg at det er noen tendenser i informantenes avvik. Det er spesielt tendenser til innsetting av vokaler og bortfall av konsonanter. Mange av avvika i disse orda fører til at stavelsesstrukturen i ordet blir likere somaliske stavelsesstrukturer. I oppgava diskuterer jeg om avviket har som mål å gjøre ordet likere somaliske stavelsesstrukturer, eller om det kan være andre grunner bak avviket. Jeg finner at det er mange mulige forklaringer, slik som transfer av ortografi, transfer av morfologi, analogi der en grammatisk regel blir overgeneralisert, og mer tilfeldige feil. Jeg finner også at bare 238 av 48 912 ord i datamaterialet mitt har avvik som *kan* skyldes transfer av stavelsesstrukturer. Likevel er det altså noen tendenser blant disse 238 orda, og i en del tilfeller er det sannsynlig at avviket har oppstått som et resultat av transfer av stavelsesstrukturer fra somali.

Abstract

This study investigates the orthography of second language learners of Norwegian with Somali as their first language. Examining 409 out of 3048 words which are not spelled in accordance with the 2005 norm of the written language Bokmål, I investigate if the errors can be explained through transfer of syllable structures from Somali.

The data has been collected from Norsk andrespråkskorpus, a corpus consisting of texts written by learners of Norwegian as a second language. The words investigated are collected from 107 texts written by 107 adult second language learners of Norwegian. Based on the methodological framework of Jarvis (2000), I aim to answer (1) whether there are certain types of mistakes that are common for many of the informants and; (2) whether the orthographic errors yield words that to a greater extent resemble Somali syllable structure.

My results show that several mistakes are common for many of the informants. Furthermore, some of the orthographic errors result in syllable structures that resemble that of Somali to a greater extent. This does not necessarily imply that these spelling errors in fact are a result of transfer of syllable structures. Instead, it could be that the spelling errors appear as a result of orthographic transfer, morphological transfer, analogy or handwriting. My findings imply that there are multiple possible reasons behind most of the errors seen in my data set, and that it is difficult to determine the cause of the errors. However, there is a subset of words where the only explanation seems to be transfer of syllable structures.

Forord

Mange må takkes for at jeg kom i mål med dette prosjektet.

Først vil jeg takke Cecilie Slinning Knudsen for at hun tok meg imot med åpne armer da prosjektet var i oppstartsfasen. Jeg vil takke både henne og Kaja Borthen for enestående veiledning. Takk for alle litteraturforslag, omhyggelige gjennomlesninger, uvurderlige tilbakemeldinger og fine Zoom-samtaler. Men viktigst av alt: Takk for at dere trodde på meg og prosjektet mitt.

En takk går til Nina Hagen Kaldhol som fatta interesse for prosjektet mitt, og som delte av sin kunnskap om somali. Takk til Hilde Johansen, som svarte på spørsmål om Norsk andrespråskorpus, takk til Jardar Eggesbø Abrahamsen, som delte av sin kunnskap om IPA-tastaturer, og takk til Ivar Berg, som ga meg nyttige litteraturtips. Språkrådet skal takkes for at de støtta prosjektet mitt med et raust stipend. Jeg er takknemlig for at de så potensial i skissa mi.

Mange skal takkes for at jeg holdt motet oppe når pandemi møtte masterinnspurt. Takk til Eivind og Håvard for fine stunder i kollektivet, og takk til dere som har tatt telefonen når jeg har trengt noen å snakke med. Spesielt skal Amanda, Morten, Ingrid og familien takkes. Mamma må i tillegg få en ekstra takk for at hun gjorde samtlige figurer i denne oppgava leservennlige, og Ingrid må få en ekstra takk for at hun hjalp undertegnede nordist med engelske formuleringer.

Trondheim, 1. juni 2020

Malin Bakke Frøystadvåg

Innhold

Figurer.....	xi
Tabeller	xii
Forkortelser og tegnforklaringer	xiii
1.0 Innledning	15
1.1 Somaliere og somali.....	15
1.2 Problemstilling og forskningsspørsmål.....	16
1.3 Oppgavas videre oppbygging	17
DEL I: Teoretisk bakteppe.....	19
2.0 Stavelsen	21
2.1 Stavelser danner ord.....	23
2.2 Stavelsegrensener	25
2.3 Stavelsestemplater.....	26
2.3.1 Det norske stavelsestemplatet	26
2.3.2 Det somaliske stavelsestemplatet.....	28
2.4 Reparasjonsstrategier	30
2.5 Universelle prinsipper for stavelседanning	33
2.5.1 Opptaktsprinsippet	33
2.5.2 Sonoritetsprinsippet	33
3.0 Skrift står for tale	37
3.1 Ortografi.....	38
3.2 Norsk ortografi.....	39
3.3 Somalisk ortografi.....	41
4.0 Mellomspråk	45
4.1 Andrespråk, målspråk, fremmedspråk, S2 og L2	46
4.2 Transfer	47
4.3 Oppsummering av kapittelet.....	49
5.0 Tidligere forskning.....	51
5.1 Forskning på transfer av stavelsesstrukturer i andrespråk	51
5.2 Forskning på ortografi i andrespråk	53

5.2.1	Forskning på ortografi i norsk som andrespråk	55
5.2.2	Stavelsesstrukturer i Hvenekilde (1985) og Vigrestad (2016).....	56
5.3	Forskning gjort på språkkontakt mellom somali og andre språk.....	59
5.3.1	Somali i møte med norsk	60
DEL II:	Studien.....	63
6.0	Metode	65
6.1	Metodologisk grunnlag	65
6.2	ASK som datagrunnlag.....	67
6.3	Om bruk av «feil» og «avvik»	69
6.5	Å søke seg fram i korpuset.....	69
6.6	Informanter	72
6.7	Utvalg av data	73
6.8	Kategorisering av data	75
7.0	Resultater, analyse og drøfting	77
7.1	Innsetting av grafem	78
7.1.1	Innsetting av vokalgrafemer ordmedialt	79
7.1.2	Innsetting av vokalgrafemer ordfinalt.....	88
7.1.3	Innsetting av konsonantgrafemer.....	93
7.2	Bortfall av grafemer.....	94
7.2.1	Bortfall av vokalgrafemer	94
7.2.2	Bortfall av konsonantgrafemer	95
7.2.3	Bortfall og innsetting av flere segmenter.....	102
7.3	Ombytting og flytting av grafemer	103
7.4	Erstatning av grafemer.....	104
7.5	Fonologisk eller ortografisk transfer?.....	105
7.6	Oppsummering.....	107
9.0	Konklusjon.....	109
9.1	Svar på problemstilling.....	109
9.2	Videre forskning	110
Litteratur	113
Appendiks 1:	datamateriale.....	117
Appendiks 2:	informanter	139

Figurer

<i>Figur 1: stavelsens oppbygging</i>	21
<i>Figur 2: polysyllabisk ord delt opp i stavelser</i>	24
<i>Figur 3: norsk stavelsestemplat</i>	26
<i>Figur 4: somalisk stavelsestemplat</i>	28
<i>Figur 5: søket mitt i ASK-korpuset</i>	70
<i>Figur 6: kode for søk med attributtene «ORT» og «somali» i ASK</i>	70
<i>Figur 7: kwic-modus i ASK</i>	71
<i>Figur 8: søk med informasjon om informanten i ASK</i>	71
<i>Figur 9: kode for søk på tekstene i sin helhet i ASK</i>	71
<i>Figur 10: liste med tekst-id vist i kwic-modus i ASK</i>	71
<i>Figur 11: å finne hele teksten i ASK</i>	72
<i>Figur 12: utsnitt av teksten i sin helhet i ASK</i>	72
<i>Figur 13: mulig analyse av <probelemer></i>	79
<i>Figur 14: stavelsesdanningsprosessen fra <ble> til <be.le></i>	80
<i>Figur 15: to mulige analyser av <ønesker></i>	81
<i>Figur 16: stavelsesdanningsprosessen fra <sker> til <es.ker></i>	82
<i>Figur 17: stavelsesdanningsprosessen fra <øns> til <øn.es></i>	83
<i>Figur 18: mulig analyse av <vestilige></i>	85
<i>Figur 19: stavelsesdanningsprosessen fra <vest> til <ves.ti></i>	85
<i>Figur 20: mulig analyse av <bedrifte></i>	88
<i>Figur 21: stavelsesdanningsprosessen fra <drift> til <drif.te></i>	89
<i>Figur 22: mulig analyse av <faktis></i>	97
<i>Figur 23: stavelsesdanningsprosessen fra <tisk> til <tis></i>	97
<i>Figur 24: mulig analyse av <hjepe></i>	98

Tabeller

<i>Tabell 1: mulige stavelsesstrukturer i norsk</i>	27
<i>Tabell 2: segmentenes og segmentgruppenes tillatte plassering i den norske stavelsen</i>	28
<i>Tabell 3: mulige stavelsesstrukturer i somali</i>	29
<i>Tabell 4: segmentenes tillatte posisjoner i den somaliske stavelsen</i>	29
<i>Tabell 5: segmenter i standard østnorsk organisert etter sonoritetshierarkiet</i>	34
<i>Tabell 6: segmenter i nordlig somali organisert etter sonoritetshierarkiet</i>	34
<i>Tabell 7: det norske alfabetet</i>	39
<i>Tabell 8: forholdet mellom segmenter og grafemer i norsk</i>	40
<i>Tabell 9: det somaliske alfabetet</i>	41
<i>Tabell 10: forholdet mellom segmenter og grafemer i somali</i>	42
<i>Tabell 11: oversikt over antall ORT i de ulike kategoriene</i>	77
<i>Tabell 12: innsetting av grafemer</i>	78
<i>Tabell 13: bortfall av vokalgrafem</i>	94
<i>Tabell 14: bortfall av konsonantgrafem</i>	96

Forkortelser og tegnforklaringer

Under følger en liste med forklaringer på forkortelser og tegn brukt i oppgava. Jeg forklarer dem også første gang jeg bruker dem i teksten.

ASK	Norsk andrespråskorpus
K	konsonant
Kj	kjerne
L1	førstespråk (engelsk betegnelse)
L2	andrespråk (engelsk betegnelse)
N2	norsk som andrespråk
ORT	ortografiske avvik fra bokmålsnormen av 2005
O	opptakt
R/r	rim
SOMALI	innlærere av norsk med somali som morsmål
S1	førstespråk
S2	andrespråk
σ	stavelse
V	vokal
<...>	grafemisk transkripsjon (slik et ord blir skrevet)
/.../	fonemisk transkripsjon (slik <i>lydsystemet</i> fungerer)
[...]	fonetisk transkripsjon (slik det faktisk blir uttalt)

1.0 Innledning

Av mars 2020 er det 28.554 innvandrere¹ fra Somalia i Norge, og somaliere utgjør med det den femte største innvandrergруппa i Norge (SSB, 2020). I tillegg finnes det en del innvandrere fra Kenya, Etiopia og Djibouti som er en del av den somaliske folkegruppen. Det finnes ingen statistikk på hvor mange innvandrere det er i Norge med somali som morsmål, men vi kan anta at veldig mange av innvandrerne fra somalisktalende områder har somali som morsmål. Det er med andre ord et stort antall innvandrere i Norge med somali som morsmål som skal lære seg, lærer seg eller har lært seg norsk. Derfor er det interessant å undersøke hvilke trekk ved norsk som kan være vanskelig å lære seg når man har somali som morsmål, og det er det jeg ønsker å undersøke i denne studien.

1.1 Somaliere og somali

Somali er en folkegruppe på Afrikas horn. En stor andel somaliere er fra Somalia, men det finnes også somaliere fra nord i Kenya, øst i Etiopia og Djibouti. Den somaliske diasporaen er omfattende fordi mange har flyktet fra den somaliske borgerkrig, som har pågått siden 1992.

Det største språket i regionen er somali, og dette er det offisielle språket i Somalia. Det er et kusjittisk språk som tilhører den afroasiatiske språkfamilien (Ethnologue: Lewis, Simons, and Fennig 2017). Som norsk har somali dialekter, som ifølge Banti (2013) kan deles inn i tre dialektgrupper: nordlig somali, benadir og ashraf. Noen regner også maay som en dialekt av somali, men det er diskutert hvorvidt dette bør regnes som et eget språk. Standard somali er basert på nordlig somali, som i tillegg til å være snakket i nord, blir snakket sør i og øst i Etiopia, tilgrensende områder sør i Somalia, nord i Kenya og i Djibouti. De andre dialektene blir snakket i sentrale deler av Somalia.

Somaliere er del av et komplekst lingvistisk landskap. I tillegg til somali og dialektene som hører med, er arabisk, engelsk, italiensk og fransk viktige språk i den somalitalende regionen på Afrikas horn. Ettersom islam står sterkt i området, er det vanlig å lære seg arabisk, som regnes som et hellig språk. Dette er også offisielt språk i Somalia. Bruken av engelsk, italiensk og fransk stammer fra kolonitida. Da var de nordlige delene av Somalia kolonisert av britene, de sørlige delene kolonisert av italienerne og det som i dag er Djibouti, kolonisert av franskmennene. Engelsk står fortsatt sterkt i dag, og er et offisielt språk i den

¹ SSB definerer innvandrere som «personer som er født i utlandet av to utenlandsfødte foreldre og fire utlandsfødte besteforeldre» <https://www.ssb.no/innvbef> (2020).

selverklærte republikken Somaliland (Banti, 2013). Det er også mange mindre stammespråk som blir snakket i Somalia, slik som det kusjittiske språket Oromo (Appleyard & Orwin, 2008).

1.2 Problemstilling og forskningsspørsmål

Somali er et språk som skiller seg fra norsk på flere måter. Det har blant annet et intrikat tonesystem som er styrende for grammatikken, og mange lyder som er ukjente for de fleste norske ører. Det har også en *stavelsesstruktur* som skiller seg veldig fra den norske, og det er dét denne studien skal dreie seg om. Somali har det vi kaller en *enkel stavelsesstruktur* der de mest komplekse stavelsene som eksisterer i språket, kun har tre lyder. Vi sier at somali har en *maksstruktur* på KVK (konsonant – vokal – konsonant). I norsk er det annerledes. Norske stavelser kan være svært komplekse, og inneholde mange lyder – ordet /felmskt/, for eksempel, inneholder hele sju lyder. I tillegg er det ulikt hvilke lyder som tillates i de ulike posisjonene i stavelen i de to språka

I denne studien er jeg interessert i å se hva som skjer når personer med somali som morsmål lærer seg norsk og somaliske stavelsesstrukturer møter norske stavelsesstrukturer, nettopp fordi de to strukturene er så ulike. Vil man kunne se spor av somalisk stavelsesstruktur i tekster skrevet av innlærere av norsk med somali som morsmål?

I studien min kommer jeg til å se på 107 tekster skrevet på norsk av innlærere med somali som morsmål (heretter kalt SOMALI). Jeg kommer dermed *ikke* til å ta for meg stavelser slik de blir uttalt, men heller slik de blir uttrykt på papiret. Tekstene er tatt fra Norsk andrespråkskorpus (ASK). I studien kommer jeg til å undersøke *ortografien* til informantene, og jeg vil se på et utvalg ord som avviker fra bokmålsnormen av 2005. Utvalget ord jeg skal forske på, har avvik fra standardortografien som kan tenkes å skyldes forskjeller i somaliske og norske stavelsesstrukturer. I studien min skal jeg, basert på teori og tidligere forskning undersøke om disse avvik faktisk kan forklares som et resultat av *transfer*, altså overføring, av somaliske stavelsesstrukturer. Problemstillinga mi får dermed følgende ordlyd:

I hvilken grad ser vi spor av somaliske stavelsesstrukturer i ortografien til innlærere av norsk med somali som morsmål?

Jeg kommer til å kategorisere noen utvalgte avvik fra bokmålsnormen av 2005, og undersøke følgende to forskningsspørsmål, basert på Jarvis (2000):

1. Er det noen avvik som går igjen blant informantene mine?
2. Har disse avvika noen fellestrekk med stavelsesstrukturen i somali?

Jeg skal undersøke alle avvik i materialet mitt, og velge ut de avvika som kan skyldes transfer av somaliske stavelsesstrukturer. Deretter skal jeg undersøke og diskutere om disse faktisk skyldes denne overføringa. Med denne problemstillinga til grunn håper jeg å bidra med ny innsikt i hva slags ortografiske avvik innlærere av norsk med somali som morsmål kan ha, og hva som kan være bakgrunnen for at de har dem.

1.3 Oppgavas videre oppbygging

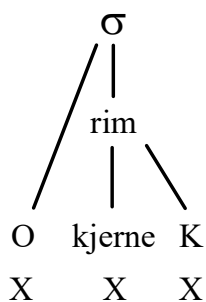
For å kunne svare på problemstillinga presentert over begynner jeg i kap. 2 med å presentere stavelsesteori. I kap. 3 snakker jeg om ortografi og forholdet mellom lyd og skrift. I kap. 3 gjør jeg rede for termene mellomspråk og transfer. I kap. 4 legger jeg fram tidligere forskning som kan være relevant for min studie. Jeg starter med å se på tidligere forskning på stavelsesstrukturer og andrespråk. Videre ser jeg på både internasjonal og norsk forskning på andrespråksortografi. I tillegg ser jeg på forskning gjort på somali i møte med andre språk. I kap. 5 presenterer jeg det metodologiske rammeverket jeg bruker, datamaterialet mitt og hvordan jeg har gått fram for å velge ut de dataene som var relevante for problemstillinga mi. Videre beskriver jeg hvordan jeg har kategorisert og undersøkt disse dataene. I kap. 6 presenterer jeg resultata mine, dvs. dataene mine kategorisert. I tillegg gjør jeg en inngående analyse av dataene, og diskuterer på grunnlag av det om avvika jeg finner faktisk kan skyldes transfer av stavelsesstrukturer. Jeg runder av med en oppsummering og forslag til videre forskning i kap. 7.

DEL I: Teoretisk bakteppe

2.0 Stavelsen

Siden denne studien skal handle om stavelser, er det naturlig å gjøre greie for hva en stavelse er. Som Husby & Kløve (2001) skriver, har det «vist seg vanskelig å definere hva en stavelse er. Likevel er det en fonologisk enhet som alle språkbrukere erkjenner eksistensen av. Alle klarer å dele språket sitt opp i stavelser» (s. 63). Dette ser vi tydelig i klassiske klappeøvelser i grunnskolen. Barna blir instruerte til å klappe mens de snakker, og dermed finne fram til stavelsene i et ord. Barna har ikke lært noe om stavelser utover dette, men de kan likevel, med den største selvsikkerhet, si at ordet <mørkegrønne>² (den mørkegrønne fargeblyanten) består av fire stavelser, og at disse er <mør>, <ke>, <grø> og <nnen>. Hva er det disse fire delene av ordet har til felles som gjør at barna identifiserer dem som enheter?

Jeg skal starte med å ta for meg stavelsens oppbygging. Stavelsen består av tre posisjoner: *kjerne*, *opptakt*, *koda*. Kjernen i stavelsen er nesten alltid fylt av en *vokal* i form av en *monoftong* eller *diftong*. Om vi ser på de fire stavelsene fra forrige avsnitt, ser vi at kjernen i <mør> er fylt med <ø>, kjerne i <ke> er fylt med <e>, kjerne i <grø> er fylt med <ø>, og kjerne i <nnen> er fylt med <e>. Disse stavelsene har altså det til felles at de har en vokal i kjerneposisjonen. Opptakt er posisjonen som kommer før kjernen, mens koda er posisjonen som kommer etter kjernen. Disse posisjonene *kan* fylles med segmenter, men de må ikke bli fylt. Segmentene må være *konsonanter* – vokaler kan ikke settes inn i opptakt eller koda. I stavelsen <mør> ser vi at konsonanten <m> er plassert i opptaktsposisjon, i <ke> har vi <k>, i <grø> har vi <r> og i <nnen> har vi <n>. Vi ser altså at opptakt kan fylles av både én og flere konsonanter. Ser vi på kodaposisjonen i de ulike stavelsene, kan vi notere oss at <mør> har <r> i koda, <ke> ikke har fylt koda, <grø> ikke har fylt koda, og <nnen> har <n> i koda. Figuren under illustrerer hvordan en stavelse er bygget opp (etter Husby & Kløve, 1998:65).



Figur 1: stavelsens oppbygging

² I denne oppgava markerer jeg skrevne ord med <...>

I denne figuren kjenner vi igjen opptakt, kjerne og koda, som jeg allerede har introdusert. Under de tre posisjonene finner vi X-er. X-ene må i opptaktsposisjon og kodaposisjon stå for konsonanter, mens kjerneposisjonen (i norsk) som oftest fylles av en vokal, men også kan fylles av enkelte konsonanter. I tillegg er det satt inn et nytt nivå, *rim*. Kjerne og koda danner til sammen rimet i stavelsen, og på et mer overordna hierarkisk nivå kan man si at en stavelse *ofte* består av opptakt og rim (det finnes opptaktsløse stavelser). På toppen av hierarkiet ser vi det greske tegnet sigma (σ), og dette står for *stavelse*.

Som en fellesbetegnelse på vokaler og konsonanter, kommer jeg til å bruke betegnelsen *segment*. Det må nevnes at segment er et begrep som *kan* være problematisk å bruke fordi enhetene i språket ikke er så rigide som segmentteori skal ha det til. Ifølge segmentteori er det nemlig slik at hvert segment i språket, for eksempel /t/³, har noen iboende trekk. I fonologien, dvs. læren om lydssystemer, er ikke bildet så enkelt. Det er ikke slik at alle segment vi identifiserer som /t/, alltid har nøyaktig samme kjennetegn. Likevel kan det være praktisk å snakke om segmenter når en undersøker visse fonologiske prosesser. Abrahamsen (2017) skriver følgende: «Ikkje minst finst fenomen som metatase, altså lydomsifting, som då kross vart til kors i delar av skandinavisk, eller når born vil ta på seg stølvane. Skal segment skifte plass på slike måtar, treng vi at segmenta faktisk finst i språkstrukturen» (s. 218). I denne oppgava *skal* jeg undersøke fenomenet metatase, og jeg kommer også til å undersøke tre andre fenomen, bortfall, epentese og erstatning, som gjør at det er praktisk å omtale enhetene som danner en stavelse, som segmenter.

Modellen i figur 1 viser en såkalt «x-slot»-modell. I denne modellen er det sentralt at det ikke skal defineres hvilke ulike segmenter som kan settes inn i de ulike posisjonene. I en teori presentert av Clements og Keyser (1983) er en lignende modell lagt til grunn, en KV-modell, men det er én viktig forskjell mellom x-slot-teori og KV-teori: I modellen til Clements og Keyser er det definert hvilke typer segmenter som kan settes inn i de ulike posisjonene, slik at det er definert at kjernen må dannes av en vokal, og opptakt og koda må fylles av konsonanter. Et eksempel på at denne definisjonen av segmenter er uheldig, finner vi i norsk. I norsk kan konsonantene /n/ og /l/ være stavelsesbærende og dermed danne kjerne i en stavelse. Slike unntak tar ikke KV-teorien høyde for. Da x-slot-teorien ble lansert av Levin (1985) var en viktig endring at det ikke skulle defineres hva slags segmenter som kan settes inn i de ulike posisjonene. I denne teorien kan X-ene byttes ut med det segmentet som passer seg i den aktuelle stavelsen i det aktuelle språket (Topintzi, 2006:13-14). Begge

³ Jeg markerer segmenter med /.../ i denne oppgava, f.eks. /t/ eller /vas/.

modellene har det til felles at de ser på hvert segment som en tidsenhet. Merk at x-ene ikke bare kan stå for segmenter. Også det prosodiske nivået kan bli representert i x-slot-teorien, slik at for eksempel toner kan bli representert med x-er (Husby & Kløve, 1998:67).

Det finnes også en tredje modell lansert av Hyman (1985), den moraiske modellen, som blir mye brukt i analyser av stavelser. I denne teorien er *moraer* sentrale, og basert på antall moraer i stavelsen, kan en stavelse bli kategorisert som tung eller lett. Moraer blir gruppert inn i det som kalles versføtter, og dette kan igjen bli knyttet til det prosodiske nivået (Topintzi, 2006:15). Dette er altså en modell som reflekterer det komplekse forholdet mellom stavelse og prosodi, og som knytter de to nivåene sammen på en god måte. For denne oppgava er det likevel ikke hensiktsmessig å bruke en slik modell, ettersom jeg ikke ser på prosodi og detaljerte lydlig forhold. Studieobjektet mitt er skrift, og jeg kommer til å konsentrere meg om segmentene de ulike tegna på papiret representerer. Derfor mener jeg det er mest ryddig å legge x-slot-teori til grunn i oppgava mi.

I figur 1 viste jeg altså hvordan en stavelse er bygget opp, og som vi skal se i 2.2, er det ulike *stavelserrestriksjoner* for ulike språk. Det vil si at stavelser i norsk må følge norske stavelserdanningsregler, og stavelser i somali må følge somaliske stavelserdanningsregler. Det er tre faktorer som kan variere fra språk til språk (Hognestad, 2019:57):

1. Hvor mange konsonanter som er tillatt i opptakt og koda.
2. Hvilke konsonanter som er tillatt i opptakt og koda
3. Hvilken rekkefølge det er tillatt at konsonantene står i

Reglene for hvordan lyder kan kombineres i et språk, kalles språkets *fonotaks*. Ulike språk har altså ulike stavelserstrukturer med ulik fonotaks. Kristoffersen (2000) argumenterer for å holde begrepene stavelserstruktur og fonotaks separert (s. 47). Jeg kommer til å skille mellom disse to begrepa når det er et poeng å få fram nyansene, men som hovedregel bruker jeg stavelserstruktur som en fellesbetegnelse på de to begrepa. Når jeg bruker betegnelsen «stavelserstruktur» i problemstillinga mi, inkluderer dette altså fonotaks.

2.1 Stavelser danner ord

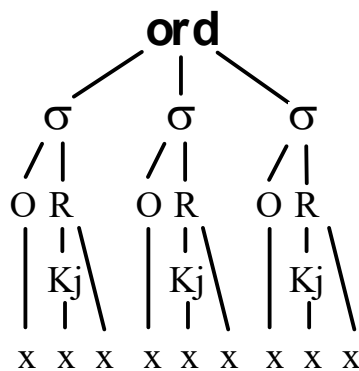
I studien min ser jeg på ord og hvordan stavelser blir realiserte i ord. Derfor er det relevant å ta for seg forholdet mellom stavelser og ord. Ord består av en eller flere stavelser.

Førstnevnte kaller vi monosyllabiske ord, og sistnevnte kaller vi polysyllabiske ord.

Eksempler er det monosyllabiske ordet /klø/ og det polysyllabiske ordet /hovar/. Noen språk, for eksempel vietnamesisk, tillater kun monosyllabiske ord (Rosén, 2001:22).

I mange tilfeller går stavelser over ordgrenser i fortløpende tale. Det vil si at den første delen av stavelsen hører til i slutten av et ord, og at den andre delen av stavelsen hører til begynnelsen av det neste ordet. For eksempel kunne vi delt opp de to ordene <kan ikke> i tre stavelser, <ka.ni.kke>⁴, der <ni> utgjør en stavelse som består av siste lyden i ordet <kan>, og første lyden i ordet <ikke>. Derfor er det viktig å se på sammenhengen når man undersøker stavelser i enkeltord, dvs. ord som står før og etter ordet som studeres.

Man kan dele opp et ord i tre konstituenter: ordinitial posisjon, ordmedial posisjon og ordfinal posisjon.⁵ Figur 2 viser et ord med tre stavelser, altså et polysyllabisk ord.



Figur 2: polysyllabisk ord delt opp i stavelser

Første stavelsen i ordet kommer jeg til å omtale som den ordinitiale stavelsen, den andre stavelsen i ordet kommer jeg til å omtale som den ordmediale stavelsen, og den siste stavelsen i ordet kommer jeg til å omtale som ordfinal. Videre kan de tre posisjonene initial, medial og final også brukes når vi snakker om segmenter. Da kan vi snakke om det ordinitiale segmentet som enten er koda eller kjerne i den ordinitiale stavelsen, og det ordfinale segmentet, som enten er kjerne eller koda i den ordfinale stavelsen. Alle segmenter mellom disse to posisjonene, er ordmediale segmenter. Det er viktig å ha betegnelser på ordinitialt og ordfinalt segment. Dette er fordi det ofte er restriksjoner på hvilke segment som kan stå i disse posisjonene. Et segment i en stavelse kan altså bli påvirket av posisjonen det har i ordet.

⁴ I denne oppgava markerer jeg stavelsegrensene med punktum.

⁵ En del litteratur omtaler den ordinitiale posisjonen som *fremlyd*, og den ordfinale posisjonen som *utlyd*, men disse betegnelsene kommer jeg ikke til å bruke i denne oppgava.

2.2 Stavelsesgrenser

I innledninga til dette kapitlet delte jeg opp ordet <mørkegrønne> i stavelserne <mør.ke.grø.nne>. Spørsmålet som da melder seg, er hvorfor jeg satt grensene der jeg satt dem. Hvorfor kunne jeg ikke for eksempel ha delt opp ordet slik: <*mør.ke.grønn.en>?⁶ Det er flere prinsipper som er styrende for stavelsgrenser, og disse blir viktige når jeg analyserer og diskuterer orda skrevet av informantene mine.

For det første må alle lydkombinasjoner i en stavelse være i tråd med norsk fonotaks (jeg omtaler regler for norsk fonotaks i kap. 2.3.1). <rk> i opptakt er ikke en tillatt segmentkombinasjon i norsk, og dermed kan ikke <rke> være en stavelse. Det neste prinsippet for stavelsgrenser, *prinsippet om maksimal opptakt*, er at man fyller opptakt med så mange konsonanter som mulig, så lenge dette ikke bryter med norsk fonotaks. Derfor sier man at <n> hører til i opptakt i siste stavelse, heller enn koda i nest siste stavelse. Det er også derfor både <g> og <r> knytter seg til den følgende <ø>-en, heller enn den foregående <e>-en – <gr> i opptakt bryter ikke med norsk fonotaks. For det tredje må konsonanter (nesten alltid) knytte seg til en vokal, som fyller kjernen i en stavelse. Av den grunn kan ikke <gr> være en stavelse i seg selv – disse konsonantene må knytte seg til en vokal. I noen tilfeller er det en konsonant som fyller kjerneposisjonen i stavelsen. I ordet <bamsen>, som gjerne blir uttalt [bamsɲ] i standard østnorsk, vil en stavelsebærende <n> fylle kjerneposisjonen i stavelsen <sn>. Vi skal se nærmere på dette i kap. 2.5.2 om sonoritetshierarkiet.

Det er likevel ingen absolutt sannhet om hvor stavelsgrenser går. For eksempel kan det argumenteres for at en bør dele <grønne> slik: <grønn.en>. I dette tilfellet knytter geminaten <nn> seg til foregående stavelse for å opprettholde morfologien. For at ikke ordet <grønn> skal miste en viktig del av stammen sin som er avgjørende for meningsinnholdet, plasseres altså <n> i koda etter kjernen <ø>. Dette kalles *det morfologiske prinsippet*. Jeg kommer ikke til å følge det morfologiske prinsippet i denne oppgava, men det er et viktig poeng at det ikke er to streker under svare på hvor en stavelsgrense går.

Et annet omdiskutert stavelsgrenseproblem er hvordan en skal analysere *totale assimilasjoner*. Når /r/ i slutten av et ord møter en /t/ i starten av neste ord vil det i østnorsk skje en assimilasjon av disse lydene, og resultatet er lyden /t/. Dette skjer for eksempel i setninga <Det er to igjen.>, der <er> og <to> blir assimilert til /et/ i tale. Om /t/ er i opptakt, og knytter seg til /o/, vil dette tilfredsstillende regelen om å fylle opptakt, men det er omdiskutert

⁶ Asterisken viser at dette ikke er grammatisk korrekt i henhold til norsk fonotaks.

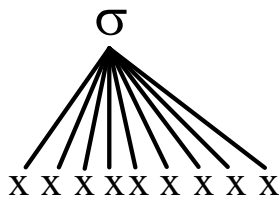
om dette er en lyd som er tillatt i opptakt med tanke på at den er jo ikke tillatt initialt i ord. Om /t/ knytter seg til den foregående /e/-en, blir norsk fonotaks tilfredsstilt, men da følger en ikke prinsippet om å fylle opptakt når det er mulig. Derfor blir disse assimilasjonene ofte ansett som *ambisyllabiske* slik at de hører til to stavelser, og både fyller koda i en stavelse, og opptakt i den neste (Knudsen & Husby, u.u.:75-78; se også Kristoffersen, 2003:23).

2.3 Stavelsestemplater

Det er som nevnt ikke slik at alle språk har samme stavelsesstruktur som den norske. Det kan vi vise gjennom språkspesifikke *stavelsestemplater*. Templater er mønstre stavelser i et språk må følge. Hvis stavelser ikke følger dette mønstret, er den altså ikke i henhold til språkets regler for stavelsesdanning. Templatet sier noe om hvor mange lyder det er lov å ha i de ulike posisjonene (opptakt, kjerne og koda) (Husby & Kløve, 2001:66-67). I 2.3.1 og 2.3.2 ser vi hvordan stavelsestemplatet ser ut i henholdsvis norsk og somali.

2.3.1 Det norske stavelsestemplatet

Figuren under viser det norske stavelsestemplatet (etter Husby & Kløve, 2001:68).



Figur 3: norsk stavelsestemplat

Av praktiske grunner viser ikke figur 3 posisjonene rim, opptakt og kjerne men den loddrette linja markerer hvor staveleskjernen er. Templatet viser at en norsk stavelse kan ha *opptil* fem segmenter i koda (etter den loddrette streken), og *opptil* tre segmenter i koda (før den loddrette streken). Det betyr altså ikke at alle stavelser *må* ha det, men at dette er maksimalt antall tillatte segmenter i disse posisjonene. Det betyr heller ikke at alle kombinasjoner av antall konsonanter i opptakt og koda er lovlig. Vi har for eksempel ikke noe ord i norsk med stavelsesstrukturen KKKVKKKKK. Under er en tabell med en oversikt over alle stavelser i norsk basert på stavelsestemplatet over, og under hver struktur følger et eksempel fra norsk på en stavelse som har denne aktuelle strukturen. Mange av eksemplene mine danner egne ord med semantisk innhold, men det er ikke obligatorisk med semantisk innhold for at en skal kalle noe en stavelse (se f.eks. /farkt/ under strukturen KVKKK).

		Koda					
		–	K	KK	KKK	KKKK	KKKKK
Opptakt	–	V /i/	VK /is/	VKK /ilt/	VKKK /olmt/	VKKKK /ulmsk/	VKKKKK /ulmskt/
	K	KV ⁷ /mo/	KVK /kor/	KVKK /form/	KVKKK /farkt/ ⁸	KVKKKK /felmsk/	KVKKKKK /felmskt/
	KK	KKV /tro/	KKVK /klok/	KKVKK /brukt/	KKVKKK /kvensk/	KKVKKKK /kvenskt/ ⁹	–
	KKK	KKKV /stri/	KKKVK /sklir/	KKKVKK /skvist/	KKKVKKK /sprelsk/	KKKVKKKK /sprelskt/	–

Tabell 1: mulige stavelsesstrukturer i norsk

Av tabell 1 kan vi se at det finnes 22 mulige stavelsesstrukturer for norske stavelser. Vi kan også se at det ikke finnes stavelser i norsk som har kombinasjonene KKKVKKKKK og KKKVKKKKK – stavelsestemplatet viser altså ikke i detalj hva slags kombinasjoner av opptakt og koda som er tillatt og ikke tillatt i norsk, noe tabell 1 kompenserer for.

I tabellen under presenterer jeg en oversikt over hvilke segmenter som kan opptre i henholdsvis opptakt, kjerne og koda i standard østnorsk. I tillegg viser jeg hvilke konsonantgrupper som kan opptre i de ulike posisjonene (etter Kristoffersen, 2000:13, 22; Husby & Kløve, 1998:82-83).

⁷ Merk at denne tabellen ikke tar høyde for at en del stavelser i norsk kan ha konsonant i kjerneposisjon.

⁸ Som i ordet /in.farkt/

⁹ Dette er ikke normert i bokmål eller nynorsk, men kan forekomme i tale hos enkelte morsmålsbrukere av norsk

	Opptakt	Kjerne	Koda
Enkeltsegmenter	/p/ /b/ /t/ /d/ /k/ /g/ /m/ /n/ /ɾ/ /f/ /v/ /s/ /ʃ/ /ç/ /h/ /j/ /l/	/i/ /y/ /a/ /u/ /ʊ/ /ɛ/ /æ/ /e/	/p/ /b/ /t/ /d/ /k/ /g/ /ŋ/ /m/ /n/ /ɾ/ /f/ /v/ /s/ /ʃ/ /j/ /l/ /t/ /ŋ/ /ɲ/
Konsonant- grupper med to segmenter	/pl/ /bl/ /fl/ /kl/ /gl/ /pr/ /br/ /tr/ /dr/ /kr/ /gr/ /fr/ /vr/ /kv/ /gv ¹⁰ /tv/ /dv/ /sv/ /sm/ /fn/ /sn/ /kn/ /gn/ /pj/ /bj/ /fj/ /tj/ /dj/ /mj/ /nj/ /sp/ /sk/ /st/	/o/ /ɔ/ /ø/ /a/ /ɑ/	/rp/ /rb/ /rf/ /rk/ /rg/ /rt/ /rd/ /rl/ /rm/ /rn/ /rv/ /rj/ /rs/ /lp/ /lb/ /lf/ /lk/ /lg/ /lt/ /ld/ /lm/ /ln/ /lv/ /lj/ /ls/ /ŋk/ /ŋt/ /ŋd/ /ŋn/ /ŋs/ /nt/ /nd/ /ns/ /mp/ /mb/ /mf/ /mt/ /ms/ /vd/ /vl/ /vn/ /ft/ /fs/ /sp/ /sk/ /st/ /gd/ /gn/ /ts/ /kt/ /ks/ /pt/ /ps/ /ʃt/ /ʃk/ /ts/
Konsonant- grupper med tre eller flere segmenter	/spr/ /skr/ /skl/ /str/ /stj/ /spl/ /spj/ /skv/		/lsk/ /lst/ /lkt/ /lmt/ /rft/ /rps/ /rkt/ /nsk/ /nst/ /rmt/ /msk/ /mst/ /mnd/ /fst/ /ʃkt/ /ŋst/ /tsk/ /ksk/ /kst/ /skt/ /psk/ /pst/ /lrnt/ /lmsk/ /lmskt/

Tabell 2: segmentenes og segmentgruppenes tillatte plassering i den norske stavelsen

Vi ser her at norsk har mange konsonantgrupper, og at konsonantgruppene kan inneholde både én, to og tre segmenter i opptakt, og opptil fem segmenter i koda. Dette er i tråd med stavelsestemplatet i figur 3. De kursiverte konsonantene bryter med sonoritetshierarkiet.

2.3.2 Det somaliske stavelsestemplatet

I figuren under ser vi stavelsestemplatet i somali (basert på Saeed, 1999: 25).



Figur 4: somalisk stavelsestemplatet

¹⁰ Dette er en sjelden kombinasjon, men konsonantgruppa opptrer for eksempel i ordet *gvervil*.

Som i figur 3, har jeg har ikke markert posisjonene rim, opptakt, kjerne og koda, men kjerneposisjonen er under den loddrette streken, og x-en før danner opptaktsposisjon, mens x-en etterpå danner kodaposisjon. I somali er det altså bare tillatt med ett segment i opptakt, og ett segment i koda. Det er kun tillatt at vokaler danner stavelser, mens opptakt og koda kan fylles av konsonanter. I tabellen under ser vi en oversikt over mulige stavelsesstrukturer i somali.

		Koda	
		–	K
Opptakt	–	V	VK
	K	KV	KVK

Tabell 3: mulige stavelsesstrukturer i somali

Av tabell 3 ser vi at det er fire mulige stavelsesstrukturer i somali (V, VK, KV og KVK), og det kommer tydelig fram at konsonantgrupper ikke er tillatt i verken opptakt eller koda.

Under presenterer jeg segmentinventar og fonotaktiske stavelsesrestriksjoner i standard somali, som altså er basert på nordlig benadiir (basert på Saeed, 1999: 7-11).

	Opptakt	Kjerne	Koda
Segmenter	/b/ /d/ /d̥/ /g/ /g/ /ʔ/ /t/ /k/ /f/ /s/ /ʃ/ /χ/ /ħ/ /ʕ/ /tʃ/ /m/ /n/ /r/ /l/ /w/ /j/	/i/ /u/ /i/ /u/ /e/ /o/ /e/ /ö/ /a/ /æ/	/b/ /d/ /d̥/ /g/ /g/ /ʔ/ /f/ /s/ /ʃ/ /χ/ /ħ/ /ʕ/ /tʃ/ /h/ /m/ /n/ /r/ /l/ /w/ /j/

Tabell 4: segmentenes tillatte posisjoner i den somaliske stavelsen

Verdt å merke seg her er at verken /t/ eller /k/ kan opptre i den somaliske kodaen, og at segmentene /p/ og /v/ ikke brukes i somali. /tʃ/ er satt opp som mulig segment i koda, men dette gjelder kun for arabiske lånord. Det finnes som sagt ikke konsonantgrupper i somali. På orddanningsnivå finnes fire regler ifølge Saeed (1999:25-26):

- (1) Det er ikke tillatt med segmentet /m/ ordfinalt, dvs. at somali ikke har ord med /m/ i koda i den ordfinale stavelsen. Hvis en /m/ havner i den ordfinale posisjonen, blir den erstatta med /n/. Det er derimot lov å ha kodaer medialt i ordet som er fylt av segmentet /m/. Det finnes visse unntak til regelen mot /m/ ordfinalt. Dette når et påfølgende suffiks som begynner på /b/, blir satt inn, f.eks. <hambályo> ‘gratulerer’.

- (2) Ord kan ikke starte med vokal. Det betyr at opptakta i den ordinitialen stavelen må fylles med en konsonant. Ofte blir opptakta fylt av en glottal plosiv, /ʔ/, for å unngå ordinitial vokal, og dette segmentet er representert av tegnet <‘>.
- (3) I ord som er trisyllabiske, vil vokalen i andre stavelse falle bort om den tredje stavelsen er et suffiks. Dette gjelder bare når andre stavelse har en kort vokal. Om ordet ser slik ut: $K^1V^1K^2V^2K^3V^3$ vil altså V^2 falle bort om denne vokalen er kort.
- (4) Prosessen *sandhi* kan også virke inn på fonologien. Sandhi er at spesifikke grammatiske endelser endre segmentene i stammen av ordet. Dette går imidlertid på morfologi som er spesifikk for somali, og av den grunn kommer jeg ikke til å undersøke transfer av sandhi, selv om det ikke er umulig at de tar med seg dette inn i

2.4 Reparasjonsstrategier

Når en andrespråksinnlærer møter målspråket sitt med morsmålets stavelsestemplat som grunnlag, kan det oppstå avvik fra målspråkets uttalenorm. Innlæreren vil kunne tilpasse målspråkets ord og stavelser til å passe morsmålstemplatet. Dette kan vi se eksempler på i Broselow (1984). Hun har studert hvordan innlærere av engelsk med arabisk som morsmål realiserer engelske ord i sitt innlærerspråk. Her følger tre eksempler fra materialet hennes:

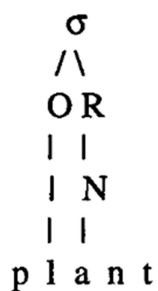
- (1) pilanti (‘plant’)
- (2) Fired (‘Fred’)
- (3) tiransilet (‘translate’)

Archibald (1998) forsøker å beskrive stavelседanningsprosessen til ordet /pilanti/. Han ser for seg at det engelske ordet /plant/ (som er et monosyllabisk ord i engelsk) blir festet til et arabisk stavelsestemplat, og at denne stavelседanninga har ulike faser. I framstillinga av prosessen under (fra Archibald, 1998) er engelsk fagterminologi brukt: R står for «rhyme», N for «nucleus», O for «onset» og C for «coda».

- (1) Vokaler blir satt inn i kjerneposisjonen, og kjerneposisjonen blir knytta til rimet.

σ
 |
 R
 |
 N
 |
 p l a n t

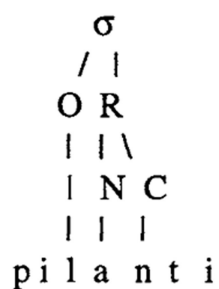
(2) Så mange opptaktskonsonanter som morsmålstemplatet tillater, blir knytta til stavelsen.



(3) Så mange kodakonsonanter som morsmålstemplatet tillater, blir knytta til rimet.



(4) Vokaler blir satt inn til høyre for konsonanter som ikke tilhører noen stavelse (/p/ og /t/).



(5) De innskutte vokalene, to /i/-er, blir plassert i en kjerneposisjon, og deretter knyttet til rimet i nye stavelser.

	σ	σ		σ			
		/					
	R	O	R	R			
			\				
	N		N	C	N		
	p	i	l	a	n	t	i

(6) Så mange opptaktskonsonanter som morsmålstemplatet tillater, blir knytta til stavelsen.

	σ	σ		σ			
	/	/		/			
	O	R	O	R	O	R	
			\				
		N		N	C		N
	p	i	l	a	n	t	i

Resultatet er altså det polysyllabiske ordet /pilanti/ som har gått fra å ha én til å ha tre stavelser, og dette har skjedd ved at det har blitt satt inn to ekstra vokaler. Arabisk stavelsestemplatet tillater kun én konsonant i opptakt og én konsonant i koda, og dermed kan ikke /p/ og /t/ knytte seg til vokalen /a/ – opptakt og koda er allerede fylt av /l/ og /n/. Nye vokaler blir dermed satt inn til høyre for disse lydene, slik at konsonantene kan danne opptakt i nye stavelser. Dette er én *reparasjonsstrategi* man kan bruke for å få et målspåksord til å passe inn morsmålstemplatet. Husby & Kløve (1998) presenterer fire reparasjonsstrategier innlæreren kan bruke for å «reparere» et målspåksord slik at det passer inn i morsmålstemplatet: *bortfall*, *epentese*, *metatese* og *erstatning* (s. 71-73).

Et bortfall er når en lyd i målspåket faller bort. For eksempel kan ordet /stram/ miste én eller flere av lydene i opptakt om morsmålstemplatet ikke tillater flere lyder i opptakt. Om for eksempel segmentet /r/ faller bort, og innlæreren realiserer målordet /stram/ som /stam/, er dette bortfall av /r/.

Epentese er synonymt med innsetting – man setter inn en lyd som ikke er der i målspåksordet. Dette kan være en annen strategi for å få et målspåksord med for stor konsonantgruppe til å passe inn i morsmålstemplatet, og det var det vi så eksempel på i ordet /pilanti/ over. Der passet ikke /plant/ inn i det arabiske stavelsestemplatet, og ekstra vokaler,

og dermed stavelser, ble satt inn slik at konsonantgruppene /pl/ og /nt/ ble brutt opp.

Metatese er når en lyd flyttes fra en posisjon til en annen. Kløve & Young-Scholten (2001) beskriver to typer metatese: *narrow* og *broad*. Førstnevnte er når to segmenter i et ord bytter plass. Sistnevnte er at segmentene i et ord endrer rekkefølge. I dette tilfellet bytter ikke to segmenter plass, det er bare ett segment som blir flyttet til en annen posisjon. Metatese, enten den er *narrow* eller *broad*, kan gjerne skje når man ikke tillater en lyd i en bestemt posisjon. For eksempel fant Kløve (1985) at innlærere med kantonesisk som morsmål uttalte /bi:l/ som /bli:/. Dette passer godt med at det ikke er tillatt med segmentet /l/ i koda i kantonesisk, og dermed har de bytta om /i/ og /l/, slik at /l/ havna i opptakta i stedet.

Den siste reparasjonsstrategien, erstatning, innebærer å bytte ut en lyd med en annen. I eksempelet vi brukte i forrige avsnitt, kunne innlærerne benyttet seg av erstatning som strategi i stedet. De kunne byttet ut /l/ med for eksempel /m/ (som er tillatt i kodaposisjon i kantonesisk), og dermed ville det passet inn i morsmålstemplatet (Husby & Kløve, 1998:81)

2.5 Universelle prinsipper for stavelsesdanning

I tillegg til at det finnes språkspesifikke regler for stavelsesdanning, finnes det universelle prinsipper, eller tendenser, for hvordan en stavelse blir dannet. Dette er regler som de fleste språk i verden følger, selv om en god del språk også bryter med dem. I mellomspråk (se kap 4) er det i mange tilfeller observert at innlæreren følger de universelle prinsippene for stavelsesdanning selv om målspråket ikke følger dem. For min oppgave er de to viktigste universelle prinsippene *opptaktsprinsippet* og *sonoritetsprinsippet*.

2.5.1 Opptaktsprinsippet

Opptaktsprinsippet går ut på at stavelser tenderer til å ha opptakt, altså at de som oftest har fylt opptaktsposisjonen med en lyd. Som Kløve & Husby (1998) sier, er det «en universell tendens til å foretrekke opptakt hvis det er mulig» (s. 74). Det er altså ikke slik at alle stavelser følger denne regelen; frasen «i dag», for eksempel, starter jo med stavelsen «i», som ikke har opptakt. Likevel er det slik at de fleste stavelser i verdens språk har opptakt, og etter det vi vet, har alle språk stavelser med strukturen KV (ibid.)

2.5.2 Sonoritetsprinsippet

Sonoritetsprinsippet handler om at stavelser tenderer til å ordne lydene sine etter et sonoritetshierarki. Når en lyd er sonor, betyr det at den er klangfull. I en stavelse skal den mest sonore lyden være kjernen, og sonoriteten skal avta mot begge ender av stavelsen. Vi sier at den mest sonore lyden i stavelsen, dvs. kjernen, er *sonoritestoppen*. Lydene som er

nærmest kjernen, enten det er i opptakt eller koda, skal være så sonore som mulig, men de skal ikke være mer sonore enn kjernen. Jo lenger ut vi kommer fra stavelseskjernen, jo mindre sonore blir lydene. Om man måler sonoritet rent fysisk, kan man ordne ulike typer lyder i et sonoritetshierarki, som jeg oppgir i (1) (Selkirk, 1984).

- (1) Plosiver, frikativer, nasaler, likvider, glidelyder, **vokaler**, glidelyder, likvider, nasaler, frikativer, plosiver

I tabellen under er en oversikt over segmentene og hvilken av de ovennevnte klassene de hører til (etter Kristoffersen, 2000:13, 22; Husby & Kløve, 1998:75).

Opptakt					Kjerne	Koda				
P	F	N	L	G	V	G	L	N	F	P
p ^h	s	m	l		i y a u u		l	m	s	p ^h
b	ʃ	n	r		ε æ e o o		r	n	ʃ	b
t ^h	ç				ø a a			ŋ		t ^h
d	h			j						d
k ^h	f								f	k ^h
g	v								v	g

Tabell 5: segmenter i standard østnorsk organisert etter sonoritetshierarkiet

Under er lydene i somali organisert etter sonoritetshierarkiet (Saeed, 1999:7-11).

Opptakt					Kjerne	Koda				
P	F	N	L	G	V	G	L	N	F	P
b	f	m	r	w	i u e o a	w	r	m	f	b
t	s	n	l	j	i u e ö æ	j	l	n	s	d
d	ʃ								ʃ	d
ɖ	χ								χ	g
k	ʔ								ʔ	g
g	ħ								ħ	
ɠ	h								h	
ʔ	ç								ç	

Tabell 6: segmenter i nordlig somali organisert etter sonoritetshierarkiet

Her ser vi altså at plosivene er de minst sonore lydene, mens vokalene er de mest sonore lydene. Ut ifra sonoritetsprinsippets regler er det altså slik at stavelser helst ikke skal ha en nasal lyd før en plosiv i opptakt fordi den minst sonore lyden skal komme først. Det er

derimot greit å ha en nasal før en plosiv i koda nettopp fordi den første lyden i koda er nær stavelseskjernen, og helst skal være ganske sonor. Det finnes ingen ord i norsk, etter det jeg vet, som starter på «mp», men det finnes derimot ord som slutter på «mp», slik som «fjomp». Dermed kan vi se at norsk følger sonoritetsprinsippet på dette punktet. Det er likevel ikke slik at alle stavelser i alle språk følger dette prinsippet, og norsk har mange tillatte konsonantgrupper som bryter med sonoritetshierarkiet. Disse er markerte i kursiv i kap 2.3.1 om det norske stavelsestemplatet i tabell 2 om norsk fonotaks.

3.0 Skrift står for tale

I denne studien skal jeg studere skrift, eller skriftspråk, som det ofte kalles. Å kalle skrifta et språk er imidlertid misvisende, for skrifta er egentlig bare en måte å uttrykke det talte språket på. Vi sier at det talte språket er primært, og skrifta er sekundær. Det er like fullt interessant å se hvordan skrifta blir realisert i ulike språk, og det er stor variasjon i forholdet mellom lyd og skrift fra språk til språk.

For det første varierer det hvilket *skriftsystem* et språk har. Et skriftsystem (som jeg kommer til å omtale som alfabet i kapitlene om norsk og somalisk ortografi), er et sett symboler som blir brukt for å uttrykke det naturlige, talte språket. Eksempler på er kyrillisk, hiragana, geez, latinskrift og runer. Mange skriftsystem brukes i flere språk: kyrillisk, for eksempel, brukes til å uttrykke russisk, serbisk, ukrainsk, mongolsk og kirgisisk, for å nevne noen. Andre skriftsystem brukes bare i ett språk, for eksempel hiragana, som kun brukes for å uttrykke japansk. Det er også slik at språk kan veksle mellom ulike skriftsystemer. Serbisk kan både uttrykkes med kyrillisk og latinskrift, og japansk har tre ulike skriftsystem, hiragana, katakana og kanji, som de bruker i ulike situasjoner (Bugge, 2019:16,31).

Skriftsystem blir gjerne modifiserte, slik at de passer til det språket de skal uttrykke. Latinskrifta, for eksempel, brukes blant annet til å uttrykke norsk, polsk og fransk, og latinskriftene brukt i disse språka har mange symboler til felles. Likevel er det noen særnorske tegn (f.eks. <æ>), særpolske tegn (f.eks. <ł>) og særfranske tegn (f.eks. <ç>). I tillegg kan tegn som ser helt like ut i de ulike språka, uttrykke forskjellige lyder. Tegnet <o> står i mange tilfeller for lyden /u/ i norsk, men for lyden /o/ på fransk.

Husby & Kløve (1998:120-121) skriver at alle verdens skriftsystem faller inn under ett av følgende tre typer: logografiske skriftsystem, syllabariske skriftsystem og alfabetiske skriftsystem. Logografiske skriftsystem baserer seg på ord og uttrykk, med andre ord konsepter, og har altså ett tegn per konsept. Den kinesiske skrifta er et eksempel på et slikt system, der for eksempel det ene tegnet 池 betyr «svømmebasseng» på mandarin. Syllabariske system, derimot, baserer seg på stavelser, og har ett tegn per mulig stavelse i språket. Et eksempel er geez-skrifta, som er mye brukt for å uttrykke språk på Afrikas horn. For eksempel bruker språket *tigrinja* denne skrifta, og i tigrinja uttrykker tegnet 𐌆 stavelsen /re/. Når det gjelder alfabetiske skriftsystem, uttrykker disse enkeltsegmenter, slik som runeskrifta (mer spesifikt eldre futhark) som uttrykker segmentet /m/ med runen 𐌆 i språket *urnordisk*.

I denne studien står språka somali og norsk i sentrum. Begge disse språka uttrykkes med latinskrifta, og tar altså i bruk et alfabetisk system der tegn uttrykker segmenter. Det er likevel forskjeller mellom de to skriftsystema; de har begge modifikasjoner som er særegne for sitt språk, og de har tegn som er felles, men som uttrykker ulike lyder. I kap. 3.2 og 3.3 gjør jeg rede for forholdet mellom lyd og skrift i de to språka, men først skal vi se litt mer på termen *ortografi*.

3.1 Ortografi

I denne studien er det ortografi som er studieobjekt. Derfor er det på sin plass å redegjøre for hva ortografi er. Apel (2011) beskriver ortografi slik:

The word orthography is derived from two Greek roots: orthos, meaning correct, and graphein, meaning to write. As such, orthography literally means correct writing. Orthographic knowledge represents the information that is stored in memory that tells us how to represent spoken language in written form; borrowing from the word's etymology, it is knowledge for the correct way to write language (s. 592).

Ortografi er altså den 'korrekte' skriftlige representasjonen av tale. Hva som er en korrekt representasjon av tale, bestemmes av normer eller konvensjoner i et bestemt språk. Det varierer fra språk til språk hva som er rett skriftlig representasjon av det aktuelle språket, og en kan si at ulike språk har ulike ortografier.

For å beskrive dette forholdet mellom lyd og skrift med fagtermer, kan vi si at det talte språket består av segmenter, og at disse kan representeres med *grafemer*. For eksempel uttrykker grafemet <t> lyden /t/. Noen ganger kan ett grafem bestå av flere tegn, og slike grafemer kalles *polygrafemer*. I norsk har vi to typer polygrafemer: digrafer og trigrafer. Et eksempel på en digraf i norsk er <sj> i <sjø> som blir uttalt /ʃ/. Her ser vi altså at ett segment blir skrevet med to tegn, som danner et polygrafem. Et eksempel på en trigraf er <skj> for /ʃ/, som i ordet <skje>, uttalt /ʃe/.

Det varierer hvor *ortofone* språk er, dvs. i hvor stor grad det er et en-til-en-forhold mellom segment og grafem. Om en har høy grad av samsvar mellom lyd og tegn, der ett grafem representerer ett segment, og omvendt, kan vi si at et språk har høy grad av ortofoni, mens hvis det f.eks. er mange ulike grafem-representasjoner for en lyd, er ikke språket særlig ortofont. Et lite ortofont språk kan også ha mange ord der grafem blir skrevet, men ikke uttalt.

3.2 Norsk ortografi

a b c d e f g h i j k l m n o p q r s t u v w x y z æ ø å

Tabell 7: det norske alfabetet

Over ser vi det norske alfabetet slik det er normert i dag – en modifisert versjon latinskrifta der grafemene æ, ø og å er særskandinaviske. Når en kan alfabetet på rams, betyr det at man har oversikt over alle symbolene brukt i det norske språket – nesten. Det finnes grafemer vi har lånt inn fra andre språk, slik som <é> (t.d. i André) med *diakritiske* tegn, dvs. tilleggstejn som kan modifisere et tegn. Men de aller fleste symbolene i det norske språket er representerte blant de 29 bokstavene i alfabetet. Det er imidlertid ikke slik at alle grafemer er representerte i alfabetet vårt. Polygrafemer er ikke tatt med i alfabetet; grafemsamband som <skj> for /ʃ/ er fraværende. I noen ord kan dessuten et tegn ha en annen lydverdi enn om den står alene. Et eksempel er grafemet <j> som isolert uttales /j/, men som kan uttales /ʃ/ i lånordet «journalist».

I kap. 5.4 la jeg fram en oversikt over *segmentinventaret* i standard østnorsk, og i hvilke posisjoner i stavelsen de ulike segmentene kan stå. I tabellen under vil jeg legge fram en oversikt over hvordan *konsonantsegmentene* kan bli presenterte som enkeltgrafemer og polygrafemer (basert Husby & Kløve, 1998:127-128). Grunnen til at jeg ikke tar for meg vokalsegmentenes grafemikk, er at dette ikke er relevant for stavelsesstrukturers oppbygging eller for reparasjonsstrategier. Kolonna *Posisjon* viser hvor i stavelsen lyden kan opptre.

Segment	Grafem	Posisjon	Eksempelord
/p/	<p>	Opptakt og koda	p akkesel, g ap
/b/		Opptakt og koda	b rillesfin, l ab
/t/	<t>	Opptakt og koda	t rubadur, h at
	<th>	Opptakt og koda (navn)	T hor, L isbeth
	<dt>	Koda	b lind t
/d/	<d>	Opptakt og koda	d artbrett, t juvradd
/k/	<k>	Opptakt og koda	k opi, t rakt
	<ch>	Opptakt (navn)	C hristine
	<q>	Opptakt (navn)	D ah l qvist
/g/	<g>	Opptakt og koda	g okk, m yg g
/s/	<s>	Opptakt og koda	s neip, f jas

	<ps>	Opptakt	pseudonym
/ʃ/	<sj> <skj> <sk> <g> <j>	Opptakt og koda Opptakt Opptakt Opptakt Opptakt	sjø, æsj skjødesløs skite gelé journalist
/ç/	<kj> <k> <tj>	Opptakt Opptakt Opptakt	kjeledress kime tjafs
/h/	<h>	Opptakt	hufse
/f/	<f> <ph>	Opptakt og koda Opptakt	fistel, biff ex.phil
/v/	<v> <hv> <w>	Opptakt og koda Opptakt Opptakt (navn)	vaktel, stav hvorledes Wenche
/m/	<m>	Opptakt og koda	mammut, svinekam
/n/	<n> <nd>	Opptakt og koda Koda	nøktern, bavian land
/l/	<l>	Opptakt og koda	list, valnøtt
/r/	<r> <rd>	Opptakt og koda Koda	rakker, bærbar bord
/j/	<j> <gj> <g> <hj> <lj>	Opptakt Opptakt Opptakt Opptakt Opptakt	jerv gjev gyselig hjelp ljome
/ŋ/	<ng> <nt>	Koda Koda	bang observant
/t/	<rt>	Koda	dyrebart
/d/	<rd>	Koda	lørdag
/ŋ/	<rn>	Koda	barnehage

Tabell 8: forholdet mellom segmenter og grafemer i norsk

I denne tabellen ser vi tydelig at det ikke er et en-til-en-forhold mellom segment og grafem i det norske språket. Mange lyder kan vi uttrykke med flere ulike grafem; /ʃ/, for eksempel, kan uttrykkes med hele fem ulike skriftuttrykk, og /j/ kan uttrykkes skriftlig på fire forskjellige måter. Det finnes også flere lyder som ikke har sitt eget symbol, men som kun kan uttrykkes med polygrafem, for eksempel /ŋ/ og /d/.

Ettersom det er bokmål informantene mine har lært seg og skriver i språktestene jeg studerer, er det bokmålsnormen av 2005 som er relevant for mine elever. Når en snakker om en slik norm, handler den i stor grad om ordnivået, og i en ortografisk sammenheng hvordan

en skriver et ord korrekt etter normen. For eksempel skriver vi <land> med en <d> på slutten som ikke uttales i det hele tatt. Evt. kan en se det som at /n/ uttrykkes med <nd>. Disse *stumme* konsonantene kommer av historiske årsaker – de har blitt uttalt i norsk eller forløpere til det norske språket, før en slutta å uttale dem, men ikke endra ortografien deretter (Husby, u.d.). Andre eksempler er substantiv i bestemt for intetkjønn entall, som <uret> som ikke har uttalt /t/ og <jorde> som ikke har uttalt /d/.

3.3 Somalisk ortografi

a b c d dh e f g h i j k kh l m n o q r s sh t u w x y

Tabell 9: det somaliske alfabetet

Over ser vi en framstilling av det somaliske skriftsystemet. Dette skriftsystemet er altså basert på latinskrifta. Vi ser at det er modifisert for å passe til somali i det at de ikke har grafema <p v z>. De har selvsagt heller ikke de særskandinaviske tegna <æ ø å>. Vi ser i tillegg at de benytter tre digrafer, <dh kh sh>, som ikke brukes i det norske skriftsystemet (Husby, 2001:13).

Somali er et svært ortofont språk. Det kan vi se i tabellen under, der grafemenes representasjon av hver lyd er satt opp. Det er et en-til-en-forhold mellom segment og grafem i det at alle segment kun kan bli representert med ett grafem, og at alle grafem bare kan stå for ett segment (basert på Saeed 1999:7-11).

Segment	Grafem/polygrafem	Posisjon	Eksempelord
/b/		Opptakt Koda	bi iq ‘reddhare’ hu b ‘våpen’
/d/	<d>	Opptakt Koda	da yax ‘måne’ ge d ‘tre’
/d/	<dh>	Opptakt Koda	dhe emman ‘diamant’, fe d h ‘kam’
/g/	<g>	Opptakt Koda	Gi riig ‘Hellas’ lu g ‘bein’
/g/	<q>	Opptakt Koda	qa bow ‘kald’ ca q liga ‘fornuft’
/t/	<t>	Opptakt	tu g ‘tyv’
/k/	<k>	Opptakt	ka al ay ‘kom’
/f/	<f>	Opptakt Koda	fa l ‘verb’ iskhila af san ‘paradoksal’
/s/	<s>	Opptakt Koda	qo so l ‘latter’ a sl i ‘ekte’

/ʃ/	<sh>	Opptakt Koda	nashir ‘kringkaste’ haashtari ‘askebegeter’
/x/	<kh> (kun i arabiske lånord)	Opptakt Koda	makhaayad ‘kafé’ akhlaakh ‘etikk’
/h/	<x>	Opptakt Koda	xanjo ‘tyggegummi’ jeex ‘fragment’
/h/	<h>	Opptakt Koda	hablobaas ‘vannkopper’ leh ‘eie’
/ʒ/	<c>	Opptakt Koda	macallin ‘lærer’ kac ‘økning’
/tʃ/	<j>	Opptakt Koda (kun i arabiske lånord)	jaamici ‘akademiker’ birijj ‘bro’
/m/	<m>	Opptakt Koda (ikke finalt)	ilmo ‘barn’ sambab ‘lunge’
/n/	<n>	Opptakt Koda	nabadeed ‘fredelig’ san ‘nese’
/r/	<r>	Opptakt Koda	rabshoole ‘bråkmaker’ firfircoon ‘iherdig’
/l/	<l>	Opptakt Koda	nolól ‘liv’
/j/	<y>	Opptakt Koda	yar ‘liten’ Nayjeeriya ‘Nigeria’
/w/	<w>	Opptakt Koda	war ‘nyheter’ gaw ‘presis’
/ʔ/	<’>	Opptakt Koda	gogo’an ‘oppstykket’ lo’ ‘storfe’

Tabell 10: forholdet mellom segmenter og grafemer i somali

I framstillinga kan vi se at grafemene <c j q w y x> har andre lydverdier enn i norsk, og de fleste av disse representerer også lyder som ikke finnes i det østnorske segmentinventaret. Samtidig kan vi se at det kun er ett grafem per lyd, og at situasjonen er en ganske annen enn i norsk; her har vi å gjøre med et svært ortofont språk. Vi ser i tillegg at nesten alle grafemene er inkludert i alfabetet. Unntaket er <’>. Morsmålsbrukere av somali som har lært seg å lese og skrive, er med andre ord vant til at de får en (så å si) fullstendig oversikt over mulige grafemer i skrift når de lærer alfabetet.

Det at somali er et såpass ortofont språk, kan nok tilskrives at en offisiell somalisk ortografi ikke ble innført før i 1972. Før den tid hadde kolonispråka italiensk og engelsk, i tillegg til arabisk vært språka en skrev. Dette endret seg imidlertid fort med innføringa i 1972 – en stor kampanje ble satt i gang for å få flest mulig til å skrive somali, og overalt kunne en lese slagordet «Haddad taqaan bar, haddaanad oqoon barod» (‘Hvis du kan det – lær det bort,

hvis du ikke kan det – lær det'). I det offentlige erstattet skriftlig somali de andre språka som tidligere hadde regjert. Når en ortografi er såpass ny, er det mange færre tilfeller av «stumme lyder», som vi så at norsk hadde mye av, nettopp fordi skriftbildet er tilpasset det nåværende lydsystemet, og ikke et gammelt lydsystem som har hatt god til på å utvikle seg (Andrzejewski, 1974:4-7).

4.0 Mellomspråk

I oppgava mi ser jeg på språket til innlærere av norsk med somali som morsmål. Dermed blir det viktig å gjøre rede for og diskutere termen *mellomspråk*. Det var Selinker (1972) som introduserte *mellomspråkhypotesen*. Han skriver at mellomspråk er:

[...] the existence of a separate linguistic system based on observable output which results from a learner's attempted production of a TL¹¹ norm. This linguistic system we will call «interlanguage» (IL¹²) (s. 214).

Han beskriver det som et unikt språkssystem som bør studeres for seg, og ikke sees på som avhengig av målspråk eller morsmål. Dette er et kognitivistisk motsvar til behavioristene, som har morsmålet og målspråket som målestokk, og som mener at det er innputten innlæreren får fra målspråket, som er avgjørende for hvordan man tilegner seg et andrespråk. I motsetning til behavioristene ser Selinker på andrespråkstilegnelse som «ein kognitiv prosess med kreativ konstruksjon» (Nistov & Nordanger, 2017:266). Det er altså flere faktorer enn innputt som spiller en rolle.

I Selinkers artikkel legges det vekt på at «voksne bruker en latent *psykologisk* struktur for å tilegna seg eit andrespråk» (Nistov & Nordanger, 2017:263). Han legger videre fem prosesser for innlæring til grunn:

1. overføring fra morsmålet
2. overføring fra undervisning og undervisningsmateriale
3. læringsstrategier
4. kommunikasjonsstrategier
5. overgeneralisering av regler i målspråket

Disse fem kjennetegnene mener Selinker er felles for all andrespråkstilegnelse, og punkt 1 og 5 kan være med på å forklare den store utfordringa til andrespråksinnlærere, nemlig *fossilisering*. Dette er at enkelte språktrekk aldri mestres i samme grad som de ville blitt mestra av en førstespråksinnlærer.

Med Selinker blir også performansen sentrum for forskninga, og han mener at man må studere faktisk språkbruk eller «the surface structure of IL sentences» (Selinker, 1972:35 i Nistov & Nordanger, 2017:266), dette i motsetning til generativistene, som i stor grad

¹¹ TL står for *target language*, dvs. målspråk

¹² IL står for *interlanguage*, dvs. mellomspråk

studerer det mentale systemet i hodet (ibid.) Samtidig har han det til felles med generativistene at han ser for seg en slags innebygd læreplan som inneholder en læringsløype en må følge for å tilegne seg et nytt språk. Dette bygger på Corder (1967), som la stor vekt på en slik «built-in-syllabus», og Corders tankesett var igjen i stor grad inspirert av Chomskys universalgrammatikk.

Selinkers mellomspråksbegrep var grensesprengende da det kom, og det har vært et sentralt bidrag på andrespråksfeltet. Samtidig har mye skjedd på forskningsfeltet siden Selinkers artikkel kom ut, og det er nyere teorier som utfordrer deler av hypotesen hans. For det første er den regelbaserte tilnærminga hans kritisert. Det at språk er basert på regelgenerering, mener mange er en forenkling av det store bildet. En kan tenke seg at en *eksemplarbasert tilnærming* er vel så viktig i en innlærersituasjon. Når en lærer seg et eksemplar, altså en konstruksjon i språket, danner dette representasjoner i minnet, og for hver gang dette ene eksemplaret blir repetert, vil også representasjonen i minnet bli forsterka (Nistov & Nordanger, 2017:278)

For denne oppgava er kanskje den viktigste utviklinga fra Selinker til i dag, den rådende *kompleksitetsteorien*. I tradisjonen etter Selinker og Corder har det ofte blitt lagt vekt på *interlocking systems* – sammenfletta systemer –, og i dette ligger det at det er «ein serie av system som avløyser kvarandre, men som likevel er samanfletta» (Nistov & Nordanger, 2017:279) Her ser vi altså et lineært syn på læring der en er innom ett system før kan komme til neste ‘steg’ i innlæringsprosessen. De som støtter seg til kompleksitetsteorien, mener det ikke er så systematisk som så. De tenker at språkinnlæringsprosessen er mer dynamisk, og at det nivået man er på, blir påvirka av interaksjon og situasjon. Språk *må* være i kontinuerlig endring, men dette blir ikke styrt av en fast læringsrekkefølge. I stedet er systemet *adaptivt* og endrer seg i takt med omgivelsene, uansett om dette måtte være nærmere en målspråksform eller lengre unna en målspråksform (ibid.)

4.1 Andrespråk, målspråk, fremmedspråk, S2 og L2

I resten oppgava kommer jeg til å bruke termen *andrespråk*. Det betyr ikke nødvendigvis at det språket som blir lært, er det andre språket en person lærer – det kan også være det tredje, fjerde, eller åttende for den sakens skyld. Det viktige er at dette språket blir lært *etter* at morsmålet, eller førstespråket. Det er også mulig å ha flere førstespråk, og det er ingen klar grense for hvor gammel en språkinnlærer må være før det skal regnes som andrespråkstilegnelse i stedet for førstespråkstilegnelse. Det har imidlertid blitt foreslått at de språka en lærer seg før treårsalderen, er å regne som førstespråk, mens de en lærer etter denne

alderen, er å regne som andrespråk. Dette er på grunnlag av at innlæringsprosessen før og etter denne alderen ser ut til å være ulik (Viberg, 1978). Mye litteratur bruker termene S1 (språk én) og S2 (språk to) om førstespråk og andrespråk, og dette er også termer jeg kommer til å bruke i oppgava mi. De engelske variantene er L1 (language one) og L2 (language two), som vi kommer til å møte på når jeg referer til engelsk forskning på feltet. I Norge er vi også godt vant med fagtermen *fremmedspråk*, og denne termen har i stor grad blitt brukt om andrespråk en lærer i skolen. Berggreen & Tenfjord (1999) beskriver forskjellen på andrespråk og fremmedspråk slik: «Mens fremmedspråksinnlæreren ofte kan klare seg med en passiv forståelse og lite automatiserte ferdigheter, må andrespråksinnlæreren kunne hente fram kunnskapene sine raskt, omstille seg til stadig nye situasjoner og produsere språk selv i mye større grad» (s. 17). Jeg vil kalle informantene mine andrespråksinnlærere fordi de bor i Norge, er en del av det norske samfunnet, og ofte trenger å hente fram de norske språkferdighetene sine i møte med det norsktalende samfunnet. De lærer altså ikke norsk på en 'skolsk' måte, slik en fremmedspråksinnlærer ville gjort. Jeg vil også referere til norsk som *målspråket* til disse informantene, og jeg vil referere til den sannsynlige norske ekvivalenten til et ord som *målspråksord*

4.2 Transfer

I problemstillinga mi spør jeg meg om man kan finne spor av somaliske stavellesstrukturer i mellomspråket til innlærere av norsk som somali som morsmål. Studien min er med andre ord en transferstudie. Derfor er det naturlig å gjøre rede for hva transfer er. Odlin (1989) gir følgende definisjon:

Transfer er den innflytelsen som kommer som et resultat av likheter og forskjeller mellom målspråket og hvilket som helst annet språk som innlæreren tidligere (og kanskje ufullstendig) har tilegnet seg (oversatt av Golden & Tenfjord, 2015:183).

Som nevnt i forrige kapittel nevner Selinker (1972) språklig overføring fra morsmålet som et kjennetegn ved andrespråksinnlæringsprosesser, og dette er en del av definisjonen til Odlin (1989). Men Odlin passer også på å nevne at det ikke bare er morsmålet som trenger å være involvert i transfer – en kan se transfer også fra andre andrespråk en har tilegna seg. Når en snakker om transfer fra ulike språk en har vært i kontakt med, bruker man gjerne fagtermen *tverrspråklig innflytelse* for å markere hvor sammensatt innflytelsen i mellomspråket kan være (Golden & Tenfjord, 2015:183). I denne studien undersøker jeg i hovedsak morsmålstransfer, men jeg kommer til å diskutere om andre språk informantene har lært, kan

spille en rolle for stavellesstrukturene mellomspråket deres.

Det er i dag bred enighet om at førstespråket har noe å si for hvordan man tilegner seg et andrespråk, og innenfor andrespråksfeltet «blir transfer forstått som eit psykologisk fenomen som eksisterer på individnivå» (Gujord & Ragnhildstveit, 2017:134). Ulike teoretiske retninger innenfor andrespråksforskninga har ulike måter å tolke transfer på, men de har alle til felles at de tenker at språk som en allerede har lært seg, er «ein del av dei samla ressursane som andrespråksinnlæreren rår over. Dette er først og fremst implisitt kunnskap» (ibid.:136). Så er det altså slik at det må ligge til grunn en prosess for at de ressursene man har fra tidligere språk, aktiveres når man lærer seg et andrespråk. Dette er først og fremst en ubevisst og implisitt prosess som innlæreren er lite bevisst. Cummins (2013) beskriver det slik:

[...] there is an underlying proficiency or knowledge base that is common across languages. This common underlying proficiency makes possible the transfer of concepts, literacy skills, and learning strategies from one language to another (s. 291-292).

Det er altså denne ressursbanken, som er satt sammen av alle språk innlæreren har lært helt eller delvis, som gjør transfer *mulig*. Men det er ikke slik at alle trekk fra tidligere lærte språk har like stor sannsynlighet for å bli overført, og et viktig begrep her er *overførbarhet*. Noen språktrekk har egenskaper som gjør dem mer overførbare enn andre trekk. Dette er ofte trekk som er *umarkerte*, dvs. trekk som finnes i mange språk, og disse blir lettere tatt i bruk i mellomspråket enn et språktrekk som er veldig markerte og dermed sjeldne. Det er også tendenser til at transfer oftest skjer tidlig i innlæringsfasen av et andrespråk, men det er eksempler på at enkelte trekk ikke viser seg som transfer før man har kommet til en mer viderekommen fase av andrespråksinnlæringa. Det er også mange faktorer som kan påvirke om det oppstår transfer eller ikke, for eksempel hvor mye språkinntak man får fra målspråket, modning, og andre individuelle faktorer (Gujord & Ragnhildstveit, 2017:137). Transfer er også et fenomen som gjerne ikke «manifesterer seg som klare døme på ikkje-meistring eller meistring» (ibid.:138). Dermed kan det være varierende om transfer er konsekvent for den enkelte språkbruker eller ikke. I studien min er informantene på et relativt høyt nivå i norsk, og det blir interessant å se om dette gjør at det er liten grad av transfer, eller om vi finner høy grad av transfer.

I Odlin (1989) sin definisjon av transferbegrepet, blir det også nevnt at transfer kan være overføring av *både* trekk som er like i S1 og S2, og trekk som er ulike i S1 og S2. En overføring av et trekk som er felles i de to språka, kaller vi *positiv transfer*, mens vi kaller

overføring av et trekk fra morsmålet som ikke finnes i målspråket, for *negativ transfer* (Kulbrandstad & Ryen, 2017:37). I denne oppgava forsker jeg på negativ transfer ved at jeg undersøker om en struktur fra somali viser seg i ord som har en annen struktur i målspråket norsk. En annen betegnelse som beskriver studien min, er *forward transfer* – fremovertransfer. Dette er et begrep på transfer som ser på førstespråks innflytelse på andrespråk. Det finnes studier der man undersøker transfer av språktrekk fra andrespråk til førstespråk, og også studier der man undersøker hvordan de involverte språka i en tilegnelsesprosess kan påvirke hverandre, men det kommer ikke jeg til å undersøke i denne studien (Gujord & Ragnhildstveit, 2017:138).

4.3 Oppsummering av kapittelet

I dette kapittelet har jeg presentert mellomspråksbegrepet til Selinker og hvordan synet på mellomspråk har endret seg i fagmiljøet siden 1972. Spesielt er endringa til å se på mellomspråk som dynamiske språk et viktig poeng. I denne oppgava kommer jeg til å analysere språket til informantene mine som et mellomspråk, altså et språk som står på egne ben og *i seg selv* er interessant å studere. Jeg kommer også til å ha ideen om dynamiske språk i bakhodet når jeg ser på språket til informantene mine.

Jeg har også definert noen sentrale termer som jeg kommer til å bruke mye videre i oppgava. Disse er andrespråk, mellomspråk, mellomspråksord, S1, S2, L1 og L2.

Transferbegrepet har jeg også presentert, og jeg har nevnt at jeg kommer til å se på negativ transfer. Det vil si at jeg kommer til å undersøke om strukturer fra somali er å finne i mellomspråket til informantene mine. Strukturene jeg undersøker *finnes* i norsk, men jeg vil undersøke om de blir brukt i mellomspråksord der de ikke ville blitt brukt i norsk. Vi har vært innom begrepet tverrspråklig innflytelse, og denne termen gir rom for at mer enn bare morsmålet kan påvirke språket – også andre ord informantene har lært, *kan* ha noe å si for strukturene jeg undersøker. Jeg kommer i hovedsak til å undersøke morsmålstransfer, men i diskusjonsdelen vil jeg også komme inn på hva andre språk kan ha å si for mellomspråket. Jeg har også nevnt at jeg undersøker framovertransfer, altså hvordan morsmålet, somali, kan påvirke mellomspråket som oppstår når informantene lærer seg norsk.

5.0 Tidligere forskning

I dette kapitlet vil jeg presentere tidligere forskning på stavelser og ortografi i en andrespråkssammenheng. Jeg starter med å presentere forskning på stavelser i andrespråkstilegnelse, og gjør rede for forskning på dette i Norge. Jeg fortsetter med å se på ortografi i andrespråk, og presenterer en studie i britisk kontekst, og de to eneste studiene som ser på norsk andrespråksortografi. Jeg kommer også til å gjøre greie for det jeg kjenner til av forskning på språkmøtet somali-norsk, i tillegg til forskning på lånord i somali og stavelens innvirkning på ortografien. Målet med kapitlet er å lage en oversikt over studier som har likhetstrekk med min studie, og se på hva slags funn som har blitt gjort i disse studiene. På den måten kan jeg plassere studien min inn i et forskningslandskap og gjøre rede for hva forskninga mi kan bidra med. I tillegg kan jeg undersøke om det er lignende funn i datamaterialet mitt, og hva som eventuelt skiller seg ut i dataene mine.

5.1 Forskning på transfer av stavelserstrukturer i andrespråk

Jarvis & Pavlenko (2008) legger fram en god oversikt over transferstudier gjort på stavelserdannning. De skriver at «most of the work on CLI effects¹³ in this area has dealt with problems that L2 learners have in producing consonant clusters in syllable onsets and codas (i.e., at the beginnings and ends of syllables)» (s. 67). Som i min studie, har altså studiene på dette feltet konsentrert seg om komplekse stavelser med flere konsonanter i opptakt og koda, og hvordan innlærere håndterer disse. Det har blitt undersøkt innlærere med mange ulike morsmålsbakgrunner og mange ulike målpråk, og i det følgende presenterer jeg to studier på feltet: én finner reparasjonsstrategiene erstatning, epentese og metatase og en som finner reparasjonsstrategien metatase.

Hansen (2001) undersøker kodaer i den muntlige produksjonen til innlærere av engelsk med kinesisk som morsmål. Kodaene hun undersøker består av henholdsvis én, to og tre segmenter. Hun finner at det er tendenser til at innlærerne bruker erstatning¹⁴ som metode i kodaer med ett segment, epentese som metode i kodaer med to segment og bortfall i kodaer med tre segment. Hun diskuterer om avvika fra engelsk målpråk kan skyldes transfer, sonoritetsprinsippet eller markerthet, og konkluderer med at alle disse prinsippene spiller inn i ulik grad stavelserdanninga hos innlærerne (s. 362).

Kløve & Young-Scholten (2001) ser nærmere på reparasjonsstrategien metatase. Her

¹³ Her står CLI effects for «cross linguistic influence effects», dvs. tverrspråklig innflytelse.

¹⁴ *Feature change* er termen brukt i Hansen (2001).

presenteres studiene til Kløve (1985, 1997) og Young-Scholten, Akita & Cross (1999). Kløve er en av få studier av stavellesdannning i N2. Hun undersøker metatase i produksjonen til innlærere av norsk med kantonesisk som morsmål. Ettersom hun, i likhet med meg, undersøker innlærere av norsk, er det interessant å se på noen eksempler fra studien hennes:

- (1) S2 [vist] – S1 [vits]
- (2) S2 [tu:ɤ] – S1 [tɕu:]
- (3) S2 [gɿv] – S1 [gɿlv]

I (1) ser vi at [t] har byttet plass med [s], og vi ser eksempel på en *smal metatase* (jf. kap. 2.4) der to konsonanter har bytta plass. I (2) ser vi et eksempel på samme type metatase, men her har en vokal og en konsonant bytta plass. I (3) ser vi eksempel på *bred metatase* slik at et segment har flyttet seg til en annen posisjon i ordet. I (1) foreslår forskeren at metatesen skjer fordi [st] i koda ikke er i tråd med sonoritetshierarkiet. (2) kan fint forklares med at kantonesisk ikke tillater mer enn én konsonant i opptakt. Som Kløve påpeker, bryter (3) med det kantonesiske stavellestemplet i det at metatesen resulterer i en kompleks opptakt: «liquids are fronted, even when they result in consonant clusters disallowed in the learner's L1s» (s. 117). I (3) er [l] en likvid lyd. Hun forklarer metatase av likvider slik: «... the more resonant a sound is, the more susceptible it is to metathesis, which explains the disproportionately high (and widespread) frequency of occurrence of liquids in metathesis» (s. 108).

Young-Scholten et al. (1999) undersøker tekstlig produksjon av polsk til informanter med japansk og engelsk som morsmål. De finner systematisk bruk av metatase, og konkluderer med at dette er en strategi på lik linje med epentese, bortfall og erstatning, i motsetning til hva man trodde tidligere. Videre finner de tendenser i begge studiene på at innlærerne produserer enklere stavelser i mellomspråket sitt enn i morsmålet, men at de også kan produsere mer komplekse stavelser enn morsmålstemplet tillater når de prøver å unngå å bryte med andre universelle restriksjoner som prinsippet om maksimal opptakt. De finner også at KV-strukturer er mer utbredt enn VK-strukturer, som er i tråd med opptaktsprinsippet presentert i kapittelet om reparasjonsstrategier.

En annen studie, Eilertsen (1990), ser på innlærere av norsk med vietnamesisk som morsmål, og hvordan de realiserer konsonantgrupper i norske ord. Eilertsen finner at innlærerne tar i bruk både bortfall og epentese. Under ser vi noen ord fra materialet, der målspråksordet /elv/ blir realisert på ulike måter.

(4) S2 /ɛw/ – S1 /elv/

(5) S2 /ɛlɔv/ – S1 /elv/

(6) S2 /ɛlvə/ – S1 /elv/

I (4) ser vi et eksempel på bortfall. Segmentet /l/ faller bort, slik at innlæreren ikke må uttale en konsonantgruppe. I (5) ser vi en epentetisk vokal som blir satt inn mellom de to konsonantene i koda, og en ny stavelse blir dannet. I (6) ser vi at en epentetisk vokal blir satt inn i slutten av ordet. Innsetting i av vokal i slutten av et ord kalles *paragoge*, og gir en «åpen utgangstavelse» (Berggreen & Tenfjord, 1999:231).

Om en ser samlet på studiene på feltet, kommer det tydelig fram at «learners who speak different L1s have different phonotactic problems and use different strategies for reducing consonant clusters that are difficult for them to produce» (Jarvis & Pavlenko, 2008: 67). I tillegg til at avvik er knytta til hva slags morsmål en har, er det også funn som tyder på at de kan være knytta til hva slags *dialekt* man snakker. Broselow (1992) fant i sin studie av arabisktalende fra henholdsvis Egypt og Irak at disse to innlærergruppene gjorde ulike feil i stavelsesdanninga, og selv om disse har samme morsmål, snakker de altså ulike varieteter av arabisk med ulike morsmålstemplat.

Det kommer også fram av studier gjort på transfer innen stavelsesdanning, at transfer at strukturene gjerne følger universelle prinsipper for stavelsesdanning relatert til for eksempel markerthet og sonoritet (Jarvis & Pavlenko, 2008:68), slik vi har sett eksempler på i Hansen (2001) og Kløve & Young-Scholten (2001)

5.2 Forskning på ortografi i andrespråk

Figueredo (2006) gjør en metastudie av 27 studier som ser på hvordan innlærere av engelsk med ulike morsmålsbakgrunner staver ord på engelsk. I studiene finner han både positiv og negativ transfer, og sistnevnte type transfer er bevis for at mange avvik er morsmålsspesifikke. Videre finner han at innlærerne i stor grad baserer seg på fonologien og ortografien i morsmålet sitt tidlig i innlæringsprosessen, men at det blir mindre transfer jo mer informantene lærer seg om engelsk ortografi. I tillegg finner han transfer av kognitive prosesser, lingvistiske prosesser, litterasitet og metalingvistiske kunnskaper. Vi skal se litt på de konkrete funna i noen av studiene Figueredo presenterer. Dette er studier gjort på voksne, i likhet med min studie.

Ibrahim (1978) undersøker en gruppe innlærere av engelsk med arabisk som morsmål på universitetsnivå, og deres realisering av engelske ord. Under følger noen av avvika som kan være relevante for problemstillinga mi:

- (7) S2 <blaying> – S1 <playing>
 S2 <Jaban> – S1 <Japan>
- (8) S2 <hapit> – S1 <habit>
 S2 <clup> – S1 <club>
- (9) S2 <communisem> – S1 <communism>

I (7) ser vi at informantene har brukt reparasjonsstrategien erstatning ved å sette inn for <p>. Dette kan forklares med at arabisk ikke har lyden /p/, og at de velger den lyden i det arabiske lydinventaret som har mest til felles med /p/, nemlig /b/, og skriver ned tegnet som betegner lyden /b/, altså .

I (8) ser vi helt motsatt prosess: én har blitt erstattet med en <p>. Forfatteren forklarer dette med fenomenet *hyperkorreksjon*. Han skriver at innlærerne mest sannsynlig har blitt gjort oppmerksomme på /p/-lyden og at mange innlærere med arabisk som morsmål sliter med denne lyden. Et resultat av det er at innlærerne kan komme til å overdrive bruken av dette segmentet, både i lyd og i skrift.

Forfatteren forklarer en ekstra <e> i (9) som en epentetisk vokal – <e> blir satt inn mellom <s> og <m>, slik at innlæreren unngår en kompleks koda som ikke er tillatt i morsmålet. Jeg vil imidlertid foreslå sonoritetshierarkiet som en annen mulig forklaring. /sm/ i koda er ikke i tråd med hierarkiet, da /s/ er en frikativ og /m/ er en nasal, og dette er markert. For å unngå en slik markert koda, kan informantene sette inn en epentetisk vokal for å bryte opp konsonantgruppa. Det er med andre ord en mulighet for at det heller er universelle prinsipper som spiller en rolle her.

Cook (1997) undersøker en rekke innlærergrupper av engelsk med ulike morsmålsbakgrunner, og sammenligner feilene deres med feil gjort av S1-innlærere av engelsk. Under ser vi noen realiseringer av engelske ord til innlærere med japansk som morsmål.

- (10) S2 <courese> – S1 <course>
 S2 <Engilish> – S1 <English>
 S2 <subejcts> – S1 <subjects>

Her ser vi et eksempel på at innlærere med morsmålstemplatet KVK¹⁵ har satt inn en epentetisk vokal i komplekse stavelser. Dette viser oss at epentese også kan vises i skrift, og

¹⁵ De fleste japanske stavelser har KV-struktur, og det er kun segmentet /n/ som kan opptre i kodaposisjon (Husby & Kløve, 1998:86).

at det ikke begrenser seg til uttale. Spesielt er eksempelet med <English> interessant, for her presenterer jo grafemet <ng> det enkelte segmentet /ŋ/. Dermed er det på et lydlig nivå ikke nødvendig med en reparasjon her om de to tegna står for det ene segmentet. I studien min kommer jeg til å diskutere lignende typer innsetting som i <English>.

5.2.1 Forskning på ortografi i norsk som andrespråk

Forskning på skriving i norsk som andrespråk har blitt beskrevet som «relativt sparsom siden selve andrespråksfeltet vokste fram på slutten av 1970-tallet» (Golden & Hvistendal, 2013:70). Enda mer fraværende er forskning på ortografi i N2 – Vigrestad (2016) beskriver forskninga som «nærmest ikke-eksisterende» (s. 4). Vi skal likevel ta oss gjennom det som finnes av relevant skriveforskning i N2. Golden & Hvistendal (2010) gir en god oversikt over forskning på skriving i et andrespråks-perspektiv i tidsrommet 1980 til 2010, og Gujord og Ragnhildstveit (2017:142-147) gir en oversikt over forskning gjort på transferfeltet i perioden 2006-2017. Basert på disse artiklene og søk på skriving i andrespråk og ortografi i andrespråk, finner vi at det er gjort to ordentlige studier på ortografi i norsk som andrespråk.

Et viktig bidrag gjort innen forskning på andrespråksortografi, er gjort av den første professoren i norsk som andrespråk, Anne Hvenekilde. Hun publiserte fem artikler mellom 1985 og 1987 der hun undersøkte ortografi i NOMIS-materialet (Norsk Mellomspråk I Skolestiler), som hun selv hadde samla inn. Dette materialet består av tekster skrevet i 1984 av elever i grunnskolen med morsmåla tyrkisk, urdu/hindi, arabisk, berberisk, tigrinja, vietnamesisk og tagalog. Elevene kunne velge mellom å skrive teksten *Min første skoledag* eller *Hva skal du gjøre til sommeren?*

Hvenekilde (1985) undersøker to informantgrupper: én med tyrkisk som morsmål og én med vietnamesisk som morsmål. Datamaterialet hennes er 188 stiler skrevet av informantgruppe én, og 98 stiler skrevet av informantgruppe to. Hun undersøker ortografi i disse stilene, og finner at de tyrkiske innlærerne har andre ortografiske avvik enn de vietnamesiske. På bakgrunn av dette argumenterer hun for at morsmål og bakgrunn har noe å si for tilegninga av norsk som andrespråk. Hun forsøker å forklare de ulike mellomspråksortografiene med morsmålenes struktur, leseopplæring i hjemlandene og elevenes familiebakgrunn (s. 41).

Hvenekilde (1986) undersøker fire av de tyrkiske informantene ett år etterpå for å se på utviklingen deres. Hun ser på syntaks, morfologi, ortografi, vokabular og tekstsammenheng. Hva som hadde blitt forbedret, varierte, men en av elevene hadde fått

forbedret ortografi vesentlig.

Hvenekilde (1987) undersøker vokalgrafemene i tekstene til to informantgrupper: 58 elever med arabisk som morsmål og 89 elever med berberisk som morsmål. Disse blir sammenlignet med hva slags feil elever med norsk som morsmål gjør, og satt opp mot deres respektive morsmål. Hun fant at noen av feilene til informantene også er vanlige blant de med norsk som morsmål, f.eks. forveksling av <å> og <o>. Andre feil, derimot, er unike for innlærerne, f.eks. forveksling av <i> og <e>. Hvenekilde mener det er problematisk å gjøre en nøyе undersøkelse av forskjellene mellom feilene morsmålstalere av berberisk og arabisk, da informantene har varierende alder og oppholdstid i Norge, og unnlater dette.

Vigrestad (2016) gjør en lignende undersøkelse som Hvenekilde i masteroppgava si. Hun undersøker ortografien til to informantgrupper, innlærere av norsk med henholdsvis vietnamesisk og bosnisk-serbisk-kroatisk som morsmål. Hun bruker ASK-korpuset for å få tilgang på 99 tekster fra hver informantgruppe, og problemstillinga går ut på å se om det er noen forskjell på avvika i de to ulike gruppene. Hun lager seg også seks forskningsspørsmål basert på hvilke feil hun tror informantene kommer til å gjøre. Her ser hun på informantens morsmål og hva slags trekk som finnes i dette språket. Forskningsspørsmålene handler om konsonantgrupper, lyder som kan bli representerte med mange ulike grafemer i norsk, innsetting av vokaler, dobbeltkonsonanter, monoftonger/diftonger og substitusjonsfeil. Vigrestad finner at det er visse forskjeller mellom de to informantgruppene, og at disse går på konsonantgrupper, dobbeltkonsonanter, og substitusjonsfeil. De tre andre gruppene med skrivefeil er ikke ulike i de to informantgruppene.

5.2.2 Stavelsesstrukturer i Hvenekilde (1985) og Vigrestad (2016)

Vi har allerede introdusert Hvenekilde (1985), som undersøkte «ordbildene», det jeg vil kalle ortografien, til elever med tyrkisk og vietnamesisk som morsmål. Vi kan lære mye av denne studien. Hvenekilde kategoriserer alle skrivefeil etter hvor i ordet de oppstår, og hvor i stavelsen den oppstår. Videre lager hun en såkalt *konkordanseliste*, en liste over alle skrivefeil, men i kontekst. Dette er viktig fordi skrivefeil i mange tilfeller ikke bare kan forklares på grunnlag av et enkeltord – noen ganger må vi se forbi ordgrensene, og undersøke orda som står før og etter ordet. Hun deler funnene sine inn i tre feiltyper: «feil i konsonantforbindelser», «feil i vokalgrafemer i stiler skrevet av tyrkiske elever» og «sammenskrevne ord i stiler skrevet av tyrkiske elever». For studien min er det mest relevant undersøke hva hun fant i kategorien «feil i konsonantforbindelser». Dette er interessant fordi både tyrkisk og vietnamesisk, på lik linje med somali, har maksstrukturen

KVK, og dermed har ikke informantene konsonantforbindelser i sine morsmål.

I materialet finner Hvenekilde reparasjonsstrategiene epentese, bortfall, metatese og erstatning, som jeg presenterte i kap. 2.4. Hun undersøker både initiale, mediale og finale konsonantforbindelser. Det hun finner mest av, er innsetting av vokaler dvs. epentese, i komplekse opptakter. Under viser jeg noen eksempler fra datamaterialet hennes.

(11) S2 <sipise> – S1 <spise>

(12) S2 <sotol> – S1 <stol>

I (11) har altså <i> blitt satt inn mellom grafemene <s> og <p>, som til sammen danner en kompleks opptakt. I (12) skjer det samme, bare med andre grafem: <o> blir satt inn mellom <s> og <t>. Ved bruk av reparasjonsstrategien epentese unngår altså informanten å danne komplekse stavelser – en ny stavelse kommer til innsetning, og den første konsonanten kan knytte seg til denne vokalen i stedet for den andre. Det er 44 reparasjoner av denne typen i initiale konsonantforbindelser, og 8 stk. i mediale og finale forbindelser, der samtlige slike feil er gjort av de tyrkiske informantene. Fraværet av denne typen feil hos de vietnamesiske informantene forklarer hun med at vietnamesisk ikke har polysyllabiske ord. Epentese innebærer innsetting av ny stavelse, og da får vi nødvendigvis et flerstava ord. Det er ikke i tråd med vietnamesiske stavelsesregler.

Videre finner Hvenekilde ni bortfall av konsonanter i konsonantforbindelser initialt, og 39 stk. medialt og finalt. Det ser vi eksempler på under:

(13) S2 <bomster> – S1 <blomster>

(14) S2 <blyan> – S1 <blyant>

I (13) har <l> falt bort i opptakta, slik at vi får en enkel opptakt i stedet for en kompleks. I (14) har <n> falt bort, og også her unngår vi den komplekse kodaen.

Vi finner også fire tilfeller av erstatning initialt og 47 medialt og finalt, der en konsonant er byttet ut med en annen. Under følger to eksempler fra materialet:

(15) S2 <skun> – S1 <stund>

(16) S2 <smømme> – S1 <svømme>

I (15) har <t> blitt erstattet med <k>, og i (16) har <v> blitt erstattet med <m>. Denne typen feil regner hun riktignok ikke med i materialet sitt fordi det ikke innebærer forenkling av konsonantforbindelser; det er jo fortsatt like mange konsonanter i opptaktene til disse orda. Hun finner heller ingen forklaring på hvorfor lydene har blitt byttet ut.

Vi finner også metatese i dataene til Hvenekilde, noe vi ser eksempler på under:

- (17) S2 <stirk> – S1 <strikk>
(18) S2 <førken> – S1 <frøken>

I (17) har grafemene <i> og <r> bytta plass. I (18) er det <r> og <ø> som har bytta plass. I (1) unngår ikke informanten den komplekse stavelsen (det er fortsatt to konsonanter i opptakt), men opptakta blir i det minste *mindre* kompleks. I (18) blir resultatet en enkel opptakt med bare én konsonant, og informanten får dermed en opptakt som stemmer overens med stavelsesstrukturen i morsmålet.

Basert på det hun finner i det innsamlede materialet, hevder hun at mange av de ortografiske avvika til elevene kan spores tilbake til stavelsesstrukturene i de respektive morsmåla, men det er også en del ortografiske feil hun tror skyldes avvik fra norsk grammatikk, heller enn stavelsesstrukturen i morsmålet. Hun finner også at det er mange flere feil blant de tyrkiske informantene, og foreslår lese- og skriveopplæring i de respektive hjemlanda som en mulig forklaring på forskjellen i feilprosent.

Vigrestad (2016) ble også presentert i forrige kapittel, og det er spesielt spørsmålet hun stiller om konsonantgrupper som blir relevant for min studie. Hun finner både bortfall og epentese i materialet sitt. Hun finner spesielt mye bortfall av konsonantgrafem blant innlærerne med vietnamesisk som morsmål, og ganske få tilfeller av epentese. Dette passer godt med at vietnamesisk er et monsyllabisk språk som ikke tillater konsonantgrupper, og at bortfall derfor er en mer nærliggende reparasjonsstrategi enn epentese. Informantene med bosnisk/kroatisk/serbisk som morsmål har også en god del avvik som skyldes bortfall, selv om det i dette morsmålet er tillatt med mer komplekse stavelsesstrukturer. Det er få forekomster av epentese også hos denne innlærergruppa.

Både Hvenekilde (1985) og Vigrestad (2016) har sammenligna to innlærergrupper av norsk i studiene sine, og sett på forskjeller og likheter i ORT mellom de ulike gruppene. I studien min kommer jeg ikke til å sammenligne to innlærergrupper på samme måte. Jeg kommer til å konsentrere meg om én innlærergruppe, nemlig SOMALI, og kartlegge og diskutere bakgrunnen for deres ORT. For å finne ut hvorvidt ORT til informantene mine er

særegne for SOMALI, vil jeg sammenligne resultatene mine med funna i Hvenekilde og Vigrestad, men dette er ikke hovedpoenget i studien min.

5.3 Forskning gjort på språkkontakt mellom somali og andre språk

For å lære mer om hvordan somali oppfører seg i møte med andre språk, kan vi ta en titt på Cardona (1988) som tar for seg asiatiske lånord i somali. Somalia har hatt mye kontakt med en rekke asiatiske land gjennom handel på Det indiske hav, og forfatteren undersøker hvilke lingvistiske konsekvenser dette har fått. Materialet han baserer seg på, er en ordbok på over 25.000 ord som nylig var gitt ut da han skrev artikkelen. Han baserer seg altså ikke på materiale direkte produsert av morsmålsbrukere, men legger til at ordboka er «carefully checked for accuracy and real use» (s. 22).

I artikkelen har forfatteren en liste over 120 lånord fra asiatiske språk, og gjør rede for opphavet til orda – hvilket språk ordet stammer fra, hvilke språk det har gått via, og hvordan det ser ut i de ulike språka. Under gir jeg eksempler på noen lånord (med opphav) jeg mener kan være relevante for å se hvordan stavelsesstrukturen i somali kan virke på ord fra andre språk. Dette er, i likhet med mine data, skrevne ord.

- | | | | |
|------|--------------|-------------------|---|
| (19) | <Qolfaad> | ‘tette’ | ← arabisk <i>qalfaṭ</i> |
| | <Jaalabbuud> | ‘båt’ | ← urdu <i>jālībōt</i> ← engelsk <i>jolly-boat</i> |
| (20) | <Babbaay> | ‘papaya’ | ← urdu ← malayisk <i>papaya</i> |
| | <Suubeehi> | ‘politikonstabel’ | ← persisk <i>sipāhī</i> |
| (21) | <Qulaan> | ‘knekt’ | ← persisk <i>ghulām</i> |
| (22) | <Baraf> | ‘snø, is’ | ← persisk <i>barf</i> |
| (23) | <Cananaas> | ‘ananas’ | ← urdu ‘ <i>ayn un-nās</i> |

Vi kan observere at begge orda i (19) er lånt inn fra urdu, og at de begge følger mønsteret *d* ← *t*, slik at ord med <ṭ> i urdu vil ende opp som <d> i somali. Videre viser (20) eksempler på overgangen *b* < *p*. I (21) ser vi overgangen *n* ← *m*. I (22) ser vi et eksempel på et ord som har fått inn en ekstra vokal <a> som ikke var der i lånordspråket persisk. (23) viser et eksempel på en innsetting av <c> (som representerer /ʕ/) foran vokalen <a>.

I disse eksemplene observerer vi altså at ord fra ulike språk går gjennom visse endringer når de blir lånt inn i somali, og vi kan se at dette gjøres i tråd med

stavelsesdanningsprinsippene i somali. $d \leftarrow t$ er naturlig, ettersom /t/ i utlyd ikke er tillatt i somali. $b < p$ går også opp med det fonologiske systemet i somali: De har ikke lyden /p/. Det er heller ikke tillatt med /m/ i slutten av somaliske ord, og dermed gir det mening at <m> blir erstattet med en annen, lignende (nasal) lyd. For å unngå en kompleks koda med sekvensen <rf> i ordet *barf*, har den somaliske versjonen av ordet fått inn en ekstra vokal mellom konsonantene. Innsetting av <c> foran vokal passer også godt med somali-systemet; ord kan ikke starte med vokal i somali. I studien min blir det interessant å se om orda i det somali-norske mellomspråket har noen lignende ortografiske trekk som orda over-

5.3.1 Somali i møte med norsk

Til tross for det store antallet somaliske innvandrere i Norge er det nesten ingen studier av hva som skjer når somali møter norsk. De to studiene jeg kjenner til, er Gujord (2013), som analyserer tempusbruk i mellomspråket til innlærere av norsk med vietnamesisk og somali som morsmål, og Kaldhol (2017), som analyserer norske innlånsord i somali hos somalitalende i Norge. Vi skal se nærmere på sistnevnte studie, som i likhet med min studie, ser på fonologiske prosesser.

Kaldhol (2017) undersøker prosodien i norske lånord i somali slik det er snakket av informantene hennes i området rundt Oslo. Selv om Kaldhols konklusjoner om prosodi ikke er direkte relevante for min stavelsesbaserte problemstilling, er det en del andre sider ved datamaterialet hennes som er interessant for meg. Det er nemlig en del av de norske ordene som endrer stavelsesstruktur når de blir lånt inn i somali. Vi skal se på noen eksempler fra datamaterialet hennes (se «Appendix B: Word lists»), og analysere dem i lys av kap. 2.3.2 om somaliske stavelsesrestriksjoner og kap. 2.4 om reparasjonsstrategier.

(24) [biblijutéeg-ga] ← [bib|ijju¹te:k] (<bibliotek>)
 [regtoor-ád-da] ← [¹rɛktur] (<rektor>)

(25) [ʃaalóom] ← [²ʃ[a:|ɔm] (<slalåm>)
 [kiteériŋ-kái] ← [kvi¹te:riŋ] (<kvittering>)

(26) [teebáan] ← [¹te:|ba:nə] (<t-bane>)
 [komúun-ka] ← [ku²mü:nə] (<kommune>)

(27) [bereév] ← [¹bre:v] (<brev>)
 [birivaád-ka] ← [pri¹va:t] (<privat>)

(28) [noʃk-é] ← [¹noʃk] (<norsk>)
 [hélq-e] ← [¹hɛlg] (<helg>)

Vi kan her se at mange ord blir uttalt annerledes enn de ville blitt uttalt på standard østnorsk, og vi kan tenke oss at de norske lånordene blir tilpasset den somaliske stavelsesstrukturen og endret deretter. Suffiksene *-ka* og *-ád-da* er her brukte for å plassere orda inn i ulike substantivklasser, men her er jeg mest interessert i å se på hvordan *stammene* i orda oppfører seg. I (24) ser vi to eksempler på reparasjonsstrategien *erstatning*. Der det opprinnelig skulle vært [t], er konsonanten endret til [d]. Dette er i tråd med de somaliske stavelsesrestriksjonene: [t] i koda er ikke tillatt, og blir omgjort til [d] i slike tilfeller. (25) viser eksempler på bortfall av konsonanter. Der vi i norsk ville uttalt [ʔ] i [²ʃ[ɑ:][ɔm], er den her fjernet. Det kan tenkes at dette har skjedd til fordel for KVK-strukturen i somali – stavelsen [²ʃ[ɑ:] har jo strukturen KKVK, og her blir den ene konsonanten fjernet for å unngå en kompleks stavelse. Vi møter en annen type bortfall i (26), der stavelseskjernen i slutten av ord blir fjerna. Dette er vanskeligere å forklare. Som vi har sett, tillater somali polysyllabiske ord, og bortfallet kan dermed ikke forklares av at somaliske ord foretrekker å være monosyllabiske.

I (27) ser vi eksempler på epentese medialt i ordet ved at stavelseskjerner blir satt inn. Ved å benytte seg av denne strategien, unngår informantene komplekse stavelsesstrukturer som er i konflikt med maksstrukturen KVK i somali. Tar vi ordet [bereév], for eksempel, ser vi at et ord som opprinnelig har KKVK-struktur, har fått inn en stavelsesbærende vokal, [e], mellom de to K-ene, slik at vi ender opp med et tostava ord med strukturen KV.KVVK, og dermed kommer innlæreren unna den forbudte komplekse opptakta. Et annet eksempel på epentese ser vi (28), der en vokal har blitt satt inn i slutten av ordet. Kaldhol plasserer disse inn i en deklinasjonsklasse og forklarer endelsen med dette. I studien min kommer jeg til å undersøke om lignende tilfeller i materialet mitt kan skyldes reparasjonsstrategier.

Reparasjonsstrategiene til Kaldhols informanter er med andre ord i tråd med hvordan en ville forvente at norske ord oppfører seg når de er bundet av somaliske stavelsesregler.

Dette er imidlertid en studie av *talespråket*, og det sier ikke noe om hvordan de ville *skrevet* ordene. Mitt studieobjekt er ortografien, det skrevne språket, og jeg er interessert i å se om lignende prosesser skjer når SOMALI tar pennen fatt.

Det er viktig å nevne at studiene til Cardona (1998) og Kaldhol (2017) har et annet utgangspunkt enn meg: Mens de ser på enkeltord lånt inn i somali, ser jeg på et mellomspråk som oppstår hos somaliere med et mål om å lære seg *norsk*. Det er ikke nødvendigvis et mål at lånordene i somali skal høres norske ut, tvert imot. I den konteksten de blir ytret, er det heller slik at det i mange tilfeller passer best om de blir ytret på en mest mulig somalisk måte. Vi skal dermed være forsiktig med å si at den samme kognitive prosessen ligger bak innlån og mellomspråk. En person med et mellomspråk har jo gjerne et *ønske* om å uttale og skrive ordene i målspråket på en så normlik måte som mulig. Likevel kan det være visse underliggende likhetstrekk mellom de to språkformene, og det er dermed interessant å se om det er noen likhetstrekk mellom mine data og dataene til Cardona og Kaldhol.

DEL II: Studien

6.0 Metode

I dette kapittelet gjør jeg rede for metoden jeg har brukt for å samle inn, kategorisere og analysere dataene mine. Jeg presenterer det metodologiske grunnlaget. Så gjør jeg rede for ASK-korpuset, som jeg har brukt som datagrunnlag. Videre beskriver jeg hvordan jeg søkte meg fram i korpuset. Jeg gjør også rede for informantene mine og deres bakgrunn. Så gjør jeg rede for datautvalget mitt, og forklarer jeg hvordan jeg har klassifisert disse dataene.

6.1 Metodologisk grunnlag

I kap 4.2 la jeg fram teori om transfer. I dette kapittelet gjør jeg greie for *metoden* jeg baserer meg på når jeg studerer dataene mine. Jarvis (2000) kan sees på som vendepunktet for transferforskning både internasjonalt og i Norge. Studier på transfer før dette punktet hadde i stor grad gitt sprikende resultater, og det var vanskelig å se mønstre. Selv om Jarvis ser muligheten for at det kan skyldes at det ikke er mønstre å finne, og at konseptet transfer ikke har fellestrekk fra språkbruker til språkbruker, mener han at det heller skyldes mangel på en felles metode for transferforskerne. Han legger derfor fram tre metodekrav for transferforskeren (her oversatt av Golden & Tenfjord 2015:187):

1. en teorinøytral definisjon av transfer
2. en presis beskrivelse av hvilke typer bevis man må ha for å kunne argumentere for eller mot transfer
3. kontroll over en rekke variabler, både sosiale (bl.a. utdanning) psykologiske (bl.a. motivasjon) og språklige (bl.a. distansen mellom morsmål og målspråk)

For å svare på det første kriteriet ovenfor foreslår han følgende arbeidsdefinisjon av transfer:

S1-innflytelse refererer til ethvert tilfelle av innlærerdataba der en kan vise til at det eksisterer en statistisk signifikant korrelasjon (eller probabilitetsbasert relasjon) mellom noen trekk i innlæreres mellomspråksprestand og S1-bakgrunnen deres (Jarvis 2010:252) (oversatt av Golden & Tenfjord, 2015:187).

Definisjonen sier noe om hvordan en kan identifisere transfer, og legger vekt på statistisk metode som fremgangsmåte. Gujord (2013) ser svakhetene ved denne definisjonen, og skriver følgende: «Using this definition of transfer requires skills in statistics, which in fact can pose a problem, or at least a challenge, in that many researchers in humanities are not educated as statisticians» (s. 121).

For å svare på kriterie (2) beskrevet ovenfor, foreslår Jarvis fire bevis for at et

språktrekk skyldes transfer som forskeren kan legge fram (her etter Golden & Tenfjord (2015:187)):

1. Innlærere som tilhører samme morsmålsgruppe, viser likheter i språkbruken i mellomspråksperformansen når det gjelder et spesielt språktrekk.
2. Innlærere med ulike morsmål, men samme målspråk, viser ulikhet i språkbruken når det gjelder det samme språktrekket.
3. Forskeren kan påvise samsvar mellom trekk i innlærernes mellomspråk og deres morsmålsperformanse.
4. Det finnes to ulike konstruksjoner i morsmålet, og den samme kontrasten kommer fram i mellomspråket deres.

I tillegg legger Jarvis vekt på at informanten bør være på et visst ferdighetsnivå. I studien min studerer jeg tekstene til 107 personer med samme målspråk (norsk) og samme morsmål (somal). Jeg sammenligner disse informantene, og ser på hvordan de realiserer noen spesielle språktrekk i mellomspråket sitt, dvs. ortografiske representasjoner av stavelser. På den måten kan jeg undersøke om noen språktrekk går igjen blant de 107 informantene. Dette er i tråd med bevis 1.

Jeg kommer ikke til å utforske punkt to i Jarvis' rammeverk systematisk. Dette er ikke en oppgave der sammenligning mellom ulike innlærergrupper av norsk står sterkt. Likevel kommer jeg til å trekke inn tidligere studier der ortografi i norsk som andrespråk har blitt utforska, og andre innlærergrupper har blitt forsket på. Gruppene jeg sammenligner med, har vietnamesisk, bosnisk/kroatisk/serbisk og tyrkisk som morsmål. Det er vanskelig for meg å sammenligne systematisk med disse studiene fordi de ikke har sett på alle de samme typene avvik som meg. De har for eksempel ikke sett på innsetting av vokal ordfinalt, og dermed er det ikke noe sammenligningsgrunnlag her. Jeg kommer likevel til å referere til disse studiene der det er mulig å sammenligne.

Det tredje beviset til Jarvis er viktig i denne studien. Jeg skal undersøke og diskutere om dataene mine, tatt fra mellomspråket til SOMALI, har noe til felles med stavelsesstrukturene i morsmålet somali. Dette er ikke informasjon om hvordan alle de 107 individene jeg forsker på, snakker eller skriver somali, men en beskrivelse av den varieteten som er mest brukt i Somalia. Siden fonologi er et trekk ved språket som er ganske stabilt, er det sannsynlig at innlærerne mine følger de fonologiske reglene presentert i kap. 2.3.2 når de snakker somali, så fremt de snakker den varieteten av somali jeg har gjort rede for.

Det siste mulige beviset Jarvis (2000) legger fram, er ikke relevant for oppgava mi.

Det handler om at en kan undersøke om en kontrast som finnes i morsmålet, også finnes i målspråket. En slik kontrast finner vi ikke i stavelsesstrukturen i somali.

Jeg kommer altså til å analysere dataene mine basert på Jarvis' bevis én og tre men jeg lar det være opp til leseren å regne på signifikansnivået til resultatene mine. Agendaen min for denne studien er å undersøke om det finnes tendenser i materialet mitt, og hva som *kan* være grunnen til tendensene. Å hevde at et trekk i mellomspråket skyldes transfer basert på matematiske utregninger, er ikke et mål i studien min.

Når det gjelder det tredje kravet til Jarvis (2000), nemlig å ha kontroll på sosiale, psykologiske og språklige faktorer, har jeg det til en viss grad. Jeg har *tilgang* på en del informasjon om informantene mine, slik som hvor lenge de har gått på norskkurs, alderen deres, og hvilke andre språk enn norsk og somali de snakker. Det er likevel mange faktorer jeg ikke har kontroll på – informantene mine er individer med ulike bakgrunner, tankesett og identiteter, og dette gjør at det blir umulig å ha kontroll på alle faktorene som kan påvirke læringssituasjonen deres. Informasjonen som ligger i korpuset kan bidra til å forklare mellomspråket til individene, men den kan ikke forklare hele den komplekse prosessen det er å lære et nytt språk. Det er heller ikke et mål med denne studien å snakke mye om hvordan sosiale faktorer kan påvirke læringsprosessen – dette er et viktig spørsmål som man kan reflektere over i det vide og det brede, men jeg velger her å fokusere på det strukturelle ved språket. Jeg kommer likevel til å reflektere litt over flerspråkligheten til informantene.

6.2 ASK som datagrunnlag

I denne studien har jeg hentet dataene mine fra Norsk andrespråkskorpus (heretter forkortet ASK). Det er Universitetet i Bergen som står bak dette korpuset. I korpuset ligger tekstene til 1735 informanter, der hver informant har skrevet én tekst hver. Alle informantene er voksne andrespråksinnlærere av norsk. De har til sammen ti ulike morsmålsbakgrunner: bosnisk/serbisk/kroatisk, vietnamesisk, somali, spansk, nederlandsk, russisk, albansk, polsk, engelsk og tysk. Det er rundt 100 informanter fra hver av disse morsmålsbakgrunnene.

Tekstene er samla inn i forbindelse med at informantene tok Språkprøven i norsk eller Test i norsk – høyere nivå. De kunne velge å svare på ulike oppgaver under disse prøvene. Dataene ble samla inn i 2006. Under opprettelsen av korpuset ble alle orda grammatisk tagget med Oslo-Bergen-taggingen, og alle avvik fra norsk rettskriving ble feilkodet. I korpuset ligger også informasjon om bakgrunnen til hver informant, men alle informanter er anonymiserte.

De to språkprøvene informantene i ASK har tatt, er på ulike nivå. *Test i norsk – høyere nivå*, ofte kalt Bergenstesten, var i utgangspunktet rettet mot innlærere med

akademisk utdanning eller studenter som ønsket å ta høyere utdanning i Norge. Med denne testen kunne de bevise at de hadde tilstrekkelig kompetanse i norsk til å jobbe eller studere i Norge. *Språktest i norsk* ble introdusert for at innlærere på et lavere nivå i norsk kunne få bevis på norskkompetansen sin. Denne prøven er i dag erstattet med *Norskprøven A1/A2*, *Norskprøven A2/B1* og *Norskprøven B1/B2*, som i stor grad følger retningslinjene for det *Det felles europeiske rammeverket for språk*, men dette rammeverket var ikke til stede i like stor grad i språktestene på starten av 2000-tallet. Derfor er det vanskelig å si nøyaktig hvilket nivå testene var på, men det er altså slik at en gruppe informanter har tatt en vesentlig vanskeligere språkprøve enn den andre gruppa (Golden & Monsen, 2015:202-203).

Tekstene til informantene mine er skrevet for hånd og digitalisert i forbindelse med opprettelsen av ASK. Dette kan være en feilkilde, ettersom grafemer eller ord ved et uhell kan gå tapt, bli lagt til eller stokkes om i denne prosessen. Derfor er det en mulighet for at avvik jeg analyserer, faktisk bare er tastefeil fra transkribenten sin side. Det er likevel usannsynlig at det er mange av disse feila i korpuset, da det var erfarne transkribenter som stod bak digitaliseringa og feilkodinga av tekstene.

Korpuset ligger i infrastrukturen CLARINO, og er gjort søkbart gjennom korpushåndteringssystemet Korpuskel. Ved hjelp av Oslo-Bergen-taggeren har alle orda i ASK blitt grammatisk «tagget». Det vil si at informasjon om lemma, bøyingsform, ordklasse og lignende, ligger inne for alle ord. I korpuset ligger også en god del informasjon om hver informant og bakgrunnen til vedkommende, slik som hjemland, morsmål, utdanning, antall timer personen har lært norsk osv.

Korpuset er også *feilkoda*. Det vil si at alle avvik i korpuset har blitt plassert i en feilkode etter hva slags type feil en finner. Feilkodinga har blitt gjort manuelt, og for at koderne skulle kode så likt som mulig, har de fulgte tre prinsipper. For det første har ikke koderne analysert resultata, og dermed ikke tatt stilling til hva et avvik skyldes. De har hatt en så deskriptiv tilnærming som mulig. For det andre har de lagt et *sannsynlighetsprinsipp* til grunn, der en legger konteksten til grunn og antar den mest sannsynlige målforma for et ord sett i lys av konteksten. Det tredje prinsippet koderne har fulgt, er prinsippet om minimal endring. Om det kan være flere alternativ til målpråksord eller målpråkssetning, velger koderne den setninga som endrer den originale setninga minst mulig.

Feilkoden som gjelder for problemstillinga mi, er ORT – ortografiske avvik. I kodeboka skriver de at «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»¹⁶ Det kan altså være ord tagget med ORT som egentlig burde hatt en annen feilkode, samtidig som det kan være ord som hadde passet inn under ORT-taggen, men som ikke er tagget som dette. Det blir også spesifisert at ORT kun gjelder for lemma, og at manglende bøyingsendinger ikke skal inn i denne kategorien. Dette diskuterer jeg i kap 6.7.

Det finnes også et *parallellkorp* der alle tekstene i ASK er retta til normert bokmål, og det går an å søke i dette samtidig som man søker i ASK, slik at en kan sammenligne. Det er viktig å påpeke at dette ikke nødvendigvis representerer det informantene mente å skrive, men basert på sammenhengen har man gjort en antagelse på hva informantene mente med det hen skrev.

ASK har restriksjoner på hvem som får tilgang til korpuset, og man må søke om tilgang. Det gjør man ved å sende en beskrivelse av forskningsprosjektet sitt via nettsidene til ASK.

6.3 Om bruk av «feil» og «avvik»

Forskningsobjektet mitt er ord med en ortografi som avviker fra bokmålsnormen av 2005. På folkemunne blir dette kalt «skrivefeil», men vi skal være forsiktige med å kalle det «feil». Vi må ikke gå i det Tenfjord, Hagen & Johansen (2009) kaller *sammenligningsfella*. Innlærere har et mellomspråk, et mentalt ortografisk system med egne regler, og orda som avviker fra bokmålsnormen, kan være helt riktige i deres mellomspråksortografi. Samtidig er det et faktum at målet til disse innlærerne er å lære seg norsk, og dermed er det et mål å tilegne seg en så bokmålsnær ortografi som mulig. Av den grunn er det naturlig å sammenligne det informantene produserer, med målspråket, og kalle noe for feil eller avvik i en slik sammenheng gir mening. Derfor vil jeg betegne dataene mine som avvik, vel vitende om at dette ikke er avvik i mellomspråksortografien til informantene mine. Jeg kommer til å bruke termen ORT som forkortelse for «ortografiske avvik fra bokmålsnormen av 2005», som er forkortelsen ASK bruker for ortografiske avvik.

6.5 Å søke seg fram i korpuset

Over har jeg gjort rede for hvordan tekstene i ASK-korpuset er grammatisk tagget, feiltagget og knyttet til hver enkelt informant. Dette gjør jobben enklere for forskeren som bruker ASK; det blir mulig å søke etter attributter, dvs. kjennetegn ved ord eller setninger, og på den måten

¹⁶ Kodeboka finnes her: <http://korpus.uib.no/humfak/ask/brukerveiledning/feilkoder.htm#feilkodingsprosessen>

få opp alle ord eller setninger som har trekk til felles. Dette har jeg fått kunnskap om fra ASKs brukerveiledning.

I studien min ønsket jeg å se på ORT gjort av innlærere av norsk med somali som morsmål. Dermed var to attributter relevante for meg:

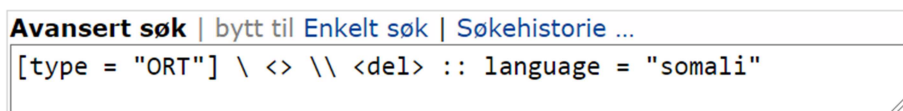
1. Morsmål: somali
2. Feilkode: ORT

Jeg søkte etter disse attributtene i korpuset ved å bruke søkegrensesnittet i Korpuskel. Her kunne jeg hake av for attributtene jeg ville at dataene mine skulle ha. Under er en skjermdump av søket mitt, samt koden for søket.



The image shows two side-by-side screenshots of the ASK search interface. The left screenshot shows a search filter for 'Morsmål = "somali"'. Below this, there is a section for 'Additional constraints (click to choose):' with a list of options: 'Tekst-id', 'Språktest', 'CEFR', and 'Tema'. The right screenshot shows a search filter for 'Feiltype = "ORT"'. Below this, there is a dropdown menu for 'gjentakelse:' set to '1'. Underneath, there is a section for 'Additional constraints for this position:' with a list of options: 'Ord' and 'Lemma'.

Figur 5: søket mitt i ASK-korpuset



The image shows a screenshot of the ASK search code editor. At the top, there are navigation links: 'Avansert søk | bytt til Enkelt søk | Søkehistorie ...'. Below this, there is a text input field containing the search query: `[type = "ORT"] \ <> \\ :: language = "somali"`. The input field has a small icon in the bottom right corner.

Figur 6: kode for søk med attributtene «ORT» og «somali» i ASK

I kap. 2.1 gjorde jeg rede for at det er viktig å se på ord i sammenheng når en analyserer stavelsene i et ord. Av den grunn var det viktig for meg å bruke en visningsmodus som ga meg god oversikt over ord før og etter ORT. Dette kommer tydelig fram i modusen *kwic*. Kwic står for «key word in context», og fremhever de orda jeg har søkt etter, samtidig som jeg får opp sammenhengen ordet står i. Under er en skjermdump av resultata av søket vist i kwic-modus.

	treff	
I. De som vil ikke støtter slike systemer blir ofte fengslet og		toturert. Noen blir drept. Derfor flykter mange politikere fra
I. De som vil ikke støtter slike systemer blir ofte fengslet og	toturert	. Noen blir drept. Derfor flykter mange politikere fra hjemla
n vil ikke støtter slike systemer blir ofte fengslet og toturert		. Noen blir drept. Derfor flykter mange politikere fra hjemla
1 av dem må flytte slik at de kan bo sammen. Som tidligere	nevt	nevt er Norge blitt en stor reisemål for innvandrere. De nor
1 av dem må flytte slik at de kan bo sammen. Som tidligere		er Norge blitt en stor reisemål for innvandrere. De norske n
lem må flytte slik at de kan bo sammen. Som tidligere nevt		er Norge blitt en stor reisemål for innvandrere. De norske n
: at mange ønsker å bosette seg her i landet. Norge er også	opptat	opptat med å hjelpe å løse mange konflikter i verden. Så de
: at mange ønsker å bosette seg her i landet. Norge er også		med å hjelpe å løse mange konflikter i verden. Så den er bl
ngje ønsker å bosette seg her i landet. Norge er også opptat		med å hjelpe å løse mange konflikter i verden. Så den er bl
t man kan finne et alternativ for å leve. Så lenge det finnes	diktaktorer	diktaktorer i verden, kan vi regne med at de ofrene vil forsø
t man kan finne et alternativ for å leve. Så lenge det finnes		i verden, kan vi regne med at de ofrene vil forsøke å finne
inne et alternativ for å leve. Så lenge det finnes diktaktorer		i verden, kan vi regne med at de ofrene vil forsøke å finne

Figur 7: kwic-modus i ASK

Når man har søkt, er det også mulig å legge til kolonner med informasjon om informanten som har skrevet ordet. Jeg haket av for at korpuset skulle vise informasjon om hjemland, 3. og 4. språk, engelskferdigheter, antall kurstimer i norsk og alder. Under er et resultatord i kontekst med informasjon om informant som har skrevet ordet.

treff	Språktest	Språk 4	Språk 3	Engelskferdigheter	Kurstimer	Alder	Hjemland
krokodille	hestflo og mange andre dyr som jeg ikke kan nevne her. I	Språkprøven		grunnivå	1501-2000	30	Somalia

Figur 8: søk med informasjon om informanten i ASK

Denne informasjonen er ikke essensiell for å svare på problemstillinga mi, men kan likevel være relevant når jeg diskuterer mulige grunner til en ORT. Vi kan merke oss kolonnene «Språk 3» og «Språk 4» er tomme – det er ikke alltid informantene har oppgitt all informasjon om bakgrunnen sin.

For å få opp en komplett liste over alle tekstene skrevet av SOMALI, slik at jeg fikk en oversikt over bakgrunnen til samtlige av informantene, har jeg haket av for *morsmål:somali* og skrevet inn koden <document> i starten av koden i søkefeltet.

Avansert søk | bytt til Enkelt søk | Søkehistorie ...

```
<document>[ ] \ <> \\ <del> :: language = "somali"
```

Figur 9: kode for søk på tekstene i sin helhet i ASK

Resultatet til søket ser da slik ut i korpuset:

Tekst-id		treff
h0684	</document>	Oppgave A Migrasjon t
h0690	</document>	OPPGAVE B Stadig fler
h0691	</document>	OPPGAVE C Et samfun
h0694	</document>	OPPGAVE A Å være kj

Figur 10: liste med tekst-id vist i kwic-modus i ASK

I enkelte tilfeller har jeg ønsket/hatt behov for å se et enkeltord i en større sammenheng enn den vi ser i kwic-modus. Da har jeg gått inn i den enkelte teksten. Dette gjør man ved å gå inn trykke på et ord i teksten, og velge *vis kontekst* øverst i venstre hjørne.



Figur 11: å finne hele teksten i ASK

Visningsmodus: ID: h0684

|| Oppgave A

|| **Migrasjon** fra en del av verden til en annen har alltid vært veldig mange nordmenn til Amerika på 1800- og 1900- taller Blant annet reiste mange bort fra sine hjemlandet på grunn i knyttet med livsforhold. || I dag fortsetter mange folk i verdi

Figur 12: utsnitt av teksten i sin helhet i ASK

I figur 12 ser vi et utsnitt av teksten i sin helhet. Tekstens ID er vist øverst på sida, i tillegg til at overskrifta viser hvilken oppgave informanten har svart på.

6.6 Informanter

For å svare på problemstillinga mi var det visse krav til hvem informantene mine er. Det første kravet var at de har norsk som andrespråk. I tillegg var det et krav at informantene var i voksen alder, ettersom barn og voksne lærer seg språk på ulike vis. I ASK har alle informantene norsk som andrespråk, og alle er voksne eller i slutten av tenåra. Informantene mine har en gjennomsnittsalder på 29,8 år.

Det er essensielt at informantene mine har somali som morsmål, ettersom jeg skal undersøke om jeg finner transfer fra somali i mellomspråk. I korpuset er det 107 informanter med somali som morsmål, og det er deres tekster jeg har brukt som datagrunnlag.

De fleste av informantene, 100 stk., har tatt Språkprøven, mens sju stk. har tatt Høyere språkprøve i norsk. En kan argumentere for at en ikke burde sammenligne informanter fra disse to informantgruppene, da de er på ulike nivå og ulike punkt i innlæringsfasen. Likevel er andrespråklæring dynamisk, og alle typer feil kan oppstå på alle innlæringsnivå, som nevnt i kap 4. Derfor har jeg valgt å se på ORT til begge gruppene – de ORT jeg forsker på,

finner en også hos informanter på høyere nivå.

Det er også viktig å nevne at alle informantene i ASK-korpuset, inkludert mine, har bestått språkprøvene de tok. Dette er i henhold til kravet til Jarvis (2000) om at informanten bør være på et generelt ferdighetsnivå før man undersøker transfer i mellomspråket, jf. kap. 6.1 om metodologisk grunnlag.

Som vi så i forrige delkapittel, har jeg hatt tilgang til, og samlet inn, informasjon om bakgrunnen til informantene mine. En oversikt over dette er å finne i appendiks 2. Dette er informasjon om hjemland, alder og 3. og 4. språk. Spesielt er det verdt å merke seg at det bare er de som har tatt Høyere språkprøve i norsk, som har oppgitt informasjon om 3. og 4. språk, og jeg har dermed ikke informasjon om hvilke andre språk enn somali og norsk de andre informantene har kompetanse i, og ikke full oversikt over flerspråkligheten til disse informantene. Det er likevel rimelig å anta at de har kunnskap om flere språk enn somali og norsk med tanke på det komplekse lingvistiske landskapet i Somalia (jf. kap. 1.1). Jeg kommer til å ta opp dette i diskusjonsdelen.

Til sist vil jeg kommentere på tidspunktet datamaterialet ble samlet inn. Tekstene jeg studerer, er samlet inn for 14 år siden. En kan spørre seg om det mellomspråket SOMALI produserer i dag vil være annerledes enn det var for 14 år siden med tanke på at norsk og somali har endret seg litt siden den gangen. Jeg tror ikke forskjellen påvirker min studie, ettersom jeg ser på fonologi og ortografi. Det tar svært lang tid for fonologiske trekk å endre seg, og det er lite trolig at stavelsesstrukturen i somali har endret seg siden den gang. Ortografien i norsk og somali er bestemt av statlige normer, og disse normene har ikke endret seg siden materialet ble samlet inn. Derfor mener jeg, i den grad jeg kan generalisere, at generaliseringer jeg gjør basert på en analyse av disse 14 år gamle tekstene, også burde gjelde for SOMALI i dag.

6.7 Utvalg av data

Datamaterialet består av 48 912 løpeord¹⁷ skrevet av SOMALI, og det er totalt 3048 ord tagget med feilkoden ORT. Jeg ser ikke på alle disse i studien min – det er bare visse ORT som er relevante for min stavelsesbaserte problemstilling. I dette delkapittelet gjør jeg rede for utvelgelsesprosessen av dataene mine og hvilke avgrensninger jeg har gjort.

I utvelgelsesprosessen har jeg valgt en deskriptiv tilnærming. Jeg har forsøkt å ikke anta for mye om hva som ligger til grunn for en bestemt ORT. I stedet har jeg tatt med alle

¹⁷ Unike ord, der man også tar med ord som er skrevet likt.

ORT som *kan*, rent deskriptivt, skyldes stavellesstrukturer. Med det mener jeg at opptakt og/eller koda i en stavelse ser annerledes ut enn det ville gjort om en skrev ordet etter normen, eller at en stavelse har blitt satt inn eller tatt bort fra ordet. En oversikt feil som kan inngå i en slik definisjon følger under.

1. Avvik i opptakt eller koda
 - 1) Bortfall av grafem
 - 2) Ombytting av to grafem
 - 3) Innsetting av konsonantgrafem
 - 4) Ett konsonantgrafem erstatter et annet
2. Innsetting av stavelse
 - 1) Innsetting av vokal
 - 2) Innsetting av vokal + innsetting av tilhørende konsonant i opptakt eller koda
3. Bortfall av stavelse
 - 1) Bortfall av vokal
 - 2) Bortfall av vokal + bortfall av tilhørende konsonant i opptakt eller koda

Det skal nevnes at termene jeg bruker for å beskrive ORT, ligner veldig på de jeg bruker når jeg snakker om reparasjonsstrategier. Det er likevel ikke det jeg snakker om her – det er først i diskusjonskapittelet at jeg kommer til å vurdere om det er en *strategi* som ligger bak et bestemt ORT. I denne delen nøyer jeg meg med å poengtere at det er et avvik i stavellesstrukturen.

I studien min har jeg ikke sett på ORT som går på vokalgrafem. Dette er ikke noe som er vesentlig i stavellesens *struktur*. Stavellesens vil ha samme stavellesstruktur og fonotaks selv om vokalen er endret. Problemstillinga mi rommer derfor ikke denne typen ORT. Jeg har heller ikke tatt med ORT som går på dobbeltkonsonant for enkeltkonsonant og enkeltkonsonant for dobbeltkonsonant. I stavellestemplatene jeg har lagt til grunn i kap 2.3.2 er ikke denne feilen noe som angår stavellesstrukturen – de skal bare markere om en stavelse er lang eller kort. I en oppgave som hadde lagt en moraisk stavellesstruktur til grunn, kunne dette vært relevant å se på, men spørsmålet om dobbeltkonsonant/enkeltkonsonant faller utenfor denne oppgavens mål.

Jeg har som sagt forsøkt å være deskriptiv i utvalget mitt av ORT. Likevel har det vært nødvendig å gjøre noen avgrensninger med hensyn til oppgavas omfang. Jeg har valgt å ikke analysere ORT som ut ifra kontekst sannsynligvis skyldes avvik i *morfologi*. Dette gjelder spesielt for ORT der konteksten tilsier at ordet i målspråket ville stått i presens og blitt

realisert med presensending, men at denne presensendinga ikke er realisert. Det er likevel ikke slik at jeg mener disse typene ORT ikke *kan* skyldes transfer av stavelsesstrukturer, men i denne oppgava kommer jeg ikke til å diskutere slike ORT, og hva de skyldes. Jeg har også valgt å se bort fra ORT som går på særskrivning, store vs. småbokstaver, tegnsetting og tallord under 12 skrevet med sifre.

Datamaterialet mitt består av enkeltord. Det betyr imidlertid ikke at jeg ikke har sett på sammenhengen orda står i. Som jeg redegjorde for i kap. kan sammenhengen orda står i, være essensiell fordi stavelser kan krysse ordgrenser. Derfor har jeg vurdert om dette gjelder for stavelsene jeg undersøker, og tatt det med i diskusjonen der ord før og etter studieobjektet mitt kan ha spilt en rolle.

6.8 Kategorisering av data

I forrige delkapittel gjorde jeg rede for hvordan jeg valgte ut dataene mine. Basert på utvalget har jeg endt opp med å undersøke 409 ord som kan skyldes transfer av stavelsesstrukturer. I dette delkapittelet beskriver jeg hvordan jeg har kategorisert ORT i disse orda.

I kategoriseringsprosessen tok jeg sikte på å være deskriptiv, slik som i utvalgsprosessen. Derfor gjelder mange av de samme prinsippa som i utvelgelsesprosessen. Jeg har kategorisert orda etter tre kjennetegn:

1. Hvor i ordet det er avvik
2. Hvor i stavelsen det er avvik
3. Hvilke(t) grafem som er involvert i avviket

Videre har jeg samla ord i grupper om de har ett eller flere av disse trekka til felles, slik at jeg har fått en oversikt over hvilke feil som går igjen.

I tillegg har jeg gjort detaljerte transkripsjoner alle orda. I tillegg til å vise hvordan ordet blir realisert i det skriftlige mellomspråket, har jeg gjort transkripsjoner av hvordan det ville blitt realisert muntlig etter norske ortografiske prinsipper. Videre har jeg oppgitt hvordan målspråksordet ville blitt uttalt i standard østnorsk. Jeg har også oppgitt lemmaet og bøyingsformen til målspråksordet skrevet på normert bokmål, samt en stavelsestranskripsjon av realiseringa til informanten. Alle data, transkripsjoner og klassifikasjoner er å finne i appendiks 1.

7.0 Resultater, analyse og drøfting

I dette kapittelet presenterer jeg resultatene av undersøkelsen min. Dataene jeg har undersøkt består av en rekke mellomspråksord hentet fra sammenhengende tekst. Disse har ortografiske avvik som gjør at de har en annen stavelsesstruktur enn den vi finner i målspråksordet. I analysen har jeg undersøkt hvilke av disse avvika som fører til at stavelsesstrukturen i mellomspråksordet blir likere den somaliske stavelsesstrukturen. Basert på undersøkelsen diskuterer jeg om avvika som *kan* være transfer av stavelsesstrukturer, *sannsynligvis* er det. Jeg diskuterer altså hvorvidt avvika er et resultat av reparasjonsstrategier med mål om å tilpasse ordet til stavelsesstrukturen i somali, eller om det kan være andre grunner bak avvika.

Datamaterialet mitt består av 48 912 løpeord hvorav 3048 er tagget som ORT. Av disse undersøker jeg til sammen 409 ORT (jf. kap 6.7 om utvalg og kategorier). Under følger en tabell med oversikt over hvor mange ORT det er i de fire kategoriene *grafeminnsetting*, *grafembortfall*, *grafemerstatning* og *grafemombyting*, og hvilke av disse som *kan* skyldes transfer av stavelsesstrukturer. Tabellen viser også *hvordan* de ulike avvika gjør at stavelsen blir tilpassa somalisk stavelsesstruktur (jf. 2.3.2 om somaliske stavelsesstrukturer).

Ortografiske avvik fra bokmålsnormen av 2005	Grafem-innsetting	Grafem-bortfall	Grafem-erstatning	Grafem-ombyting	Total
<i>Antall ORT</i>	98	238	53	23	412
Resulterer i en enklere konsonantgrafemgruppe	51	113	1	1	166
Resulterer i at man unngår <m> ordfinalt	0	1	1	0	2
Resulterer i at man unngår opptaktsløs stavelse ordinitialt	1	4	0	0	5
Resulterer i at man unngår <t> i koda	0	32	6	0	38
Resulterer i at man unngår <k> i koda	0	0	4	0	4
Resulterer i bortfall av vokalkjerne i andre stavelse i trisyllabiske ord med suffiks	0	23	0	0	23
Resulterer <i>ikke</i> i at stavelsen blir likere somalisk stavelsesstruktur	47	65	41	22	172
Antall ORT som <i>kan</i> være resultat av en reparasjonsstrategi	52	173	12	1	=238

Tabell 11: oversikt over antall ORT i de ulike kategoriene.

I tabell 11 ser vi at det er flest grafeminnsettinger og grafembortfall i dataene mine. Det er altså i disse kategoriene det er flest ORT, og derfor er det interessant å undersøke disse tendensene spesielt nøye. Jeg skal også omtale ORT som går under kategoriene grafemerstatning og grafemombyting, men disse kommer ikke til å bli viet like mye plass som de to første kategoriene, da det er langt færre forekomster her.

Vi ser at 165 av 237 ORT som *kan* skyldes transfer av stavelsesstrukturer, resulterer i at grupper av konsonantgrafemer blir mindre komplekse, og av den grunn diskuterer jeg også denne kategorien mer inngående enn de andre kategoriene. ORT som resulterer i at man unngår <t> i koda og ORT som resulterer i at den midterste stavelsen i et trisyllabisk ord faller bort, vil også bli diskutert mer inngående, fordi det også er visse tendenser til dette. Vi ser også at 172 av 407 avvik *ikke* gjør at stavelsesstrukturene blir mer like somaliske stavelsesstrukturer. Jeg kommer til å omtale disse avvika og gjøre rede for hvorfor de ikke kan skyldes transfer av somali, men disse vil ikke bli drøfta like inngående som orda som *kan* skyldes transfer.

7.1 Innsetting av grafem

Under er en oversikt over innsetting av grafemer i datamaterialet mitt som *kan* være resultat av en reparasjonsstrategi. Her har jeg altså ikke inkludert de 47 avvika i datamaterialet mitt som *ikke* kan forklares med transfer av stavelsesstrukturer, men disse vil også bli omtalt underveis i kapittelet.

Plassering	Vokalgrafem	Konsonantgrafem	Antall informanter
Ordmedialt	36	1	19
Ordfinalt	14	0	13
Ordinitialt	0	1	1
<i>Sum:</i>	<i>50</i>	<i>2</i>	<i>33</i>

Tabell 12: innsetting av grafemer

Tabell 12 viser oss at det er flest eksempler at vokalgrafeminnsettinger fører til at målspråksordet får likere struktur som somaliske strukturer. Vi ser at det bare er ett tilfelle der innsetting av konsonantgrafem kan være et resultat av en reparasjonsstrategi. Under følger tre eksempler på innsetting fra materialet mitt.

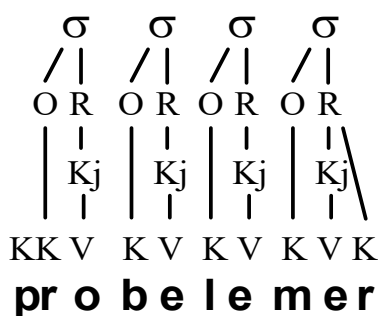
- (1) N2 <probelemer> – N1 <pro.ble.mer>
- (2) N2 <bedrifte> – N1 <be.drift>
- (3) N2 <helge> – N1 <elg>

I (1) ser vi mellomspråksordet <probelemer> med målspråksord <problemer>. Der blir vokalgrafemet <e> satt inn mellom to konsonanter, og <l>. I (2) ser vi også innsetting av vokalgrafem, men her skjer dette helt til sist i målspråksordet <bedrift> og danner dermed en ny ordfinal stavelse. En mulig analyse av både (1) og (2) er at vokalinnsettingene er reparasjonsstrategier som ikke tillater konsonantgrupper for at stavelsene skal bli mer like somaliske stavelser. I (3) ser vi også eksempel på en vokalinnsetting ordfinalt som kan ha som mål å bryte opp en gruppe av konsonantgrafemer. I tillegg ser vi et eksempel på innsetting av et konsonantgrafem ordinitialt, og det er mulig det er benyttet en reparasjonsstrategi for å unngå vokal ordinitialt. I 7.1.1 gjør jeg rede for ord som i (1), der et vokalgrafem har blitt satt inn mellom to konsonanter, i 7.1.2 snakker jeg om ord som i (2), der et vokalgrafem har blitt satt inn *etter* to eller flere konsonantgrafemer og i 7.1.3 snakker jeg om konsonantinnsetting, som i (3).

7.1.1 Innsetting av vokalgrafemer ordmedialt

Med 38 tilfeller av vokalgrafeminnsetting ordmedialt, er dette den hyppigste typen ORT i kategorien grafeminnsetting. I det følgende diskuterer jeg om alle disse avvika kan skyldes reparasjonsstrategien epentese, og dermed transfer av stavelsesstrukturer, eller om det kan være andre grunner til avvika.

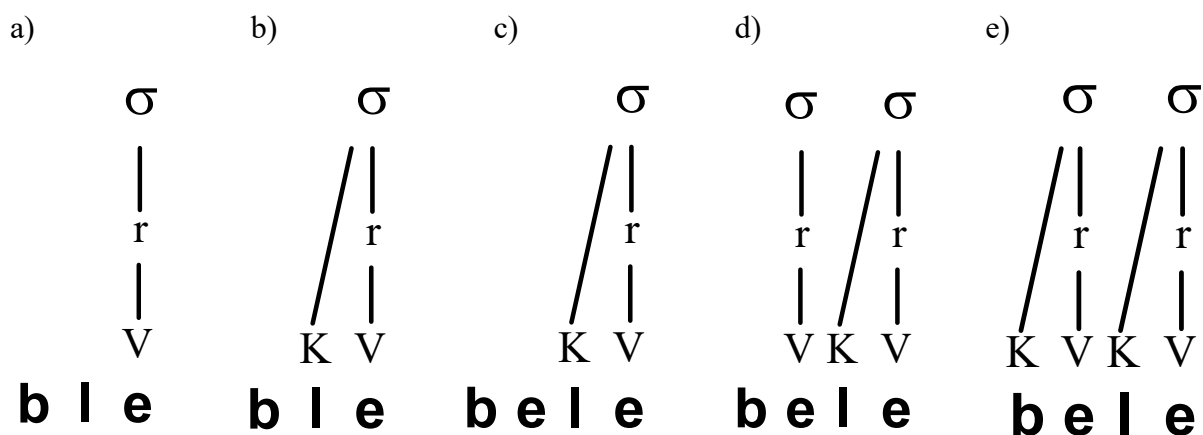
Alle innsettingene av vokalgrafemer medialt skjer mellom to konsonantgrafemer. Denne typen innsetting ser vi eksempel på i ordet mellomspråksordet <probelemer> med målspråksformen <problemer>. I figur 13 er <probelemer> analysert.



Figur 13: mulig analyse av <probelemer>

I figuren ser vi at mellomspråksordet har fire stavelser. Om vi sammenligner med målspråksordet <pro.ble.mer>, ser vi at målspråksordet har tre stavelser, altså én stavelse mindre enn mellomspråksordet. I modellene i resten av kapittelet, kommer jeg ikke til å benevne verken opptakt, kjerne eller koda for at figurene skal være mer leservennlige, men

alle disse tre posisjonene er altså en del av alle stavelsene vi skal se på. I det følgende skal vi se på hvordan den ordmediale stavelsen <ble> blir til <be.le>. Basert på Archibald (1998) sin analyse av stavelседanninga til ordet /pilanti/ (jf. kap 2.4), kan vi tenke oss følgende stavelседanningsprosess for <be.le>:

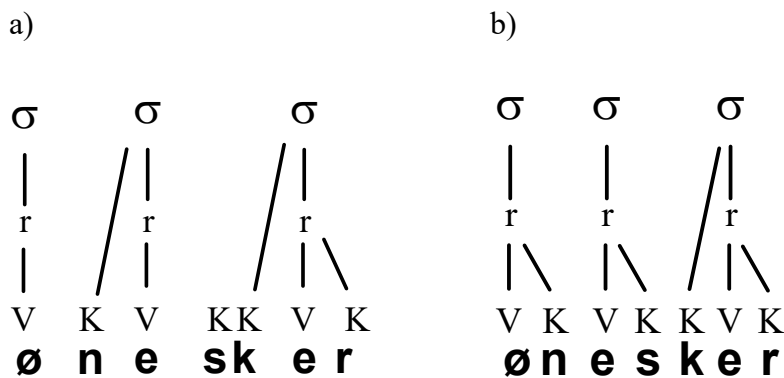


Figur 14: stavelседanningsprosessen fra <ble> til <be.le>

I figur 14 ser vi altså hvordan informanten kan ha tilpasset stavelsen <ble> slik at den blir i tråd med somaliske stavelsesstrukturer i mellomspråket. I figur 14a ser vi at stavelseskjernen <e> knytter seg til et rim, som igjen knytter seg til en stavelse. Figur 14b viser neste steg i stavelседanningsprosessen der mulige konsonanter kan knytte seg til stavelseskjernen og danne opptakt eller koda. Om informanten hadde lagt norske stavelsesstrukturer til grunn, hadde det ikke vært et problem at to konsonanter, både og <l>, hadde fylt opptaktsposisjonen til <e>. Likevel ser vi i figur 14b at det kun er <l> som knytter seg til stavelsen, mens fortsatt trenger en stavelse å knytte seg til. Dette kan man altså tolke som at det *ikke* er norske stavelsesstrukturer som ligger til grunn for ordformer <probelemer>, men heller somaliske. Somali har en maksstruktur på KVK, og tillater dermed ikke flere enn én konsonant i opptaktsposisjon. Om en legger somaliske stavelsesstrukturer til grunn, gir det altså mening at ikke blir knyttet til opptakt. Om en i tillegg tar i betraktning at KV-strukturer er den vanligste oppbygginga av stavelser i verdens språk, er dette også i tråd med universelle prinsipper. For at skal få en stavelse å knytte seg til, blir en vokal satt inn til høyre for den stavelsesløse lyden, slik vi ser i figur 14c. Den samme prosessen finner vi også i eksempelet med /pilanti/ der /i/ blir satt inn til høyre for /p/. På lignende vis kan knyttes til en *opptaktsposisjon* i stedet for en kodaposisjon, og prinsippet om maksimal opptakt blir fulgt (jf. kap. 2.5.1). I 14d blir den innsatte vokalen

knyttet til et rim, som blir knyttet til en stavelse, og i figur 14e kan knytte seg til opptakta i den nydannede stavelsen. Resultatet blir altså to separate stavelser som ikke bryter med maksstrukturen i somali, i stedet for én enkelt stavelse som bryter med maksstrukturen. Denne typen innsetting kjenner vi igjen på lydnivå fra Broselow (1972) i eksemplene /Fired/ for /Fred/ og /tiransilet/ for /translate/. Også Eilertsen (1990) finner eksempel på dette med målspråksordet /elv/ som blir uttalt /eləv/. I studiene av skrift finner vi det i Cook (1997) der målspråksordet <course> blir stavet <courese> (jf. kap. 5.2).

I figur 14 blir altså vokalgrafemet satt inn til *høyre* for den stavelsesløse konsonanten slik at den nye stavelsen ender opp med å følge prinsippet om maksimal opptakt. For de fleste av orda er det likevel slik at vokalgrafemet ser ut til å bli satt inn til *venstre* for den stavelsesløse konsonanten. Under ser vi mellomspråksordet <ønesker> fra datamaterialet mitt, analysert på to forskjellige måter.

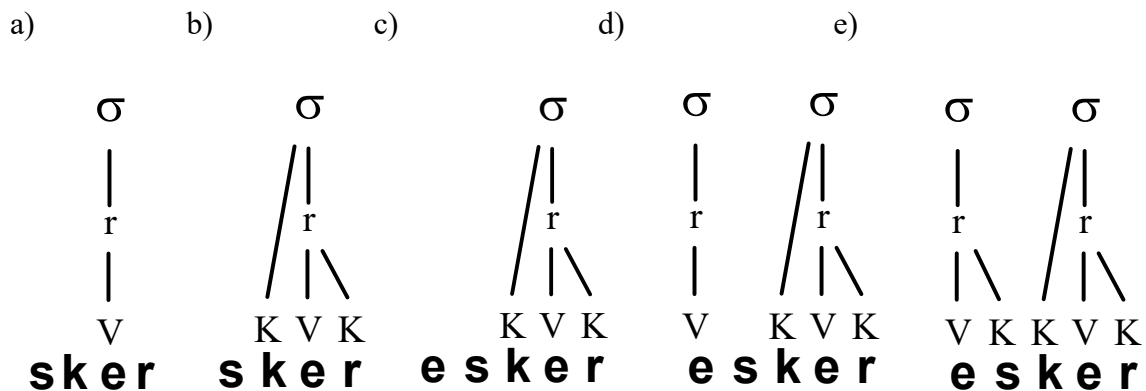


Figur 15: to mulige analyser av <ønesker>

Om vi analyserer målspråksordet <ønesker> etter norsk fonotaks, blir det naturlig å dele det opp slik: <øn.sker>. Vi kunne tenkt oss at den innsatte vokalen <e> var knytta til <n>, og at <e> dermed var satt inn til høyre for en konsonant, slik analysen er i figur 15a. I denne analysen følger alle stavelser prinsippet om maksimal opptakt. Likevel er det flere grunner til at dette er en problematisk analyse. For det første har vi i kap 2.3.2 om somaliske stavelsesstrukturer sett at det *er* tillatt med én konsonant i koda. I målspråksordet <øn.sker> er det bare én konsonant i koda i stavelsen <øn>, og det er dermed ikke behov for å reparere stavelsen til å ligne mer på somaliske stavelsesstrukturer – stavelsen er allerede i tråd med somalisk struktur. For det andre er det en annen stavelse i ordet, <sker>, med strukturen <KKVK>, som *faktisk* bryter med somaliske stavelsesstrukturer. Stavelsen har to konsonanter i opptakt, og danner dermed en konsonantgruppe som ikke er tillatt i en somalisk opptakt. Om en legger analysen <øn.sker> til grunn, er det altså opptakta i den siste stavelsen

som trenger en reparasjon for å passe med den somaliske stavelsen, og ikke koda i den første stavelsen. Analysen <ø.ne.sker> kan med andre ord ikke regnes for å være en reparasjonsstrategi – det er ingen stavelse som blir mer lik den somaliske strukturen.

I 15b ser vi en analyse der det innsatte vokalgrafemet <e> heller får <s> knytta til seg. Vi kan tenke oss følgende stavelserdanningsprosess for ordet <ø.nes.ker>:



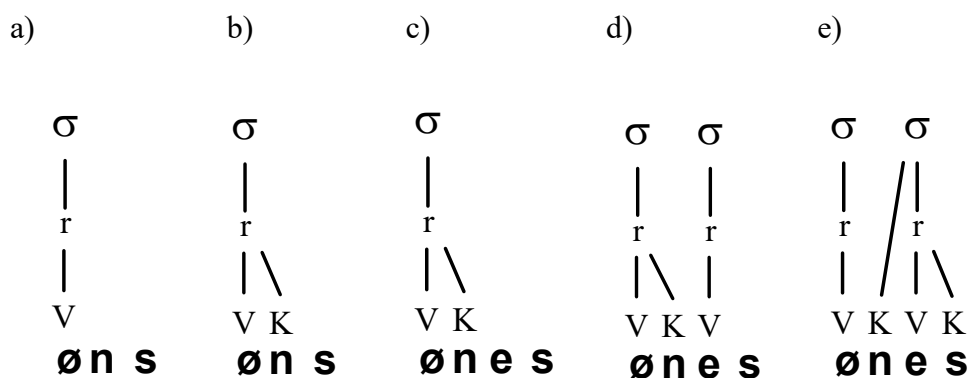
Figur 16: stavelserdanningsprosessen fra <sker> til <es.ker>

I 16 ser vi prosessen fra målspråksstavelsen <sker> til mellomspråksstavelsene <es.ker>. I figur 16a danner <e> kjerne i stavelsen, og blir knytta til et rim, som igjen blir knytta opp mot en stavelsernode. I figur 16b knytter mulige konsonanter seg til stavelsen: <k> fyller opptaktsposisjonen og knytter seg direkte opp mot stavelsernoden, mens <r> fyller kodaposisjon og knytter seg opp mot rimet. <s> blir ikke knyttet til stavelsen. I figur 16c ser vi at <e> blir satt inn til venstre for den stavelserløse konsonanten <s>, og i figur 16d danner denne vokalen kjerne i en helt ny stavelse. I figur 16e blir <s> knyttet til den nye stavelsen, og fyller kodaposisjonen i denne stavelsen. Vi går altså fra å ha én stavelse med kompleks opptakt, til to stavelser uten konsonantgrupper. I denne analysen kan det innsatte vokalgrafemet tolkes som å være et resultat av en epentetisk reparasjonsstrategi, ettersom innsettinga gjør at alle stavelser i ordet er i tråd med somaliske stavelserstrukturer. Argumentet *mot* denne analysen, er at det ikke er i tråd med prinsippet om maksimal opptakt å sette inn en epentetisk vokal til *venstre* for den stavelserløse konsonanten. Hadde mellomspråket fulgt dette prinsippet, ville det vært naturlig at mellomspråksordet var <øn.se.ker>, slik at den epentetiske vokalen ble satt inn til *høyre* for den stavelserløse konsonanten.

Det er mulig å legge en analyse til grunn der innsettinga av <e> fungerer som en

reparasjonsstrategi i *koda*. I de to foregående analysene har jeg lagt norsk fonotaks til grunn, men det er godt mulig at informantene mine analyserer målspråksordet etter en mellomspråksfonotaks. Som vi så i kap 4.0 om mellomspråk, er mellomspråk å regne som egne språkssystem med egne regler, og dermed er det naturlig å tenke at mellomspråk både har en unik stavelsesstruktur og en unik fonotaks. Det er derfor mulig at informantene mine setter stavelsesgrensa et annet sted enn en nordmann vanligvis ville gjort, og vi kan tenke oss analysene <øns.ker> eller <ønsk.er>. Den første analysen kan motiveres av at det ordet unngår konsonantgruppa <sk> i opptakt. <sk> bryter med sonoritetshierarkiet, og hvis mellomspråket vil være i tråd med dette universelle prinsippet, er analysen <øns.ker> mulig. Den andre mulige løsningen, <ønsk.er>, er også en analyse som er i tråd med sonoritetshierarkiet, men denne er mindre sannsynlig fordi stavelsen her bryter med prinsippet om maksimal opptakt.

Om en legger analysen <øns.ker> av målspråksordet til grunn, er det altså en konsonantgruppe i koda i den ordinitiale stavelsen. Denne må brytes opp om stavelsen skal være i tråd med somaliske strukturer. I mellomspråksordet <øn.es.ker> blir den brutt opp. Vi kan tenke oss følgende stavelsesdanningsprosess:



Figur 17: stavelsesdanningsprosessen fra <øns> til <øn.es>

Figur 17 viser altså hvordan den ene stavelsen <øns> blir til de to stavelserne <ø.nes> gjennom en reparasjonsprosess. I figur 17a ser vi at <ø> danner kjernen i stavelsen og knytter seg til et rim, som knytter seg til stavelsesnode. I figur 17b knytter mulige konsonanter seg til stavelsen, og <n> fyller koda, mens <s> ikke knytter seg til stavelsen. I figur 17c blir en <e> satt inn til *venstre* for den stavelsesløse sibilanten <s>, og denne danner en ny stavelse i figur 17d. I figur 17e blir <s> knytta til stavelsesnoden, og <n> går fra å fylle koda i den ordinitiale stavelsen, til å fylle opptakt i den ordfinale stavelsen. Dette er for å følge prinsippet om maksimal opptakt, men vi kan også tenke oss analysen <øn.es>, der <n> fortsetter å tilhøre

kodaposisjon i den ordinitiale stavelsen og dermed *ikke* følger prinsippet om maksimal opptakt – som nevnt er det nærliggende å tenke at det er en helt egen fonotaks i mellomspråket, og universelle regler for stavelsesdanning *må* ikke bli fulgt i denne mellomspråksfonotaksen. Uansett om vi legger analysen i figur 17 eller analysen <øn.es> til grunn, ender vi opp med to stavelser der ingen posisjoner i noen av stavelsene har konsonantgrupper. I denne analysen fungerer altså innsettinga av <e> som en reparasjonsstrategi som gjør at stavelsen blir i tråd med maksstrukturen KVK i somali. Noen eksempler på datamaterialet mitt som må analyseres samme måte som <ønesker> i figur 16 eller 17, følger for at innsettinga skal bli tolka som en epentetisk reparasjonsstrategi.

- (4) N2 <elesket> – N1 <el.sket>
- (5) N2 <islamiske> – N1 <islam.ske>
- (6) N2 <blomester> – N1 <blom.ster>

Her må altså den innsatte <e>-en få en konsonant knytta til koda om vi skal tolke det som en reparasjonsstrategi. I (4) og (5) kan vi tolke det som at opptakta <sk> blir reparert når vokalgrafemet blir satt inn og danner ny kjerne, og <s> kan danne koda i denne nye stavelsen. I (6) skjer det samme med <st>. Verdt å merke seg er at mange av stavelsene i datamaterialet mitt som ser ut til å repareres, inneholder frikativene <s>, og i disse eksemplene bryter <st> og <sk> med sonoritetsprinsippet ved at plosivene <t> og <k> kommer *etter* <s>, selv om <s> er mer sonor og skulle vært nærmere stavelseskjernen om stavelsen hadde fulgt sonoritetsprinsippet. Det er altså mulig at akkurat disse stavelsene repareres oftere enn andre stavelser fordi mellomspråket helst vil følge sonoritetsprinsippet. Likevel er ikke /s/ + konsonant en veldig markert kombinasjon i opptakt, og mange språk bryter med sonoritetsprinsippet ved å ha /s/ + konsonant i sitt språk (van Oostendorp, Ewen & Rice, 2011:4)

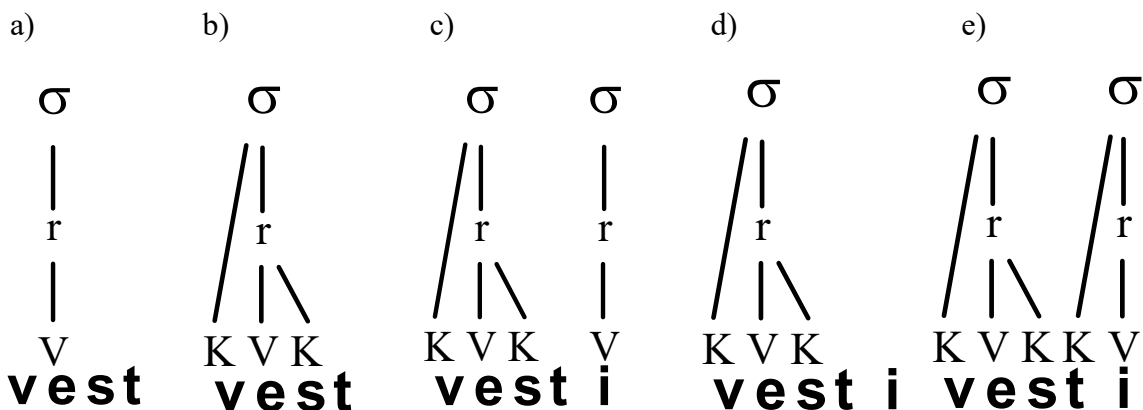
I eksempelet <probelemer> en vokal satt inn *mellom* to konsonanter som danner en konsonantgruppe i opptakt. I eksempelet <ønesker> har jeg vist to mulige analyser der en konsonantgruppe blir brutt opp: Legger vi analysen <øn.sker> til grunn, kan vi si at den komplekse opptakta <sk> blir brutt opp ved at en setter inn en epentetisk vokal *før* den komplekse konsonantgruppa. Legger vi analysen <øns.ker> til grunn, blir den epentetiske vokalen satt inn mellom to konsonanter som danner en kompleks stavelse, akkurat som i <probelemer>. Det finnes også ord i datamaterialet mitt der en epentetisk vokal blir satt inn *etter* en konsonantgruppe. Under er en analyse av ordet <vestilige>.



Figur 18: mulig analyse av <vestilige>

I 18 er mellomspråksordet <vestilige> analysert som <ves.ti.li.ge>. Merk at <g> ikke er uttalt i målordet <vestlige> i standard østnorsk, og siden <g> ikke har noen lydlig representasjon, er den heller ikke knyttet til en stavelse i analysen min. Men en må huske på at jeg, på bakgrunn av mine skriftlige data, ikke kan si noe om hvordan denne informanten ville *uttalt* ordet. Det er absolutt mulig at <g> i mellomspråksordet <vestilige> har en lydlig representasjon, og dermed har annen ortografi enn norsk. Men i denne analysen av mellomspråksordet har ikke <g> en lydlig representasjon.

Om en analyserer målspråksordet <vestlige> etter norsk fonotaks, er det naturlig at stavelsesgrensa går mellom <t> og <l>, slik: <vest.lige>. I denne analysen ser vi at det er en kompleks koda med to konsonanter i den ordinitiale stavelsen, <sk>, og denne strukturen strider med somaliske stavelsesstrukturer som kun tillater én konsonant i koda. Basert på analysen <vest.lige> kan en tenke seg at en innlærer kan benytte seg av en reparasjonsstrategi for å tilpasse ordet til somaliske stavelsesstrukturer. Under viser jeg en mulig stavelsesdanningsprosess av <ves.ti>.



Figur 19: stavelsesdanningsprosessen fra <vest> til <ves.ti>

Figur 19 viser altså hvordan målspråksstavelsen <vest> blir omdannet til de to stavelserne <ves.ti> i mellomspråksordet. I figur 19a danner <e> kjernen i en stavelse, og i figur 19b knyttes mulige konsonanter opp mot denne stavelsen. Her fyller <v> opptakt og <s> koda, mens <t> ikke knytter seg til stavelsen. I figur 19c blir et vokalgrafem, <i>, satt inn til høyre for den stavelsesløse konsonanten <t>, og denne vokalen danner kjernen i en ny stavelse i figur 19d. Til slutt kan den stavelsesløse <t>-en fylle opptakta i den nye stavelsen. Gjennom denne prosessen har vi gått fra én stavelse med en konsonantgruppe i koda, til to stavelser der verken opptakt eller koda i de to stavelserne inneholder en konsonantgruppe.

Selv om det er mange ord der det er sannsynlig at informanten har satt inn et vokalgrafem for å reparere ordet, ser vi mange tilfeller av at ikke *alle* konsonantgrupper i ordet blir brutt opp. Vi har allerede sett på eksempelet <probelemer> der konsonantgruppa <bl> blir brutt opp med en epentetisk <e>. Om vi ser på den ordinitiale stavelsen, derimot, ser vi at opptakta er fylt av *to* konsonanter, og denne konsonantgruppa bryter med somaliske stavelsesstrukturer. Det er altså slik det er flere konsonantgrupper i målspråksordet, men i mellomspråksordet har kun den ene blitt reparert. Andre eksempler der én konsonantgruppe har blitt reparert, mens en annen står urørt, er:

- (7) N2 <blomester> – N1 <blomster>
- (8) N2 <flyktning> – N1 <flykting>
- (9) N2 <ekstremister> – N1 <ekstremister>

Både (37) og (38) har en initial konsonantgruppe som ikke er reparert, til tross for at andre konsonantgrupper i ordet har blitt reparert med epentese. I (39) er denne konsonantgruppa medial. Det må påpekes at (39) har fått brutt opp konsonantgruppa <str> *litt* ved at vokalgrafeminnssettinga fører til at <s> kan knytte seg til koda i den nye stavelsen. Selv om ordet fortsatt inneholder konsonantgrupper, er disse mindre komplekse enn de var før reparasjonen. Også i Hvenekilde (1985) finner vi forekomster av slike hybridord. For eksempel skriver en av informantene hennes med tyrkisk som morsmål <bolomster>. Her kan innsettinga av vokalgrafemet fungere som en reparasjonsstrategi, men den samme strategien er ikke brukt for konsonantgruppa <st> senere i ordet. Hybridorda i både min og Hvenekildes studie er klare eksempler på at mellomspråk er dynamiske (jf. kap. 4 om mellomspråk). Det er ikke slik at en regel i mellomspråket er rigid, og alltid må følges – språkbrukeren vurderer hva hen vil bruke av den språklige ressursbanken sin i den enkelte situasjon, og ulike stavelsesdanningsprosesser kan skje, også innenfor samme ord.

Det er en del vokalgrafemer i materialet mitt som settes inn mellom

konsonantgrafemer, men som også kan ha andre forklaringer. Under er tre eksempler på slike innsetninger.

- (10) N2 <ungegutt> – N1 <unggutt>
(11) N2 <ungedom> – N1 <ungdom>
(12) N2 <hjemmeland> – N1 <hjemland>

I disse eksemplene er <e> satt inn mellom to ord som danner et sammensatt ord. Derfor er det mulig dette er avvikende bruk av sammensetningsformativ, og ikke en reparasjonsstrategi.

Om vi gjør en analyse av målspråksordet til (10), <unggutt>, kan vi etter norsk fonotaks dele det opp som <ung.gutt>, <KVK.KVK>. Strukturen til begge stavelsene i dette ordet er altså i tråd med somaliske stavelsesstrukturer, for somali tillater ett segment i både opptakt og koda, og dermed er det ikke noe behov for en reparasjon.

En annen gruppe ord med innsatte grafemer som muligens kan ha andre forklaringsfaktorer enn reparasjonsstrategier, er ord der <e> blir satt inn mellom grafemene <d> og <r>, <t> og <r> og <r> og <n>. Under følger noen eksempler på slike ord:

- (13) N2 <foreldere> – N1 <foreldre>
(14) N2 <brødere> – N1 <brødre>
(15) N2 <søstere> – N1 <søstre>
(16) N2 <utfordere> – N1 <utfordre>
(17) N2 <forendere> – N1 <forandre>
(18) N2 <vinnerene> – N1 <vinnerne>

Vi ser at alle nesten alle innsettingene bryter opp konsonantgruppene <dr> og <tr>, og at stavelsene som et resultat av dette er i henhold til somaliske stavelsesstrukturer ved at de bryter opp konsonantgrupper. Derfor kan det hende at denne innsettinga er en reparasjonsstrategi. <rn> blir også brutt opp i (18), men dette er en digraf som står for ett segment, og dermed er det på lydnivå egentlig ingen konsonantgruppe å bryte opp. Det kan likevel være at dette er en strategi for å unngå dette segmentet, da somali ikke har segmentet /ŋ/ i lydinventaret sitt. Ved å bryte opp digrafen <rn> blir det i stedet to monografer <r> og <n> som står for segmenter som finnes i somali.

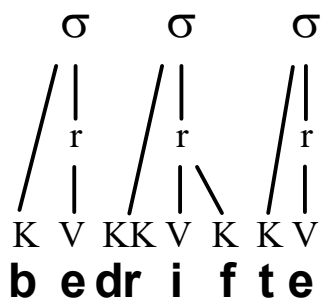
Samtidig som dette kan være en reparasjonsstrategi, er det også mulig at dette er analogi. Mange ord i det norske språket har endelsen -ere og -ene. Det kan tenkes at informantene følger hovedregelen også i disse orda, selv om dette er ord som får bortfall av vokal og ikke følger denne regelen.

Vokalgrafem kan altså settes inn medialt for å reparere stavelser slik at en unngår en kompleks stavelse. Vokalgrafemene kan settes inn mellom konsonantene i konsonantgruppa, etter konsonantgruppa, eller før konsonantgruppa. Vi ser at en del ord har brutt opp én konsonantgruppe, men at andre konsonantgrupper i samme ord ikke har blitt brutt opp. I en gruppe ord kan også morfologisk analogi være forklaringa bak innsettinga, selv om dette også kan være en reparasjonsstrategi, og i annen gruppe ord er det mulig at avvikende bruk av sammensetningsformativ er forklaringa heller enn reparasjonsstrategier. I materialet mitt er det likevel 22 konsonantinnsetninger medialt i ordet som vanskelig kan skyldes noe annet enn reparasjonsstrategier.

7.1.2 Innsetting av vokalgrafemer ordfinalt

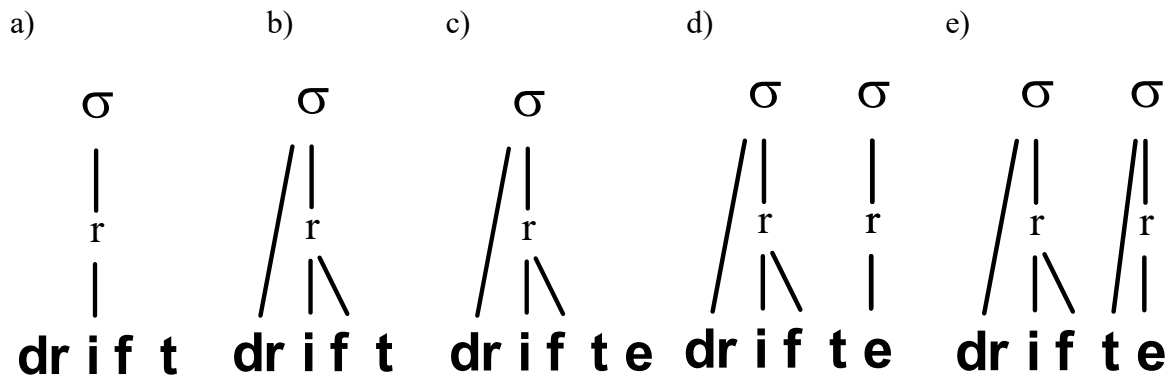
Som vi ser av tabell 12 er det 14 tilfeller av vokalgrafeminnsetting ordfinalt som *kan* skyldes reparasjonsstrategier. Under drøfter jeg *om* disse innsettingene er reparasjonsstrategier, eller om det kan være andre grunner til innsettingene.

I forrige seksjon drøftet vi innsetting av vokaler ordmedialt. Disse ble stort sett satt inn *mellom* to konsonantgrafem, og her er det lett å se at dette er en reparasjonsstrategi. Men innsetninger kan også fungere som reparasjonsstrategier når man setter dem inn *etter* en gruppe av konsonantgrafemer, slik vi så med eksempelet <ves.ti.lige>. Under er en mulig analyse av et ord fra datamaterialet mitt som har fått vokalgrafeminnsetting ordfinalt.



Figur 20: mulig analyse av <bedrifte>

I figur 20 er en mulig analyse av ordet <bedrifte> med målspråksord <bedrift>. En kan tenke seg denne mulige mellomspråksprosesseringen av stavelen <drift> til mellomspråket:



Figur 21: stavelserdanningsprosessen fra <drift> til <drif.te>

I figur 21 ser vi hvordan målspråksstavelen <drift> kan bli til mellomspråksstavelene <drif.te>. I 21a blir vokalen <i> knytta til en stavelse, i 21b blir alle mulige konsonanter knytta til koda. Her får konsonantgruppa <dr> fylle opptakt, mens koda bare blir fylt av <f>. <t> forblir stavelsesløs. I 21c blir en vokal satt inn til høyre for den stavelsesløse konsonanten, og i 21d danner denne vokalen en stavelse. I 21e kan <t> knytte seg til opptakta i den nye stavelsen. Slik blir kodaen <ft> reparert til å passe med somaliske stavelserstrukturer. Vi finner den samme typen reparasjon på lydnivå i Eilertsen (1990), der en informant sier /ɛlvə/ for målspråksordet /elv/ (jf. kap 5.1 om tidligere forskning på transfer av stavelserstrukturer).

Det interessante med denne analysen er det kun tillates ett segment i koda, mens to segmenter kan fylle opptakt. Dette kan være, som diskutert over, et hybridord der noen konsonantgrupper blir brutt opp, mens andre ikke blir det. Et annet alternativ er at det ligger en annen analyse til grunn der <d> knytter seg til den foregående stavelsen som har en ledig kodaposisjon. Da blir stavelsergrensene heller slik: <bed.rif.te>. Andre innsettinger ordfinalt fra datamaterialet er:

(19) N2 <folke> – N1 <folk>

(20) N2 <helge> – N1 <helg>

(21) N2 <valge> – N1 <valg>

I (19) er det altså en kompleks koda i målspråksordet <folk>, og konsonantgruppa <lk> må brytes opp om ordet skal passe med den somaliske kodaen. Når vokalgrafemet <e> blir satt inn etter <k>, får <k> en ny stavelserkjerne å knytte seg til som har ledig plass i opptakt. Det

er den samme prosessen i (20) og (21). Konsonantgruppa <lg> bryter med den somaliske kodastrukturen, og en epentetisk <e> blir satt inn til høyre på for <g>, slik at <g> kan knytte seg til en annen stavelseskjeerne.

Et annet ord fra datamaterialet mitt som også ser ut til å følge dette mønsteret, er <gange> ‘gang’. Det er jo også her en gruppe med konsonantgrafemer i koda som tilsynelatende strider mot den somaliske strukturen. Det er imidlertid en viktig forskjell, for mens <ft>, <lk> og <lg> ovenfor er grupper av grafemer som står for *ulike* lyder, er <ng> en digraf som står for den *ene* lyden /ŋ/. Dermed er det i utgangspunktet ikke noe behov for at språkbrukeren reparerer denne stavelsen, for lydlig sett bryter den ikke med somali. En mulig forklaring på at språkbrukeren likevel har satt inn et vokalgrafem etterpå, er at hen legger somalisk ortografi til grunn når hen skriver ordet. I somali har de ikke lyden /ŋ/, og sekvensen <ng> ville på somali representere de to separate segmentene /n/ og /g/. Om en legger dette tankesettet til grunn, betyr jo det at det er to segmenter i koda i dette mellomspråksordet, og det strider med somaliske regler. Dermed gir det mening å sette inn en epentetisk vokal etter monografemet <g>. Dette kan støttes opp av at det norske alfabetet ikke inneholder digrafen <ng> eller noen representasjon av lyden /ŋ/. Når en lærer alfabetet får en altså kun vite om <n> og <g>, og dette kan bidra til at digrafen <ng> som representerer en helt annen lyd, blir tolket som de to separate lydene som finnes i alfabetet.

Et annet eksempel på dette er <sporte> ‘spurt’, der en kan analysere det som at <rt> blir tolka som de separate segmentene /r/ og /t/ i stedet for at det er en digraf som står for /ʁ/, slik ortografien er i målspråksordet. Dermed kan en tenke seg at <t>-en trenger å knytte seg til en ny stavelse for å unngå kompleks koda, og en epentetisk vokal kommer til unnsetning. Det er imidlertid viktig å huske på at dette avviket også kan komme av en mellomspråks*morfologi*. Det kan være vanskelig å skille mellom preteritum og presens perfektum når man lærer norsk som andrespråk, og veien fra formen <(har) spurt>, som konteksten tilsier at det skal være, til formen <sporte> er kort. Det vanskelig å avgjøre om det er fonologi eller morfologi som ligger til grunn her.

Et annet ord i datamaterialet mitt som *kan* skyldes at det er somalisk ortografi som ligger til grunn for avviket, er <lande>. Har kan vi tenke oss at informanten har behandlet <n> og <d> som to separate segmenter der <n> står for /n/ og <d> står for /d/. I kap 3.2 om norsk ortografi skrev jeg om denne sammensetninga, og at det er mulig å tolke den på to måter: enten kan man si at det er ‘en stum d’ ordfinalt, og at vi dermed ikke uttaler /d/-en som en skulle tenke seg at <d> representerer. Eller så kan man si at konsonantsambandet <nd> står for lyden /n/. Dermed kan det enten hende at informanten ikke er klar over, eller har

glemt, at <d>-en ikke uttales, eller at informanten ikke har fått med seg at når <n> og <d> står sammen ordfinalt, fungerer de som en representasjon for lyden /n/. Og når <n> og <d> representerer to separate segmenter i mellomspråksortografien, påvirker dette stavelsen. Da overstiger koda maksimumsgrensa på én konsonant, og det gir mening å sette inn en epentetisk vokal som det siste segmentet, <d>, kan knytte seg til. Ordet er reparert fra <land> (KVKK) til <lan.de> (KVK.KV).

En annen gruppe ord fra datamaterialet mitt der ortografien i somali kan ha spilt en rolle, følger under:

- (22) N2 <gikke> – N1 <gikk>
 (23) N2 <fjelle> – N1 <fjell>
 (24) N2 <musikke> – N1 <musikk>

Vi ser at målspråksorda i alle disse eksemplene er ord med geminanter, altså lange konsonanter, i uttale, og at de blir representert i skrift med dobbeltkonsonant, slik hovedregelen er i det norske språket. Mellomspråksorda har også to like konsonantgrafemer på rad i disse orda. Om vi skal tolke den ordfinale <e>-en som en måte å tilpasse somaliske stavelsesstrukturer til norsk, kan vi ikke tolke disse doble konsonantene som et uttrykk for lang konsonant. I stedet må de tolkes som to separate segment, slik en måtte ha tolket dem om en skulle lagt somalisk ortografi til grunn. I somali uttrykker de nemlig ikke geminanter med dobbeltkonsonant – i stedet praktiserer de dobbeltskriving av vokaler når de skal markere lang vokal, og enkeltskriving av vokaler når de skal markere lang konsonant.

Det finnes imidlertid en annen måte å analysere disse orda på der de *ikke* blir reparert til å passe inn i somalisk stavelsesstruktur. Dette er at dobbeltkonsonantene faktisk representerer geminanter, men at geminanten har fått en ny stavelse å støtte seg på for at geminanten skal havne i opptakt, heller enn i koda. I stedet for somaliske stavelsesprinsipper, er dette i tråd med universelle stavelsesprinsipper der tendensen til maksimal opptakt blir tilfredsstilt. Dette kan også være forklaringa i de følgende eksemplene:

- (25) N2 <bade> – N1 <bad>
 (26) N2 <hale> – N1 <hall>
 (27) N2 <være> – N1 <vær>
 (28) N2 <live> – N1 <liv>
 (29) N2 <tide> – N1 <tid>
 (30) N2 <våre> – N1 <vår>

- (31) N2 <mine> – N1 <min>
 (32) N2 <framtida> – N1 <framtid>

I (25)-(32) har jeg analysert orda som at det siste segmentet i ordfinal koda i målspråksordet, har knyttet seg til den innsatte <e>-en i mellomspråksordet. Om vi ser målspråksorda i lys av somaliske stavelsesstrukturer, er det ikke noe *behov* for å reparere disse orda, for ingen av orda strider mot somalisk struktur (med unntak av den ordinitiale stavelsen i (32), men denne blir ikke forsøkt reparert). Likevel kan altså dette være en måte å få mellomspråket til å følge det universelle prinsippet om maksimal opptakt. Den mest plausible forklaringa for innsettinga i disse orda er imidlertid heller at informanten surrer med bøyingsendinger eller blander med lignende ord. I (30) og (31) er det sannsynlig at informanten rett og slett har blanda entall og flertall av disse pronomenene. I (25) er det sannsynlig at informanten blander substantivet <bad> med verbet <bade>, noe som også kan ha skjedd i (27). I (26) er det mulig at skrivemåten til to ulike substantiv, <hall> og <hale>, er blanda. I (28) og (30) finnes dativformene <tide> og <live> i de faste uttrykkene «i live» og «i tide», og det skal lite til å blande med disse. I (32) kan informanten ha blandet ubestemt og bestemt form.

To ord i datamaterialet mitt der mange av de samme forklaringene som over kan ligge til grunn, er <dårlige> og <veldige>. Her er det imidlertid mange faktorer som spiller inn. For det første uttaler man ikke <g>-en i målspråksorda på standard østnorsk. Likevel *kan* det hende at denne <g>-en faktisk ville blitt uttalt i mellomspråket – det vet vi ikke. Men om dette er tilfellet, kan vi forklare innsettinga av <e> med at <g> helst vil være i opptakt framfor koda om mellomspråket følger universelle tendenser. Ellers er det også mulig å forklare disse innsettingene med sammenblanding av entallsending og flertallsending.

Det siste tilfellet av innsetting ordfinalt som jeg finner i datamaterialet, er <økonomie>. Dette ordet er vanskelig å forklare som en reparasjonsstrategi. Her er verken konsonantgrupper som må brytes opp til fordel for somaliske stavelsesstrukturer eller enkeltkonsonanter som må flyttes til fordel for universelle stavelsesprinsipper. Det er også vanskelig å forklare det med sammenblanding av former. En mulig forklaring kunne vært at dette er transfer av skrivemåte fra et annet språk der ordet skrives relativt likt, men har noen forskjeller fra den norske stavemåten, for eksempel fransk <économie>, og blant somaliere i Djibouti er det vanlig å kunne fransk, i tillegg til at informanten kan ha vært i kontakt med dette språket andre steder enn i hjemlandet.

Det er viktig å påpeke at det er vanskelig å konkludere noe om de ordfinale innsettingene fordi de blir satt inn til slutt i ordets lemma. Dermed er det vanskelig å vite om

innsettingene faktisk bare er bøyingsendinger eller avledningssuffiks, eller om de skyldes stavelsesstrukturer. Som vi allerede har diskutert, kan de være et resultat av at en ikke mestrer morfologien i norsk, og dermed bruker feil form i den aktuelle konteksten. Et annet viktig poeng er at det kan være transfer av morfologi fra morsmålet, som Kaldhol (2017) påpeker (jf. kap 5.3.1). I innlånsorda hun undersøker, er det også et knippe ord som får innsetting av /e/ ordfinalt. Hun tolker det til at disse orda tilhører en somalisk deklinasjonsklasse der alle ord slutter på /e/. Dette kunne også vært tilfelle for orda fra mitt datamateriale

7.1.3 Innsetting av konsonantgrafemer

Det er først og fremst enkelttilfeller av konsonantinnsetting i materialet, og få av innsettingene fører til at orda får en stavelsesstruktur som er mer i tråd med somali. Under er de to innsettingene som *kan* tenkes å være reparasjonsstrategier.

(33) N2 <helge> – N1 <elg>

(34) N2 <optimistisk> – N1 <optimistisk>

I (33) ser vi mellomspråksordet <helge> som har fått satt inn en konsonant ordinitialt slik at ordet ikke starter på vokal. Dette passer godt med somaliske orddanningsregler, ettersom det ikke er tillatt med en vokal ordinitialt i somali – en konsonant må fylle den ordinitialt opptakta. Dette ligner på eksempelet med <Cananaas> vi så i kap 5.3 om asiatiske lånord i somali. Her har det malayiske ‘*ayn un-nās* blitt lånt inn i somali, og for å unngå vokal ordinitialt, har ordet fått satt inn konsonanten <C> (som står for lyden /ʃ/, jf. kap. 3.3 om somalisk ortografi).

I (34) har en ekstra <k> blitt satt i en ordmedial koda i målspråksordet <optimistisk>. En *kan* tolke det som at dette er for å unngå <t> i koda. Legger analysen <op.ti.mist.isk> av målordet til grunn, er det en <t> i koda i stavelsen <mist>. Somali tillater ikke <t> i koda, og innsettinga av <k> kan gjøre at heller <k> her fyller koda, mens <t> blir flytta til opptakt i neste stavelse. Dette er likevel lite sannsynlig, for somali tillater jo heller ikke <k> i koda. Men det er ikke umulig at <k> i koda er å foretrekke over <t> i koda i mellomspråket.

7.2 Bortfall av grafemer

7.2.1 Bortfall av vokalgrafemer

Oversikten under viser bortfall av vokalgrafemer i dataene.

Grafem	Ordinitialt	Ordmedialt	Ordfinalt	Total
<e>	0	5	26	31
<i>	2	2	1	5
<y>	0	1	0	1
<a>	0	3	1	4
<i>sum:</i>	2	11	28	41

Tabell 13: bortfall av vokalgrafem

I tabellen at det er en del bortfall av vokal ordfinalt, men også noe ordmedialt og ordinitialt.

Det er til sammen 41 tilfeller av vokalgrafembortfall. Eksempler fra denne kategorien er:

(35) N2 <kvinn> – N1 <kvinne>

(36) N2 <årsken> – N1 <årsaken>

(37) N2 <slam> – N1 <islam>

I (35) ser vi eksempel på vokalgrafembortfall ordfinalt, i (36) ser vi eksempel på bortfall av vokalgrafembortfall ordmedialt, og i (37) ser vi eksempel på bortfall av <i> ordinitialt.

Det er interessant å se at det skjer mange bortfall av vokaler ordfinalt. Likevel er det vanskelig å relatere dette til somaliske stavelsesstrukturer. I mange tilfeller vil bortfall av vokal ordfinalt resultere i at konsonantgrupper oppstår eller at konsonanter blir mer komplekse. Et eksempel er ordet <gansk> med målspråksordet <ganske> der den finale <e>-en faller bort. I målspråksordet kan jo sekvensen <sk> knytte seg til stavelsen den finale <e>-en danner, men i mellomspråksordet må hele sekvensen <nsk> knytte seg til den ordinitiale stavelsen, og vi får en veldig kompleks koda. Kaldhol (2017) finner lignende mønster, og forklarer det med at disse orda blir satt inn i en somalisk substantivklasse, deklinasjon 2, der /e/ ikke kan forekomme ordfinalt. Det kan hende at det Kaldhol finner i sine innlånsdata også kan gjelde for mine mellomspråksdata. Dette kommer jeg riktignok ikke til å diskutere mer i denne oppgava, siden dette ikke er avvik som kan forklares med transfer av stavelsesstrukturer.

Det som *kan* føres tilbake til somaliske *fonotaks*, er de ordmediale bortfalla av vokaler i datamaterialet mitt. Under følger eksempler på dette:

(38) N2 <årsken> – N1 <årsaken>

- (39) N2 <mosken> – N1 <moskeen>
 (40) N2 <eiren> – N1 <eieren>

(38-40) er eksempler på ord som har tre stavelser, og de har alle fått et bortfall av vokalgrafem i stavelse nummer to. Dette er i tråd med at den midterste stavelsen faller bort i somali når den tredje stavelsen er et suffiks. I (38) faller den stavelsesbærende <a>-en bort, i (39) faller den stavelsesbærende <e>-en bort, og i (40) faller den stavelsesbærende <e>-en bort.

Det skal også nevnes at det er en del bortfall ordmedialt i datamaterialet som trolig skyldes forholdet mellom lyd og skrift i norsk. Eksempler er <kjøkknet> ('kjøkkenet') og <vidregående> ('videregående'). I disse orda er <e> spesielt i spontan tale ikke eller så vidt uttalt, og dermed er det fort gjort å glemme dem i skrift. Noen ord er i grenseland til denne kategorien. For eksempel kan vi si at <nøvendig> ('nødvendig') uttales uten <d> av en del personer som snakker standard østnorsk, men mange uttaler den også. Det er altså vanskelig å avgjøre om dette skyldes transfer av stavelsesstrukturer fra somali eller påvirkning fra muntlig norsk.

7.2.2 Bortfall av konsonantgrafemer

Tabellen viser hvor mange ganger et konsonantgrafem faller bort i datamaterialet mitt. Med 'faller bort' mener jeg at et konsonantgrafem som ikke var der i målpråksordet, er det i mellompråksordet. Kategoriene er plassering i stavelse, plassering i ord og hvilket grafem som faller bort. Når det gjelder stavelseskategorien, er den med slik at jeg kan lokalisere om det er konsonantgrafembortfall knyttet til <t> eller <k> i koda, ettersom disse ikke er tillatte i den somaliske stavelsen. Ordkategorien er med slik at jeg skal få en oversikt over antall bortfall knytta til <m> i utlyd. Grafemkategorien er med slik at jeg får en oversikt over hvilke tendenser som finnes i materialet mitt.

	Ordinitialt	Ordmedialt		Ordfinalt		
Grafem	Opptakt	Opptakt	Koda	Koda	Total	Mulig reparasjonsstrategi
<d>	0	1	9	1	11	9
<g>	0	0	2	5	7	4
<k>	0	11	2	1	14	10
<l>	0	2	7	1	10	8
<m>	0	0	2	1	3	2
<n>	0	1	9	0	10	12
<r>	0	3	13	2	18	15
<s>	2	5	16	2	25	30
<t>	0	0	10	26	36	36
<j>	0	9	0	0	9	9
<v>	0	1	1	0	2	2
<h>	2	0	0	0	2	2
<p>	1	0	0	0	1	1
<i>Sum:</i>	<i>5</i>	<i>30</i>	<i>70</i>	<i>37</i>	<i>148</i>	<i>139</i>

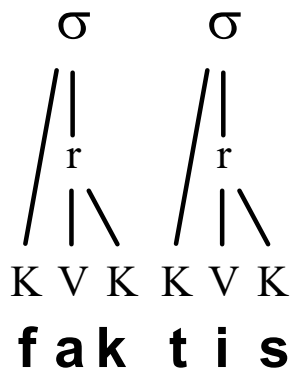
Tabell 14: bortfall av konsonantgrafem

Av tabellen ser vi at det er en stor andel av bortfalla, 139 av 148 som *kan* skyldes transfer av stavelsesstrukturer. Det er imidlertid en stor andel av de 139 bortfalla som også kan ha andre forklaringer. I det videre gjør jeg rede for mulige analyser av orda og diskuterer hvor sannsynlig det er at de ulike typene bortfall er reparasjonsstrategier med mål om å forenkle konsonantgrupper eller unngå visse segmenter i forbudte posisjoner. Under følger noen eksempler fra materialet mitt.

- (41) N2 <nevt> – N1 <nevnt>
- (42) N2 <ektra> – N1 <ek.stra>
- (43) N2 <spørålet> – N1 <spørs.mål.et>
- (44) N2 <kansje> – N1 <kan.skje>
- (45) N2 <biblioteke> – N1 <bi.bli.o.te.ket>
- (46) N2 <so> – N1 <som>

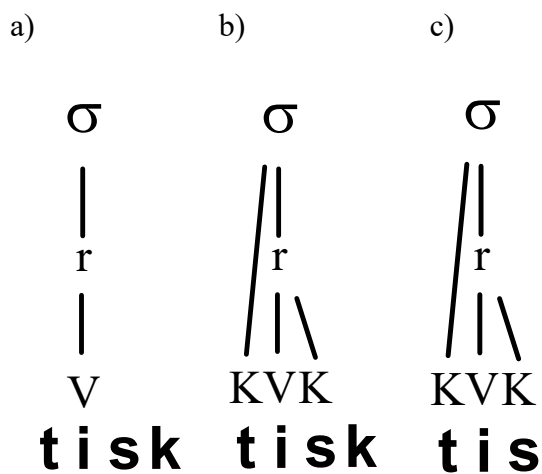
I (41) har konsonanten <n> falt bort i kodaposisjon der det ville være <n> i målspråksordet. I (42) det vært et bortfall av <s> i opptaktsposisjon. I (43) har <s> falt bort i kombinasjonen <rs>, som i norsk ortografi står for segmentet /ʃ/. I (44) har <k> falt bort i opptakt i sekvensen <skj>, som òg står for segmentet /ʃ/ i norsk. I (45) har <t> falt bort ordfinalt, og denne er stum i norsk ortografi. I (46) har grafemet <m> falt bort ordfinalt. I det følgende kommer jeg å diskutere ord som i (41-46) og reflektere over om de ulike typene bortfall er reparasjonsstrategier.

Mange av konsonantbortfalla i materialet mitt ser ut til å ha den hensikt å gjøre en konsonantgruppe mindre kompleks. Under ser vi en analyse av et mellomspråksord dette gjelder for:



Figur 22: mulig analyse av <faktis>

I figur 17 ser vi en analyse av mellomspråksordet <faktis> med målspråksord <faktisk>. Her har det altså skjedd et bortfall av <k> i koda. Basert på analysen kan vi se for oss stavelsesdanningsprosessen under.



Figur 23: stavelsesdanningsprosessen fra <tisk> til <tis>

I figur 23 er altså en mulig reparasjonsprosess fra målspråksstavelsen <tisk> til mellomspråksstavelsen <tis> illustrert. I 23a blir vokalen <i> knyttet til et rim, som igjen blir knyttet til en stavelsesnode. I 23b blir alle tillatte konsonanter knytt til stavelsen, slik at <t> fyller opptakt, og <s> fyller koda. <k> forblir uten stavelsek, og det ser ut som denne stavelsen altså ikke tillater konsonantgrupper i koda. I 23c faller den stavelsesløse <k>-en bort. Det er stor sannsynlighet for at bortfallet av <k> er en reparasjonsstrategi, da det resulterer i at ordet unngår den komplekse kodaen <sk>. Den komplekse kodaen tillates ikke i somali, og ved hjelp av reparasjonsstrategien unngår informanten å bryte med dette

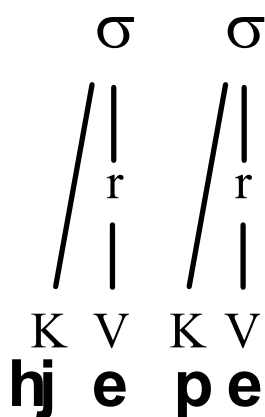
prinsippet. <k> er i tillegg forbudt i koda i somali, og derfor kunne bortfallet være et resultat av denne restriksjonen. Dette er det imidlertid mindre sannsynlig, da <k> forekommer i koda i den første stavelsen, men absolutt ikke umulig – vi har allerede sett flere eksempler på hybridord der enkelte stavelser er reparerte, mens andre ikke er det. Andre ord som har samme type bortfall som <faktis>, er:

(47) N2 <eksamber> – N1 <eksempler>

(48) N2 <hell> – N1 <helt>

I (47) har den ordfinale stavelsen i målspråksordet en kompleks opptakt med de to segmentene <p> og <l>. Dermed er det nærliggende å tolke bortfallet av <l> som en reparasjonsstrategi for å gjøre opptakta mindre kompleks. I tillegg kan vi merke oss at <p> har blitt til i dette ordet, og det skyldes trolig at segmentet /p/ og representasjonen <p> ikke er del av det somaliske segmentinventaret. I (48) har målspråksordet en kompleks koda med to konsonanter, og en kan tenke seg at informanten har unngått konsonantgruppa ved å fjerne et segment. Det er viktig å merke seg at dette *kan* skyldes at informanten rett og slett glemte å sette en strek over den siste <l>-en, og at personen egentlig mente å skrive <t>.

<faktis> er eksempel på et ord der det er tydelig hvilken stavelse det bortfalte konsonantgrafemet hører til. For flesteparten av konsonantbortfalla er ikke stavelsegrensa like tydelig. Figur 18 viser et slikt ord fra datamaterialet mitt.



Figur 24: mulig analyse av <hjeppe>

Her ser vi en analyse av ordet <hjeppe> med målspråksordet <hjeppe>. Det har altså falt bort en <l>. Spørsmålet man da kan stille seg, er hvor denne <l>-en var plassert før den falt bort. Om vi analyserer ordet med norsk fonotaks til grunn, ville det gitt mening å plassere <l> i koda, og ikke i opptakt sammen med <p>, da sekvensen <lp> i opptakt ikke er tillatt i norsk. Samtidig må det nesten være en analyse som <hje.lpe> som ligger til grunn, om det skal gi

mening at bortfallet er en reparasjonsstrategi. Somali tillater jo konsonantgrupper over stavelsesgrenser, og dermed er det ikke nødvendig med en reparasjon i analysen <hjel.pe> der strukturen er <KVK.KV>. Et alternativ kan også være at vedkommende har analysert ordet som <hjelp.e>, og trosset universaliet om maksimal opptakt. I den situasjonen *er* det faktisk en kompleks konsonantgruppe i koda, og da er bortfall av <l> en mulig reparasjonsstrategi. Om det er tilfellet, kan en også tenke seg at mellomspråksanalysen blir <hjep.e>, men det er også fullt mulig at <p> har blitt flytta til opptakt i denne prosessen, slik <hjelp.e> → <hjep.e> → <hje.pe>.

Som nevnt i kapittelet om mellomspråk, og som diskutert i kapittelet i vokalbortfall, er mellomspråk egne språkssystem, og dermed er det også naturlig å tenke at de har sin helt egen mellomspråksfonotaks. Av den grunn må altså verken norsk fonotaks eller somalisk fonotaks ligge til grunn når en innlærer av norsk med somali som morsmål avgjør hvor stavelsesgrensa i et ord går, og derfor kan analysene <hje.lpe> og <hjelp.e> være mulige etter mellomspråksfonotaksen. Samtidig er det universelle prinsipper som ikke taler for disse analysene. <lp> i opptakt strider med sonoritetshierarkiet, og som tidligere nevnt, er schwa, som det er svært sannsynlig at <e>-en ordfinalt står for, en lyd som er 'skjør', og dermed er det usannsynlig at denne danner en helt egen stavelse.

Det er mange ord i datamaterialet mitt som har samme dilemma som ordet <hjepe>, for eksempel:

- (49) N2 <kvadrameter> – N1 <kvadratmeter>
- (50) N2 <kutur> – N1 <kultur>
- (51) N2 <fratida> – N1 <framtida>
- (52) N2 <øsket> – N1 <ønsket>
- (53) N2 <vaskelig> – N1 <vanskelig>
- (54) N2 <adre> – N1 <andre>

I alle disse eksemplene er konsonanten som har falt bort, alene i kodaposisjon. Slik er det i alle fall om en skal analysere målordet etter norsk fonotaks. Om det hadde vært tilfelle, hadde det imidlertid ikke vært noen grunn til å reparere orda (49-54) til å passe inn i den somaliske stavelsesstrukturen, for somali tillater jo konsonantgrupper over stavelsesgrenser. For at bortfalla skal tolkes som reparasjonsstrategier som gjør stavelsesstrukturene likere somaliske stavelsesstrukturer, må en legge andre analyser til grunn enn <kva.drat.me.ter> og <kul.tur> - i stedet må en nesten legge til grunn analysene <kult.ur> eller <ku.ltur>. Men førstnevnte strider med prinsippet om maksimal opptakt, og sistnevnte har en opptakt som strider med

sonoritetshierarkiet. En mulig analyse for ord som i (1-3) kan dermed være at bortfallet *ikke* skyldes transfer av stavellesstrukturer, men at informantens mellomspråk rett og slett ikke takler konsonantgrupper veldig godt, selv ikke når de går over stavellesgrenser.

Mellomspråket har jo sitt eget system, og det kan fint hende at dette er et kjennetegn ved mellomspråket selv om konsonantgrupper er tillatt over stavellesgrenser i somali.

I (52-54) kan mye av den samme analysen som for <ønesker> ligge til grunn. I målspråksordet <øn.sker> er det en kompleks opptakt i den ordfinale stavelsen. En kan tenke seg at et naturlig bortfall vil være at <s> eller <k> faller bort. Det er imidlertid <n> i koda i den ordinitialen som faller bort, enda denne <n>-en ikke er en del av en konsonantgruppe. Legger vi en mellomspråksfonotaks til grunn der <ønsker> blir analysert som <øns.ker>, derimot, kan vi tenke oss at bortfallet av <n> er en reparasjonsstrategi. En annen mulig analyse er at det *er* analysen <øn.sker> som ligger til grunn. Vi kan tenke oss denne prosessen: <øn.sker> → <ø.sker> → <øs.ker>. Her har altså <n> falt bort, og vi står igjen med en ordinitial stavelse uten koda. Dermed har segmentet <s> i den påfølgende opptakta mulighet til å knytte seg til <ø> sin koda. På den måten blir konsonantgruppa <sk> brutt opp.

En annen type bortfall i datamaterialet mitt er bortfall av konsonantgrafemet <s>. I noen tilfeller er dette bortfall av en <s> som står for segmentet /s/, slik som i <ekperimenter> med målspråksord <eksperimenter> og <vigtigte> med målspråksord <viktigste>, men i de fleste tilfellene av denne typen bortfall, er <s> en del av digrafen <rs> som står for /ʃ/. Under følger eksempler på dette:

(55) N2 <forkur> – N1 <forkurs>

(56) N2 <spørålet> – N1 <spørsmålet>

I (55) har <s> falt bort ordfinalt og i koda. I (56) har <s> også falt bort i koda, men ordmedialt. I målspråksorda <forkurs> og <spørsmål> står kombinasjonen <rs> for lyden /ʃ/, og denne retroflekse lyden finnes ikke i det somaliske segmentinventaret. Dermed kan det tenkes at dette er en måte å unngå denne lyden på. En annen tolkning kan være at <r> og <s> står for ulike lyder i disse mellomspråksorda. Om de gjør det, danner disse to lydene til sammen en konsonantgruppe, og i den situasjonen kan man tolke bortfallet av <s> som en reparasjonsstrategi for å gjøre konsonantgruppa mindre kompleks.

Dette skjer ofte med de retroflekse lydene. Under er eksempler:

(57) N2 <urettfadig> – N1 <urettferdig>

(58) N2 <toturert> – N1 <torturert>

I (57) møter vi i målspåksordet digrafen <rd>, som står for den retroflekse lyden /d/. Merk at disse grafememe også *kan* uttales som to separate segmenter /r/ og /d/. Også her kan <r> og <d> representere unike segmenter i mellomspråket, og dermed bli tolka som konsonantgrupper. I (58) har målspåksordet digrafen <rt>, som står for lyden /t/, men også her kan de to tegna representere ulike segmenter, og i den situasjonen kan bortfallet av <r> være en reparasjonsstrategi for å bryte opp konsonantgruppa.

Jeg omtalte eksempler som inkluderer stumme konsonanter i kap 7.1.2 om vokalinnsetting finalt. Denne problematikken finner vi igjen i konsonantgrafembortfallskategorien. Under følger eksempler.

(59) N2 <spennede> – N1 <spennende>

(60) N2 <gjeder> – N1 <gjelder>

I (59) ser vi et bortfall av <n> fra grafemgruppa <nd>. I målspåksordet ville man ikke uttalt <d> på standard østnorsk, og lydlig sett er det dermed ingen grunn til å bruke en reparasjonsstrategi, for på det fonologiske nivået er det ingen konsonantgruppe. Om <d> i mellomspråksordet derimot representerer en lyd, gir det mening å kalle bortfallet en reparasjonsstrategi, for da unngår man en gruppe av konsonanter. Analysen der <d> ikke blir tolket som stum i mellomspråket forsterkes ytterligere av at det er <n> og ikke <d> som faller bort i mellomspråket. Etter norsk fonotaks ville man analysert ordet som <spe.nnen.de>, men i denne situasjonen er det ingen grunn til å reparere ordet, for <n> og <d> tilhører da forskjellige stavelser, og dermed er det ingen konsonantgruppe som trenger å reduseres. Dermed kan analysene <spe.nnend.e> eller <spe.nne.nde> ligge til grunn her. (60) har samme tilfelle som (59) der en konsonant faller bort i en gruppe av konsonantgrafemer, og det grafemet som står igjen er en stum <d>. Også her hører <l> og <d> til ulike stavelser, og en må legge en alternativ analyse til grunn for at innsettinga skal ansees som en reparasjonsstrategi.

Det er nesten ingen tilfeller av konsonantbortfall av <k> eller <m>. /k/ er jo ikke tillatt i kodaposisjon i somali, men ser ut til å være tillatte i koda i mellomspråket. Det eneste tilfellet som kan tolkes som bortfall av /k/ i koda, er <faktis>, men som foreslått tidligere i kapitlet, er dette mer sannsynlig en måte å unngå en kompleks koda på. I somali tillates ikke /m/ og representasjonen <m> ordfinalt, men det finnes bare ett tilfelle der <m> er fjernet, og dette er <so> med målspåksord <som>, og etter somaliske fonologiske regler er

det også heller å forvente at <m> blir erstattet med <n> i stedet for å falle bort. Det er derimot ganske mange bortfall av <t> i koda, og /t/ i koda er jo ikke tillatt i somali. Derfor kan dette bli tolka som reparasjonsstrategi. Samtidig er de fleste bortfalla av <t> ordfinale, og de fleste bortfalla skjer der man ikke ville uttalt /t/ på standard østnorsk. En annen analyse er altså at den muntlige norske formen blir tatt med inn i skriftbildet.

Det er også noen ord som endrer endelsen <sk> til <k> eller <kk>. Disse er <pessimistikk>, <fantastik>, <klassikk> og <dramatikke>. En kunne tolket dette som en måte å unngå komplekse stavelsesstrukturer på, men dette er ord som er lånt inn i norsk fra engelsk, og det er mulig at dette er transfer av engelsk ortografi, heller enn somaliske stavelsesstrukturer.

I tillegg er det en god del bortfall av <s> <k> og <j> ordfinalt. Dette kan nok knyttes til segmentet /ʃ/, og de mange mulige representasjonene dette segmentet kan ha i norsk ortografi, og dette er jo også vanlige skrivefeil blant nordmenn nettopp fordi det er så mange grafem som kan presentere lyden. Men en *kan* tenke seg at dette er en reparasjonsstrategi. Om informantene ikke har fått med seg at for eksempel <skj> eller <sk> er polygrafem, kan de tolke det som at <s> <k> og <j> er monografem og presenterer hver sin lyd. Om tolkninga er slik, vil det jo oppstå en konsonantgruppe i mellomspråket, og det kan være grunn til å bruke en reparasjonsstrategi, slik at opptakta blir mindre kompleks.

7.2.3 Bortfall og innsetting av flere segmenter

Vi finner noe eksempler på at stavelser med konsonanter i opptakt eller koda, faller bort.

Under er eksempel på noen av orda som har bortfall av stavelse.

- | | | |
|------|----------------|--------------------|
| (61) | N2 <alkol> | – N1 <alkohol> |
| (62) | N2 <giff> | – N1 <gifte> |
| (63) | N2 <diskurer> | – N1 <diskuterer> |
| (64) | N2 <koren> | – N1 <koranen> |
| (65) | N2 <konkurrer> | – N1 <konkurrerer> |

Vi finner ett eksempel på innsetting av stavelser der det er konsonant i opptakt eller koda. Dette er <sa> i ordet <tradisasjoner>, målspråksform <tradisjoner>. Alle disse er ORT som vanskelig kan forklares som en reparasjonsstrategi – mellomspråksordet bryter ikke med somaliske stavelsesstrukturer, men det gjør ikke målspråksordet heller. Unntaket er (64) som har fått bortfall av vokal og kodakonsonant i et trisyllabisk ord med suffiks i tredje stavelse. Jeg kan ikke forklare hvorfor konsonanten forsvinner, men det er mulig at vokalen har falt

bort som resultat av regelen om bortfall.

7.3 Ombyttning og flytting av grafemer

Det finnes 23 tilfeller av ombyttning mellom to grafemer i materialet mitt. Det er nesten ingen av disse som fører til at strukturen i det aktuelle ordet blir likere somaliske strukturer, og det er noen av ombyttningene som resulterer i *mer* komplekse stavelsesstrukturer enn det er i målspråksordet. Under ser vi eksempler på det.

(66) N2 <flyte> – N1 <fyl.te>

(67) N2 <triveds> – N1 <triv.des>

I (66) var det altså ingen kompleks opptakt eller koda i målspråksordet, men når <l> og <y> bytter plass, får mellomspråksordet en kompleks ordinitial opptakt. I (67) er det heller ingen konsonantgrupper, men når <d> og <e> bytter plass, blir det en kompleks koda, og konsonantgruppa <ds> strider med sonoritetshierarkiet. Det er imidlertid en ombyttning som fører til at en konsonantgruppe blir brutt opp, og denne er presentert under

(68) N2 <torisim> – N1 <turisme>

Her kan vi tolke det som at målspråksordet *turisme* har byttet <m> med <e>, og at <e> deretter har endret vokalkvalitet til <i>. En annen analyse er at dette er transfer fra engelsk <tourism>, og at det har skjedd en *innsetting* i stedet for ombyttning. Uansett hvilken av de to analysene vi legger til grunn, har det uansett ført til at konsonantgruppa <sm> har blitt brutt opp. Dermed er dette en reparasjonsstrategi, enten det er metatese eller epentese.

Dette ordet minner oss på at vi må huske på hvor kompleks språkbakgrunnen til en andrespråksinnlærer kan være. Jeg har tidligere snakket om at innlærere benytter seg av en slags ressursbank når de lærer et nytt språk, og at denne inneholder all språkkunnskap man har. Dette inkluderer erfaringer fra morsmål og målspråk, men det er et viktig poeng at alle andre språk en har vært innom i løpet av livet, også blir en del av denne ressursbanken, uansett om dette er et språk man ikke kan veldig godt. I innledningen til denne oppgava gjorde jeg rede for mulige språk informantene mine kan ha vokst opp med. Men informantene mine har i tillegg en historie bak seg der de har flytta på seg, og har kanskje vært innom mange ulike språksamfunn før de kom til Norge. Det er derfor viktig å huske på at ikke alt kan forklares av morsmål, og i en mer omfattende analyse enn denne, kunne en tatt utgangspunkt i flere enn ett språk i en undersøkelse av transfer.

Til slutt tar jeg for meg flytting. Det finnes ett to eksempel på flytting i datamaterialet mitt. Dette er flytting av <r> i <forbrede> med målspråksordet <forbedre>. Her oppstår konsonantgruppa
 som følge av flyttinga, og denne flyttinga fungerer ikke som reparasjonsstrategi.

7.4 Erstatning av grafemer

Som vi så i kap 2.4 om reparasjonsstrategier er erstatning en strategi man kan bruke når et segment ikke er tillatt i et språk, enten det er i visse posisjoner eller generelt ikke er en del av segmentinventaret. Da blir det aktuelle segmentet erstatta av et annet segment som faktisk finnes i språket, og de to segmentene har så mange lydlige trekk til felles som mulig

I materialet mitt ser vi får tendenser i denne kategorien. Likevel er enkelte erstatninger som går igjen blant flere informanter. En kan merke seg at det er mye erstatning av <p> med , men også tendenser til erstatning av med <p>. Under ser vi eksempler på disse ombyttingene.

(69) N2 <benger> – N1 <penger>

(70) N2 <pud> – N1 <bud>

(71) N2 <viske> – N1 <fiske>

I (69) har <p> blitt erstatta med en , i (70) er erstatta med <p> og i (71) er <f> erstatta med <v>. Tendensene vi ser til veksling mellom <p> og , er interessante, men de kan ikke skyldes transfer av stavelsesstrukturer. I stedet er det sannsynlig at de skyldes mangelen på segmentet /p/, og også grafemet <p>, i somali. Dermed bytter de ut <p> med det segmentet i segmentinventaret sitt med flest like trekk som <p>, og dette er . Når det omvendte skjer, og blir til <p>, skyldes dette trolig hyperkorleksjon, og dette er nøyaktig samme ORT som vi så eksempel på i Broselow (1976) – også der manglet informantene segmentet /p/ i lydinventaret sitt, og også der ble både <p> erstatta med , og erstatta med <p>. Vi finner også to eksempler på at <f> blir byttet ut med <v> (for eksempel at målspråksordet <fiske> blir til <viske>), og dette kan forklares med hyperkorleksjon, da /v/ eller <v> ikke er en del av segmentinventaret i somali, og de dermed kan ha blitt litt ekstra oppmerksomme på denne nye lyden, og overgeneraliserer.

Erstatninga <p> → og → <p> er på segmentnivå, og ikke stavelsesnivå. Derfor kan det heller ikke sies å skyldes transfer av stavelsesstrukturer. I det videre skal jeg undersøke erstatninger som *kan* skyldes overføring av stavelsesstrukturer. Da er det først og fremst erstatning av <k>, <t> og <m> som er interessant, ettersom de to førstnevnte ikke er

tillatte i koda, og den sistnevnte ikke er tillatt ordfinalt. Vi finner fire eksempler i datamaterialet mitt der <k> har blitt byttet om til henholdsvis <q>, <g>, <r> og <t>, og det er kun de tre sistnevnte erstatningene som skjer i koda. De ser vi under:

- | | | |
|------|----------------|------------------|
| (72) | N2 <viktigste> | – N1 <viktigste> |
| (73) | N2 <språr> | – N1 <språk> |
| (74) | N2 <beherst> | – N1 <behersket> |

I 72 har altså en <k> blitt til en <g> i koda, og dette kan fint regnes som en reparasjonsstrategi nettopp fordi den blir erstatta med en <g>, som er den stemte motparten. Dette er helt i tråd med somaliske stavelsesdanningsregler som ikke tillater <t> i koda og dermed erstatter denne med sin stemte motpart om den havner i kodaposisjon. Dette har vi også sett flere eksempler på i lånordsstudiene til Kaldhol (2017) og Cardona (1988), mens denne prosessen nesten er totalfraværende i mellomspråket til informantene jeg studerer, men unntak av dette ene eksempelet. I (73) *har* <r> erstatta <k> i koda, og etter denne erstatninga blir stavelsen i tråd med somalisk fonotaks, som ikke tillater <k> i koda. I (73) kan vi analysere ordet som at det har skjedd bortfall av <e> i tillegg til erstatning av <k> til <t>. Når <e> faller bort i dette ordet, havner <sk> i koda i stedet for opptakt, og ettersom <k> ikke er tillatt i denne posisjonen, er det logisk at det skjer en erstatning. Segmentene som erstatter <k> er likevel ikke å forvente. Ingen av disse har mange fellestrekk med /k/, og etter somaliske stavelsesprinsipper ville det gitt mer mening om grafemet <g>, den stemte motparten til /k/, ble satt inn.

Når det gjelder <t>, finnes det få eksempler på dette grafemet blir erstatta med sin stemte motpart, men ett eksempel finner vi: ordet <sammenligned> med målspråksord <sammenlignet>. Her unngår altså språkbrukeren å få <t> i kodaposisjon ved å bytte ut lyden med <d>, som har flest felles trekk med <t>. Det skjer også kun én erstatning av <m>, og denne er ikke ordfinal, og kan dermed ikke forklares med orddanningsprinsipper i somali.

Erstatning av segmenter som er forbudt i en stavelsesposisjon, skjer altså i svært liten grad, og i de tilfellene det skjer, er det vanskelig å se et mønster i datasettet. Det er stor sannsynlighet for at det skyldes tilfeldigheter eller andre faktorer enn transfer av stavelser.

7.5 Fonologisk eller ortografisk transfer?

I forrige kapittel så vi at bortfall og innsetting ofte skjer i posisjoner der en ikke skulle forvente at det er nødvendig med en reparasjon. Dette gjelder for eksempel i ordet <kultur> med målspråksordet <kultur>. Målspråksordet vil etter norsk fonotaks og prinsippet om

maksimal opptakt bli analysert som <kul.tur>. Ingen posisjoner i de to stavelser har her komplekse konsonantgrupper, og derfor er det ut ifra denne analysen ikke noe behov for en reparasjon, for stavelsen er jo i tråd med somaliske stavelsesprinsipper.

En mulig forklaring på dette har jeg allerede vært inne på i analysen. Mellomspråk regnes, som vi så i kap 4 som selvstendige språkssystem som bør studeres på lik linje som morsmål eller målspråk. Det er system som styres av egne regler, og derfor bør man være forsiktig med å gå i sammenligningsfella der målspråket er målestokken. Derfor har jeg foreslått alternative analyser av ord som dette der en egen mellomspråksfonotaks er gjeldende. Legger vi en annen fonotaks til grunn kan målspråksordet <hjelper> bli analysert som <hje.lper> eller <hjelp.er>, og i begge disse tilfellene er det en kompleks konsonantgruppe i koda. Med en slik analyse til grunn gir det altså mening å reparere stavelsen ved å bruke bortfall eller epentese som reparasjons-strategier. Denne analysen kan også støttes av at det ikke finnes konsonantgrupper i somali. Dermed har ikke innlærere med somali som morsmål den samme intuisjonen som en nordmann om hvor en stavelsegrense går. Siden det er vanlig med konsonantgrupper i norsk, utvikler en også en intuisjon om hvor de ulike segmentene i en konsonantgruppe hører til. Vi kan tenke oss at noen med somali som morsmål ikke utvikler denne intuisjonen fordi de aldri trenger å fordele flere enn to segmenter, og da er det naturlig at det første segmentet hører til koda, og det andre segmentet hører til opptakt. Dette legger et grunnlag for at fonotaksen i mellomspråket til informantene mine kanskje ikke er like streng som i norsk, og at det er litt mer tilfeldig hvordan en konsonantgruppe fordeler seg mellom koda og opptakt.

Det er imidlertid en annen måte å analysere disse uforutsette reparasjonsstrategiene på, og det er om vi tar i betraktning det *skriftbildet* som finnes i somali. Som jeg presenterte i kap 3.3 er somali et ganske ortofont språk der hver lyd blir representert av ett grafem. Det er i tillegg svært få digrafer i språket – de to som finnes er <kh> og <dh>. Informantene mine er altså kjent med konseptet digraf fra morsmålet sitt, men man møter ikke på det så ofte i somali som man gjør i norsk. Det er altså ikke så vanlig med lange sekvenser med konsonantgrafemer i somali. Siden det er lov med strukturen KVK.KVK i somali, kan det danne seg konsonantgrupper på tvers av stavelsegrenser, og dermed er det ikke så uvanlig å se to konsonantgrafemer på rad i somali. Sekvenser som er lenger enn dette er sjeldnere å se, men det finnes altså sekvenser der et ord har grafemrekka <KVKK.KV>, og her står altså de to K-ene i den ordinitiale kodaen for digrafen <kh> eller <dh>, og ikke for to ulike segmenter.

Et annet poeng er at digrafene <kh> og <dh> ikke finnes i norsk, og ingen av

digrafene som finnes i norsk, f.eks. <rs>, er å finne i somali. Polygrafemer som har flere enn to tegn, f.eks. <skj> finner man heller ikke i somali. Man finner derimot som oftest de enkelte tegna som polygrafemene består av. For eksempel har somali både grafema <n>, som står for /n/ og <g> som står for /g/, men de har ikke polygrafemet <ng> som står for /ŋ/. Dermed er det nærliggende å tenke at de vil tolke disse som to selvstendige grafem som representerer to ulike segmenter. Som vi har sett kan dette resultere i at informantene tolker to tegn som en konsonantgruppe i samme stavelsesposisjon, og det kan igjen resultere i at informantene prøver å reparere stavelsen.

Uansett om informantene mine tolker disse som ulike segmenter eller ikke, kan en tenke seg at informantene mine er vare på grupper av konsonantgrafemer, og at de lange sekvensene som finnes i norsk strider med det skriftbildet de kjenner fra somali. Vi kan for eksempel ta ordet <vaskelig> fra datamaterialet mitt. Her er målspråksordet <vanskelig> med følgende grafemsekvens: <KVKKKVVK>. I ordet finner vi altså en grafemsekvens på tre konsonanter. Denne sekvensen er relativt sjelden i somali. Om vi legger til grunn en analyse der avviket grunner i ortografisk transfer fra somali til mellomspråket, kan vi tenke oss at en innlærer vil prøve å bryte opp sekvensen <nsk>. Dette kan være uten tanke på stavelser – det som er viktig, er at det er færre konsonantgrafemer på rad, og det spiller liten rolle *hvor* i sekvensen det skjer en reparasjon, bare det blir færre konsonantgrafem mellom hvert vokalgrafem.

Basert på det jeg har lagt fram av teori i del 1 av denne avhandlingen, mener jeg at det er hold i begge analysene. Det er absolutt mulig at informantene mine har et mellomspråk der stavelsgrensene er mer flytende enn i norsk, og at det ligger en annen fonotaks til grunn enn den norske. Samtidig er det en god mulighet for at det er skriftbildet informantene kjenner fra somali, som spiller inn her, og at lange sekvenser av konsonantgrafemer må brytes opp i mellomspråket. Dette kan også være uavhengig av stavelsesstrukturer. Begge analysene legger uansett til grunn at det har skjedd en reparasjon, enten det gjelder rekker av konsonantgrafemer som er fundert i skriftbildet, eller om det er fundert i et mer lydlig nivå der konsonantgrafemer som *representerer* en sekvens av konsonanter er i fokus.

7.6 Oppsummering

I dette kapitlet har vi sett at det kan mange mulige grunner bak et avvik i datamaterialet mitt. Avvika kan for eksempel skyldes ortografisk transfer, morfologisk transfer, analogi eller mer tilfeldige skrivefeil. Likevel er det en del ord i materialet mitt som høyst sannsynlig skyldes transfer av stavelsesstrukturer. Dette gjelder ord som ser ut til å få brutt opp

konsonantgrupper gjennom reparasjonsstrategiene epentese og bortfall, slik som <ektra>, som får bortfall av <s> og <kekestara>, som får innsetting av <e> og <a>. I materialet mitt er det 13 bortfall jeg mener høyst sannsynlig skyldes at man prøver å unngå konsonantgrupper, og 20 innsetninger som høyst sannsynlig er epentetiske reparasjonsstrategier med samme hensikt. Dette er fordi jeg vanskelig kan finne en annen forklaring for hvorfor disse tendensene forekommer i materialet mitt. Det er imidlertid mange av de andre avvika som også *kan* være reparasjonsstrategier, men for de andre avvika finnes jeg også andre forklaringer. Jeg har i dette kapittelet talt for at informantene mine har en egen mellomspråksfonotaks, og at de ikke nødvendigvis må sette stavelsesgrensene likt som i norsk. Med en annen fonotaks til grunn, kan mange avvik, spesielt bortfall, kunne forklares som reparasjonsstrategi.

Som vi har sett var det 238 ord i et datamateriale på 48 912 ord som *kunne* forklares med transfer av stavelsesstrukturer fra somali til norsk. Informantene legger med andre ord som oftest norske stavelsesstrukturer til grunn når de danner stavelser i ord. Men både informanter som tok språkprøven og informanter som tok høyere test i norsk, har avvik som går på stavelsesstrukturer, og dette er altså ikke avvik som kun oppstår i tidlige innlærerfaser. Det er noen tendenser i materialet mitt som gjør at disse avvika ikke virker tilfeldige. Det er eksempler på at ulike informanter har avvik i samme målpråksord, og at avviket er helt likt. For eksempel er det tre ulike informanter som skriver mellomspråksordet <ønesker>. Videre blir samme målpråksord reparert på andre måter, som i <øsker>, der bortfall har blitt brukt, heller enn epentese. Dette viser hvor unik hver enkelt språkbruker er, og at en ikke kan forutse hvordan et mellomspråksord blir sendt ut. Mellomspråk er, som diskutert, dynamiske språkssystem, og det kan variere fra person til person hvordan det blir sendt ut, selv om man har samme morsmål og samme målpråk. Det kan også variere for den enkelte personen om det blir benyttet en reparasjonsstrategi. Det kan bli gjort i én setning, men ikke i den neste. Og det kan bli gjort i én stavelse, men ikke i den neste, slik vi for eksempel har sett i <blomester>.

9.0 Konklusjon

Vi er ved veis ende. Resultata mine er lagt fram og diskuterte, og jeg har dermed grunnlag til å svare på den overordnede problemstillinga mi. Basert på det jeg har funnet ut i studien min, og det jeg har funnet ut om tidligere forskning gjort på feltet, har jeg også noen forslag til hva en kan forske videre på innen ortografi i norsk som andrespråk.

9.1 Svar på problemstilling

I denne studien har jeg undersøkt følgende problemstilling:

I hvilken grad ser vi spor av somaliske stavelsesstrukturer i ortografien til innlærere av norsk med somali som morsmål?

For å kunne svare på dette spørsmålet, har jeg undersøkt et utvalg tekster fra Norsk andrespråkskorpus og stilt to forskningsspørsmål: a) Er det noen ortografiske avvik knytta til stavelsesstrukturer som går igjen blant informantene mine? b) Har disse avvika noen fellestrekk med stavelsesstrukturen i somali?

Når det gjelder hvilke ortografiske avvik som går igjen blant informantene mine, er dette i hovedsak at vokalgrafem blir satt inn, at vokalgrafem faller bort og at konsonantgrafem faller bort. Vokalgrafemer blir stort sett satt inn i grupper av konsonantgrafemer ordmedialt, mens det ligger i sakens natur at ordfinale vokaler blir satt inn etter én eller flere konsonanter. Vokalgrafema som blir fjerna medialt, fjernes gjerne i andre stavelse, men ordfinale vokalgrafembortfall er den hyppigste typen vokalbortfall. Bortfall av konsonantgrafem er en type avvik som går igjen, og dette skjer ofte når konsonantgrafemet står ved siden av et annet konsonantgrafem.

Av avvika som går igjen, kan følgende typer avvik føre til at et mellomspråksord passer inn i den somaliske strukturen:

- Vokalinnssetting ordmedialt → Eksempel: <ekestara> ('ekstra')
- Vokalinnssetting ordfinal → Eksempel: <bedrifte> ('bedrift')
- Vokalbortfall ordmedialt → Eksempel: <eiren> ('eieren')
- Konsonantbortfall ordmedialt → Eksempel: <ektra> ('ekstra')
- Konsonantbortfall ordfinalt → Eksempel: <faktis> ('faktis')

Noen av innsettingene eller bortfalla er helt klart reparasjonsstrategier for å redusere konsonantgrupper. De bryter opp konsonantgrupper som tilhører samme stavelse og samme posisjon i stavelsen. Dette er imidlertid når en analyserer stavelsen med norsk fonotaks til

grunn. I mange tilfeller av bortfall eller innsetting kan ikke disse forklares som reparasjonsstrategier om en legger norsk fonotaks til grunn. Dette er fordi de ikke blir satt inn i grupper av konsonantgrafemer som tilhører samme stavelse, men mellom grafemer som tilhører *ulike* stavelser. Dermed må stavelsgrensene være annerledes i mellomspråket enn de er i norsk om vi skal tolke disse orda som reparasjonsstrategier, for somali tillater konsonantgrupper på tvers av stavelsgrenser.

En annen måte å forklare disse bortfalla på, er at ortografisk transfer ligger til grunn for avvika, heller enn fonologisk transfer. Her kan ordbildet i somali være avgjørende for innsettingene og bortfalla som skjer. Informantene er vant til å se ordbildet <KVK.KVK> i språket sitt, der monografer regjerer – det finnes kun to digrafer. Disse to digrafene finnes ikke i norsk. Norsk er et språk med svært mange polygrafem, som <skj> og <rs>, og det kan tenkes at somaliere tolker de ulike tegna i polygrafemene som egne segmenter som dermed danner større grupper av konsonantgrafemer. Det kan også tenkes at dette ikke er feil som knyttes til et fonologisk nivå, men at det rett og slett bryter med skriftbildet <KVK.KVK>, og at konsonanter som gjør dette skriftbildet mer komplisert, kan bli forsøkt fjerna.

Basert på det jeg har svart på de to forskningsspørsmåla, *er* det stor sannsynlighet for at de avvika som går igjen kan forklares med transfer av somaliske stavelsesstrukturer. Et alternativ er at (noen av) avvika kan forklares med transfer av somalisk *ortografi*. Det er ganske liten prosentandel av de 48 912 orda som kan skyldes transfer av stavelsesstrukturer eller ortografi, men en liten andel ser ut til å kunne det.

9.2 Videre forskning

I denne studien har jeg konsentrert meg om stavelsesstrukturer og hvordan de kan påvirke ortografien til innlærere av norsk med somali som morsmål. Det er mange andre perspektiver en kunne hatt for å finne ut mer om ortografien til denne innlærergruppa. For det første ville en naturlig fortsettelse av min studie være å gjøre en systematisk lik undersøkelse av en gruppe med norsk som andrespråk som har et *annet* morsmål enn somali. Disse kunne hatt et morsmål der stavelsesstrukturene var likere de norske, og på den måten kunne en sammenligna de to gruppene og sett om dette er noe som er spesifikt for informanter med språk der komplekse stavelsesstrukturer ikke er tillatt.

Jeg har også nevnt flere avvik som går igjen, men som jeg ikke har tatt for meg i studien min. Disse er:

- Avvik i vokalgrafemer
 - Tendenser til vokalharmoni

- Avvik i dobbelt- og enkeltkonsonanter
- Avvik i på segmentnivået

I videre studier kunne en altså undersøkt tre punkter: En kunne undersøkt om vokalgrafemavvik kan skyldes transfer av vokalinventaret i somali, eller eventuelt transfer ortografien i somali. En kunne også undersøkt om avvik i dobbelt- og enkeltkonsonanter kan komme av skriftsystemtransfer, slik at avvika skyldes at innlærerne følger de somaliske reglene for markering av lang og kort lyd. Som nevnt i diskusjonsdelen, finner jeg også tendenser til vokalharmoni der det er epentese av vokal. Det hadde vært interessant å undersøke om dette er transfer av somalisk vokalharmoni i mellomspråket. Vi har også sett at det er avvik på segmentnivå som trolig skyldes at /p/, representert som <p> og /v/ representert som <v> ikke er en del av det somaliske konsonantinventaret, og dermed er det også avvik på segmentnivå det kan være interessant å studere nøyere.

Som det kommer fram i framstillinga mi av tidligere forskning på ortografi i norsk som andrespråk, er det skrint med studier. Etter det jeg kjenner til, er dette den tredje studien gjort på feltet, og inkludert min studie har det blitt gjort undersøkelser av innlærere med tyrkisk, vietnamesisk, bosnisk/serbisk/kroatisk og somali. Studiene som har blitt gjort her i Norge, og studier av mellomspråk i andre land, tyder på at morsmålet kan ha mye å si for de ortografiske avvik som en finner. Derfor er det interessant å undersøke innlærere med så mange morsmålsbakgrunner som mulig. Ved å gjøre det kan vi kartlegge hvilke feil som går igjen, få mer kunnskap om typiske avvik for de ulike innlærergruppene, og hva som er årsaken bak dem. Da blir dette lettere for norsklæreren å sette seg inn i ulike innlærereres tankesett, og rette opp avvika på en effektiv måte.

Av studiene som finnes, er det én studie som ser på ortografi hos barn, og to som ser på ortografi hos voksne. Mer forskning på forskjeller mellom aldersgrupper kan være fruktbart. Da kan vi få mer kunnskap om det er ulike ortografiske avvik som går igjen hos innlærere i ulike aldre, og om noen aldersgrupper lærer raskere enn andre aldersgrupper. Internasjonal forskning tyder på at yngre innlærere lærer talespråk fortere, men det kan være interessant å undersøke om dette også gjelder for ortografien, og, ikke minst, om dette gjelder for norsk ortografi.

Som nevnt i kap. 4 om mellomspråk, er det viktig å huske på at hver enkelt andrespråksinnlærer har sin unike grammatikk. Det er alltid interessant å se tendenser til avvik som går igjen blant innlærere med samme morsmålsbakgrunn og samme målspråk, og disse kan hjelpe oss med å hjelpe innlærergrupper. Samtidig vil jeg slå et slag for mer

kvalitative studier av enkeltinnlæreres ortografi. Disse kan minne oss på hvor forskjellige mellomspråksortografier kan være, og at det tross alt er individer vi møter i klasserommet.

Litteratur

- Abrahamsen, J. E. 2017. Når verktya endrar seg: hundre år etter segmentet. *Norsklæraren* 1/2017, s. 213-219
- Andrzejewski, B. W. 1974. The introduction of a national orthography for Somali. *African Language Studies*. 15. 199–203
- 1978. The development of a national orthography in Somalia and the modernization of the Somali language. *Horn of Africa*. 1. 39–46.
- Apel, K. 2011. What Is Orthographic Knowledge? *Language, Speech, and Hearing Services in Schools*. Vol. 42.
- Appleyard, D., Orwin, M. 2008. The Horn of Africa: Ethiopia, Eritrea, Djibouti, and Somalia, I *Language and national identity in Africa*. Red. Simpson, A. Oxford: Oxford University Press.
- Archibald, J. 1998. Second language phonology, phonetics, and typology. *Studies in second language acquisition*, 20, s. 189 – 211.
- ASK – brukerveiledning: <http://korpus.uib.no/humfak/ask/brukerveiledning/brukerveiledning.htm>
- ASK – kodebok: <http://korpus.uib.no/humfak/ask/brukerveiledning/feilkoder.htm>
- Banti, G. 2013. Strata on loanwords from Arabic and other Semitic languages in Northern Somali. . I *African Arabic: Approaches to Dialectology*. Red. M. Lafkioui. Berlin/Boston: De Gruyter Mouton. Chap. 6, pp. 185–210.
- Berggreen, H., Tenfjord, K. 1999. *Andrespråklæring*. Oslo: Ad Notam Gyldendal.
- Broselow, E. 1992. Nonobvious transfer: On predicting epenthesis errors. I S. Gass & L. Selinker (red.), *Language transfer in language learning* (s. 71–86). Amsterdam: Benjamins
- Bugge, E. 2019. *Skriftspråkleg mangfald som ressurs i opplæringa. Ei kort innføring*. Oslo: Universitetsforlaget
- Cardona, G. R. 1988. Somalia and the Indian Ocean: Cultural and Linguistic contacts. I *Proceedings of the Third International Congress of Somali Studies*. Red. A. Puglielli, s. 22-35.
- Clements & Keyser. 1983. *CV Phonology*. MIT press, Cambridge, Mass.
- Cook, V. J. 1997. L2 users and English spelling. *Journal of Multilingual and Multicultural Development*, 18, 474–488.

- Cook, V. J., Bassetti, B. 2005. An Introduction to Researching Second Language Writing Systems. I Vivian Cook & Benedetta Bassetti (red.) *Second Language Writing Systems*. Clevedon: Multilingual matters
- Corder, P. 1967. The significance of learners' errors. *International review of Applied Linguistics*, 5(2-3), 161-167.
- Eilertsen, S. 1990. *Konsonant-grupper – en barriere. En undersøkelse av uttalen av konsonantgrupper hos vietnamesere som lærer norsk*. (Hovedfagsoppgave). UiB, Bergen.
- Figueredo, L. 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. *Reading and Writing*. Vol. 19. Nr. 8.
- Golden, A., Hvistendahl, R. 2010. Skriveforskning i et andrespråksperspektiv. *NOA – Norsk som andrespråk*, vol 26, nr. 2, 36-66.
- 2013. Inndeling av et landskap: Hvordan kategorisere studier av skriving på andrespråket? *NOA – Norsk som andrespråk*, vol. 29, nr. 1.
- 2015. Forskning på andrespråksskriving i Skandinavia, med vekt på de norske studiene. I Anne Golden & Elisabeth Selj (red.), *Skriving på norsk som andrespråk. Vurdering, opplæring og elevenes stemmer* (1. utg., s. 53-68). Oslo: Cappelen Damm.
- Golden, A., Tenfjord, K. 2015. Tverrspråklig innflytelse i skriftlige innlærertekster. I Anne Golden & Elisabeth Selj (red.), *Skriving på norsk som andrespråk. Vurdering, opplæring og elevenes stemmer* (1. utg., s. 53-68). Oslo: Cappelen Damm.
- Gujord, A. K. H. 2013. *Grammatical encoding of past time in L2 Norwegian. The roles of L1 influence and verb semantics*. (Doktorgradsavhandling). UiB, Bergen.
- Gujord, A. K. H., Ragnhildstveit, S. 2018. Tverrspråkleg påverknad – status og trendar i norsk andrespråksforskning. I Ann-Kristin H. Gujord og Gunhild Tveit Randen (red.) *Norsk som andrespråk – perspektiver på læring og utvikling*. Oslo: Cappelen Damm
- Hansen, J. G. 2001. Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics*, 22, 338–365.
- Husby, O., Kløve, M. H. 1998. *Andrespråksfonologi. Teori og metodikk*. (1. utg.). Oslo: Gyldendal Norsk Forlag.
- Husby, O. (u.d.) *Norsk som andrespråk – Uttaleundervisning*, upublisert manuskript, Institutt for språk og kommunikasjon, NTNU

- Hvenekilde, A. 1985. Ordformer i norske stiler skrevet av tyrkiske og vietnamesiske elever. *NOA – Norsk som andrespråk*, nr. 1, 17-76. DOI?
- Hyman, Larry. 1985. *A Theory on Phonological Weight*.
- Ibrahim, M. H. 1978. Patterns in Spelling Errors. *English Language Teaching Journal*, 32, 207–212
- Jarvis, S. 2000. Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language learning* 50.
- Jarvis, S., Pavlenko, A. 2008. *Crosslinguistic Influence in Language and Cognition*. New York/London: Routledge.
- Kaldhol, N. H. 2017. *The Oslo Dialect of Somali. Tonal adaptations of Norwegian loanwords* (Masteroppgave). UiO, Oslo.
- Kløve, M. H., Young-Scholten, M. 2001. Repair of L2-syllables through metathesis. *International Review of Applied Linguistics in Language Teaching* Vol 39: Nr. 2
- Kristoffersen, G. 2000. *The phonology of Norwegian*. Oxford: OUP Oxford.
- Kulbrandstad, L. A., Ryen, E. 2018. Hvordan læres et andrespråk? Overordnede perspektiver. I Ann-Kristin H. Gujord og Gunhild Tveit Randen (red.) *Norsk som andrespråk – perspektiver på læring og utvikling*. Oslo: Cappelen Damm
- Lewis, M. P., Simons G. F., Fennig C. D. 2017. *Ethnologue: Languages of the World*, 20th ed. Dallas, Texas: SIL International.
- Nistov, I., Nordanger, M. 2018. Kva er mellomspråk? Gjensyn med eit sentralt omgrep i andrespråkforskninga. I Ann-Kristin H. Gujord og Gunhild Tveit Randen (red.) *Norsk som andrespråk – perspektiver på læring og utvikling*. Oslo: Cappelen Damm
- Odlin, T. 1989. *Language Transfer. Cross-linguistic influence in language learning*. Cambridge: Cambridge university press
- Rosén, Victoria. 2001. *Vietnamesisk. En kontrastiv og typologisk introduksjon*. Trondheim: Tapir akademisk forlag
- Saeed, J. 1999. *Somali*. Amsterdam: John Benjamins Publishing Co.
- Selinker, L. 1972. Interlanguage, *International Review of Applied Linguistics*, 10, 67-92.
- Selkirk, E. 1984. On the Major Class Features and Syllable Theory. I Aronoff, M. & Oehrle, R.T. (Red.): *Language Sound Structure*. MIT Press, Cambridge, Mass.

- SSB. 2020. Innvandrere og norskfødte med innvandrerforeldre. Hentet fra <https://www.ssb.no/innvbef>
- Tenfjord, K., Hagen, J. E., Johansen, H. 2009. Norsk andrespråkskorpus (ASK) – design og metodiske forutsetninger. *NOA – norsk som andrespråk, vol. 25, nr. 1*, 52-81.
- Topintzi, N. 2006. *Moraic Onsets* (Doktorgradsavhandling). University of London, London. 13-15
- Vigrestad, T. 2015. Andrespråklæring og ortografi. I Anne Golden & Elisabeth Selj (red.), *Skrijving på norsk som andrespråk. Vurdering, opplæring og elevenes stemmer*. (1. utg., s. 53-68). Oslo: Cappelen Damm
- Vigrestad, T. 2016. *Ortografi i norsk som andrespråk. En undersøkelse av ortografien i norsk hos voksne innlærere med vietnamesisk og bosnisk/kroatisk/serbisk som morsmål* (Masteroppgave). UiO, Oslo.

Appendiks 1: datamateriale

I denne tabellen er de 409 orda jeg har klassifisert etter hvilken posisjon i ordet og stavelsen det er avvik samt hvilke(t) grafem som er involvert i avviket. Her blir konsonantgrupper forkortet med **K**, og der et vokalgrafem møter konsonantgrupper, sier jeg om dette er før, i eller etter gruppa. Jeg har også klassifisert det etter om det er innsetting av grafem, bortfall av grafem, ombytting av to grafem eller erstatning av et grafem med et annet. I tabellen forkorter jeg innsetting til **I**, bortfall til **B**, ombytting til **O** og erstatning til **E**. Enkelte ord i tabellen forekommer i flere kategorier fordi ordet har flere ulike typer avvik i samme ord. Med klassifiseringene gjort her, har jeg funnet at det er 238 ord som *kan* skyldes transfer av stavelsesstrukturer. I studien min diskuterer jeg hvorvidt disse orda er resultat av transfer.

Mellomspråksord	Lemma	Målspråksord	Avviks-type	Avvik: ord, stavelse, grafem
<elesket>	elske	elsket	I	medialt, <e> før K, kjerne
<ønesker>	ønske	ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesker>		ønsker	I	medialt, <e> før K, kjerne
<ønesket>		ønsket	I	medialt, <e> før K, kjerne

<vestilige>	vestlig	vestlige	I	medialt, <i> før K, kjerne
<ekstremister>	ekstremist	ekstremister	I	medialt, <i> før K, kjerne
<islamiske>	islamsk	islamske	I	medialt, <i> før K, kjerne
<flyktning>	flyktning	flyktning	I	medialt, <i> før K, kjerne
<problemer>	problem	problemer	I	medialt, <e> i ¹⁸ K, kjerne
<blomster>	blomst	blomster	I	medialt <e> før K, kjerne
<kontrollen>	kontroll	kontrollen	I	medialt <o> før K, kjerne
<kvinner>	kvinne	kvinne	I	medialt <i> i K, kjerne
<ekstera>	ekstra	ekstra	I	medialt <e> før K medialt <a> i K, kjerne
<Sri Lanka>	Sri Lanka	Sri Lanka	I	medialt <i> i K, kjerne
<derfor>	derfor	derfor	I	medialt <e> i K, kjerne
<fremst>	fremme	fremst	I	medialt <e> mellom ¹⁹ K, kjerne
<viktigst>	viktig	viktigst	I	medialt, <e> før K, kjerne
<viktigst>		viktigst	I	medialt, <e> før

¹⁸ Med ‘i konsonantgruppe’ mener jeg mellom to konsonantgrafemer som tilhører samme stavellesposisjon. Her ligger norsk fonotaks til grunn for at jeg vurderer to grafemer som del av samme stavele, men jeg problematiserer å legge norsk fonotaks til grunn i diskusjonskapittelet mitt.

¹⁹ Med ‘mellom’ mener jeg mellom to konsonantgrafemer som hører til ulike stavelser i henhold til norsk fonotaks, og der ingen av disse konsonantgrafemene inngår i en konsonantgruppe. Det er altså ikke noen konsonantgruppe å bryte opp i disse tilfellene om vi legger norsk fonotaks til grunn.

				K, kjerne
<viktigst>		viktigst	I	medialt, <e> før K, kjerne
<viktigst>		viktigst	I	medialt, <e> før K, kjerne
<foreldere>	forelder	foreldre	I	medialt, <e> i K, kjerne
<foreldere>		foreldre	I	medialt, <e> i K, kjerne
<forenderer>	forandre	forandrer	I	medialt, <e> i K, kjerne
<forendere>		forandrer	I	medialt, <e> i K, kjerne
<utfordere>	utfordre	utfordre	I	medialt, <e> i K, kjerne
<vinnerene>	vinner	vinnerne	I	medialt, <e> i mulig K, kjerne
<bedere>	bra	bedre	I	medialt, <e> i K, kjerne
<bedere>		bedre	I	medialt, <e> i K, kjerne
<bedere>		bedre	I	medialt, <e> i K, kjerne
<underer>	undre	undrer	I	medialt, <e> i K, kjerne
<søstere>	søster	søstre	I	medialt, <e> i K, kjerne
<brødere>	bror	brødre	I	medialt, <e> i K, kjerne
<hushjelp>	hushjelp	hushjelp	I	finalt, <e> etter K, kjerne
<valge>	valg	valg	I	finalt, <e> etter K, kjerne

<fjelle>	fjell	fjell	I	finalt, <e> etter mulig K, kjerne
<fjelle>		fjell	I	finalt, <e> etter mulig K, kjerne
<lande>	land	land	I	finalt, <e> etter mulig K, kjerne
<veldige>	veldig	veldig	I	finalt, <e> etter mulig K kjerne
<gange>	gang	gang	I	finalt, <e> etter mulig K kjerne
<gange>		gang	I	finalt, <e> etter mulig K kjerne
<folke>	folk	folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<folke>		folk	I	finalt, <e> etter mulig K kjerne
<helge>	elg	elg	I	finalt, <e> etter K kjerne
<bade>	bad	bad	I	finalt, <e> kjerne
<bade>		bad	I	finalt, <e> kjerne
<bade>		bad	I	finalt, <e> kjerne
<bade>		bad	I	finalt, <e> kjerne
<huse>	hus	hus	I	finalt, <e> kjerne
<live>	liv	liv	I	finalt, <e> kjerne
<tide>	tid	tid	I	finalt, <e> kjerne

<være>	vær	vær	I	finalt, <e> kjerne
<hale>	hall	hall	I	finalt, <e> kjerne
<bedrifte>	bedrift	bedrift	I	finalt, <e> kjerne
<sporte>	spørre	spurt	I	finalt, <e> etter mulig K kjerne
<gikke>	gå	gikk	I	final, <e> etter mulig K kjerne
<vindue>	vindu	vindu	I	final, <e> kjerne
<økonome>	økonomi	økonomi	I	final, <e> kjerne
<ferieer>	ferie	ferie	I	medial, <e> kjerne
<pleiede>	pleie	pleie	I	medial, <e> kjerne
<skjent>	kjent	kjent	I	initial, <s> i <skj> opptakt
<skjedelig>	kjedelig	kjedelig	I	initial, <s> opptakt
<slektninger>	slektning	slektninger	I	initial, <s> opptakt
<skjøtt>	kjøtt	kjøtt	I	initial, <s> opptakt
<Skilden>	kilde	kilde	I	initialt, <S> opptakt
<besklytter>	beskytte	beskytte	I	medial, <j> opptakt
<besklytte>		beskytte	I	medial, <j> opptakt
<beskymre>	bekymre	bekymre	I	medial, <s> og <j> opptakt
<heldihvis>	heldigvis	heldigvis	I	medial, <h> opptakt
<arrangerer>	arrangere	arrangerer	I	medial, <j> opptakt

<helge>	elg	elg	I	ordinitialt, <h> danner opptakt i ordinitial stavelse opptakt
<optimiskisk>	optimistisk	optimiskisk	I	medial, <k>, koda
<diktatorer>	diktator	diktator	I	medial, <k>, koda
<positivet>	positiv	positive	I	final, <t>, koda
<prokdusere>	produsere	produsere	I	medial, <k>, koda
<retorikk>	retorikk	retorikk	I	medial, <h> opptakt
<Ethiopia>	Etiopia	Etiopia	I	medial, <h> opptakt
<hverden>	verden	verden	I	initialt, <h> opptakt
<ferierne>	ferie	feriene	I	medial, <r> koda
<brøykte>	røyke	røykte	I	initialt, opptakt
<hart>	ha	hatt	I	medialt, <r> koda
<I fjord>	i fjor	i fjor	I	finalt, <d> koda
<kollisjon>	kollisjon	kollisjon	I	medial, <g> koda
<interlligente>	intelligent	intelligente	I	medial, <r> koda
<veigen>	vei	veien	I	medial, <g> opptakt
<andet>	annen	annet	I	medial, <d> opptakt
<siger>	si	sier	I	medial, <g> opptakt
<grubber>	gubbe	gubber	I	medial, <r> opptakt
<svage>	svak	svake	I	medial, <g> opptakt
<bakgrunden>	bakgrunn	bakgrunn	I	medial, <d> opptakt

<grund>	grunn	grunn	I	final, <d> koda
<miljøfarlig>	miljøfarlig	miljøfarlig	I	medial, <r> koda
<syklonger>	syklon	sykloner	I	medial, <g> koda
<nevnt>	nevnt	nevnt	B	medial, <n> koda
<tekt>	tekst	tekst	B	medial, <s> koda
<adre>	andre	andre	B	medial, <n> koda
<ande>		andre	B	medial, <r> opptakt
<kopliserte>	komplisert	kompliserte	B	medial, <m> koda
<ekstra>	ekstra	ekstra	B	medial, <s> opptakt
<øsket>	ønske	ønsket	B	medial, <n> koda
<viktigte>	viktig	viktigste	B	medial, <s> opptakt
<vaskelig>	vanskelig	vanskelig	B	medial, <n> koda
<vaskellig>		vanskelig	B	medial, <n> koda
<vaskellig>		vanskelig	B	medial, <n> koda
<midt>	mild	mildt	B	medial, <l> koda
<menneker>	menneske	mennesker	B	medial, <s> opptakt
<eksamber>	eksempel	eksempler	B	medial, <l> opptakt
<ekperimenter>	eksperiment	eksperimenter	B	medial, <s> opptakt
<faktis>	faktisk	faktisk	B	final, <k> koda
<vedig>	veldig	veldig	B	medial, <l> koda
<kultur>	kultur	kultur	B	medial, <l> koda
<hjepepleier>	hjelpepleier	hjelpepleier	B	medial, <l> koda
<hjepe>	hjelpe	hjelpe	B	medial, <l> koda
<fratida>	framtida	framtida	B	medial, <m> koda
<kvadratmeter>	kvadratmeter	kvadratmeter	B	medial, <t> koda
<slekninger>	slektning	slektninger	B	medial, <t>, koda
<utvikler>	utvikle	utvikler	B	medial, <l>

				opptakt
<forsøge>	forsørge	forsørge	B	medial, <r> koda
<namere>	nærme	nærmere	B	medial, <r> koda
<hell>	hel	helt	B	final, <t>, koda
<vigtigte>	viktig	viktigste	B	medial, <s> opptakt
<toturert>	torturere	torturere	B	medial, <r> i <rt> i koda
<urettfædig>	urettferdig	urettferdig	B	medial, <r> i <rd> i koda
<forsatt>	fortsatt	fortsatt	B	medial, <t> i <rt> i koda
<forsette>	fortsette	fortsette	B	medial, <t> i <rt> i koda
<forssette>		fortsette	B	medial, <t> i <rt> i koda
<forsatte		fortsatte	B	medial, <t> i <rt> i koda
<forsetter>		fortsetter	B	medial, <t> i <rt> i koda
<forsetter>		fortsetter	B	medial, <t> i <rt> i koda
<forsatt>		fortsatt	B	medial, <t> i <rt> i koda
<forretning>		forretning	forretning	B
<balkon>	balkong	balkong	B	final <g> i <ng> koda
<dønet>	døgn	døgn	B	medial <g> koda
<engan>	engang	engang	B	final <n> i <ng> koda
<legte>	lengte	lengte	B	medial <n> i <ng> koda
<spørsmålet>		spørsmålet	B	medial <s> i <rs>

	spørsmål			koda
<spørsmålet>		spørsmålet	B	medial <s> i <rs> koda
<spørsmål>		spørsmål	B	medial <s> i <rs> koda
<forkerne>	forsker	forskerne	B	medial <s> i <rs> koda
<noske>	norsk	norske	B	medial <r> i <rs> koda
<nosk>		norsk	B	medial <r> i <rs> koda
<nork>		norsk	B	medial <s> i <rs> koda
<forkur>	forkurs	forkurs	B	final <s> i <rs> koda
<gjere>	gjerne	gjerne	B	medial <n> i <rn> koda
<kansje>	kanskje	kanskje	B	medial <k> i <skj> opptakt
<kansje>		kanskje	B	medial <j> i <skj> opptakt
<kanske>		kanskje	B	medial <j> i <skj> opptakt
<kanske>		kanskje	B	medial <j> i <skj> opptakt
<kanske>		kanskje	B	medial, <j> i <skj> opptakt
<kanske>		kanskje	B	medial, <j> i <skj> opptakt
<kankje>		kjanskje	B	medial, <s> i <skj> opptakt
<sikkelig>		skikkelig	skikkelig	B

<forskellige>	forskjellig	forskjellige	B	medial <j> i <skj> opptakt
<forsjellige>		forskjellige	B	medial <k> i <skj> opptakt
<forsjellige>		forskjellige	B	medial <k> i <skj> opptakt
<forsjellige>		forskjellige	B	medial <k> i <skj> opptakt
<forsjellig>		forskjellig	B	medial <k> i <skj> opptakt
<forsjellig>		forskjellig	B	medial <k> i <skj> opptakt
<forsjellig>		forskjellig	B	medial <k> i <skj> opptakt
<forsjellig>		forskjellig	B	medial <k> i <skj> opptakt
<kjønnte>	skjønne	skjønnte	B	initial <s> opptakt
<kjønner>		skjønner	B	initial <s> opptakt
<kilt>	skilt	skilt	B	initial <s> opptakt
<kjønnsmønsteret>	kjønnsmønster	kjønnsmønsteret	B	medial <s> koda
<omsorgfull>	omsorgsfull	omsorgsfull	B	medial <s> koda
<tilfredsstillende>	tilfredsstillende	tilfredsstillende	B	medial <s> koda
<forskningstasjonen>	forskningsstasjon	forskningsstasjon	B	medial <s> koda
<jordbruksforskninginstitutt>	jordbruksforskninginstitutt	jordbruksforskninginstitutt	B	medial <s> koda
<sportaktiviteter>	sportsaktivitet	sportsaktiviteter	B	medial <s> koda
<oppholdstillatelse>	oppholdstillatelse	oppholdstillatelse	B	medial <s> koda
<aldergrense>	aldergrense	aldergrense	B	medial <s> koda
<sykehjem>	sykehjem	sykehjem	B	medial <e> kjerne
<overraskelse>	overraskelse	overraskelse	B	medial <r> koda

<overasket>	overasket	overasket	B	medial <r> koda
<overasket>		overasket	B	medial, <r> koda
<uttrykk>	uttrykk	uttrykk	B	medial <t>, koda
<sykepleierutanning>	sykepleier- utdanning	sykepleierutdannin g	B	medial <d> koda
<kalt>	kald	kaldt	B	medial <d> koda
<rund>	rund	rundt	B	final <t> koda
<manglene>	manglende	manglende	B	medial <d> opptakt
<normann>		nordmann	B	medial <d> koda
<Normenn>	nordmann	nordmenn	B	medial <d> koda
<nodmennene>		nordmennene	B	medial <r> koda
<spennene>	spennende	spennende	B	medial <d> koda
<unervisning>	undervisning	undervisning	B	medial <d> opptakt
<gjelder>	gjelder	gjelder	B	medial <l> koda
<spennede>	spennende	spennende	B	medial <n> koda
<spennede>		spennende	B	medial, <n> koda
<nøvendig>	nødvendig	nødvendig	B	medial, <d> koda
<varandre>	hverandre	hverandre	B	medial, <h> opptakt
<vært>	hvert	hvert	B	initial, <h> opptakt
<his>	hvis	hvis	B	medial, <v> opptakt
<help>	hjelp	hjelp	B	medial, <j> opptakt
<helper>	hjelper	hjelper	B	medial, <j> opptakt
<helper>		hjelper	B	medial, <j> opptakt
<Kjærlighet>	kjærlighet	kjærlighet	B	medial, <r> koda
<kjærlighet>		kjærlighet	B	medial, <r> koda

<sykiske>	psykisk	psykiske	B	medial, <p> opptakt
<de>	det	det	B	ordfinal, <t>, koda
<de>		det	B	final, <t>, koda
<de>		det	B	final, <t>, koda
<sykehuse>	sykehus	sykehuset	B	final, <t>, koda
<sykkehuse>		sykehuset	B	final, <t>, koda
<hjemlande>	hjemland	hjemlandet	B	final, <t>, koda
<miljøe>	miljø	miljøet	B	final, <t>, koda
<biblioteke>	bibliotek	biblioteket	B	final, <t>, koda
<hjemme>	hjem	hjemmet	B	final, <t>, koda
<hjemme>		hjemmet	B	final, <t>, koda
<hjemme>		hjemmet	B	final, <t>, koda
<badde>	bad	badet	B	final, <t>, koda
<språke>	språk	språket	B	final, <t>, koda
<ansikte>	ansikt	ansiktet	B	final, <t>, koda
<politie>	politi	politiet	B	final, <t>, koda
<løpe>	løp	løpet	B	final, <t>, koda
<løpe>		løpet	B	final, <t>, koda
<svare>	svar	svaret	B	final, <t>, koda
<samfunne>	samfunn	samfunnet	B	final, <t>, koda
<håre>	hår	håret	B	final, <t>, koda
<veldi>	veldig	veldig	B	final, <g> koda
<ferdi>	ferdig	ferdig	B	final, <g> koda
<fardi>		ferdig	B	final, <g> koda
<muligheter>	mulighet	muligheter	B	final, <g> koda
<heldihvis>	heldigvis	heldigvis	B	final, <g> koda
<hellidager>	helligdag	helligdager	B	final, <g> koda
<me>	med	med	B	final, <d>, koda
<asma>	astma	astma	B	medial, <t>, koda
<anneledes>	annerledes	annerledes	B	medial, <r> koda

<alle>	aller	aller	B	final, <r> koda
<elle>	eller	eller	B	final, <r> koda
<elle>		eller	B	final, <r> koda
<sosia>	sosial	sosial	B	final, <r> koda
<Grüneløkka>	Grünerløkka	Grünerløkka	B	ordmedial, <r> koda
<hvi>	hvis	hvis	B	ordfinal, <s> koda
<ska>	skal	skal	B	ordfinal, <l> koda
<klassikk>	klassisk	klassisk	B	ordmedial, <s> koda
<dramatikke>	dramatisk	dramatiske	B	medial, <s> koda
<fantastik>	fantastisk	fantastiske	B	medial, <s> koda
<årsken>	årsak	årsaken	B	medial, <a> kjerne
<mosken>	moske	moskeen	B	medial, <e> kjerne
<mosken>		moskeen	B	medial, <e> kjerne
<mosken>		moskeen	B	medial, <e> kjerne
<koren>	Koranen	Koranen	B	medial, <a> kjerne
<koren>		Koranen	B	medial, <a> kjerne
<Sogn og Fjordne>	Sogn og Fjordane	Sogn og Fjordane	B	medial, <a> kjerne
<begnte>	begynne	begynte	B	medial, <y> kjerne
<begnte>		begynte	B	medial, <y> kjerne
<begnte>		begynte	B	medial, <y> kjerne
<rikelig>	rikelig	rikelig	B	medial, <e>

				kjerne
<bleir>	bleie	bleier	B	medial, <e> kjerne
<intersert>	interessert	interessert	B	medial, <e> kjerne
<mennsker>	menneske	mennesker	B	medial, <e> kjerne
<mensker>		mennesker	B	medial, <e> kjerne
<kjøkknet>	kjøkken	kjøkkenet	B	medial, <e> kjerne
<Ethopia>	Etiopia	Etiopia	B	medial, <i> kjerne
<muligheter>	mulighet	muligheter	B	medial, <e> kjerne
<brømt>	berømt	berømt	B	medial, <e> kjerne
<utsende>	utseende	utseende	B	medial, <e> kjerne
<innbyggere>	innbygger	innbyggere	B	medial, <e> kjerne
<månder>	måned	måneder	B	medial, <e> kjerne
<måndene>		månedene	B	medial, <e> kjerne
<diskutere>	diskutere	diskutere	B	medial, <e> kjerne
<ekspermentene>	eksperiment	eksperimentene	B	medial, <i> kjerne
<famlie>	familie	familie	B	medial, <e> kjerne
<gærnt>	gæren	gærent	B	medial, <e> kjerne
<komitèn>	komité	komitéen	B	medial, <e> kjerne

<medføllse>	medfølelse	medfølelse	B	medial, <e> kjerne
<fredlige>	fredelig	fredelige	B	medial, <e> kjerne
<vidergående>	videregående	videregående	B	medial, <e> kjerne
<vidergående>		videregående	B	medial, <e> kjerne
<familien>	familie	familien	B	medial, <e> kjerne
<aldern>	alder	alderen	B	medial, <e> kjerne
<flinkre>	flink	flinkere	B	medial, <e> kjerne
<noklen>	nøkkel	nøkkelen	B	medial, <e> kjerne
<gånde>	gå	gående	B	medial, <e> kjerne
<interssant>	interessant	interessant	B	medial, <e> kjerne
<vidre>	videre	videre	B	medial, <e> kjerne
<vidre>		videre	B	medial, <e> kjerne
<stedenfor>	istedenfor	istedenfor	B	initial, <i> kjerne
<slam>	islam	islam	B	initial, <i> kjerne
<slam>		islam	B	initial, <i> kjerne
<slam>		islam	B	initial, <i> kjerne
<seksti>	seksti	seksti	B	final, <i> kjerne
<Somali>	Somalia	Somalia	B	final, <a> kjerne
<karrier>	karriere	karriere	B	final, <e> kjerne
<karrier>		karriere	B	final, <e> kjerne
<karrier>		karriere	B	final, <e> kjerne

<grens>	grense	grense	B	final, <e> kjerne
<gansk>	ganske	ganske	B	final, <e> kjerne
<mang>	mange	mange	B	final, <e> kjerne
<klass>	klasse	klasse	B	final, <e> kjerne
<skol>	skole	skole	B	final, <e> kjerne
<skol>		skole	B	final, <e> kjerne
<rothin>	rutine	rutine	B	final, <e> kjerne
<vist>	vite	visste	B	final, <e> kjerne
<verst>	verst	verste	B	final, <e> kjerne
<mennesk>	menneske	menneske	B	final, <e> kjerne
<kvinn>	kvinne	kvinne	B	final, <e> kjerne
<modern>	moderne	moderne	B	final, <e> kjerne
<mang>	mange	mange	B	final, <e> kjerne
<mang>		mange	B	final, <e> kjerne
<kaff>	kaffe	kaffe	B	final, <e> kjerne
<grupp>	gruppe	gruppe	B	final, <e> kjerne
<kunn>	kunne	kunne	B	final, <e> kjerne
<allen>	alene	alene	B	final, <e> kjerne
<allen>		alene	B	final, <e> kjerne
<nest>	neste	neste	B	final, <e> kjerne
<flyvend>	flyvende	flyvende	B	final, <e> kjerne
<flyvend>		flyvende	B	final, <e> kjerne
<enest>	eneste	eneste	B	final, <e> kjerne
<enest>		eneste	B	final, <e> kjerne
<lei>	leie	leie	B	final, <e> kjerne
<konkurrer>	konkurrere	konkurrerer	B	final, stavelse ²⁰ , <er>
<ant>	annen	annet	B	medial, stavelse, <ne>
<kaffen>	kafé	kaféen	B	medial, stavelse <é>

²⁰ Med stavelse mener jeg at flere segmenter fra samme stavelse har falt bort.

<tenk>	tenke	tenkte	B	final, stavelse <te>
<gift>	gifte	gifte	B	final, stavelse <te>
<beherst>	beherske	behersket	B	final, stavelse <ke>
<sivilasjon>	sivilisasjon	sivilisasjon	B	medial, stavelse
<diskurer>	diskutere	diskuterer	B	medial, stavelse
<pafle>	bøffel	bøffel	O	<l> og <e>
<eksempel>	eksempel	eksempel	O	<l> og <e>
<torisim>	turisme	turisme	O	<i> og <m>
<triveds>	trives	trivdes	O	<e> og <d>
<fikser>	fiske	fisker	O	<s> og <k>
<tilt>	litt	litt	O	<l> og <t>
<Sinddab>	Sindbad	Sindbad	O	<d> og
<sinll>	snill	snill	O	<n> og <i>
<flyte>	fylle	fylte	O	<l> og <y>
<sølning>	løsning	løsning	O	<l> og <s>
<sølning>		løsning	O	<l> og <s>
<blakong>	balkong	balkong	O	 og <l>
<blakongen>		balkongen	O	 og <l>
<forskjellig>	forskjellig	forskjellig	(O)	<k> og <s> bytter
<jgøres>	gjøre	gjøres	O	<g> og <j>
<forbrede>	forbedre	forbedre	(O)²¹	<r> flytter
<griet>	greit	greit	O	<i> og <e>
<samarbiedspartnere>	samarbeidspart ner	samarbeidspartner	O	<i> og <e>
<seir>	si	sier	O	<i> og <e>
<seir>		sier	O	<i> og <e>
<Gaints>	giant	giants	O	<i> og <a>
<sestør>	søster	søster	O	<e> og <ø>

²¹ Jeg regner dette under kategorien *ombyting og flytting* i analysen.

<Giello>	Geilo	Geilo	O	<e> og <i>
<sunnfunnsfag>	samfunnsfag	samfunnsfag	E	medial, <m> → <n> i koda
<persom>	person	person	E	final <n> → <m> i koda
<språk>	språk	språk	E	final <k> → <r> i koda
<viktigste>	viktig	viktigste	E	medial <k> → <g> i koda
<vigtigte>		viktigste	E	medial <k> → <g> i koda
<eksistere>	eksistere	eksistere	E	medial <k> → <x> i koda
<sammenligned>	sammenligne	sammenlignet	E	final <t> → <d> i koda
<overnatten>	overnatte	overnattet	E	final <t> → <n> i koda
<tilknyter>	tilknytte	tilknyttet	E	final <t> → <r> i koda
<resultal>	resultat	resultat	E	final <t> → <l> i koda
<bott>	bo	bodd	E	final <d> → <t> i koda
<arbeitsoppgave>	arbeids- oppgave	arbeidsoppgave	E	medial <d> → <t> i koda
<of>	og	og	E	final, <f> → <g>, koda
<noer>	noen	noen	E	final, <n> → <r>, koda
<annerleder>	annerledes	annerledes	E	final, <s> → <r>, koda
<stutt>	slutt	slutt	E	medial, <l> → <t>, opptakt

<alkolol>	alkohol	alkohol	E	medial, <h> → <l> opptakt
<van>	være	var	E	final, <r> → <n> koda
<hal>	ha	har	E	final, <r> → <l> koda
<sharf>	skjerf	skjerf	E	initial, <skj> → <sk> opptakt
<physik>	fysikk	fysikk	E	initial <f> → <ph> opptakt
<lillatelse>	tillatelse	tillatelse	E	initial <t> → <l> opptakt
<misforståelse>	misforståelse	misforståelse	E	medial <t> → <l>, opptakt
<misforståelser>		misforståelser	E	medial <t> → <l>, opptakt
<kjenner>	tjene	tjener	E	initial <t> → <k>, opptakt
<kjenner>		tjener	E	initial <t> → <k>, opptakt
<vallegball>	volleyball	volleyball	E	medial <y> → <g> koda
<qvalifiser>	kvalifisere	kvalifiserer	E	initial <k> → <q> opptakt
<kiraf>	giraff	giraff	E	initial <g> → <k>, opptakt
<angre>	andre	andre	E	medial <d> → <g>, opptakt
<guller>	gutt	gutter	E	medial <t> → <l>, opptakt

<omsørgefylt>	omsorgsfull	omsorgsfullt	E	medial, <s> → <e> etter K, kjerne
<pafle>	bøffel	bøffel	E	initial → <p>, koda
<prosjyrer>	brosjyre	brosjyrer	E	initial → <p>, opptakt
<prosjyrer>		brosjyrer	E	initial → <p>, opptakt
<puda>	bud	buda	E	initial → <p>, opptakt
<pud>		bud	E	initial → <p>, opptakt
<pud>		bud	E	initial → <p>, opptakt
<pe>	be	be	E	initial → <p> opptakt
<petyr>	bety	betyr	E	initial → <p> opptakt
<kompinere>	kombinere	kombinere	E	medial → <p> opptakt
<ampulanse>	ambulanse	ambulanse	E	medial → <p> opptakt
<blas>	plass	plass	E	initial <p> →
<blas>		plass	E	initial <p> → opptakt
<flyblaser>	flyplass	flyplasser	E	medial <p> → opptakt
<flyblasen>		flyplassen	E	medial <p> →

				 opptakt
<benger>	penge	penger	E	initial <p> → opptakt
<benger>		penger	E	initial <p> → opptakt
<eksamber>	eksempel	eksempler	E	medial <p> → opptakt
<hvertval>	hvertfall	hvertfall	E	medial <v> → <t> opptakt
<viske>	fiske	fiske	E	initial <v> → <t> opptakt

Appendiks 2: informanter

Her følger to tabeller med informasjon om alle de 107 informantene mine. Den første tabellen er en oversikt over de sju informantene som har tatt norskprøve på høyere nivå (jf. kap. 6.2) Her oppgir jeg alder, hjemland og tredje- og fjerdespråk.. Den andre tabellen viser informasjon om de hundre informantene som har tatt Språkprøven. Her oppgir jeg alder og hjemland. Attributtene er ulike i de to tabellene fordi det er ulikt hvilken informasjon de to gruppene har gitt om seg selv. De som har tatt norskprøve – høyere nivå, har angitt tredje- og fjerdespråk, men ikke innlæringstid. De som har tatt Språkprøven, har latt feltet om tredje- og fjerdespråk stå åpen. I tabell 2 har jeg bøtt på mangelen av tredje- og fjerdespråksinformasjon ved å legge til informasjon om hjemlandet. Selv om vi ikke vet nøyaktig hvilke andre språk enn norsk og somali informantene behersker, kan hjemlandet gi oss en viss idé om hva slags språksituasjon de har vokst opp i. Siden det et krav at alle informantene mine har somali som morsmål, har jeg ikke spesifisert dette i tabellene.

Tabell 1: Høyere nivå

Nr.	Alder	Hjemland	3. og 4.-språk
1	35	Somalia	Arabisk
2	31	Somalia	Arabisk
3	26	Somalia	Arabisk og fransk
4	25	Somalia	Arabisk
5	30	Somalia	Arabisk og ungarsk
6	45	Somalia	Arabisk og italiensk
7	47	Somalia	Arabisk og russisk

Tabell 2: Språkprøven

Nr.	Alder	Hjemland	Nr.	Alder	Hjemland
8	44	Somalia	58	19	Somalia
9	24	Somalia	59	30	Somalia
10	25	Somalia	60	24	Somalia
11	23	Somalia	61	42	Somalia
12	30	Somalia	62	32	Somalia
13	30	Somalia	63	24	Somalia

14	35	Somalia	64	33	Somalia
15	25	Somalia	65	38	Somalia
16	16	Somalia	66	35	Etiopia
17	23	Somalia	67	31	Somalia
18	20	Somalia	68	32	Somalia
19	27	Somalia	69	42	Somalia
20	25	Somalia	70	23	Somalia
21	39	Somalia	71	34	Somalia
22	22	Somalia	72	34	Somalia
23	25	<i>Ikke oppgitt</i>	73	50	Somalia
24	27	Somalia	74	36	Somalia
25	20	Somalia	75	22	Somalia
26	25	Somalia	76	23	Somalia
27	28	Somalia	77	19	Somalia
28	32	Somalia	78	27	Somalia
29	29	Somalia	79	34	Somalia
30	32	Somalia	80	18	Somalia
31	24	Somalia	81	23	Somalia
32	51	Somalia	82	19	Somalia
33	26	Somalia	83	17	Somalia
34	32	Somalia	84	25	Somalia
35	27	Somalia	85	38	Somalia
36	34	Somalia	86	31	Somalia
37	36	Somalia	87	38	Somalia
38	48	Somalia	88	26	Somalia
39	32	Somalia	89	48	Somalia
40	24	Somalia	90	27	Somalia
41	26	Somalia	91	28	Somalia
42	35	Somalia	92	37	Somalia
43	30	Somalia	93	20	Somalia
44	42	Somalia	94	29	Somalia
45	28	Somalia	95	35	Somalia

46	30	Somalia	96	29	Somalia
47	26	Somalia	97	24	Somalia
48	22	Somalia	98	44	Somalia
49	51	Somalia	99	25	Somalia
50	28	Somalia	100	25	Somalia
51	31	Somalia	101	26	Somalia
52	26	Somalia	102	35	Somalia
53	21	Somalia	103	18	Somalia
54	30	Somalia	104	26	Somalia
55	31	Somalia	105	27	Somalia
56	25	Kenya	106	39	Somalia
57	18	Somalia	107	24	Somalia

Informantene mine har en gjennomsnittsalder på 29,8 år. Aldersspennet er fra 16 til 51 år.

