

Pórhildur Oddsdóttir

MA-speciale i dansk

Ord for ord

Ordforråd i dansk hos ni islandske gymnasieelever

Háskóli Íslands

Januar 2004

Vejledere:

Auður Hauksdóttir

Nina Møller Andersen

Inholdsfortegnelse

| | |
|---|-----------|
| 1 Indledning..... | 3 |
| 1.1 Formål med projektet..... | 3 |
| 1.2 Problemformulering..... | 4 |
| 1.3 Hvad kan undersøgelsen bruges til? | 4 |
| 1.4 Specialets disposition..... | 5 |
| 2 Teorier om fremmedsprogsindlæring – tilegnelse af ordforråd..... | 7 |
| 2.1 Hvad er et ordforråd – og hvor mange ord skal man kunne?..... | 7 |
| 2.2 Historisk oversigt..... | 9 |
| 2.3 Leksikalsk kompetence. At kunne ord..... | 13 |
| 2.4 Ordforråd og læsning..... | 19 |
| 2.5 Hypotesemodellen..... | 22 |
| 2.6 Kognitiv teori..... | 23 |
| 2.7 Eksplicit viden - implicit viden..... | 25 |
| 2.8 Leksikaliserede helheder..... | 28 |
| 2.9 Receptiv viden - produktiv viden..... | 33 |
| 2.10 Sammenfatning..... | 34 |
| 3 Anvendte metoder til måling af produktivt ordforråd..... | 36 |
| 3.1 Metoder..... | 36 |
| 3.2 Leksikalsk tæthed (LT)..... | 36 |
| 3.3 Leksikalsk originalitet (LO)..... | 36 |
| 3.4 Leksikalsk raffinement (LR)..... | 37 |
| 3.5 Leksikalsk variation (LV). Type/Token Ratio (T/T)..... | 37 |
| 3.6 Leksikalsk frekvensprofil (LFP)..... | 38 |
| 3.7 Kritik – Nye metoder..... | 39 |
| 4 Empirisk del | 41 |
| 4.1 Undersøgelsens gennemførelse..... | 41 |
| 4.1.1 Ideen bag undersøgelsen..... | 41 |
| 4.1.2 Valg af skole..... | 42 |
| 4.1.3 Informanternes baggrund..... | 43 |
| 4.1.4 Kriterier for valg af informanter..... | 43 |
| 4.1.5 Indsamling af data..... | 44 |
| 4.1.6 Behandling af data..... | 46 |
| 4.1.7 Hvornår tælles et ord som et ord?..... | 47 |
| 4.2 Afprøvning af metoder I..... | 47 |
| 4.2.1 Overvejelser angående fremgangsmåden..... | 47 |
| 4.2.2 De to opgaver – den første og den sidste..... | 48 |
| 4.2.3 Leksikalsk tæthed (LT)..... | 49 |
| 4.2.4 Leksikalsk originalitet (LO)..... | 51 |
| 4.2.5 Leksikalsk raffinement (LR)..... | 54 |
| 4.2.6 Konklusion..... | 54 |
| 4.3 Afprøvning af metoder II..... | 55 |
| 4.3.1 Størrelse og inddeling af informanternes ordforråd..... | 56 |
| 4.3.2 Leksikalsk variation. T/T ratio..... | 64 |

| | |
|---|------------|
| 4.3.3 Ni informanter – to opgaver..... | 67 |
| 4.3.3 Lige lange opgaver..... | 71 |
| 4.3.5 Konklusion..... | 74 |
| 4.4 Afprøvning af metoder III..... | 75 |
| 4.4.1 Leksikalsk frekvensprofil (LFP)..... | 76 |
| 4.4.2 LFP af hele ordforrådet..... | 76 |
| 4.4.3 LFP af enkelte opgaver..... | 78 |
| 4.4.4 Sammenligning af frekvenskategorier i to opgaver..... | 82 |
| 4.4.5 Tre informanter – oversigt..... | 85 |
| 4.4.6 Kerneord..... | 90 |
| 4.4.7 Diskussion og konklusioner..... | 91 |
| 4.5 Andre evalueringsmetoder..... | 92 |
| 4.5.1 Receptive prøver..... | 92 |
| 4.5.2 Adjektiver..... | 99 |
| 4.5.3 Udtryk og vendinger..... | 100 |
| 5 Sammenfatning og konklusion..... | 106 |
| 6 Litteratur..... | 112 |
| 7 Oversigt over tabeller og figurer..... | 117 |
| 8 Bilag..... | 121 |

1 Indledning

1.1 Formål med projektet

Dette projekt er et MA-speciale ved Islands Universitet. Formålet med specialet er at undersøge det produktive ordforråd i dansk hos en gruppe islandske elever i løbet af deres første semester i en islandsk integreret gymnasieskole¹. Meningen med undersøgelsen er at kortlægge sprogproduktionen i dansk og på den måde at få indsigt i den ordforrådstilegnelse der finder sted under forløbet.

I gennem de ca. 25 år jeg har været lærer, har jeg været med til mange kurser i fremmedsprogpædagogik, deltaget i udarbejdelse af undervisningsmateriale og prøvet lidt af hvert i forsøg på at lave om på undervisningen for at gøre den mere målrettet og effektiv især med henblik på inddragelse af kommunikative og interaktive metoder.

Den skole jeg har været ansat på de sidste 16 år, har i løbet af den tid gennemgået en større omvæltning fra at være et forholdsvis traditionelt gymnasium med klassesystem til at være en integreret² skole med modulsystem. Stort set alle fag kører i ét semesters kurser med seks timer om ugen. I al den tid er en stor del af lærerens energi blevet brugt på at følge med i udviklingen og holde trit med forandringerne. Der har aldrig rigtigt været tid til virkelig at undersøge om al den omvæltning har gjort noget positivt for eleverne, om progressionen har tjent det formål den skulle. Det er naturligvis umuligt at skrue tiden tilbage, men nu gives der lejlighed til at vurdere situationen – i håb om en mulighed for at danne et grundlag for det videre arbejde.

¹ Islandske gymnasieskoler er stort set enten skoler med traditionelt klassesystem eller integrerede skoler med modulsystem. Den skole, hvor mine informanter går, tilhører den sidste gruppe. Alle fag er inddelt i moduler. Et modul kan som regel defineres som seks timer om ugen i et semester (14 uger) med en afsluttende eksamen eller løbende evaluering. Elever på naturfaglig og samfundsfaglig linie skal have to moduler dansk, mens elever på sproglig linie skal have den tredje med. Elever, der går på det samme hold i dansk, går ikke nødvendigvis sammen i andre fag.

De traditionelle gymnasier med klassesystem fordeler det samme antal undervisningstimer på to skoleår.

² Udover gymnasielinie er skolen fagskole for hotel- og levnedsmiddelfag og byder desuden på forskellige kortere erhvervsrettede uddannelser.

1.2 Problemformulering

Hvordan forholder det sig med udviklingen af elevernes sprogproduktion (ordforrådstilegnelse) i dansk i gymnasietidens første semester?

- Hvor stort ordforråd har eleverne i begyndelsen af semestret?
- Hvor stort ordforråd har eleverne i slutningen af semestret?
- Hvad slags ordforråd er der tale om; kerneord, specialiseret ordforråd, akademisk ordforråd? Frekvensen af de ord der bliver brugt.
- Kan der spores en udvikling i ordforrådet?
- Er der tale om udvikling i brug af kollokationer?
- Er der forskel på elevernes ordforråd, når de udtrykker sig mundtligt og skriftligt?
I så fald, hvilken forskel er der tale om?
- Hvilke strategier bruger eleverne, når de ikke kender det rigtige ord for det de vil udtrykke?
- Hvilken rolle spiller opgavernes forskellighed?

1.3 Hvad kan undersøgelsen bruges til?

Hovedformålet med denne undersøgelse er at finde ud af, om der kan konstateres nogen udvikling i ordforråd i dansk hos eleverne i løbet af deres første semester i gymnasiet.

I løbet af de sidste år er der lagt op til en hel del forandringer i skolesystemet, ikke mindst i gymnasieskolerne. Skolerne har fået en ny læseplan, hvor der lægges vægt på kommunikativ undervisning, IT, elevdifferentiering og elevautonomi i alle fag. Desuden er kravene til optagelse på skolernes studenterlinier blevet skærpet. På den skole informanterne kommer fra, er der desuden sket det indenfor danskfaget, at i stedet for at eleverne har dansk i tre semestre med 4 timer om ugen, har de nu dansk i to semestre med seks timer om ugen. Dette gælder naturfaglig og samfundsfaglig linie, men sproglig linie har yderligere et semester med seks timer om ugen.

Da jeg skulle finde et emne til mit speciale, faldt valget på at undersøge ordforrådsindlæring og elevernes udvikling i brug af ordforråd over et semester. Ved at

undersøge og analysere elevernes produkter åbnes der nye muligheder for at opdage og udpege svagheder og huller i deres formåen samt lejlighed for lærerne til at pejle sig frem til en forbedring af undervisningen eller tilrettelæggelsen af den.

Efter min mening er det en fordel at kunne indsamle materiale uden at eleverne bliver påvirket af en udefra kommende forsker. Det materiale, jeg har, er udelukkende opgaver der indgik i elevernes semesterplan, og det har derfor ikke krævet andres tilstedeværelse end lærerens. Ved semestrets start gav alle informanterne deres dansklærer lov til at kopiere deres produkter mod at de blev brugt anonymt. Derudover det har disse elever arbejdet uden indblanding udefra og er hverken blevet udsat for ekstra pres eller opfordring.

Da jeg selv har været med til at udarbejde det undervisningsforløb, der bliver fulgt, og de opgaver, som bliver brugt på skolen, er der en vis fare for at jeg vil komme til at lægge en lidt for subjektiv synsvinkel på denne undersøgelse. Derfor er det et bevidst valg fra min side ikke at komme i direkte kontakt med eleverne og på den måde prøve at lægge lidt afstand mellem min rolle som lærer og min rolle som specialeskriver.

1.4 *Specialets disposition*

Denne opgave er inddelt i tre hovedafsnit. I teoridelen vil jeg begynde med spørgsmålet om, hvad ordforråd egentlig er, for derefter at give en kort historisk oversigt over, hvilken stilling ordforrådet har haft i fremmedsprogsforskningen.

Hvad vil det sige at kunne et ord, – er et afgørende spørgsmål i sprogtileningsforskningen. Jeg vil gøre rede for forskellige aspekter af det område og komme ind på noget af den omfangsrige forskning, der har været i gang i løbet af de sidste 20-30 år. Blandt de grundlæggende teorier i forskningen er teorien om hypotesedannelse og –afprøvning, samt teorien om sprogtilenings kognitive processer. Forskning indenfor området angående hvorvidt eksplicit viden kan gøres implicit og hvilke processer kan aktiveres i den forbindelse til at gøre indlæring mere effektiv, er et af de problemer man forsøger at finde en løsning på. Teorier om den rolle leksikaliserede helheder spiller i færdigheden at beherske et sprog, har åbnet en ny dimension i sprogtileningsforskningen. Receptiv vs. produktiv viden er endnu et område,

sprogforskningen har beskæftiget sig med. Dette område har også været aktuelt i undervisningssituationen, ikke mindst i forbindelse med evaluering af sprogfærdigheder.

For det andet vil jeg gøre rede for anvendelse, fordele og kritik af de metoder, der er blevet brugt i forsøget på at måle og evaluere skriftligt og mundtligt ordforråd.

For det tredje vil jeg afprøve disse metoder på fem (fire skriftlige og én mundtlig) opgaver i dansk lavet af ni islandske elever på første semester i et islandsk gymnasium. Opgaverne er skrevet i løbet af 13 uger. Formålet er at finde ud af, hvilke metoder der bedst egner sig til at belyse henholdsvis ordforrådets omfang og kvalitet hos eleverne. Desuden vil jeg prøve at finde ud af, om det er muligt, ved brug af en eller flere af disse metoder at konstatere en udvikling hos eleverne i løbet af semestret og i så fald, vil jeg give bud på hvori den består.

Andre evalueringsmetoder bliver også brugt til at belyse informanternes ordforråd. Informanterne blev testet i receptivt ordforråd med den hensigt at få indblik i forholdet mellem deres receptive og produktive viden om sprog. Til sidst vil jeg se lidt nærmere på sammensætningen af informanternes ordforråd, især brug af adjektiver på den ene side og brug af faste udtryk og vendinger på den anden. Talesprog og strategier får en beskedent opmærksomhed til allersidst.

Til slut vil jeg opsummere og konkludere angående de resultater, jeg har nået frem til under bearbejdningen af undersøgelsens materialer.

Jeg vil hermed bruge lejligheden til at sige tak til alle de personer, der har gjort det muligt for mig at skrive dette speciale. Hjertelig tak til eleverne i MK, der gav mig lov til at bruge deres opgaver, til kollegaerne Þórdís og Magnús, til 'Specialegruppen' på Engelsk Institut, til familien og alle de andre der har måttet bære over med mig. Sidst, men ikke mindst, tak til mine vejledere, Auður Hauksdóttir og Nina Møller Andersen for deres tålmodighed, overbærenhed, opmuntring og inspiration.

2 Teorier om fremmedsprogsindlæring – tilegnelse af ordforråd

2.1 Hvad er et ordforråd – og hvor mange ord skal man kunne?

Ordforråd udvikler sig hurtigt hos børn, som er i gang med at tilegne sig deres modersmål. Et fem år gammelt barn har sandsynligvis et ordforråd, som består af omkring 2000 ord, 17-18 år gamle kender cirka 40.000 ord (måske dobbelt så mange, hvis egennavne, ordsprog og faste vendinger tæller med). Hvis det er rigtigt, lærer børn og unge mindst 3000 og muligvis helt op til 5000 ord om året. I alderen fra 5-18 år lærer de gennemsnitligt op til 13 ord hver eneste dag. Det er klart, at de unge ikke bliver undervist i alle disse ord og at de som regel ikke bliver specielt trænet i tilegnelsen af ordene. Mange forskere mener dog, at læsning spiller en stor rolle i den sammenhæng. (Pind, 1997:259)

Nation antager, at en studerende på college kan 58.000 basisord, 155.000 ord når afledte ord tælles med (Nation 2001:365).

Anglin (1993) foretog ordforrådsmålinger hos 6, 8 og 10 år gamle børn:

| Alder | Antal basisord (grundform) | Antal afledte ord |
|-------|----------------------------|-------------------|
| 6 år | 3.000 | 2.000 |
| 8 år | 4.500 | 5.500 |
| 10 år | 7.500 | 16.000 |

Anglins resultater viser, at ordforrådet vokser med stigende hastighed i forhold til børnenes alder. Antal afledte ord stiger langt mere end basisord; det skyldes dels, at de førstnævnte er flere og dels det, at modenhed og større horisont gør det muligt for barnet at tilegne sig mere udviklet ordforråd.

Danske eller islandske tal angående ordforrådstilegnelse eksisterer ikke eller har ikke været tilgængelige. I engelsk taler man om, at en uddannet modersmålstalende kender omkring 20.000 ordfamilier.³ Nation (2001) mener, at den vurdering er for lav, fordi

³ En ordfamilie består af et basisord, afledte former og nært beslægtede ord.

mennesker lærer nye ord hele tiden, op til 1000 ordfamilier hvert år i deres voksenliv. Han mener, at den mængde ord er opnåelig for mennesker, der lærer engelsk som andetsprog⁴, men at den er uden for rækkevidde for dem, som lærer engelsk som fremmedsprog. For den modersmålstalende har alle nye ord samme værdi. Det kan man til gengæld ikke sige om nyerehvervede ord i fremmedsproget. I den sammenhæng skal man spekulere over ordenes hyppighed i sproget. For flestes vedkommende gælder det, at man i begyndelsen af fremmedsprogsindlæringen har mest brug for ord af høj frekvens, for det er jo den slags ord der bliver brugt i hverdagens kommunikation angående helt basale ting i tilværelsen.

Linnaruds (1986) sammenlignede skriftlige engelsk opgaver skrevet af 17 år gamle svenske elever, som havde lært engelsk i 9 år, med skriftlige opgaver fra jævnaldrende engelsktalende. Hendes resultater viste afgørende forskel på sproget hos modersmålstalende og andetsprogs-/fremmedsprogsindlærere:

| Engelsktalende | Svenskere |
|--|---|
| <ul style="list-style-type: none"> • skrev længere tekster med færre fejl • mere varierende brug af de mest frekvente ord (mest i adjektiver og adverbialer) | <ul style="list-style-type: none"> • manglede variation i sprogbrug • var mindre originale i brug af ordforråd • mindre brug af idiomer og kollokationer |

Nation taler om at for den, som stiler mod en akademisk uddannelse, skal sætte sig for at lære de 2000 mest højfrekvente ord. Det antal ord vil dække ca. 80% af de faglige teksters ordforråd, men ud over det, er det nødvendigt at kende 570 ordfamilier af den akademiske ordliste, foruden fagspecifikke ord, for at dække de sidste 20% af ordforrådet. (Nation, 2001:14-18)

En læser skal kunne forstå 95% af ordforrådet i en tekst for at få det fulde udbytte af teksten (Gimbel, 1995).

⁴ *Andetsprog*: Et sprog som ikke er indlærerens modersmål, men som han støder på i sin hverdag og har brug for at kunne anvende til en vis grad. *Fremmedsprog*: Et sprog som indlæreren studerer uden for det land eller område sproget bliver brugt. Nogle forskere mener, at begrebet '*andetsprog*' også indebærer begrebet '*fremmedsprog*'. Læseren må derfor ud fra konteksten afgøre, hvilket af de to begreber det drejer sig om.

I engelsk er der blevet lavet lister over højfrekvente ord, lavfrekvente ord og ord, der bruges inden for forskellige områder (for special purposes). Den slags lister findes, mig bekendt, endnu hverken i dansk eller islandsk. Derimod har Ruus (1995) udarbejdet en liste over 1.117 ord, som hun definerer som kerneord i det danske sprog. Den liste har jeg benyttet mig af (jf. 4.4.6) I islandsk er der blevet lavet specifikke ordlister over nye ord inden for forskellige fagområder (íðorðalistar).

I 1943 gjorde Richards rede for en basisordliste i engelsk på 850 ord. Den mente han var nok til at kunne udtrykke sig på engelsk. Nation (1990) påpegede, at med disse 850 ord kunne der gives udtryk for 12.425 meninger, hvilket igen ville give problemer med at udvælge den 'basale' eller den mest 'praktiske' mening af hvert ord (Carter og McCarthy, 1988:3).

2.2 Historisk oversigt

Hvordan lærer man fremmedsprog? Hvordan underviser man i fremmedsprog? Disse to spørgsmål indeholder uendeligt mange underspørgsmål eller delspørgsmål, og mange af dem vil der nok aldrig kunne gives et entydigt svar på. Nogle af de mest udforskede spørgsmål de sidste 10-15 år går på den ene side ud på elevernes tilegnelse af ordforråd. På den anden side opstår spørgsmål angående det at undervise i ordforråd og om det overhovedet kan lade sig gøre at undervise i ordforråd. Et stort og varieret ordforråd er en helt uundværlig del af vores kommunikative kompetence; nogle mener ligefrem at det er den vigtigste del.

I forordet til Carters og McCarthys bog '*Vocabulary and Language Teaching*' (1988:vii) udtrykker Candlin, hovedredaktøren af bogen, sin forbavselse over, at det først efter midten af 80'erne skete, at *ordforråd* blev identificeret som et hovedemne (key topic) i den ellers brede vifte af emner inden for sprogpedagogikken. Man har behandlet ordforrådet som et underordnet emne inden for læsning eller skrivning eller som en del af andetsprogstilegnelsen. Redaktøren prøver at give forskellige forklaringer på denne anskuelse, blandt andet at lingvistisk teori og undersøgelser af intersproget har overskygget ordforrådet, men slutter sig enig med det selve bogen understreger, nemlig, at det at studere ordforråd er selve kernen i sprogundervisning og sprogstilegnelse. Hvad

enten det så angår læseplan, evaluering af elevernes præstationer, undervisningsmaterialer eller den iøjnefaldende ting, at det er på den måde de fleste elever ser sprog eller vanskeligheder ved at lære sprog. Hvis han har ret, har *ordforråd* som selvstændigt og videnskabeligt emne mindre end 20 års historie.

Ifølge Richards & Rodgers var ordforrådets rolle indenfor 'The oral approach'⁵ til undersøgelse i 20'erne og 30'erne. På den ene side var der den antagelse Palmer og nogle andre specialister i sprogundervisning havde, at ordforråd var et af de vigtigste aspekter inden for fremmedsprogsundervisningen. På den anden side mente mange at ordforråd var en vigtig komponent til læsefærdighed. Det førte til arbejdet med de ordlister i engelsk, som har dannet grundlaget for udviklingen af undervisningsmateriale i engelsk som andet- og fremmedsprog. Lingvister og sproglærere er nu begyndt at sætte fokus på ordforråd efter årtiers ignorering (Richards & Rodgers 2001:36).

Ordforrådet kan ikke adskilles fra de øvrige dele af sprogundervisningen og indlæringen. Vi bruger ord i alle former for kommunikation, hvad enten vi taler, lytter, læser eller skriver. I den forbindelse støder vi på nye ord eller vi opdager nye betydninger af ord vi allerede kender. Men hvordan kan ordforrådsindlæring eller undervisning i ordforråd gøres til en målbevidst del af fremmedsprogstilægnelsen? Hvad har forskere og lærere givet sig af med i den forbindelse og hvor langt er man nået? I hvilken retning peger forskningen nu og hvilken lære kan en fremmedsproglærer drage af den? Carter og McCarthy har lyttet til lærere og studerende og har lavet en liste over nogle af de spørgsmål som tit bliver stillet.

- Hvor mange ord består et brugbart ordforråd af i et fremmed sprog?
- Hvilke ord er mest formålstjenligt at lære først?
- Er nogle ord mere gavnlige for indlæreren end andre i begyndelsen af fremmedsprogsindlæringen?
- Er nogle ord sværere at lære end andre? Kan ord inddeles i sværhedsgrader?
- Hvad er den bedste måde at huske nye ord på?
- Er det mest praktisk at lære ord som enkelte ord på en ordliste, i par eller i kontekst?
- Hvad med ord, der har forskellige betydninger? Skal de undgås? Hvis ikke, skal nogen betydninger læres før andre?

⁵ Britisk udspring fra den direkte metode.

- Er der nogen ord, som indlæreren sandsynligvis møder i talesprog frem for skriftsprog? Hvis det er tilfældet, kender vi dem så?

(Carter og McCarthy 1988, s.1-2)

Der gives ingen absolutte svar på disse spørgsmål, men alle markerer de en holdning til fremmedsprogsindlæring. Spørgsmålene er stadigvæk aktuelle 15 år efter bogens udgivelse. Nogle af disse spørgsmål vil blive drøftet og sat i sammenhæng med andre spørgsmål angående det nyeste i forskningen på dette område.

Efter 1945 og hen imod slutningen af tresserne dominerede det syn Fries lagde for dagen i sin bog *'Teaching and Learning English as a Foreign Language'*. Han mente at det ikke var ordforrådet der var så svært, men selve beherskelsen af lydsystemet og den grammatiske struktur. Eleven havde kun brug for et basisordforråd for at kunne praktisere den syntaktiske struktur. Når dette stadium først var nået, kunne eleven bevæge sig til de næste tre stadier:

- at lære ordforråd til produktion
- udvidelse af ordforrådet til genkendelse
- at lære ordforråd angående specielle områder og erfaring

(Fries 1945:39)

Amerikanske strukturalister fokuserede på fonologi og syntaktiske mønstre. Parallelt var det den kontrastive analyse og så adfærdspsykologien, som dannede grobund for den audiolingviale metode, som ekko af Fries. Deres forhold til undervisning i ordforråd kan udtrykkes i det Rivers skrev:

- Ingen vidste, hvilket ordforråd elever ville få brug for
- For megen interesse for ordforråd kunne give eleven det indtryk, at det bare drejede sig om en masse ord
- Tøven i fremkaldelsen af et ord skyldtes for megen undervisning i ordforråd
- Tilegnelse af L1⁶ skulle foregå med et lille ordforråd indtil strukturerne var på plads og kunne beherskes fuldt ud.

(Rivers 1968:208)

⁶ L1 er en betegnelse for et menneskes første sprog, som oftest modersmål. L2 er en betegnelse for det målsprog fremmedsprogsundervisningen/-indlæringen sigter mod.

Wilkins var af en anden mening og gav udtryk for at det at kunne ordforråd var mindst lige så vigtigt, som det at beherske grammatiske regler. Han mente at leksikalsk semantik kunne hjælpe os til at:

- forstå oversættelsesprocessen
- organisere ordforrådet – hvilket er nødvendigt fordi ord ikke læres isoleret
- finde ud af ordenes fulde betydning – knyttet til den udviklede betydningstilegnelse (dvs. betydning, mening, forståelse af et ord – hvilket ofte høres i den påstand at nogen har 'lært' et ord på et meget tidligt stadium).

(Wilkins 1972:130-32)

Twaddell (1973) fremsatte den opfattelse i begyndelsen af 70'erne, at det ville være umuligt at undervise i alle de ord indlærerne havde brug for. Derfor ville det være vigtigt at hjælpe dem til at udvikle gættestrategier, så de kunne tackle ukendte ord og stort set frigøre sig fra ordbøger. Han blev dermed en af de første til at se ordforrådsindlæring, som en form for '*skill*', og for at føre ansvaret for indlæringen fra læreren til eleven. Beherskelse af gættestrategier blev senere set som det centrale i læseforståelsen.

Ifølge Henriksen førte fejlanalyser i elevsprog i 70'erne til konstateringen af, hvor stor indflydelse modersmålet har på ordforrådstilegnelsen. Når modersmål og målsproget er nært beslægtede sprog, får man mange ord forærende, ikke mindst receptivt, som ved læsning.

Grundforskningen har skaffet viden som medfører, at syntaksen ikke er så dominerende mere. Forøget interesse for kulturformidling i sprogundervisningen kræver mere viden om ord og ordenes betydning. For at kunne snakke flydende og nuanceret har indlæreren brug for et varieret ordforråd. Det er i undervisningen, som viden om sprog (betydning) og sprogindlæring (psykolingvistikken), formidles og gør eleverne mere kompetente i tilegnelsen af ordforråd. (Henriksen 1999a:75)

2.3 Leksikalsk kompetence. At kunne ord.

I det foregående afsnit har jeg fokuseret på det historiske aspekt af ordforrådstilegnelsen. Et naturligt næste skridt vil være, at se nærmere på, hvad man forstår ved leksikalsk kompetence.

Hvad vil det sige at kunne et ord? Hvad betyder det i grunden at 'lære' ordforråd? Richards (1976) anså det for at være 'noget af det kendskab, som kan regnes til det der kaldes leksikalsk kompetence'. Han har sammenfattet en karakterisering af (modersmåls lignende) leksikalsk kompetence til otte brede 'antagelser':

1. Modersmålstalende bliver ved med at udvikle deres ordforråd hele deres voksenliv (men der sker ikke stor udvikling i syntaksen efter at man er blevet voksen). Alt fra 20.000 til 100.000 ord kunne være en persons receptive ordforråd.
2. At kende et ord er at kunne bruge det (i skrift og i tale) og vide hvilket/hvilke ord det som oftest knytter sig til (frekvens og kollokationer).
3. At kende et ord er at vide, hvornår ordet ikke kan bruges, (funktions- og situationsbundet begrænsning for brug af det).
4. At kende et ord er at kende dets syntaktiske brug (transitive verber, kasus).
5. At kende et ord er at kende dets underliggende former og afledte ord.
6. At kende et ord er at vide, hvor det passer ind i et netværk af associationer med andre ord i sproget.
7. At kende et ord er at kende dets semantiske værdi.
8. At kende et ord er at kende dets forskellige betydninger (polysemi)

(Richards, J.C. 1974 efter Carter og McCarthy, 1988:44-45)

Ifølge Meara er Richards' liste over 'antagelser' blevet brugt af mange forskere som model eller ramme for beskrivelse af ordforrådsviden ('vocabulary knowledge framework') i over 20 år. Han standser op ved 'antagelse' nummer seks og kommer til det resultat, at den rummer et antal skjulte antagelser:

- hvert enkelt ord i sproget kommer ind i et netværk af associationer med andre ord i sproget
- netværket er stort set ens for alle der taler sproget
- det er muligt for os at specificere, hvad dette netværk er
- en der taler sproget, 'kender' netværket
- netværket er endeligt og stabilt
- et netværk af associationer er et primært billede af et ordforråd – frem for et sekundært fænomen, der adskiller sig fra andre dybere strukturerede egenskaber.

- et tosproget ordforråd er ikke mærkbart anderledes end ordforrådet hos etsprogede mennesker.

(Meara, 1999; 3-4)

Meara giver to bud på alternativer til Richards' 'antagelser'

1. De tre følgende dimensioner bør række til kategorisering af indlærerens leksikalske kompetence.

- størrelsen af ordforrådet
- graden af automatisk indlemmelse af nye ord og ordforbindelser i ordforrådet
- en enkelt form for måling af i hvor høj grad leksikalsk struktur lænker ord i et ordforråd sammen

Kun i den første dimension har man forsket i dybden og i den tredje er nogle forskere allerede i gang. Afprøvning af ordforråd indenfor de tre dimensioner kunne give et godt billede af den enkeltes ordforråd. I stedet for at se på hvert enkelt ord, vil det være muligt at se indholdet af leksikonet som en helhed.

2. En fremgangsmåde hvor ordet er i centrum, ikke indlæreren; hvor man prøver at identificere alle de stadier ordene skal igennem indtil de bliver fuldt ud integreret i den talendes ordforråd. Modeller af denne karakter er ofte inkluderet i arbejdet med andetsprogets leksikalske tilegnelse, men sjældent udført i detaljer. Hvor de er udført, går man ud fra den ide at den underliggende metafor er en slags kontinuum. Det at lære ordforråd er blevet sammenlignet med et bjerg, som kan vokse både i højde og areal.

(Meara, 1999: 5)

Meara synes det er ansvarsløst at tale om et kontinuum i tilegnelsen af ordforråd, når det drejer sig om overgang fra passiv til aktiv, for det kræver at mindst en dimension forandres hele tiden, hvilket er svært at forestille sig i forbindelse med en model for ordforråd. Hvad ændres og hvordan fører disse ændringer til produktion af den påkrævede effekt? Meara synes ideen om, at ord skal forbi eller over eller igennem et antal diskrete stadier mere lovende end ideen om et kontinuum. Om der er stadier af passivitet eller stadier af aktivitet mellem de to poler, står ikke klart.

Folk, der beskæftiger sig med ordforråd, har endnu ikke udviklet ideen til mere end en grov metafor – med to undtagelser:

Wesche og Paribakht (1996) har arbejdet ud fra ideen om 'vocabulary knowledge scale' og lavet en skala med fem stadier af kendskab til ordforråd. Eleverne afgør selv, hvor godt de kender til et specielt ord:

- Jeg kan ikke komme i tanker om at have set det ord før
- Jeg har set ordet før, men jeg ved ikke, hvad det betyder
- Jeg har set ordet før og jeg tror det betyder _____
- Jeg kender ordet. Det betyder _____
- Jeg kan bruge ordet i en sætning. Fx _____

Wesche og Paribakht går ud fra, at et ord på stadium fem er bedre integreret end ord på stadium tre eller fire. Mearas kritik går ud på, at det er svært at afgøre, hvor stor fremgang der er mellem de forskellige stadier og at det er muligt for en indlærer at skrive en korrekt sætning, som inkluderer ord hun ikke kender betydningen af. Der er heller ingen beviser for, at et ord, der engang er nået fra det ene stadium til det andet, forbliver der. Derfor har han lavet en 'multi-stadie model', hvor et ord kan bevæge sig på tværs af de forskellige stadier, både frem og tilbage. Et ord kan altså gå fra stadie 0 til 4 eller fra stadie 4 til 3. En ting, der gør ideen om et kontinuum tvivlsom. Endvidere begrænser denne metode det antal ord, der kan testes hver gang.

Meara lavede en test sammen med Sanchez (Meara og Sanchez 1993), hvor de kun brugte stadier 2-5⁷ og lod eleverne klassificere et stort antal ord i disse fire kategorier med to ugers mellemrum. De fandt ud af, at mange af ordene ikke blev kategoriseret på samme måde første og sidste gang. Nogle af ordene faldt ud, andre delvis kendte kom op på et højere stadium. Sanchez brugte resultaterne til at vurdere og forudsige ændringer af ordforråd over længere tid.

Meara afslutter sin artikel med at understrege, at vi ikke har så meget brug for at forstå ord i detaljer, som at opnå en meget dybere forståelse af ordforråd. I hans øjne er helheden mere interessant end summen af delene.

Det ser ud til, at der hersker to hovedretninger indenfor forskning i tilegnelsen af ordforråd. Der er dem, som ser på tilegnelsesprocessen for enkelte ord og der er dem,

⁷ Lige mange stadier hos W&P på den ene side og Meara & Sanchez på den anden side. De førstnævnte bruger tallene 0-4 og de sidstnævnte 1-5 til at nummerere stadiene.

som mener at tilegnelsen af ordforrådet som helhed og dets stabilitet, er det mest interessante.

Begge dele er lige vigtige; i undervisningen tænker man i ord fra den ene dag til den anden, men på længere sigt er det væksten i ordforrådet der tæller. Derfor er det vigtigt, at både lærer og indlærer er bevidste om disse to aspekter af ordforrådsindlæringen (Meara, 1999).

Nation har ændret og udvidet sin liste over delaspekter i leksikalsk kompetence i bogen *'Learning Vocabulary in a Foreign Language'*.

Tabel 2.3.a *Hvad vil det sige at kunne et ord (Nation 2001:27)⁸*

| | | | |
|-----------|-----------------------------------|---|---|
| Form | talt | R | Hvordan lyder ordet? |
| | | P | Hvordan udtales ordet? |
| | skrevet | R | Hvordan ser ordet ud? |
| | | P | Hvordan skrives og staves det? |
| | dele af ordet | R | Hvilke dele af ordet er genkendelige? |
| | | P | Hvilke dele af ordet bruges til at udtrykke meningen? |
| Betydning | form og betydning | R | Hvilken betydning signaliserer denne ordform? |
| | | P | Hvilken ordform kan bruges til at udtrykke denne mening? |
| | koncept og referencer | R | Hvad er indbefattet i koncepten? |
| | | P | Til hvilke ting /ideer refererer koncepten? |
| | associationer | R | Hvilke andre ord minder det om? |
| | | P | Hvilke ord kunne vi bruge i stedet for? |
| Brug | grammatiske funktioner | R | I hvilke strukturer kan ordet bruges? |
| | | P | I hvilken struktur må vi bruge ordet? |
| | kollokationer | R | Hvilke ord eller typer af ord optræder sammen med dette ord? |
| | | P | Hvilke ord eller typer af ord må vi bruge sammen med dette ord? |
| | ... i brug registrering, frekvens | R | Hvor, hvornår, og hvor ofte kunne vi regne med at møde ordet? |
| | | P | Hvor, hvornår og hvor ofte kan vi bruge dette ord? |

Listen viser, at fuldstændig leksikalsk kompetence indeholder en lang række delelementer, som hvert for sig kan beherskes både receptivt og produktivt. For at kunne et ord skal man vide noget om ordets form, betydning og anvendelsesmuligheder. Med hensyn til ordets form skal man fx vide, hvordan ordet udtales og skrives og hvilke dele det er sammensat af. For at forstå ordets betydning skal man kunne knytte ordets form og

⁸ R: *receptiv kunnen*, P: *produktiv kunnen*

betydning sammen, kende ord der er beslægtede og ord der har samme eller lignende betydning. Endvidere skal man vide, hvornår ordet kan bruges og i hvilken sammenhæng.

Denne opdeling af komponenter hos Nation harmonerer godt med Aitchisons (1994) beskrivelse af L1-kompetence:

'form og betydning' svarer til *'labelling'*, hvor ord får betydning

'koncept og referencer' svarer til *'packaging'*, hvor ordets betydning afgrænses

'associationer' svarer til *'network building'*, hvor ordet sættes i relation til andre ord

Derudover kommer så alt, der vedrører anvendelsesaspektet, som den fuldt kompetente sprogbruger må beherske.

Ved at dele den leksikalske kompetence op i delaspekter kan man bedre analysere enkeltdelene, som utvivlsomt vil være en fordel både i undervisningen og ved at udvikle tests, der skal måle udvalgte dele af den leksikalske kompetence.

Wilkins (jf. s. 6), Nation, Aitchison og Henriksen har alle beskrevet leksikalsk kompetence i tre dimensioner og disse beskrivelser ligner hinanden meget. Wilkins var der 20 år før Aitchison, men hun er den eneste der går ud fra L1 i sin definition. Nation fremsætter sin analyse, som en del af større sammenhæng mens Henriksen indsnævrer og fordyber dimensionerne.

Henriksens tre dimensioner er:

- a) delvis til præcis forståelse,
viden om en leksikalsk enhed er færdigheden i at oversætte den leksikalske enhed til L1 at vælge den rigtige definition i en flervalgsprøve
at danne par af L1 og L2 ord
- b) dybde af viden,
kvaliteten af indlærerens ordforråd, en række forskere (heriblandt Ellis og Nation) har understreget, hvor kompleks ordforrådsviden er og de mange former for viden, der skal indgå i den fulde forståelse eller hvor forskellig betydning et ord kan have
- c) udvikling fra receptiv til produktiv kontrol,
færdigheder i overføring fra receptivt til produktivt brug

(Henriksen, 1999b: 303)

Det er et velkendt faktum, at kun en del af det receptive ordforråd nogen sinde bliver produktivt. Henriksen deler Melkas (1997) opfattelse, når hun ikke vil se på det receptive og det produktive, som to adskilte ordbøger eller to sæt ordforråd. Hun foretrækker at se på det som et kontinuum. Desuden er det uklart, hvorfor man skulle lave en linje mellem

det receptive og det produktive. Findes der et stadium, hvor man kan sige at viden ikke længere er receptiv men produktiv, eller hvor ligger skillelinjen for at receptiv viden konverteres til produktiv viden? Hun udpeger, at noget taler for, at der ikke er tale om et 'enten-eller', når der tales om reception og produktion. I forbindelse med viden kan nogle aspekter være blevet produktive mens andre bliver ved med at være receptive. Relationer mellem de to første dimensioner er i grunden kontinuerlig viden og dimension tre er et kontrol kontinuum, som beskriver adgangsniveauer eller brugsfærdigheder og som kan aktiveres gennem forskellige typer af receptive og produktive opgaver.

Når det tages i betragtning, hvordan relationer mellem de tre dimensioner udvikles, opstår der to spørgsmål:

- a) er dybdeforståelse en forudsætning for udvikling af præcis forståelse?
- b) Er præcis viden og dybdeforståelse en forudsætning for at beherskelsen af et ord bliver produktiv?

Henriksen arbejder ud fra Aitchisons beskrivelse af L1 ordforrådsindlæring: labelling, packaging og network building. At etablere en model for leksikalsk udvikling til vejledning i undersøgelser af ordforrådstilegnelse kræver mere præcis specifikation af de forskellige dimensioner af leksikalsk kompetence, indre relationer mellem dem og hvordan de harmonerer med de processer der skal til at lære og bruge et ord.

Meara vil ikke tale om et kontinuum i receptiv versus produktiv ordforrådsviden. Undersøgelser har vist, at et ord kan 'flyde' frem og tilbage mellem forskellige stadier (jfr. s. 17), når en indlærer bliver testet to gange i de samme ord med et kort mellemrum. Ordforrådet har måske den samme størrelse, men indholdet er ikke helt det samme.

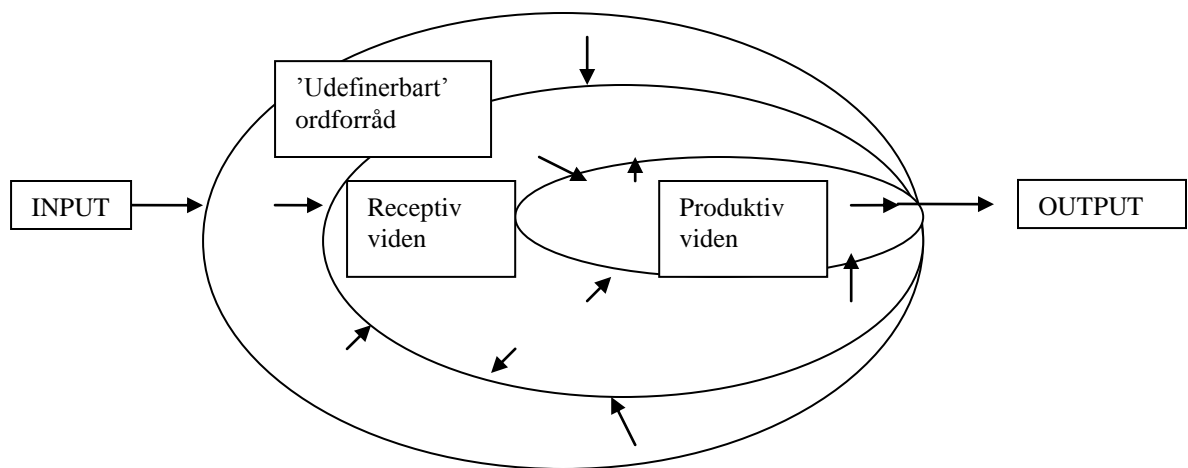
Henriksen tænker på et kontinuum (1999b), hvor receptiv viden bliver produktiv, at der ikke er nogen skillelinje, hvor man kan sige, at viden ikke længere er receptiv, men er blevet produktiv.

Det er en spændende mulighed, at forene begge ideer under en model, hvor man kan tænke på al input som en 'levende masse udefinerbart/uanalyseret ordforråd' i første omgang. Når et ord får en betydning, er der tale om receptiv viden og så produktiv viden, når ordet kan bruges i tale eller skrift (output). Hele tiden kan ordet 'flyde' mellem de tre stadier, som gennem en 'halvgennemtrængelig hinde', men langtidsprocessen er, at både den receptive og den produktive viden vokser og det 'udefinerbare' ordforråd bliver

mindre og mindre, forudsat at indlæreren arbejder med sit nye sprog. Outputtet kommer fra den produktive del af 'inputmassen'. Da det produktive ordforråd bliver ved med også at være receptivt, må den produktive viden være en del af den receptive viden, som også er en del af 'inputmassen'.

Meget forenklet kunne min illustration se sådan ud:

Figur 2.3.a Viser en forenklet ide om, hvordan ord kan 'flyde' frem og tilbage mellem forskellige stadier, samtidigt med at først receptiv viden og så produktiv viden bliver større.



Afslutningsvis kommer en definition af den kompetente sprogbruger samt et klassisk spørgsmål, som endnu ingen har svar på.

En kompetent sprogbruger:

- kender mange ord (kvantitet)
- ved meget om enkelte ord (kvalitet)
- kan bruge dem i kommunikation (kontrol)

En sprogbruger kender 20.000 ord aktivt i sit modersmål. Hvor mange og hvilke er centrale for en givet fremmed- eller andetsprogsindlærer? (Henriksen 1995: 13)

2.4 Ordforråd og læsning

Mange forskere har spekuleret i forholdet mellem læsning og ordforråd. Fx hvordan læsning forøger ordforrådet og hvor stor effekt viden om ordforrådet har på læsning.

Sætninger er læselige, hvis de indeholder højfrekvente ord og er korte frem for lange. Hvis teksten ikke indeholder for mange nye ord, kan de nye ord læres ud fra konteksten.

Nation og Coday mener, at der er mange faktorer, der spiller ind i afgørelsen om hvorvidt en tekst er let eller svær for en bestemt læser; fx om den er konkret eller abstrakt. Ved at udskifte 15% af højfrekvente ord i en tekst med lavfrekvente ord kan det føre til en signifikant formindsket læseforståelse. En afgørende ting er, hvor lang teksten er, hvor i teksten ordene bliver skiftet ud; og om det er i vigtige dele af teksten eller mindre vigtige. (Nation og Coday 1988:98)

Henriksen (1995a) er blandt andre (fx Nation 2001) af den mening at 2000 ord er et minimumskrav for oparbejdelse af tilstrækkelig læsefærdighed på målsproget, men nok til at mere end 80% af en almindelig tekst kan læses og forstås. Der er dog en lang vej fra at forstå 85% af en tekst (de mest hyppige ord) til de 95% som Gimbel (1995) nævner som et minimum for at kvaliteten af ens læsning bliver god nok.

Førfagligt ordforråd ligger mellem det mest grundlæggende ordforråd og det helt specifikke ordforråd, som er lang mindre hyppigt og som de fleste danske børn heller ikke behersker. Ifølge norske og danske undersøgelser er fagord og førfagord den del af skolebøgernes læsetekster, som elever af en anden etnisk baggrund har sværere ved at tilegne sig end de som har norsk/dansk som modersmål. Fagordene er svære fordi der er næsten ingen overlapning mellem to fag. Førfagord er ord, som erfarne lærere mener det er nødvendigt at forklare for alle elever. (Gimbel, 1995).

Her er der klart en forbindelse; førfagord og fagord udgør muligvis de 15 procentpoint af teksten, som dækker intervallet fra 80-95%. En del af eleverne taler sikkert flydende norsk/dansk og derfor er deres mangler i ordforrådet ikke synlige i det daglige. Jeg har lagt mærke til, at islandske unge, som har boet i udlandet i mange år, ofte har svært ved at følge med i forskellige fag i skolen. Selvom de taler flydende islandsk, rækker deres ordforråd ikke når læsebøgernes ordforråd bliver mere specifikt.

Nattinger påstår, at det er indholdsord, der volder problemer, funktionsord er forholdsvis få og forekommer ofte og læres derfor hurtigt. Det største problem er substantiver, verber, adjektiver og adverbier, som står for abstrakte koncepter, ord som ikke kan ignoreres, fordi de bærer nødvendige oplysninger for forståelse af teksten, men er samtidig lavfrekvente og svære at huske. (Nattinger, 1988: 64)

Gode elever kan gætte sig frem til en stor del af ukendte ord, alt op til 60 eller 80%, hvis tætheden af ukendte ord ikke er for stor.

Nation lægger vægt på, at det er vigtigt at lære strategier, som kan bruges ved indlæring af ordforråd. Han peger på tre store kategorier af ordforrådsstrategier, som indebærer strategier af forskellig kompleksitet:

1. Planlægning af ordforrådsindlæring – vælge, hvad der skal fokuseres på og hvornår
2. Kilder – finde oplysninger om ord, analysere dele af ordet, bruge konteksten, konsultere ordbøger / ordlister / andre elever / læreren / en indfødt
3. Processer – huske ordforråd og have det parat til brug; være opmærksom på, genkalde, lave, skabe.

(Nation, 2001: 218-222)

Haastrup (1994) brugte *tænke-højt* introspektion og retrospektion til at studere L2-indlæreres procedurer til at finde frem til eller gætte sig til betydning af et ord, hvilke hjælpemidler de bruger, hvordan de kombinerer forskellig viden fra forskellige kilder. Haastrup klassificerede den viden eleverne benyttede sig af ved at bruge Cartons tre kategorier:

1. intersproglige: cues (ledeord) baseret på L1, låneord i L1 eller viden om andre sprog
2. intrasproglige: cues (nøgleord) baseret på viden om engelsk
3. kontekstuelle: cues baseret på teksten eller informanternes viden om verden

Laufer og Nation siger, at et vigtigt mål med undervisning i ordforråd er at få indlærerne til at bruge deres ordforrådsviden kommunikativt. Laufer lavede en undersøgelse i 1991, som viste at elevernes brug af ordforråd i skriftlige opgaver voksede i løbet af en periode på 14 uger og 28 uger. Det er muligt at måle vækst i ordforråd over tid, men det er ikke til at afgøre om vækst i ordforråd er et resultat af at lære et nyt ordforråd, aktivering af allerede kendt ordforråd eller det at eleven lægger bedre mærke til ordforråd, som et resultat af at være blevet bedre til at skrive. Derfor er det nødvendigt at se, om vækst i skriftligt ordforråd er relateret til en mere direkte måling af ordforrådets omfang. (Laufer og Nation, 1995)

Der er behov for forskning, der målrettet undersøger størrelsen af ordforrådet hos forskellige grupper og som forsøger at fastslå mere præcist, hvor stort et ordforråd der er nødvendigt for at en sprogbruger bliver i stand til at løse bestemte opgaver. (Henriksen, 1995a:15)

2.5 Hypotesemodellen

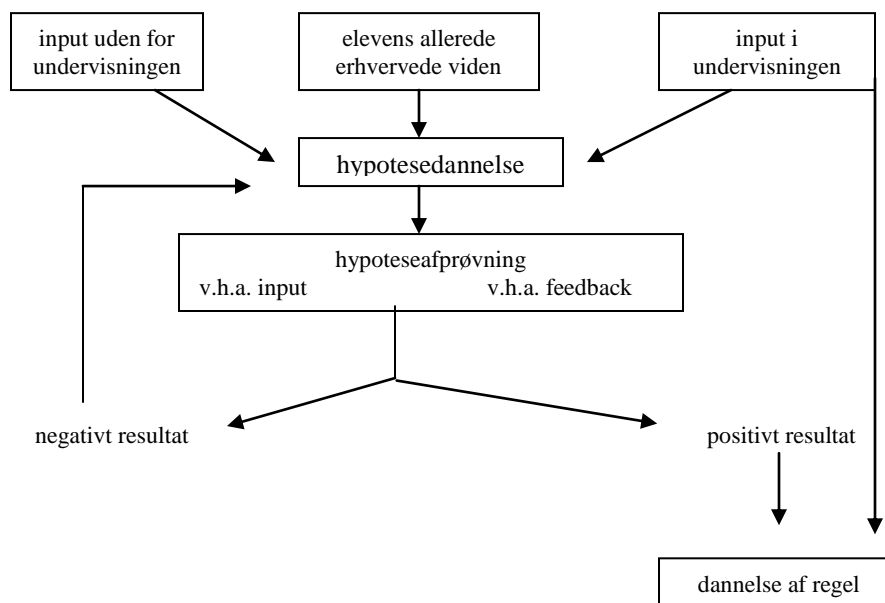
I fremmedsprogsundervisningen er det ikke nok, at eleverne får input. Der sker ikke ret meget medmindre eleverne arbejder med inputtet indtil det bliver en del af deres sprog. Holmen (1988) er blandt de forskere, der mener at hypotesedannelse og hypoteseafprøvning er den vigtigste proces i tilegnelsen af målsproget. Indlæreren laver, mere eller mindre bevidst, hypoteser ud fra den viden hun har om fremmedsproget, men det er ikke sikkert hun kan formulere sine hypoteser. Sprog, som i første omgang læres uanalyseret, bliver også genstand for hypoteseafprøvning.

Modellen viser tre forskellige kilder til hypotesedannelse. Hypoteseafprøvning er, overordnet set, en teoristyret deduktiv proces hvor indlæreren afprøver sin teori om et system med det input hun udsættes for. Hypoteseafprøvning kan gå ad flere veje:

1. Receptivt: indlæreren sammenholder sine hypoteser med input
2. Produktivt - feedbackbaseret: indlæreren justerer sine hypoteser efter den feedback hun får på sin sprogbrug
3. Metalingvistisk: indlæreren søger råd hos lærer, grammatik og andet
4. Interaktionelt: indlæreren søger hjælp hos fx samtalepartner

(Færch og Kasper, 1983)

Figur 2.5.a Indlæringsmodel Færch et al. (1984:200)⁹ Modellen viser, hvordan input og viden bliver genstand for hypotesedannelse, og relationer overfor udfaldet af hypoteseafprøvningen.



⁹ Holmen (1988: 83) oversatte modellen fra Færch, Haastrup og Philipson (1984): Learner Language and Language Learner.

Indlæreren er uafbrudt i gang med afprøvning og revidering af hypoteser mens sprogtilægnelsen varer ved og ophører først, når hun ikke kan eller vil stræbe efter større færdigheder eller kompetence. Tit sker det, at indlæreren overforbruger et sprogligt fænomen, når hun prøver at bruge en hypotese eller en regel i en ny sammenhæng. En vigtig ting er, at hvis indlæreren ikke er parat til at fokusere på et sprogligt fænomen, så tager hun ikke notits af rettelsen. Elevens egen indsats er drivkraften i sprogtilægnelsen, hvilket indebærer, at sprog læres, når man har brug for det og gør fremskridt, når man tør afprøve sit sprog i et trygt miljø. (Lund, 1997: 118)

Output er tegn på at fremmedsprogstilegnelse har fundet sted. Mængden og kvaliteten af outputtet er så en målestok på, i hvor høj grad eller på hvilket stadium denne tilegnelse er. Fluency og accuracy er to forskellige dimensioner af sprogbrug. Sprogbrug, hvad enten den foregår som led i undervisning eller i andre sammenhænge, fører til fluency, men ikke nødvendigvis til accuracy.

Det sproglige output har tre funktioner, der kan tjene til at fremme indlærerens korrekthed.

- På det første stadium oplever hun et gab mellem det hun gerne vil sige og det hun kan sige.
- På det andet stadium, når hun til hypoteseafprøvning, hvor hun prøver sine egne hypoteser angående forståelse af det hun hører/læser eller en lingvistisk formulering af det hun vil sige.
- På det tredje stadium har outputtet en metalingvistisk funktion, hvor indlæreren reflekterer over sin sprogbrug med det formål at kunne kontrollere det på en bevidst måde.

(Swain, 1995)

2.6 Kognitiv teori

McLaughlin siger, at det at lære et fremmed sprog kræver tilegnelse af en indviklet kognitiv viden – men også tilegnelse af en indviklet lingvistisk viden. Den specifikke fokus indenfor den kognitive teori er indlæringsprocessen – det at tilegne sig en indviklet kognitiv viden.

At lære er en kognitiv proces, som involverer internal repræsentation, som er baseret på sprogsystemet, og inkluderer procedurer til at vælge *ordforråd*, *grammatiske regler* og *pragmatiske konventioner* i sprogbrug. Ved performance, dvs. når indlæreren prøver at udtrykke sig, sker der en ustandselig restrukturering, hvor indlæreren forenkler, udvider

og opnår forøget kontrol over den interne repræsentation – automatisering og restrukturering er det centrale i den kognitive teori.

Der foregår aktivering af visse steder i hukommelsen hver gang passende input er til stede. Det fører så igen til bearbejdning af en struktur eller viden. Automatisering sker når formulering af en eller flere strukturer er blevet indarbejdet så godt, at det ikke kræver nogen energi at fremkalde den bestemte struktur på hvilket som helst tidspunkt, hvor den passer ind i udtryksituationen.

Når automatiske processer er blevet etablerede på et stadium i udviklingen, hvor indviklede oplysninger bliver bearbejdet, kan kontrollerede processer flyttes fra det nemme og enkle til det mere svære og indviklede.¹⁰

Kontrolleret processing kræver mere tid fordi kun en sekvens kan aktiveres ad gangen, men er derimod nemme i brug. Kontrolleret proces styrer strømmen af oplysninger fra korttidshukommelsen til langtidshukommelsen, og viden læres og bliver automatisk.

Kapaciteten af arbejdshukommelsen er begrænset hos indlærerne og derfor kan de kun tænke på nogle få processer ad gangen.

Restrukturering og omorganisering af 'gammel' viden åbner muligheder for at træne og indøve nye færdigheder. Fremmedsprogstilegnelse er ikke lineær og voksende – men tilbagegang og tab af former, som så ud til at være mestret. Nye former overdækker de allerede lærte, og det kræver omstrukturering. (McLaughlin, 1987: 134-7)

Ellis er af den mening, at sproglig transfer mere og mere skal forstås som en kognitiv proces; dvs. fremmedsprogsindlærere bruger modersmålet strategisk ved indlæring af fremmedsproget og i processen ved at forstå og producere beskeder på fremmedsproget. Andre aspekter i fremmedsprogstilegnelsen kan ses som en mental proces, som involverer brug af strategier og forklarer, hvordan et system af fremmedsprogsviden bliver udviklet og brugt i kommunikation.

Kognitiv sprogstilegnelsesteori ser lingvistisk viden ikke som noget forskelligt fra andre typer af viden og ser strategier, som er ansvarlige for dens udvikling som generelle af natur, relateret til og involveret i anden slags indlæring.

¹⁰ Efter Shiffrin and Schneider (1977) i McLaughlin (1987)

Det er muligt at tale om at klassificere teorier i L2 tilegnelse som enten 'kognitive' eller 'lingvistiske', hvor begge perspektiver er inddraget.

Her tales der om kognitive sider af fremmedsprogsindlæringen på den ene side og på den anden om fremmedsprogskommunikation. Der skelnes mellem teorier, som forklarer, hvordan indlærere konstruerer deres mentale repræsentationer af fremmedsproget (hvordan viden om regler og ting som udgør fremmedsproget er udviklet) og teorier, som forklarer, hvordan indlærere bruger deres viden i aktuelt sprogbrug (hvordan fremmedsprogsforståelse og -produktion hænger sammen).

Denne skelnen mellem 'tilegnelse' og 'procedural evne' er teoretisk klar, men der opstår problemer i teorier, som ser 'tilegnelse' som sidestillet med 'brug'. En skole af kognitiv psykologi mener at sproglig viden bedst kan ses, som et netværk, der involverer et indviklet sæt af forbindelser. (Ellis, 1994:347-8)

Med psykolingvistikken forsøger man at komme tæt på de mentale processer, der foregår i sprogbrugerens hoved, når vi opfatter ord, husker eller glemmer ord og når vi leder efter et ord vi ved vi har, men ikke kan komme i tanke om. 'Det mentale leksikon', 'ordforrådslageret', det 'semantiske netværk', hjælper os til at fatte de ikke-observerbare mentale processer. (Henriksen, 1999a:73).

De fleste forskere mener, at ordforrådet i det mentale leksikon er lagret i forskellige betydningsnetværk. Sådanne relationelle kroge er vigtige for en udbygget dybdeforståelse af et ords betydning og kan støtte optagelsen af nye ord i langtidshukommelsen (Henriksen, 1995b:).

2.7 Eksplicit viden - implicit viden

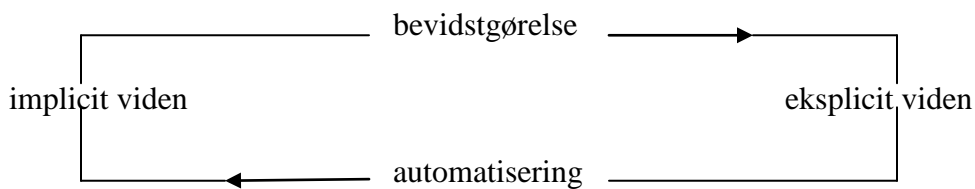
De fleste forskere er enige om, at sprogindlærere besidder to slags viden, eksplicit og implicit. Den første kan indlæreren formulere, men ikke nødvendigvis anvende i praksis mens den anden kan anvendes, men ikke formuleres.

Den amerikanske sprogforsker Stephen Krashen skelner mellem sprogtilegnelse og sprogindlæring. For Krashen er de implicite indlæringsprocesser centrale, idet han mener at sprogbrugereren trækker på sin implicite viden i den sproglige kommunikation. Den eksplicite viden, som er et resultat af de eksplicite indlæringsprocesser, er derimod kun

brugbar som 'monitor' (dvs. til at rette og korrigere det sproglige produkt), når og hvis kommunikationssituationen levner tid til sproglig refleksion. (Henriksen, 1998)

Den mere traditionelle undervisning drejede sig for det meste om den eksplicite viden, men den kommunikative undervisning lægger op til en stadig udbygning af elevernes implicite viden. Bevidstgørelsen er ikke længere et mål i sig selv, men en integreret del af indlæringsprocessen og kun formålstjenlig, hvis denne bringes videre. Målet er, at den eksplicite viden automatiseres til rygmarvsviden gennem brug af sproget (Holmen, 1988).

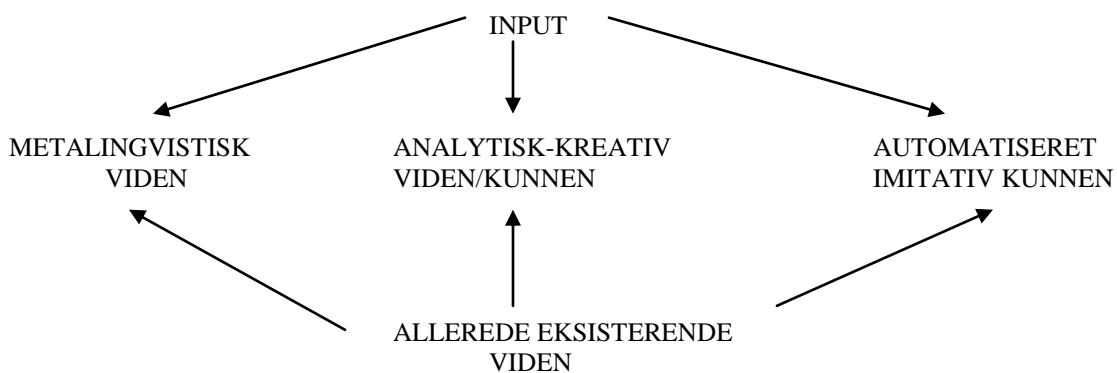
Figur2.7.a I denne model er de to processer føjet til hypotesedannelsen og -afprøvningen som elementer i indlæringen. (Holmen, 1988)



Ellis taler om to typer af implicit viden, formular viden og regelbaseret viden. Den første består af færdiglavede helheder. Den sidste består af generaliserede og abstrakte strukturer, som er blevet internaliseret. I begge tilfælde er denne viden intuitiv og derfor for det meste skjult, indlærerne er ikke bevidste om, hvad de kan. Det kan kun konstateres i en aktuel performance. (R. Ellis,1994:355)

Dele af den eksplicite sprogbeherskelse bliver efterhånden automatiseret. Hvor meget af indlærerens sprog, der er automatiseret, vil blandt andet kunne aflæses i 'fluency'. Jo mere flydende indlæreren taler, desto flere automatiserede 'øer' har hun at hvile på. Automatisering er således et udtryk for, hvor let adgang indlæreren har til et lagret område og hvor ureflektet produktionsprocessen kan foregå. Følgende model afspejler de forskellige vidensformer og deres samspil: (Lund, 1997)

Figur 2.7.b De forskellige vidensformer og deres samspil. (Lund, 1997:119)



Henriksen (1998) siger, at mange forskere har argumenteret for et 'både-og' snarere end et 'enten-eller', dvs. der er samtidige kognitive processer, der begge spiller en central rolle for sprogindlæringen. N. Ellis (1994) mener, at hvis der kan afklares, hvilke sider af sprogfærdigheden, der påvirkes af hvilke processer, og hvilke faser i sprogindlæringsprocessen der mest karakteriseres ved den ene og /eller den anden proces, vil det være muligt at give mere præcise forslag til, hvordan indlæringsprocessen kan påvirkes mest hensigtsmæssigt i en undervisningssituation. Ellis mener endvidere, at implicit proces er en automatisk proces, hvor forhold som hyppighed i input, repetition og hukommelse for sproglige enheder spiller en afgørende rolle. Den eksplicite proces er kognitivt langt mere kompleks og påvirkes i langt højere grad af forhold så som arbejds måde, fx bearbejdningsdybde, og sprogindlærerens evne til at analysere og danne synteser. Dertil kommer, at rigtig stavemåde og udtale af et ord indlæres ganske ubevidst gennem automatiske, implicite indlæringsprocesser, som er påvirket af forhold som

- hvor hyppigt indlæreren møder ordet
- hvor længe det er siden ordet har været i inputtet eller er blevet brugt
- hvor regelmæssigt det forekommer i inputtet (dvs. forhold som tidsinterval mellem repetition af samme ord)

At lære sprog handler naturligvis ikke blot om at lære et meningstomt sprogligt udtryks element, men om at kunne koble udtryks- og indholdssiden sammen til et brugbart sprogligt tegn.

Henriksen kommer frem med den hypotese, at den skelnen mellem de implicitte og eksplicitte indlæringsprocesser, som blev nævnt for ordforrådsindlæringen, måske også gælder for flerordsfraserne (Henriksen, 1998:27-8).

2.8 Leksikaliserede helheder

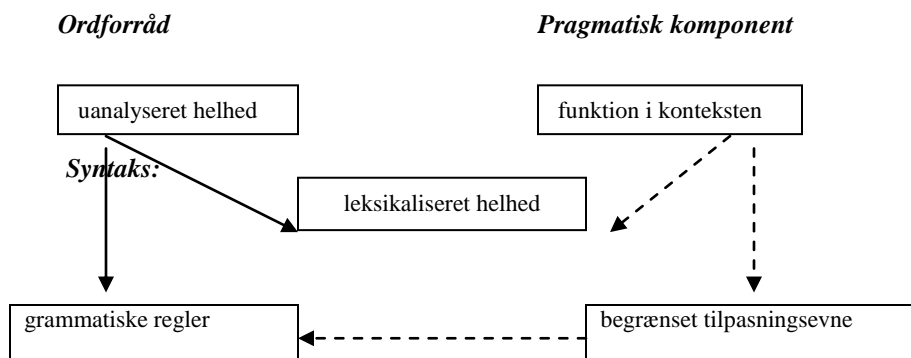
Henriksen går ud fra, at de fleste sprogfolk er enige om at sprogtilegnelse, både på modersmål og et hvilket som helst nyt sprog, er en kognitiv proces, hvor sprogindlæreren på basis af inputtet og gennem hypoteseafprøvning (jf. s. 22) og feedback gradvis oparbejder en forståelse for sprogets systematik. På basis af sit eget sprogsystem kan sprogindlæreren så forhåbentligt gå ud og bruge sine sproglige byggeklodser kreativt og hensigtsmæssigt i en række nye situationer.

Mange forskere har imidlertid peget på, at en forbløffende stor del af de ytringer, vi alle bruger, slet ikke er kreative og nye, men tilsyneladende hentes frem og genbruges fra et lager af fasttømrede, leksikaliserede udtryk. Den leksikografiske forskning, hvor store mængder af sproglige data er blevet indsamlet i store databanker, har bl.a. kunnet afsløre, hvor hyppigt de samme sproglige udtryk og konstruktioner forekommer i det indsamlede materiale. Nogle har endda foreslået, at mere end 60% af det, vi siger, består af sådanne faste holofraser eller uanalyserede helheder (Henriksen, 1998:25).

Forskerne Nattinger og DeCarrico (1992) mener, at *den leksikaliserede helhed* (the lexical phrase) vil blive en lovende nyhed i sprogundervisningen. De beskriver leksikaliserede fraser som *helheder* (chunks) af sproget i varierende længde.

Computere gør det muligt at sortere leksikaliserede helheder ud af en tekst eller et korpus i en håndevending. Tekstundersøgelser og undersøgelser i sprogtilegnelse kan forenes på den måde, da den leksikaliserede helhed er en ideel helhed til brug i sprogundervisningen.

Figur 2.8.a Samspil mellem pragmatisk komponent, ordforråd og syntaks.
(Nattinger og DeCarrico, 1992:16)



Modellen belyser, hvordan leksikaliserede fraser først læres som uanalyserede leksikaliserede helheder sammen med forbundne funktioner i konteksten. Lidt efter lidt tilegnes andre helheder, som har lignende syntaktiske strukturer, indlæreren tager strukturen ud af sammenhæng i konteksten, analyserer den og generaliserer så der dannes en regulær syntaktisk regel. Den nederste kasse til højre kompletterer repræsentationen af relationer mellem pragmatisk komponenter, ordforrådet og syntaksen. Den sorte linje beror på grammatisk kompetence, men den stiplede linje pragmatisk kompetence.

Leksikaliserede fraser er præfabrikerede leksikalsk-grammatiske enheder, som får stilling et sted mellem de traditionelle poler af ordbog og syntaks, sammensat form/funktion, som forekommer meget oftere og har mere idiomagtigt rettet mening end det sprog, som bliver sat sammen til enhver tid. Netop på grund af den stilling mellem ordforråd og grammatik har denne del af sproget været forsømt indenfor lingvistisk analyse. Leksikaliserede helheder er præfabrikeret tale og evnen til at kunne bruge dem hjælper os til at tale flydende.

Leksikaliserede helheder når ud over andre ting end *idiomer* og *klichéer* (idioms and clichés), som dog hidtil har fået den største opmærksomhed inden for lingvistikken. Idiom er et kompleks af ord med frosen syntaks, som udtrykker noget andet end summen af det ordene står for. Klicher ligner idiom til dels, idet de også er forholdsvis frosne, men meningen er som oftest umiddelbart hvad ordene siger.

Helheder kan være *non-canonical* eller *canonical*. Den første kategori indeholder faste vendinger, som ikke adlyder sprogets typiske mønstre eller strukturer. Faste

vendinger i den sidste kategori holder sig derimod til den typiske sætningsstruktur. Leksikaliserede helheder bruges til at vise visse funktioner som fx i en samtale.

Nattinger & DeCarrico har analyseret leksikaliserede helheder i engelsk både efter form og funktion og derigennem vist, hvor stor en del disse helheder udgør af det dagligt anvendte sprog. De gør meget ud af at fremmedsproglærere burde gå ud fra læren om leksikaliserede helheder i deres undervisning. Leksikaliserede helheder er ikke noget, der kun findes i engelsk, de er en vigtig del af alle sprog. Og som eksempel giver Nattinger & DeCarrico et større udvalg af leksikaliserede fraser på kinesisk, russisk og spansk i et appendiks. Leksikaliserede fraser er derfor også en vigtig del i det danske sprog såvel som det islandske (Nattinger & DeCarrico, 1992: 1-45).

Leksikalsk approach refererer ifølge Richards og Rodgers til den opfattelse, at byggematerialet i sprogindlæringen og kommunikationen ikke er grammatik, funktioner, begreb, en form for planlægning eller undervisning, men leksis, dvs. ord og ordforbindelser. Leksikalsk approach reflekterer den antagelse, at ordforråd er det centrale i sprogsystemet, andetsprogsindlæring og -sprogbrug, og specielt på flerords leksikalske helheder, som læres og anvendes som enkeltord. Lingvistisk teori har også erkendt en mere central rolle for ordforråd i lingvistisk deskription. Før var fokus primært på syntaksen, nu lægges der mere vægt på ordforrådet og hvordan ordforrådet er formateret, kodet og organiseret. (Richards og Rodgers, 2001:132-3)

Man har lagt vægt på den rolle leksikaliserede helheder har både i modersmåls og fremmedsprogsundersøgelser. Udvikling i computerstøttede sprogundersøgelser (corpus linguistics) har ført til at man i skolen/undervisningen kan have adgang til en enormt stor database – det gælder nok ikke alle sprog endnu, men muligheden er der. Blandt andet har man fokuseret på kollokationer af leksikaliserede helheder og flerordshelheder. Et stort antal leksikalsk baserede tekster og computerressourcer er nu tilgængelige til hjælp ved organisering og undervisning i leksikonet. Leksikalsk approach i sprogundervisning prøver at udvikle læseplaner og sprogundervisning grundet på det sprogsyn, at ordforrådet spiller den centrale rolle. Nation påstår at *viden om sprog er viden om kollokationer*

- grammatiske regler kan opdages gradvist i sproget, men viden om sprog og sprogbrug hænger sammen med opbevaring af leksikaliserede helheder i langtidshukommelsen og erfaring af, hvor sandsynligt det

er at visse helheder forekommer med andre visse helheder uden at behøve at referere til underliggende regler. Viden om kollokationer er essensen af sproglig viden

- al flydende og passende sprogbrug kræver viden om kollokationer den bedste måde at forklare, hvordan sprogbrugere producerer modersmålsnende sætninger og bruger sproget flydende er, at de foruden at kende sprogets regler, opbevarer de hundredetusinder af præfabrikerede fraser i deres hukommelse. Derfor er det sandsynligt, at hvert ord i sproget lagres mange gange – en gang som et enkelt ord og mange gange i leksikaliserede helheder.
- mange ord er brugt i et begrænset sæt af kollokationer – og at kende dem er en del af det, der indebærer at kende ordene. I nogle tilfælde er kollokationer så idiomatiske, at de kun kunne lagres som huskede helheder. I andre er der almene regler for kollokationer.

(Nation, 2001: 318)

Leksikaliserede helheder har den fordel, at det tager kortere tid at fremkalde dem end summen af de enkelte ord, det betyder hurtighed. Ulempen ved helhederne er lagringen. Der er langt flere helheder end de komponenter, som de bliver sammensat af. Det betyder en stor mængde enheder, som skal lagres og som igen kan betyde, at det bliver svært at fremkalde dem. Hvis de bliver lagret (i langtidshukommelsen) som uanalyserede enheder, består vanskeligheden i, at det så bliver svært at danne nye helheder, for enhederne til at lave nye helheder ikke står til rådighed.

Nation formoder, at det i øjeblikket ser ud til at højfrekvente helheder (fx un+able) er opbevaret som en enhed mens de lavfrekvente (fx un+ambigu+ous+ness) er rekreeret efter reglerne hver gang de bruges. Når helheder bliver større, bliver frekvensen lavere; digte, sange, taler bliver lagret på den måde.

Hvis det bliver accepteret, at 'viden om sprog er viden om kollokationer', betyder det, at af alle sekvenser med kollokationer, både regulære og idiomatiske, er det vigtigst at lære de højfrekvente. Det meste af indlæringen vil finde sted gennem betydningsfokuseret receptivt og produktivt sprogbrug. Og for ordforrådsindlæringens vedkommende mangler der forskning angående:

- hvad der er højfrekvente kollokationer
- hvad der er uforudsigelige højfrekvente kollokationer
- fælles strukturer i kollokationer
- udarbejdning af ordbøger til at hjælpe med lavfrekvente kollokationer

(Nation, 2001:328)

I mange af de faste fraser er der ingen eller måske få variationsmuligheder. Udtrykket er ofte fasttømret og må derfor rent artikulatorisk og genkendelsesmæssigt tilegnes som en sproglig enhed, der mere eller mindre kan læres udenad uden større analytisk bearbejdning ved hjælp af implicite indlæringsprocesser. (Nation, 2001: 318-28)

Angående den centrale kobling mellem udtryk og indhold, dvs. viden om ytringens betydning og brugsfunktion (indholdsmæssige eller pragmatiske betydning) skal sprogbrugeren derimod danne og afprøve hypoteser vedrørende dette. Ofte er der tale om store forskelle og betydning og funktion fra sprog til sprog. (Henriksen, 1998)

Lund deler den holdning med Ellis, at uanalyserede/uautomatiserede helheder gradvis bliver genstand for analyse og at de udgør en væsentlig funktion i opbygningen af den analytisk-kreative viden. Krashen derimod er overbevist om, at præfabrikerede helheder og den analytisk-kreative viden er uafhængige størrelser og at der ikke kan finde overførsel sted mellem de to former for viden. (Lund, 1997:115-16)

Ifølge Henriksen er en bestemt sproglig sammenstilling blevet accepteret i sproget som en fast enhed, ofte med en helt fast betydning eller brugsfunktion, når et udtryk er blevet leksikaliseret. Udtrykket er altså blevet institutionaliseret i det pågældende sprogsamfund. Hun henviser til Nattinger og DeCarrico, når hun nævner, at der findes mange forskellige slags faste enheder i sproget lige fra de faste sammenstillinger på et par ord til lange idiomatiske udtryk. Her giver hun overblik over nogle af typerne.

- Nogle af de mindste enheder er kollokationerne, dvs. faste ordsammenstillinger mellem fx navneord og udsagnsord eller navneord og tillægsord (*kaste med bold, skyde med pil, sur fløde, harsk smør*).
- Mindre sætningsdele med en helt fast struktur, fx udtrykkene *i det store og hele, for så vidt, lad mig nu se*. Mange af disse ord har en adverbial funktion, idet de kommenterer det sagte eller er med til at skabe struktur i en samtale eller en tekst.
- En anden type er de mange faste udtryk, som har en helt bestemt pragmatisk funktion ved fx at signalere en bestemt sproglig hensigt fra den talendes eller skrivendes side. Dette gælder i høj grad for de mange forskellige sproghandlingstyper, der oftest realiseres på helt bestemte, ritualiserede måder; *nu kan det være nok, tak for mad, det er min omgang*. Her er oftest tale om sætningsrammer, hvor der er mulighed for en vis sproglig variation; *__ ønsker __ __ tillykke med __; må __ bede om __; vær så venlig at __; velkommen til __; må __ præsentere __*.

(Henriksen, 1998:25)

Pawley og Syder (1983) fremsatte den teori, at det netop er sådanne faste udtryk, der er hemmeligheden bag den indfødte sprogbrugers evne til at udtrykke sig både flydende og idiomatisk.

Det er nødvendigt for sprogindlæreren at oparbejde et repertoire af sådanne automatiserede helheder for at kunne udtrykke sig flydende. Findes sætningsrammen allerede i det mentale lager, kan sprogbrugeren koncentrere sig om at vælge den rigtige bøjningsendelse på verbet og finde det helt rigtige ord og den korrekte numerusform for det direkte objekt. Herved øges ikke blot elevens fluency, men også dennes mulighed for at fokusere på den sproglige præcision og korrekthed.

2.9 Receptiv viden - produktiv viden

Nattinger mener, at forståelse af ordforråd bygger på strategier, som giver indlæreren lov til at forstå ord og opbevare dem, giver denne lov til at huske mens produktion beror på strategier, som aktiverer ens lager ved at genkalde disse ord fra hukommelsen og bruge dem i passende situationer. Et voksende antal forskere tror på, at forståelse bør komme før produktion i sprogundervisningen.

Nattinger lægger vægt på, at det ikke er nødvendigt at kende hvert enkelt ords betydning for at forstå tekstens mening. At forståelse af teksten er den største hjælp til at forstå de vanskelige ord. Indlærerne har heller ikke brug for at vide alle betydninger af et bestemt ord, de kan være tilfreds med at kende den mest generelle mening (indfødte kender ikke altid alle betydninger). Og så efter at have mødt et ord i dets mange kontekster, kan man nærme sig den fulde forståelse af ordets betydning. Til sidst skal indlæreren ikke kigge efter alle ord i ordbogen, men bruge forskellige teknikker for at opdage betydning – ordbogen bør være den sidste udvej (Nattinger 1988:63).

Nation siger, at grunden til at receptiv brug er større end produktiv, kan være grunden til forskellig størrelse af receptivt og produktivt ordforråd og kan gøre det svært at måle hele ordforrådet. Nogle mener, at receptivt ordforråd og produktivt ordforråd skal læres på forskellig måde, dvs. undervisningen skal foregå forskelligt. (Nation, 2001:29)

Nation gør rede for Laufers undersøgelse fra 1998, hvor hun testede receptiv og produktiv størrelse af ordforråd i engelsk hos ikke-modersmålstalende¹¹, på den ene side andetsprogsindlærere (ESL) og på den anden side fremmedsprogsindlærere (EFL).

Hendes resultater er :

- receptivt ordforråd er større end produktivt
- forholdet mellem receptivt og produktivt ordforråd er ikke konstant
- når ordforrådet bliver større, bliver den receptive andel forholdsvis højere (vokser mere)
- en stor del af højfrekvente ord kendes både receptivt og produktivt
- vækst i ordforråd med direkte målinger kan ikke nødvendigvis ses i vækst af ordforråd i brug (fx i andel af lavfrekvente ord, der bliver brugt i en skriftlig øvelse)
- receptivt og produktivt ordforråd udvikles på forskellig måde

Receptivt og produktivt ordforråd udvikles på forskellig måde hos de to grupper.

Laufer foreslår, at vækst i ordforråd er forskellig hos ESL¹²-indlærere og EFL-indlærere. Hun mener, at hos middelgode og gode EFL-indlærere ligger den receptive og den produktive størrelse af ordforrådet forholdsvis tættere på hinanden end den gør hos ESL-indlærere. Grunden kan være forskellige indlæringsmetoder. EFL-indlærere får den formfokuserede instruktion (direkte undervisning, direkte indlæring), som er så typisk for mange EFL- kurser og det kan være årsagen til den lille forskel i receptiv - produktiv score. ESL-indlærere har et stort passivt ordforråd, som følge af et langvarigt (omfangsrigt) input. Denne passive viden ser ikke ud til at kunne overføres til aktivt brug. (Laufer, 1988 efter Nation, 2001:370)

2.10 Sammenfatning

Her er der gjort et forsøg på at give oversigt over de ting, der er aktuelle når der tales om indlæring af fremmedsprog.

¹¹ Hun brugte tre forskellige prøver for at måle ordforråd hos ESL og EFL:
Passiv test – niveau tester (Nation 1983, 1990): *passiv/receptiv*
Aktiv test – Productive Levels Test (Laufer og Nation 1995): *kontrolleret aktiv*
LFP Lexical Frequency Profile: *fri aktiv*

¹² **ESL**: English as a second language. **EFL**: English as a foreign language.

I årenes løb har mange spekuleret over

- hvordan et individ tilegner sig et nyt sprog
- hvilke ting er vigtige / mindre vigtige
- hvilke mekanismer der aktiveres i hjernen under indlæringen
- hvilke processer er fælles for 'alle' (fleste)
- hvilke metoder, der er brugbare / velegnede til tilegnelse / undervisning af et andet- eller fremmedsprog
- hvad er det mindste, man skal lære for at kunne overleve i et samfund, hvor fremmed- eller andetsproget bruges
- hensigt og motivation med indlæringen
- hvor ligger grænsen for receptivt / produktivt sprogbrug?
- hvor ligger grænsen for eksplicit / implicit sprogbrug / færdigheder – kan denne grænse overskrides?

Forskning er godt i gang på mange områder, en del af resultaterne finder efterhånden vej til undervisningslokalerne. Ny teknik gør det muligt at skanne tekster / ordforråd meget hurtigt, som igen må kunne bruges til at fremme indlæring af et 'hensigtsmæssigt' ordforråd

3 Anvendte metoder til måling af produktivt ordforråd

3.1 Metoder

I dette afsnit vil jeg gøre rede for de forskellige metoder forskere har benyttet sig af for at måle ordforråd i skriftlig og mundtlig sprogproduktion. Disse metoder er især blevet brugt i forskning angående fremmedsprogstilegnelse (L2), men også til måling af ordforråd i modersmål (L1). Her bliver der også gjort rede for forskernes egne meninger om metoderne, hvad de kan bruges til og hvilke forbehold der skal tages.

3.2 Leksikalsk tæthed (LT)

Leksikalsk tæthed er en måling af, hvor stor andel af teksten består af indholdsord (som oftest er defineret som substantiver, handlingsverber, adjektiver og adverbialer, som er afledt af adjektiver). Det er almindeligt at opgive denne måling i procent af ordeksemplarer. Formlen er

$$LT = \frac{\text{antal indholdsord i teksten}}{\text{antal ord} \times 100}$$

Denne metode kræver subjektiv vurdering af hvilke ord, der skal tælles, som indholdsord og kan derfor ikke være stabil. Desuden spiller tekstens syntaks, kohæsive faktorer og komposition en rolle for resultatet af målingen. (Laufer og Nation 1995).

3.3 Leksikalsk originalitet (LO)

Leksikal originalitet er en måling, der bruges, når tekster fra en gruppe personer skal sammenlignes. Den angiver, hvor stor procent af ordene kun bruges af en person i gruppen (dvs. ordene er originale).

$$LO = \frac{\text{antal originale ord i teksten}}{\text{antal ord} \times 100}$$

Her er det indlærernes performance i forhold til gruppen der tæller, hvis gruppen ændres, ændres også grundlaget for målingen (Laufer og Nation 1995)

3.4 Leksikalsk raffinement (LR)

Leksikalsk raffinement er en måling af andelen af raffinerede eller sofistikerede ordksemplarer. De raffinerede ord må da defineres ud fra formålet med testen og personernes niveau og alder. Testen egner sig kun til at sammenligne personer indenfor en speciel gruppe, fx en skoleklasse.

$$LR = \frac{\text{antal raffinerede ord i teksten}}{\text{antal ord} \times 100}$$

Da sammenligningen beror på alder og kompetenceniveau, må kategorier defineres eller rammer sættes hver gang metoden bliver brugt.

3.5 Leksikalsk variation (LV). Type/Token Ratio (T/T)

Denne metode måler forholdet mellem forskelligheden¹³ af ord i en tekst og hvor mange ord i alt teksten indeholder.

$$LV = \frac{\text{Typer af ord i teksten}}{\text{Antal ord i teksten}}$$

T/T måling bruges mest for at sammenligne tekster, og målingen bliver mest pålidelig når teksterne er af en vis længde.

Høj LV indebærer, at der bruges mange forskellige ord og at der er småt med gentagelser i teksten. Det er nemmere at undgå gentagelser i en kort tekst end en lang en, så denne metode er ikke valid for sammenligning af tekster af forskellig længde.

Når teksten når et vist antal ord, vil T/T stabilisere sig. Hvilket niveau det sker på, afhænger af genre og emne. T/T måling af tekster i forskellige genrer vil variere, og

¹³ Generelt taler man ikke om ordfamilier i denne sammenhæng, en ordform tæller som en type ord, fx er *familie* og *familien* to forskellige ordtyper selvom de tilhører den samme ordfamilie.

mundtlige teksttyper har som regel en lavere T/T værdi end skriftlige tekster (Linnarud, 1986).

Laufer og Nation siger at T/T ratio har vist sig at være ustabil for korte teksters vedkommende og at tekster af forskellig længde påvirker resultaterne. Derfor kunne tekster af samme længde give et bedre resultat i målingerne. De mener at T/T måling først og fremmest viser hvor godt en elev kan udtrykke sig med det ordforråd han/hun kender. Denne måling skelner ikke mellem kvalitet af de forskellige ord, som defineres ud fra sjældenhed. En elev kan score højt T/T ved at bruge mange ordformer fra få ordfamilier og derved står han/hun på lige fod med en anden, der bruger ord fra mange familier. (Laufer og Nation, 1995:310)

Laufer og Nation (1995) kritiserer alle ovennævnte metoder og anbefaler LFP (Leksikalsk Frekvens Profil), som de påstår kommer uden om de andre metoders svagheder.

3.6 Leksikalsk frekvensprofil (LFP)

Den leksikalske frekvensprofil blev præsenteret 1995 af Laufer og Nation. De hævder, at LFP er en pålidelig måling, som er stabil for to tekster over 200 ord produceret af den samme indlærer. Den skelner også mellem indlærere på forskellige færdighedsniveauer. LFP viser den relative andel ord fra forskellige frekvensbånd. Ved hjælp af frekvensordbog kan ordforrådet inddeles i grupper, såkaldte frekvensbånd. Hvis frekvensbåndene har et interval på 1000 ord, vil det første frekvensbånd indeholde de tusind ord, som har højest frekvens, dvs. ordene med rang fra 1 til 1000. Det andet frekvensbånd vil dække ordene med rang fra 1001 til 2000 osv. Nation og Laufer præsenterer to slags målinger – for nybegyndere og for indlærere på et højt færdighedsniveau. Målingen består i at regne ud

1. andel ord i frekvensbånd 1 (dvs. ord med rang mellem 1 og 1000) + andel ord i frekvensbånd 2 (dvs. ord med rang mellem 1001 og 2000) + andel andre ord.
2. andel ord i frekvensbånd 2 (dvs. ord med rang mellem 1001 og 2000) + andel akademiske ord (dvs. som findes på en liste udarbejdet af Nation, 1990) + andel mindre frekvente ord (dvs. som ikke er på frekvensbånd 1, 2 eller på den akademiske liste).

Jeg bestemte mig for at udarbejde informanternes ordforråd i frekvensbånd og prøve på den måde at få et bredere billede af hvorledes informanternes ordforråd er sammensat.

Så vidt jeg ved, foreligger der ikke noget materiale der inddeler det danske ordforråd i frekvensbånd på samme måde, som man har gjort med det engelske. Dansk frekvensordbog er derfor blevet min bibel i denne henseende.

Laufer (1994) har vist, at mens informanterne er i kontakt med det engelske sprog udvikles deres LFP. Ordforråd af frekvens under 2000 bliver større og ord af frekvens over 2000 bliver færre. LFP forholder sig nogenlunde ens mellem samme slags skriveopgaver lavet af samme elever med få dages mellemrum.

3.7 Kritik – Nye metoder

Meara og Bell (2001) påstår, at erfaringen har vist at LFP har dårlige måleegenskaber og at metoden ikke skelner mellem tekster fordi den stoler på simpel tælling af ord fra kategorier 3 og 4 (jf. 2.3 s.15). De to er i færd med at udvikle en metode, der skal være velegnet til at måle korte elevtekster. De bruger frekvenslister udarbejdet af Nation som grundlag, men fremgangsmåden er en helt anden. De kalder metoden P_Lex og den består i at inddele ordforrådet i to kategorier, nemme ord og svære ord (højfrekvente og mindre frekvente). Elevteksterne bliver delt op i sekvenser af 10 ord og herefter tæller man hvor mange svære ord der forekommer i hver sekvens. Så regner man sammen hvor mange sekvenser har 0 svære ord, hvor mange har 1, 2, 3 svære ord osv. og laver et diagram over resultatet. Meara og Bell mener, at denne metode er pålidelig til måling af korte L2-tekster

Meara og Fitzpatrick (2000) mener, at en af grundene til at det er langt sværere at måle produktivt ordforråd end receptivt, er at det produktive, talt eller skrevet, er så bundet til konteksten, at det ville være svært ud fra en kort tekst at vurdere den rigtige størrelse af indlærerens produktive ordforråd. De har brugt en test, som de kalder Lex30. Der har man udvalgt 30 forholdsvis højfrekvente ord, der bliver vist et ad gangen hvert 30. sekund. Eleverne skal så helst skrive tre ord, som har association til hvert af de 30 stimulus-ord de får forevist. Responsordene bliver vurderet, bl.a. ud fra frekvens, og omregnet til points og så bliver elevernes scorer sammenlignet.

Meara og Fitzpatrick mener, at denne metode fortæller mere om elevernes produktive ordforråd end LFP-metoden. Til gengæld består produktionen næsten udelukkende af indholdsord.

De to metoder, P_Lex og Lex30, lyder spændende. De er begge to forholdsvis enkle i brug, og det kræver ikke lang tid at prøve efter disse metoder, eller finde frem til resultater.

Alle de metoder, der er blevet præsenteret her, er blevet udviklet og brugt i den hensigt at finde ud af, hvordan man kan måle og vurdere produktivt sprogbrug. Den rigtige, absolutte metode er endnu ikke blevet præsenteret. Hver for sig er begrænset, men er dog en slags milepæl i søgen efter viden.

De to metoder, jeg har bestemt mig for at bruge i min undersøgelse – Leksikalsk variation (LV) og Leksikalsk frekvensprofil (LFP) - giver hver for sig et bestemt billede af informanternes ordforråd. Nogle steder kan der drages sammenligninger, andre steder er målingerne for forskellige til at det kan lade sig gøre.

Grunden for valg af disse metoder er for det første, at jeg synes det vigtigt at bruge to forskellige metoder for at se, om de giver de samme eller lignende resultater og kan støtte hinanden – eller det modsatte. Den anden grund er, at LV er meget nem i brug og det tager kun et øjeblik at måle T/T ratioen, når teksten først er skrevet ind i computeren. Desuden kan tekster forkortes, dvs. der kan tages hvilken som helst del af tekster (begyndelse, midte, slutning) og sammenligne – det bliver dog gjort i meget beskeden grad her i specialet. Endelig synes jeg det er vigtigt, at måling/metode viser eller giver en idé om ordforrådets sammensætning eller spredning over frekvenskategorier.

4 Empirisk del

I denne del af specialet vil jeg først gøre rede for baggrunden for denne undersøgelse, ideen bag den, valg af skole og informanter, samt indsamling af data. Derefter vil gøre rede for de metoder jeg brugte til måling af ordforråd, en for en, og fortælle om fordele og ulemper. Sammenligning af enkelte opgaver, enkelte informanter, oversigt over alle opgaver, alle informanter, fokus på tre informanter – i forsøg at få så bredt et billede af informanternes sprogproduktion som muligt.

4.1 Undersøgelsens gennemførelse

4.1.1 Ideen bag undersøgelsen

Da jeg skulle til at bestemme mig for emne til mit MA-speciale stod det mig klart, at det skulle omfatte noget, som jeg som sproglærer stod overfor til daglig i mit arbejde.

Tit har jeg og nogle af mine kollegaer spekuleret over, om den undervisning og de opgaver vores elever fik, var gode nok og hvordan disse ting kunne gøres bedre. Hvordan kunne det måles? Hvad skulle - eller kunne det sammenlignes med?

Ordforråd er et område indenfor fremmedsprogsundervisningen, som der forskes meget i og vi har i vores undervisning lagt stor vægt på at gøre eleverne bevidste om adskillige metoder, de kunne bruge til at tilegne sig nyt ordforråd.

Vi har i grunden længe arbejdet efter det Wilkins sagde i 1972 (jf. s. 8) '*Uden grammatik kan man udrykke meget lidt, men uden ordforråd ingenting*'. Valget faldt som sagt på ordforråd, men i overensstemmelse med den diskussion, der havde fundet sted, skulle opgaven eller undersøgelsen tage udgangspunkt i noget der allerede var i gang i undervisningen.

I efteråret 2001 var det kun anden gang en stor gruppe, 6 hold på ca. 25 elever, skulle gennem det første modul i dansk, DAN 103, efter den nye læseplan. Det faldt sammen med mit studieorlov og så syntes jeg det ville være en god ide at prøve på en eller anden måde at kortlægge elevernes sprogproduktion i løbet af semestret.

Jeg tænkte lidt over, om jeg skulle lave nogle ensartede opgaver, som skulle forelægges eleverne først og sidst på semestret og på den måde sammenligne deres præstationer med cirka 12-13 ugers mellemrum.

Ved nærmere eftertanke besluttede jeg mig til at bruge de opgaver eleverne alligevel skulle igennem. På den måde ville jeg få et mere autentisk¹⁴ materiale at arbejde med, men også flere forskellige slags opgaver. Et egentligt sammenligningsgrundlag var dermed ude af billedet, men i stedet for var det muligt at se om brug af autentiske opgaver ville give oplysninger man ikke kunne få på en anden vis. Måske kunne man på den måde få mere ud af eleverne, da opgaverne gerne skulle stille andre og eventuelt mere omfattende krav til eleverne og på flere områder end én type opgaver ville gøre.

Arbejdet går først og fremmest ud på at finde ud af, hvor eleverne virkelig står og hvad de tilegner sig i løbet af det første semester, og at læreren kan få oplysninger herom.

Undervisningsmetoder er ikke i fokus i denne undersøgelse. Jeg håber dog på, at en analyse af informanternes opgaver giver et grundlag for at udarbejde en mere effektiv plan for opgaver og aktiviteter der går ud på

- ordforrådsindlæring
- differentiering
- elevansvar
- computerstøttet sprogundervisning

4.1.2 Valg af skole

Jeg synes det er en fordel at kunne vælge en skole, hvor man er godt kendt med alle forhold og fortrolig med både undervisningsmaterialet og undervisningsmetoderne og hele tiden kan komme og gå og følge med i, hvad der sker uden at nogen lægger særlig mærke til en.

Til trods for det behagelige og praktiske i at foretage undersøgelsen i trygge omgivelser, kan det sikkert også føre til at noget bliver taget for givet eller at noget undgår ens opmærksomhed.

¹⁴ Autentisk materiale i den forstand, at informanterne ikke skulle have fornemmelsen af at være behandlet anderledes ved at få andre eller flere opgaver at løse end de andre på holdet. Endvidere, at materialet skulle genspejle undervisningssituationen.

Skolen er i en kommune med knap 25.000 indbyggere og over 90% af eleverne på gymnasielinien kommer fra kommunens seks grundskoler.

4.1.3 Informanternes baggrund

De elever, der begyndte deres studier i Menntaskólinn i Kópavogur (MK) i efteråret 2001, tilhører en af de sidste årgange der havde dansk som første fremmedsprog. De har fulgt den gamle læseplan i folkeskolen, men kommer ind i den nye læseplan i gymnasieskolen.

Disse elever har ifølge læseplanen haft dansk i fem år i grundskolen, de første to år to timer (à 40 min.) om ugen, de næste to år tre timer om ugen og det sidste år fire timer om ugen. Overgangen fra grundskolen er lidt brat i den forstand, at eleverne får mere intensiv undervisning end før idet de har tre dobbelttimer om ugen i stedet for fire enkelttimer og for de flestes vedkommende er det deres sidste år med dansk. Desuden skal de muligvis tackle forskellige innovationer (bl.a. tekniske nyskabelser) i danskfaget på MK, samt at tilrettelæggelsen af undervisningen og elevernes arbejdsformer har gennemgået en grundig reform, som kan besigtiges på <http://utskolar.is> >kennarar>danska¹⁵, og derfor kan være en del anderledes end eleverne er vant til fra deres gamle skole.

4.1.4 Kriterier for valg af informanter

For at begrænse variabler så meget som muligt har jeg valgt elever der kommer direkte fra folkeskolen, og som ikke har nogen speciel forbindelse til Danmark, hvad angår familie eller ophold i landet udover højst et enkelt turistbesøg. Informanterne går på det samme hold¹⁶ i gymnasiet og har derfor alle fået den samme undervisning og gennemgået pensummet på samme måde i løbet af semestret.

¹⁵ En web til både lærere og elever, som indeholder materiale til hjælp eller vejledning ved brug af IT indenfor enkelte fag eller fagområder i skolerne.

¹⁶ I denne skole går eleverne ikke i klasser, men bliver opdelt i hold i starten af hvert semester. Hver elev kan i teorien derfor tilhøre ligeså mange hold, som de fag hun går til.

Elleve elever ud af 22 opfyldte kravene om at være kommet direkte fra grundskolen, de havde ingen særlige forbindelser til Danmark, og de gav lov til at deres opgaver blev brugt som undersøgelsesmateriale.

Informantgruppen består af de ni elever, tre drenge og seks piger, som afleverede alle opgaverne rettidigt.

4.1.5 Indsamling af data

4.1.5.1 Opgavetyper

Undervisningsforløbet indbefatter en række opgaver, som er inkluderet i semestrets arbejdsplan med fastlagte datoer inden semestrets start. Alle hold nybegyndere gennemgår det samme program og laver de samme opgaver inden for et så snævert tidsrum som muligt. Alle opgaver er lavet i undervisningstimerne undtagen den mundtlige, som til gengæld er optaget på videobånd, så elevernes præstationer er meget synlige.

Det materiale, jeg har valgt ud, består af fem produktive opgaver; fire skriftlige opgaver og en mundtlig, optaget på videobånd. Desuden blev der forelagt en test i receptivt ordforråd, men kun for det hold informanterne kommer fra.

De fem produktive opgaver giver et tværsnit af de opgaver eleverne laver i løbet af det første semester.

'En god vinter' ideassociationer; hvad forbinder du med denne overskrift? Bygger på paratviden, oplagret ordforråd. Eleverne fik 15 minutter uden forberedelse og hjælpemidler. Emnet vedrører privatsfæren, hvor eleverne kan fortælle om egne erfaringer og oplevelser.

'Andrea elsker mig' (roman af Niels Rohleder) – det skriftlige produkt er en slags opsummering af input og proces efter ca. en uges projektarbejde med tekstens indhold og ordforråd. Eleverne fik 40 minutter til at skrive et udkast, aflevere til læreren og en dobbeltime til at færdiggøre arbejdet. Hjælpemidler: Ordbog, grammatiknøgle, romanen, selvvalgte nøgleord fra projektarbejdet.

'Den barske virkelighed' og 'Rejse i Danmark/Island', to af semestrets fem temaer skulle udgøre baggrunden for 15 minutters samtale(r). (Det første tema går ud på at tjene penge, flytte hjemmefra og hvem skal betale hvad i forholdet forældre/unge.) Den anbefalede fremgangsmåde var at udarbejde ordblomst (mind-maps) med det stof, de skulle bruge, og improvisere ud fra nøgleord til fri produktion i samtaleform. Eleverne fik ubegrænset tid til produktion af 15 minutters samtale. Fri brug af hjælpemidler.

'Udsigt til Marmorkirken' (Novelle af Anders Bodelsen). Eleverne har arbejdet med fire noveller og skal skrive om en af dem. De ved ikke i forvejen hvilken. Opgaven går ud på, at de kender og kan gøre rede for novellens forløb, plot og ordforråd (krimi). Her går opgaven ud på at ord og indhold filtreres sammen. Hjælpemidler: Ordbog og grammatiknøgle. (30 min.)

'Sundhed' er et af semestrets fem temaer. Eleverne ved ikke i forvejen, hvilket tema de skal skrive om, men de ved at de har frit valg med hvad de forbinder med begrebet 'Sundhed' bare de bruger et passende ordforråd. De får 10 minutter til at genkalde temabundet og til en vis grad abstrakt specialordforråd.

'Den receptive prøve' i fire dele

- A. 33 ord fra 'danske kerneord'
- B. 39 ord fra pensum.
- C. 15 faste vendinger fra pensum.
- D. 10 sætninger ordret fra pensum hvor der mangler et ord – flervalgsprøve.

4.1.5.2 Semestrets kalender

Oversigt over de opgaver der arbejdes med i undersøgelsen

| | | |
|--------------------------------|-------------------------|--|
| 29. august (1. uge) | Fristil | 'En god vinter' 15 min. Uden forberedelse – ingen hjælpemidler. |
| 25. oktober (9. uge) | Stil om en roman | Romanen 'Andrea elsker mig' af Niels Rohleder 120 min. (40+80). Forløb: Læse romanen, selvstændig læsning hjemme efter en kort præsentation i klassen Gennemgang af nogle litterære begreber i klassen Flervalgsopgave (multiple-choice): tjek af læsning og forståelse 40 min. Gruppearbejde: hver gruppe med en opgave som bliver |

| | | |
|----------------------------------|-------------------------------|--|
| | | præsenteret mundtligt for hele holdet ved hjælp af overheads (kopier tilgængelige for alle) Stil: valg mellem to opgaver (ud af seks, hvad de godt vidste i forvejen) tid mellem 40 min. og de sidste 80 til yderligere forberedelse hjemme. Hjælpemidler: romanen, ordbog, grammatiknøgle, 15-20 stikord |
| 12. november (12. uge) | Samtaler på video-bånd | Den mundtlige prøve. Samtaler, dialoger, i 15 minutter. Samtalernes emne skulle dække to af semestrets temaer fra læsebogen og tiden skulle deles lige mellem de to temaer – ca. 7 ½ minut pr. elev. Forberedelse og optagelser uden for timerne. Alle hjælpemidler tilladt. |
| 22. november (13. uge) | Hurtig-skrivning | Temaet 'Sundhed' fra læsebogen og Internettet. Ingen hjælpemidler. |
| 22. november (13. uge) | Stil om en novelle | 'Udsigt til Marmorkirken' af Anders Bodelsen, 30 min. Eleverne vidste ikke i forvejen, hvilken af de noveller de havde læst, de skulle skrive om. Hjælpemidler: ordbog, grammatiknøgle, 15-20 stikord. |
| 22. november (13. uge) | Receptivt ordforråd | Danske 'kerneord' + pensumrelaterede ord og udtryk, 15 min. Eleverne vidste ikke om denne prøve i forvejen. Ingen hjælpemidler. |

4.1.6 Behandling af data

Informanternes lærer fotokopierede opgaverne fra hele holdet og gav dem til mig efterhånden som semestret skred frem. Det var først i slutningen af semestret, at det blev klart hvor stor informantgruppen ville være. Jeg ville helst have alle med som opfyldte kriterierne (jfr. 4.1.4.), men ikke alle afleverede de opgaver der blev udvalgt. Alle skriftlige tekster, samt alle samtaler fra videobåndene, blev skrevet op på computer. De skriftlige tekster er skrevet nøjagtigt af, men tekster fra videobånd er skrevet af med korrekt stavemåde, selv om udtalen ikke altid var perfekt. Grammatiske fejl blev skrevet af og ord, som tydeligvis var forkerte, blev skrevet af efter udtalen. Hvor der var tøven, ufuldendte replikker, eller hvor informanterne brugte hinandens rigtige navne, bruges Hvor udtalen var utydelig og ordene ikke kunne høres bruges xxx. Tre af informanterne havde partnere til 'Talesprog' fra deres hold, men som ikke opfyldte kriterier for at blive informanter. Programmet Clan/Childes blev brugt til at tælle ord, regne ud frekvens af ord, T/T ratio og kalkulere helhedsordforråd.

Alle informanter er anonyme. I specialet har de fået nummer fra 1-9.

4.1.7 Hvornår tælles et ord som et ord?

I opgaverne fra informanterne er der en del stavfejl og grammatiske fejl. Når der skal tælles ord, er det hensigtsmæssigt, at det samme ord staves ens hver gang. Alle de steder, hvor jeg synes det stod helt klart hvilket ord det drejede sig om, rettede jeg fejlene. (Eksempel: inge (=ingen), skedte (=skete))

Andre steder, hvor jeg ikke var helt sikker, fik jeg hjælp af nogle danske universitetsstuderende og lærere til at afgøre, om ord skulle tælle med eller kategoriseres som uforståelige eller af tvivlsom karakter. (Eksempel: fåråret (=foråret), barnelømmelse (=børnelammelse)). De steder hvor ordene blev brugt forkert, dvs. rigtige danske ord, stavet rigtigt, men brugt i en forkert sammenhæng blev kategoriseret som *tvivlsomme* (fx 'måle' i stedet for 'male'). Det er dog ikke helt entydigt. Mange skriver fx konsekvent 'so' i stedet for 'så' og der kan man ikke tage fejl af meningen. Her fandt jeg det formålstjenligt at godtage 'so' i betydningen 'så'.

4.2 Afprøvning af metoder I

4.2.1 Overvejelser angående fremgangsmåden

Da jeg havde arbejdet lidt med undersøgelsesmaterialet, opstod problemet med at ordne materialet systematisk for at give en læser et logisk overblik over, hvad det hele drejer sig om - ikke mindst spørgsmålet angående, hvad det egentlig er jeg vil med denne undersøgelse. I indledningen leger jeg med den ide at kortlægge sprogproduktionen i dansk hos en gruppe informanter i løbet af deres første semester i gymnasiet. I starten havde jeg håbet på at kunne finde en udvikling hen ad vejen. Skete der noget fra første til niende uge? Fra niende til trettende uge? Eller fra første til trettende uge? Nogle forskere (jf. Henriksen & Haastrup, 2000) mener, at et semester er for kort tid til at konstatere fremskridt/udvikling i tilegnelsen af ordforråd, dvs. at kunne måle forskel i kunnen eller færdigheder på to tidspunkter med kun tretten ugers mellemrum. Andre (Laufer 1994) har konstateret fremskridt over en periode på 14 uger.

Det står mig klart, at det ikke er muligt at gøre rede for alle de ting, som kan udarbejdes af det materiale jeg har i hænderne. Jeg må derfor udvælge det jeg tror vil give den bedste oversigt og støtte mig så vidt det er muligt til det, der kan måles for så at drage konklusioner deraf.

Det ligger i sagens natur, at det er svært at isolere enkelte træk. Der er alt for mange variabler, der kan påvirke informanten i den producerende (skrivende/talende) stund. Desuden findes der ikke nogen metode der kan trække informantens totale ordforråd ud på en gang. Det er derfor et af de problemer en undersøgelse af denne art må slås med, nemlig hvad slags opgaver er de mest hensigtsmæssige for at kunne måle ordforråd hos en informant? Og vil samme slags opgaver også virke hensigtsmæssige for alle informanter? Mine informanter har produceret fem forskellige opgaver, som kan sammenlignes i alle mulige ender og kanter, men der skal trædes varsomt når resultaterne skal fortolkes. Resultaterne kan dog muligvis give os lærere et fingerpeg om, hvad vi skal fokusere på når vi ønsker optimal sprogproduktion af eleverne.

Forskellige metoder er blevet brugt for at sætte målestok på informanternes sprogproduktion. Disse metoder er blevet introduceret i afsnit 3.1-3.6. Det er egentlig kun to metoder, som jeg synes er brugbare i denne sammenhæng og dem vil jeg argumentere for i afsnit 4.4.7. Først vil jeg starte med eksempler på, hvordan de metoder jeg **ikke** har valgt virker i praksis. I den hensigt har jeg valgt at bruge to af de skriftlige opgaver, den første og den sidste, for at vise hvordan de virker og på det grundlag argumentere for, hvorfor jeg ikke synes de er brugbare for hele produktionen.

4.2.2 De to opgaver – den første og den sidste

Disse to opgaver er lavet som hurtigskrivning, henholdsvis på 15 og 10 minutter med tolv ugers mellemrum. Opgaverne kan henholdsvis karakteriseres som ideassociationer over emnerne 'En god vinter' og 'Sundhed'. Det er nemmest og mest oplagt at sammenligne de to opgaver¹⁷ med hinanden, da de blev forelagt i hver sin ende af semestret.

¹⁷ Informanternes opgaver findes som bilag 1:i-xix

Tabel 4.2.2.a. 'En god vinter' og 'Sundhed': Antal ord hos alle elever og gennemsnit af sætningslængde i de to opgaver.

| | Inf. 1 | Inf. 2 | Inf. 3 | Inf. 4 | Inf. 5 | Inf. 6 | Inf. 7 | Inf. 8 | Inf. 9 | Gennemsnit |
|--------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------------|
| 'En god vinter' Antal ord | 87 | 120 | 165 | 140 | 136 | 126 | 149 | 158 | 211 | 144 |
| 'Sundhed' Antal ord | 100 | 124 | 257 | 141 | 107 | 136 | 144 | 206 | 177 | 155 |
| 'En god vinter' Ord/sætning | 17,4 | 17,0 | 23,6 | 10,8 | 10,5 | 12,6 | 19,8 | 17,6 | 13,2 | 15,8 |
| 'Sundhed' Ord/sætning | 9,1 | 11,2 | 19,8 | 11,8 | 7,2 | 10,5 | 15,9 | 17,2 | 9,1 | 12,4 |

Af tabel 4.2.2.a kan det tydeligt ses, at selv om opgaverne er skrevet i løbet af kort tid, er der stor forskel i længden. De længste er to en halv gang så lange som de korteste i begge opgaver. Gennemsnit for længden vokser med 11 ord (9,3%) mellem de to opgaver, fra 144 ord til 155 ord til trods for kortere tid.

Sætningslængden bliver kortere hos alle informanter undtaget en, i gennemsnit 27,4%, dvs. længden falder fra 15,8 ord pr. sætning til 12,4 ord pr. sætning. Det er dog svært at se, hvorfor det sker. Det er muligt, at informanterne har lært at formulere sig mere klart og præcist i løbet af semestret, men det kan også godt være at den sidste opgaves natur har medvirket til at sætningerne er blevet kortere. Informanterne har prøvet at få så meget materiale med i opgaven som muligt og derfor prøvet at formulere sig så præcist som muligt. Der er ikke gjort noget forsøg på at vurdere korrekthed i forhold til længde, men der er færre ord, der bliver karakteriseret som 'tvivlsomme' eller uforståelige i den sidste opgave (jf. tabel 4.4.4a)

Det er bemærkelsesværdigt, at det ikke er de samme tre informanter som har de længste tekster begge gange. Den samme informant har derimod den korteste tekst begge gange. Tre informanter producerede en kortere tekst i den sidste opgave end den første.

4.2.3 Leksikalsk tæthed (LT)

$$LT = \frac{\text{antal indholdsord i teksten}}{\text{antal ord} \times 100}$$

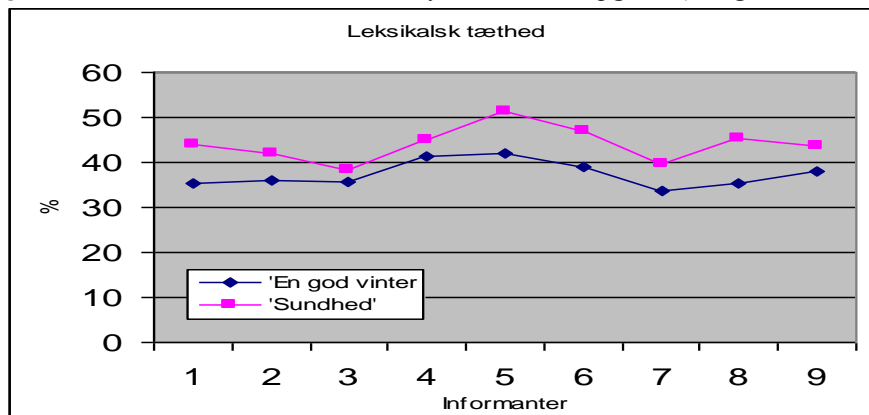
Ved brug af denne metode blev alle ord talt med som fungerede som indholdsord (jf. afsnit 3.5), selvom nogle af dem kunne regnes for at være af en 'tvivlsom' oprindelse og som ikke i alle tilfælde ville kunne forstås af mennesker med dansk som modersmål. Disse ord er dog få og har derfor ikke nogen afgørende indflydelse på resultaterne. Eksempler på 'tvivlsomme ord' er *efterdågen, *føjdigt, *gørest, *hukser, *regelse.

Tablet 4.2.3.a. Leksikalsk tæthed hos alle ni informanter i 'En god vinter' og 'Sundhed'.

| | Inf. 1 | Inf. 2 | Inf. 3 | Inf. 4 | Inf. 5 | Inf. 6 | Inf. 7 | Inf. 8 | Inf. 9 | Gennemsnit |
|--|--------|--------|--------|--------|--------|--------|--------|--------|--------|------------|
| 'En god vinter' | 35,2 | 36,1 | 35,8 | 41,4 | 41,9 | 38,9 | 33,6 | 35,4 | 37,9 | 37,4% |
| 'Sundhed' | 44,0 | 41,9 | 38,4 | 45,1 | 51,5 | 47,1 | 39,6 | 45,4 | 43,8 | 44,1% |
| 'En god vinter' og 'Sundhed' – forskel i % | 25,0 | 16,1 | 7,3 | 8,9 | 22,9 | 21,1 | 17,9 | 28,2 | 15,6 | 17,9 |

Af tabel 4.2.3.a fremgår, at alle informanter har højere værdi af leksikalsk tæthed i den sidste tekst. Der er mindre forskel på højeste og laveste værdi i den første opgave end den sidste. Forhøjelsen af LT varierer meget hos de enkelte informanter. Informant 5 ligger højest i begge opgaver og er forholdsvis langt den højeste i den sidste opgave, dvs. 9% over informant 6, som kommer nummer to. Figur 4.2.3.a viser denne variation meget tydeligt.

Figur 4.2.3.a Leksikalsk tæthed hos ni informanter i to opgaver ('En god vinter' og 'Sundhed')



På figur 4.2.3.a kan der ses en tydelig forskel på LT værdier. Det at LT er højere i den sidste opgave, indebærer at indholdsord er blevet forholdsvis flere end funktionsord. Det er dog tekstens komposition og struktur, der er afgørende for, hvordan forholdene mellem de to kategorier ord bliver.

Ved at sammenligne LT værdi med sætningslængde / tekstens komposition og struktur (jf. afsnit 4.2.2.) er det påfaldende hos informant 5, at høj LT værdi falder sammen med meget korte sætninger, nærmest telegramsprog, især i den sidste opgave. Den samme tendens kan spores hos informanter 1 og 6, men deres tekster præges mere af optællinger eller opremsninger end hos informant 5. Dette gælder dog ikke for informant 9, der også har korte sætninger, men teksten er betydeligt mere kohærent end de tre ovennævnte (jf. bilag 1).

Den store forskel i LT værdi mellem de to opgaver kan skyldes at temamæssigt og konceptmæssigt er de to opgaver af vidt forskellig karakter, som allerede er nævnt (jf. afsnit 4.2.2). Den første opgave lægger op til en fortælling, hvor lange sætninger kan være naturlige. Den sidste opgave går mere ud på at formidle oplysninger at prøve på at medtage så meget som muligt. Derfor bruges korte, præcise sætninger eller optællinger og opremsninger.

Denne metode følger i grunden ikke ret mange oplysninger til den LV (leksikalsk variation, jf. afsnit 4.3.) giver om selve ordforrådet, hverken hvad angår ordforrådets kvantitet eller kvalitet. Den siger til gengæld noget om syntaktiske og kohæsive faktorer i tekstens komposition og kan også afspejle stil (Laufer & Nation 1995).

En stor ulempe ved metoden er dog den, at det er tidskrævende at arbejde med informanternes tekster på denne måde. Teksterne skal analyseres i indholdsord / ikke-indholdsord (funktionsord) og det kræver tid og energi at gøre det manuelt. Den teknik jeg har til rådighed, omfatter ikke den mulighed at udføre arbejdet elektronisk.

4.2.4 Leksikalsk originalitet (LO)

$$LO = \frac{\text{antal originale ord i teksten}}{\text{antal ord} \times 100}$$

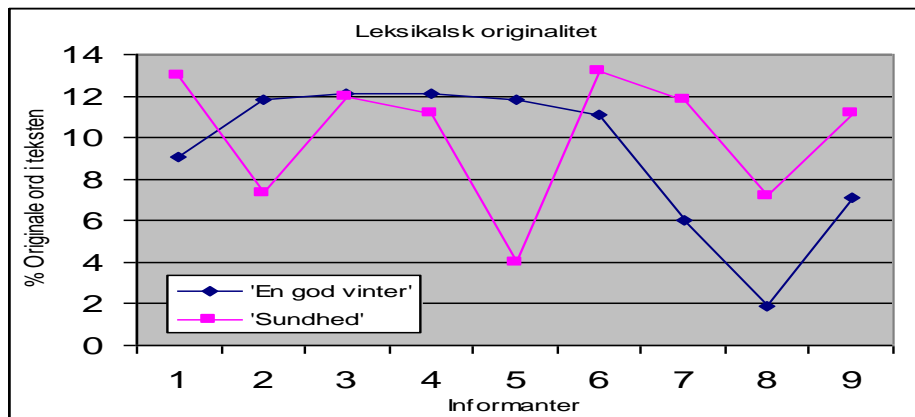
Ved afprøvning af denne metode er nogle få ord blevet valgt fra i begyndelsen, enten fordi de blev brugt forkert eller fordi de blev skrevet så forkert, at de ikke umiddelbart kunne genkendes eller accepteres som danske ord.

Tabel 4.2.4.a. Procentdel og antal **originale ord** hos alle ni informanter i to opgaver ('En god vinter' og 'Sundhed').

| | Inf. 1 | Inf. 2 | Inf. 3 | Inf. 4 | Inf. 5 | Inf. 6 | Inf. 7 | Inf. 8 | Inf. 9 | Gennemsnit |
|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------------|
| 'En god vinter' | 9,1% | 11,8% | 12,1% | 12,1% | 11,8% | 11,1% | 6,0% | 1,9% | 7,1% | 9,2% |
| | 8 | 14 | 20 | 17 | 16 | 14 | 9 | 3 | 15 | 12,9 ord |
| 'Sundhed' | 13,0% | 7,3% | 12,0% | 11,2% | 4,0% | 13,2% | 11,8% | 7,2% | 11,2% | 10,1% |
| | 13 | 9 | 31 | 16 | 4 | 18 | 17 | 15 | 20 | 15,9 ord |

Tabel 4.2.4.a giver både procentdel af originale ord i informanternes tekster og antal originale ord. Det er klart, at det ikke drejer sig om mange ord, gennemsnitstal på henholdsvis 12,9 og 15,9 ord, som dog viser at informanterne som gruppe bruger flere originale ord i den sidste opgave. Forskellen er dog ikke større end 0,9%.

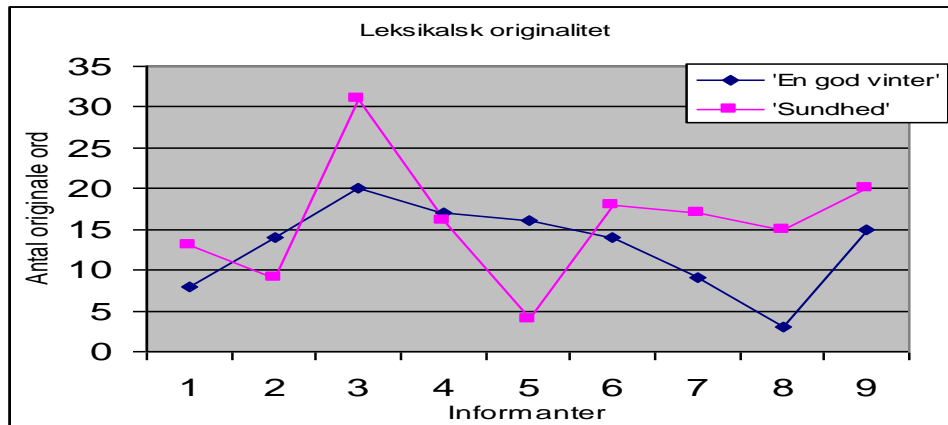
Figur 4.2.4.a Viser procentdel originale ord i de to tekster 'Vinter' og 'Sundhed'



Af tabellen fremgår der ikke noget afgørende mønster, men diagrammerne på figurer 4.2.4.a og 4.2.4.b viser et klarere billede end tabellen. Informanter 3 og 4 har forholdsvis lige mange originale ord i begge tekster, hvorimod informanter 2 og 4 har forholdsvis langt færre originale ord i den sidste tekst, især informant 5. Flere af informanterne har forholdsvis flere originale ord i den sidste tekst. Det, at informant 3 har langt højere antal

originale ord i den sidste tekst, selv om det procentvis er nogenlunde det samme, kan forklares med at informant 3 har den største forskel i længden af opgave 1 og opgave 2 (henholdsvis 165 og 258 ord, jf. tabel 4.2.2). Informant 8 kan siges at være højdespringeren mellem de to opgaver, selvom hun ligger lidt under gennemsnit i den sidste opgave. Grunden er selvfølgelig den, at hun kun har brugt 3 originale ord i den første tekst. Informanter 5 og 8 er ekstremisterne i disse opgaver, men det er svært at afgøre, hvorfor deres opgaver er så forskellige. Der kan være så mange variabler der spiller ind og derfor kan det være farligt at drage forhastede konklusioner af disse to eksempler.

Figur 4.2.4.b Viser antal originale ord i de to tekster 'Vinter' og 'Sundhed'



Selvom denne måling, dvs. procentdel af originale ord i de enkelte tekster, ikke skulle kunne påvirkes af teksternes længde, drejer det sig her om så korte tekster og som helhed set ret begrænset ordforråd, til trods for at informanterne havde forholdsvis frie hænder angående hvad de skrev. Denne måling siger ikke ret meget om, hvad slags ord det er, der regnes for at være originale. Det kan lige så godt dreje sig om ganske almindelige dagligdags ord (højfrekvente), som mere sjældne ord. Heller ikke denne metode giver oplysninger om kvaliteten af det brugte ordforråd eller om hvorvidt informanterne har gjort fremskridt i indlæring af ordforråd i perioden mellem de to opgaver.

Det ville måske være mere pålideligt at lægge hver informants ordforråd fra alle fem opgaver som grundlag for brug af denne metode, eftersom det ville dreje sig om et langt større ordforråd for alle informanter. Jeg tror dog ikke, det ville give flere oplysninger om

ordforrådets sammensætning, kvalitet eller egenskaber end den stikprøve, vi har her ovenfor.

4.2.5 Leksikalsk raffinement (LR)

$$LR = \frac{\text{antal raffinerede ord i teksten}}{\text{antal ord} \times 100}$$

Denne metode egner sig ikke til brug på de data jeg har, da den kræver, at man har en udtømmende liste over det ordforråd elever i 1.g i Island burde have stiftet kendskab med gennem undervisningsmaterialet. Denne viden foreligger ikke på nuværende tidspunkt. Det ville dog være af stor interesse at lave en oversigt over det ordforråd der præsenteres i læsebøger og supplerende materialer. Det kunne gøre det nemt for forskere og lærere at undersøge og vurdere, hvor stor en del af ordforrådet eleverne behersker på et bestemt tidspunkt og hvordan denne tilegnelse udvikler sig over en bestemt årrække efter en langt større målestok end de undersøgelser jeg hidtil har læst om. Uden tvivl ville det bidrage til, at lærerne kunne være mere målbevidste i undervisning af ordforråd – og det kunne måske også føre til, at eleverne blev mere målbevidste i deres tilegnelse af ordforråd.

Teknikken gør dette muligt, men elevdifferentiering og varierende læsestof hentet fra Internettet gør det sværere. Skulle der også tages højde for talesprog; lærerens tale, dialoger og diskussioner i klassen, ville opgaven blive endnu sværere.

4.2.6 Konklusion

De tre metoder, jeg her har gennemgået, viser klart at sproglærere og sprogforskere har søgt efter metoder til at måle elevens kunnen og viden på et givet tidspunkt på den ene side og over et tidsrum på den anden side. Alle disse metoder har krævet en masse arbejde og været tidskrævende, især før computerens tid. De to første metoder '*Leksikalsk tæthed*' (4.2.3) og '*Leksikalsk originalitet*' (4.2.4) bygger begge to på en ret mekanisk vurdering (inddeling i ordklasser og måling af originalitet), uden at give nogen afgørende resultater. For mig at se kan der ikke konstateres nogen fremskridt i tilegnelse af

ordforråd ved brug af disse metoder i de to opgaver, jeg har lagt til grundlag for afprøvning af metoderne. Resultater målt med LT afhænger af at alle informanter bruger samme slags stil i deres opgaver, men det kan selvfølgelig ikke lade sig gøre at støbe alle i den samme form. Og som det før blev sagt, måler LO kun relativ originalitet, som hænger sammen med hver enkelt opgave, men kan ikke bruges til sammenligning af forskellige opgaver.

Fra min synsvinkel er det kun den sidste metode, *'Leksikalsk raffinement'* (4.2.5), der kan bruges og udvikles til et godt måleapparat – men det kræver en stor indsats at få oparbejdet en databank, der udgør det nødvendige grundlag for at metoden kan virke. Med den databank vil forskere og lærere have det fornødne sammenligningsgrundlag til brug under de varierende forhold, der kan gøre sig gældende i de forskellige elevgrupper, eller ved de forskellige skoler eller angående de forskellige fremmedsprog. Denne metode kan spille sammen med LV (4.3) eller LFP (4.4) for at gøre det hele mere interessant!

4.3 Afprøvning af metoder II.

'Leksikalsk variation'

$$LV = \frac{\text{Typer af ord i teksten}}{\text{Antal ord i teksten}}$$

bliver her brugt på hele sprogproduktionen, alle fem opgaver fra alle ni informanter – 45 opgaver i alt. Her gives der ikke bare tal over produktionen som helhed, men også over de enkelte opgaver. Hensigten er at vise

- en oversigt over størrelsen af det ordforråd informanterne behersker og derfra gå over til mere detaljerede oplysninger, såsom
 - sammenligning af de forskellige opgaver
 - sammenligning af enkelte informanter
 - oversigt over tre informanter og to opgaver i et forsøg på at sammenligne deres ordforråd ud fra forskellige kriterier

Ved brug af denne metode tæller alle ord i teksten, undtagen tal der ikke er skrevet med bogstaver. Ord, som staves på forskellige måder, tælles som én type ord, hvor det tydeligvis fremgår af konteksten, at det drejer sig om et og samme ord. Det medfører, at der også tælles ord med, der egentlig ikke hører til det danske sprog, såsom navne på steder og personer, foruden ord som informanterne griber til eller kreerer, når deres kunnen ikke slår til. Nogle eksempler: Akureyri, Benidorm, hjemmelektier, fåråret/foråret, snø/sne, væjret/vejret, smytte/smitte, opfunder/opfinder.

Jeg har ikke kunnet finde absolutte kriterier for, hvad der skal tælles med når denne metode bliver brugt. Jeg har valgt at tælle alt med fordi ellers bliver det svært at sætte grænser mellem det der skal med, og det der ikke skal med. Det må være det vigtigste, at tællingen er konsekvent. Det skal fremhæves her, at en analyse af selve ordforrådet bliver foretaget senere i specialet.

De tal, der kan aflæses fra denne måling, skal ikke sammenlignes med noget andet og står på den måde helt for sig selv. Der findes heller ingen målestok for værdien af LV – eller T/T ratioen (type/token ratio), kun det der siges i afsnit 3.5, at ratioen for skriftligt sprog er højere end for talesprog og ligger hos en erfaren modersmålstalende omkring 0,5.

Det er længden af teksten, der er det mest afgørende for T/T ratioen, en kort tekst har en høj ratio. I første omgang bliver teksterne dog målt uanset længde for at se, om der kan findes et mønster i ratioen henholdsvis hos informanterne eller i opgaverne. Senere bliver der så lavet en sammenligning mellem den originale opgave og en forkortet version af samme opgave. På grundlag af det kan så enkelte informanter sammenlignes med hensyn til ordforrådets omfang. Endnu en gang skal det dog fremhæves, at der er mange variabler der påvirker informanternes præstationer og alle resultater skal derfor vurderes med forbehold.

4.3.1 Størrelse og inddeling af informanternes ordforråd

I dette afsnit vil der gøres rede for informantgruppens produktion i tal. Fem opgaver fra ni informanter tæller 1.729 forskellige ordformer og 15.925 ord i alt.

Tabel 4.3.1.a Antal ord, ordformer, andel i helheden og T/T ratio i alle opgaver.

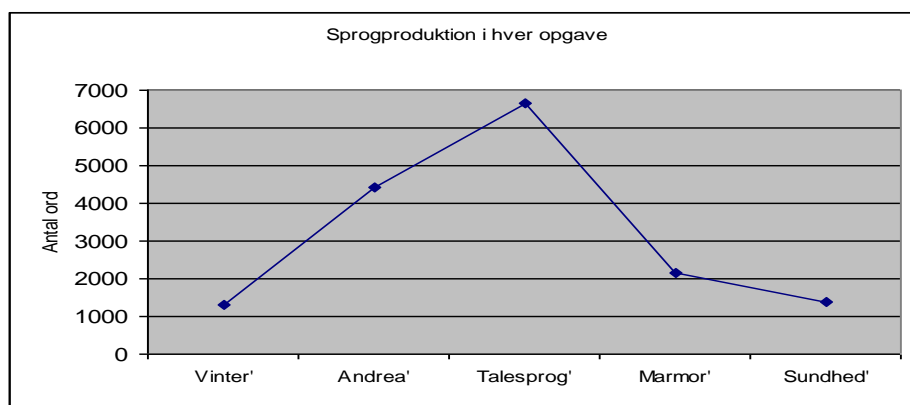
| | 'Antal ordformer' (type) | Enkelte opgavers andel i 'antal ordformer' | Antal 'ord i alt' (token) | Enkelte opgavers andel i 'antal ord i alt' | T/T ratio. | Tid til rådighed ved produktion af opgaven |
|---------------------|--------------------------|--|---------------------------|--|------------|--|
| Alle opgaver | 1.729 | 100,0% | 15.925 | 100,0% | 0,109 | |
| 'En god vinter' | 305 | 17,6% | 1.292 | 8,1% | 0,236 | 15 min |
| 'Andrea elsker mig' | 754 | 43,6% | 4.441 | 27,9% | 0,170 | 120 min. |
| 'Talesprog' | 824 | 47,7% | 6.664 | 41,8% | 0,124 | ubegrænset ¹⁸ |
| 'Marmorkirken' | 387 | 22,4% | 2.136 | 13,4% | 0,181 | 30 min. |
| 'Sundhed' | 381 | 22,0% | 1.392 | 8,7% | 0,274 | 10 min. |

Tabel 4.3.1.a viser alle tal over 'antal ordformer' og 'ord i alt' informanterne bruger i hele den produktion som undersøgelsen går ud på. Tabellen viser også, hvor lang tid informanterne havde til hver enkelt opgave.

Her angives også, hvor stor en del af hele ordforrådet, ordformer og ord i alt, informanterne bruger i hver opgave, kort sagt, oversigt over hele sprogproduktionen i tal.

Det kan tydeligt ses, at opgaverne er forskellige, ikke bare i længde men også i T/T ratio. Det skyldes til dels det faktum, at jo længere en tekst er jo lavere ratio (jf. 3.5). Derfor kan det ikke lade sig gøre umiddelbart at sammenligne opgaverne med hinanden. Opgaverne kan vurderes fra forskellige synsvinkler. En af dem er at se, om opgavernes længde kan fortælle noget om forudsætninger for sprogproduktion. En anden ting er at se efter, om det er en eller flere informanter, der producerer de længste/korteste opgaveløsninger.

Figur 4.3.1.a Hele sprogproduktionen inddelt efter opgaver i antal ord.



¹⁸ Dvs. lige så lang tid informanterne valgte at bruge på opgaven, jf. afsnit 4.1.5.1.

Figur 4.3.1.a viser, hvordan hele sprogproduktionen, 15.925 ord, fordeles på de fem opgaver. Af enkelte opgaver har 'Talesprog' langt den største andel i bruttoproduktionen (jf. bilag 1:v-xvii).

Antallet af ord skulle egentlig genspejle den tid informanterne havde til rådighed i de forskellige skriftlige opgaver, men det er ikke tilfældet. Hvis tre af opgaverne bliver sammenlignet, den første og de to sidste, viser den mellemste forholdsvis dårlige resultater.

Tabel 4.3.1.b Sammenligning af opgavers længde i forhold til tidsfaktoren.

| | <i>Tid i minutter</i> | <i>Antal 'ordformer'</i> | <i>Antal 'ord i alt'</i> |
|-----------------|-----------------------|--------------------------|--------------------------|
| 'En god vinter' | 15 | 305 | 1.292 |
| 'Marmorkirken' | 30 | 387 | 2.136 |
| 'Sundhed' | 10 | 381 | 1.392 |

Tabel 4.3.1.b viser, at informanterne har fået dobbelt så lang tid til 'Marmorkirken', men antal 'ord i alt' er langt fra dobbelt. Ordformer er kun en anelse flere end i den sidste opgave, selv om informanterne har haft tre gange så lang tid til at producere den. Sammenligner man derimod den første og den sidste opgave uden at kigge på de andre, kan man nemt komme til den konklusion, at informanterne har gjort så store fremskridt i løbet af de tolv uger, at det kan måles på den måde – flere ord på mindre tid. For at lege lidt med tallene kan man antage, at hvis informanterne havde fået det samme antal minutter til den sidste opgave som den første, ville de være nået til knap 2.100 ord. Hvorfor når informanterne så ikke at skrive mindst 2.600 ord eller snarere omkring 4.000 ord i opgaven om 'Marmorkirken'?

Det er nærliggende at antage, at der må være en forskel i selve opgaverne, som medfører, at der er andre ting informanterne har at slås med i 'Marmorkirken' end 'Sundhed'. Det ser ud til, at det er sværere for informanterne at gengive den røde tråd fra en novelle end skrive ud fra ideassociationer. Ganske vist har novellen 'Marmorkirken' en svær komposition og plottet er indviklet, men der er dog på forhånd givet både indhold og ordforråd. Man kunne regne med, at det ville hjælpe en hel del på vej i sprogproduktionen, men tværtimod ser det ud til, at den frihed ideassociationer byder på i

komposition, ikke bare gør det nemmere for informanterne at producere en forholdsvis lang tekst i løbet af kort tid, men også gør det lettere at frembringe et større antal ordformer.

Opgaver nr. 2 og 3 i rækkefølgen ('*Andrea*' og '*Talesprog*') er for det første meget længere end de tre andre, sammenlagt ca. 70% af hele sprogproduktionen (jf. fig. 4.3.1.a). For det andet er de af vidt forskellig karakter, da den ene er skriftlig og den anden mundtlig. Den ene er lavet i undervisningstimerne, den anden udenfor. I den ene fik informanterne afmålt tid, i den anden ubegrænset tid (dvs. den tid informanterne mente de havde til rådighed fra andre pligter). Den ene ('*Andrea*') bygger på en roman, og opgaven kræver et kendskab til begivenhedernes rækkefølge (=romanens indhold) foruden ordforrådet der hører til, men opgaven giver også en vis frihed i brug af synsvinkel og tilkendegivelse af egne meninger. Den anden opgave ('*Talesprog*') bygger kun på ordforråd og materialer fra to af semestrets temaer (rejser og dagligdags problemer / aktiviteter), ellers var der frit spil angående det endelige produkt.

Den første opgave, '*Andrea*' havde godt 90% af 'antal ordformer' i forhold til '*Talesprog*', men knap 70% af 'antal ord i alt'. '*Talesprog*' er den eneste opgave, der har nogenlunde samme forhold 'ordformer' og 'antal ord i alt', som andel af helhedsordforrådet i begge grupper (48% / 42%). Det kan skyldes to grunde. For det første er talesprog ikke så koncentreret som skriftsprog, der er flere gentagelser, fyldeord/tøveord, mindre sammenhæng, kropssprog og mimik eller gestus, som kan bruges, når ordforrådet ikke rækker til. For det andet gælder det, at jo længere teksten bliver, jo lavere T/T ratio. Jo kortere tekst, jo højere T/T ratio, (jf. 3.5). En anden bemærkelsesværdig ting er, at opgaven om romanen ('*Andrea*') viser et bedre resultat hos informanterne end opgaven om novellen ('*Marmorkirken*'). Ganske vist giver den en lavere T/T ratio, men hvis tidsfaktoren tages i betragtning, ser det lidt anderledes ud.

Tabel 4.3.1.c Sammenligning af opgavers længde i forhold til tidsfaktoren

| | Tid i minutter | Antal 'ordformer' | Antal 'ord i alt' |
|--------------------------|----------------|-------------------|-------------------|
| ' <i>En god vinter</i> ' | 15 | 305 | 1.292 |
| ' <i>Andrea</i> ' | 120 | 754 | 4.439 |
| ' <i>Marmorkirken</i> ' | 30 | 387 | 2.136 |

Tabel 4.3.1.c viser de tre opgaver i kronologisk rækkefølge. På forhånd ville man måske regne med, at informanterne ville kunne skrive forholdsvis meget på en halv time, men når tiden bliver længere bliver produktionen langsommere, mere tid går til spekulationer og brug af hjælpemidler. Den store forskel på en opgave på 30 minutter og 120 minutter skulle snarere berøre produktionens kvalitet, færre stavfejl, færre grammatiske fejl, bedre komposition. – Kort sagt, en anelig del af tiden ville gå med at skrive udkast og bruge hjælpemidler. Forskellen mellem en opgave på 15 minutter og 30 minutter skulle først og fremmest ligge i længden. Den forholdsvis store forskel skulle ligge mellem 'Vinter' og 'Marmorkirken', men ifølge tabellen her nedenunder er forskellen mindst mellem de to opgaver, ikke bare hvad længde angår, men også hvad antal ordformer angår. Disse resultater kunne umiddelbart friste én til at konkludere, at skriftlige opgaver omkring noveller er noget af det sværeste man kan byde sine elever, hvis man vil vurdere opgaver ud fra kriterier der angår ordforråd.

Tabel 4.3.1.d Forhold mellem tid og produktion af ordformer og ord i alt.

| | Tid til rådighed | Produktion af ordformer | Produktion af ord i alt |
|---------------|------------------|-------------------------|-------------------------|
| Vinter/Andrea | 8 x 15 min. | 1/2,5 ordformer | 1/3,4 ord i alt |
| Vinter/Marmor | 2 x 15 min. | 1/1,3 ordformer | 1/1,7 ord i alt |
| Marmor/Andrea | 4 x 30 min. | 1/1,9 ordformer | 1/2,1 ord i alt |

Tabel 4.3.1.d viser, en måde at sammenligne de forskellige opgaver på. Informanterne havde 8 gange så lang tid til at skrive opgaven 'Andrea', som de havde til opgaven 'Vinter'. Ordformer bliver dog kun 2,5 gange så mange, og ord i alt bliver kun 3,4 gange så mange. På samme måde kan man vurdere forholdene mellem 'Vinter' og 'Marmor' på den ene side og 'Marmor' og 'Andrea' på den anden side.

Tabel 4.3.1.e. Det højeste og laveste antal ordformer i hver enkelt opgave. I midten ligger gennemsnitstal for ordformer i alle opgaver.

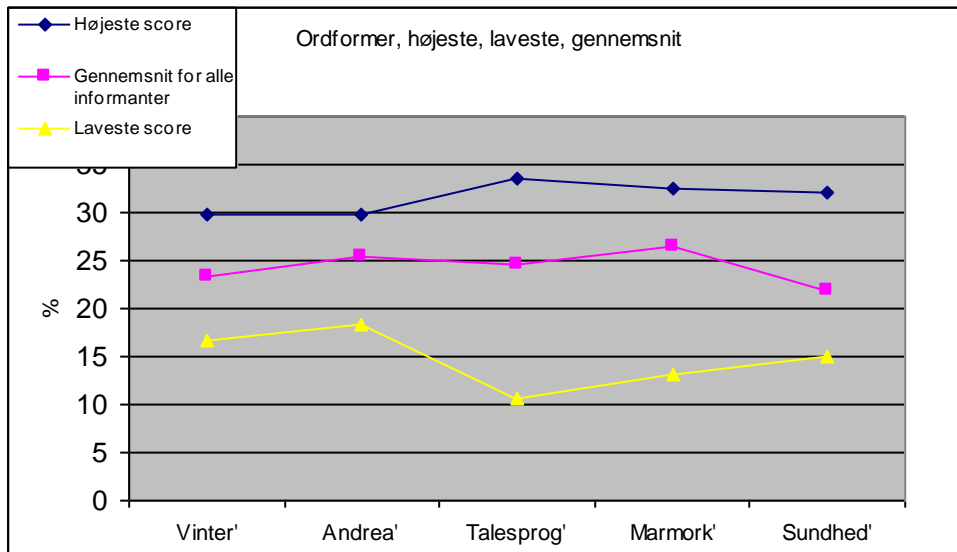
| | 'Vinter' | 'Andrea' | 'Talesprog' | 'Marmor.' | 'Sundhed' |
|--|-------------------|-------------------|-------------------|-------------------|-------------------|
| Flest ordformer / helhed | 91/305 | 225/754 | 276/824 | 127/387 | 122/381 |
| Procentdel af alle ordformer i enkelte opgaver. Højeste score. | 29,8% (Inf. 9) | 29,8% (Inf. 8) | 33,5% (Inf. 4) | 32,9% (Inf. 9) | 32,0% (Inf. 3) |
| Procentdel af alle ordformer i enkelte opgaver Gennemsnit pr. informant | 23,3% | 25,5% | 24,6% | 26,4% | 21,8% |
| Procentdel af alle ordformer i enkelte opgaver. Laveste score. | 16,7% (Inf. 1) | 18,3% (Inf. 7) | 10,6% (Inf. 6) | 13,2% (Inf. 1) | 15,0% (Inf. 5) |
| Færrest ordformer / helhed | 51/305 | 138/754 | 87/824 | 51/387 | 57/381 |

Tabel 4.3.1.e viser, hvor høj procent af det samlede antal ordformer de højest placerede informanter har, hvor lav procent de lavest placerede informanter har og hvordan de informanter, der har den højeste og laveste score, forholder sig til gennemsnittet. Dvs. forholdet mellem antal ordformer i enkelte opgaver og antal ordformer i alt i hver enkelt opgave. Det samlede antal ordformer i 'Vinter' hos alle informanter er 305, men informant 9, der har den højeste score, har fx 91 ordformer i 'Vinter', eller 29,8%.

Størst forskel er der i 'Talesprog', hvor informant 6 kun har 43,1% af gennemsnittet og kun 31,5% af det informant 4, topscoreren, har.

Det viser sig også, at der er en stor spredning blandt informanterne. Ingen af dem stiller sig både øverst og nederst. Begge steder er der en informant, som har rekord i to opgaver. Fire informanter skiftes til at stille sig øverst og andre fire nederst. Kun en informant (nr. 2) ud af de ni undgår de yderste grænser.

Figur 4.3.1.b Antal ordformer – yderste værdier og gennemsnit for hver opgave.



Figur 4.3.1.b viser fordelingen af ordformer (jf. tabel 4.3.1.e). Den højeste score går højest i 'Talesprog', men størst forskel er der i 'Sundhed', hvor gennemsnittet har den laveste værdi mens topscoreren holder sig på højt niveau. De to første opgaver 'En god vinter' og 'Andrea' ligger tættest med hensyn til højeste og laveste score i forhold til gennemsnitsværdien. Den laveste score er lavest i 'Talesprog' og længst fra

gennemsnittet af alle værdier. I 'Talesprog' og 'Marmorkirken' er der informanter, der går langt ned i forhold til gennemsnittet, men i 'Sundhed' er det gennemsnittet, der er lavest. Der er en form for balance i de to første opgaver, i de næste to ligger gennemsnittet nærmere toppen end bunden. Det viser, at de laveste er ekstremere, men det omvendte er tilfældet i den sidste opgave, hvor gennemsnittet er nærmere bunden end toppen, så er det topscoreren der er ekstremisten.

På samme måde kan der laves en oversigt for procentdel af ord i alt og på den måde sammenligne de enkelte opgaver med hinanden.

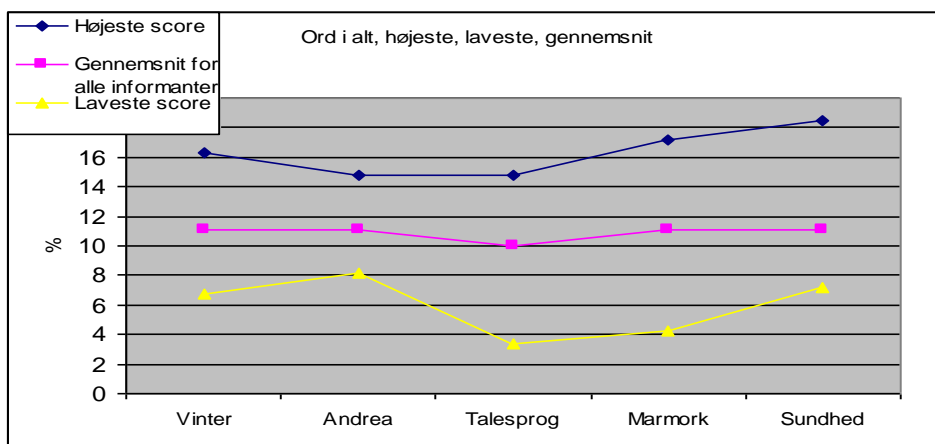
Tabel 4.3.1.f Det højeste og laveste antal ord i alt i hver enkelt opgave. I midten ligger gennemsnitstal for ord i alt i alle opgaver.

| | 'Vinter' | 'Andrea' | 'Talesprog' | 'Marmork.' | 'Sundhed' |
|--|-------------------|-------------------|-------------------|-------------------|-------------------|
| Flest ord i alt | 211/1292 | 656/4441 | 985/6664 | 367/2136 | 257/1392 |
| Procentdel af alle ord i enkelte opgaver. Højeste score. | 16,3% (Inf. 9) | 14,8% (Inf. 9) | 14,8% (Inf. 2) | 17,2% (Inf. 8) | 18,5% (Inf. 3) |
| Gennemsnit pr. informant | 11,1% | 11,1% | 10% ¹⁹ | 11,1% | 11,1% |
| Procentdel af alle ord i enkelte opgaver. Laveste score. | 6,7% (Inf. 1) | 8,1% (Inf. 7) | 3,4% (Inf. 6) | 4,2% (Inf. 1) | 7,2% (Inf. 1) |
| Færrest ord i alt | 87/1292 | 360/4441 | 229/6664 | 89/2136 | 100/1392 |

Tabel 4.3.1.f viser, hvor meget de højeste og laveste værdier afviger fra gennemsnittet. Værdierne for 'Andrea' ligger tættest op ad hinanden. Det største udsving viser sig i værdierne for 'Marmorkirken'. Ligesom i ordformer (tabel 4.3.1.e) svinger 'Sundhed' længst op over gennemsnittet.

¹⁹ Gennemsnittet er 1/9 del af hele ordforrådet undtagen i 'Talesprog' hvor der blev afleveret 10 opgaver – derfor 10% her mod 11,1% andre steder.

Figur 4.3.1.c viser højeste, laveste og gennemsnitsantal ord i hver opgave

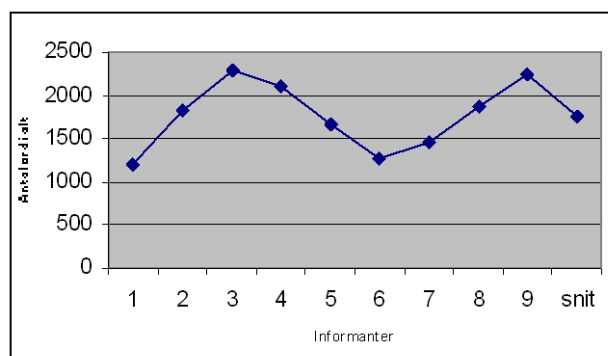
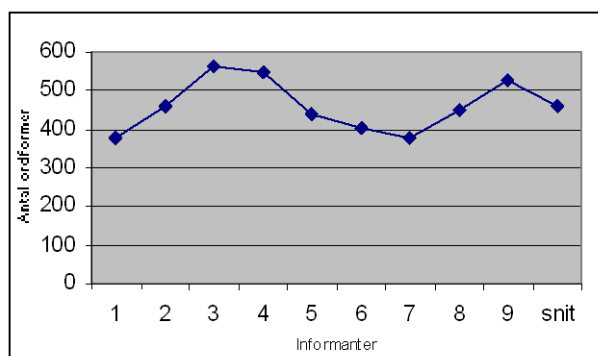


Tabellen (4.3.1.f) og figur 4.3.1.c viser i grunden, hvordan ekstremisterne holder sig til gennemsnittet. Det siger dog ikke noget om, hvor mange der ligger i nærheden af gennemsnittet eller om det kun er en informant, der afviger ekstremt meget.

Igen her er det fire informanter, der deler toppladserne, men kun tre ligger i bunden. Informant 1 er den, der producerer mindst tekst i det hele taget jf. tabel 4.3.1.f

Dette diagram minder en del om det med ordformerne, tallene er dog lavere og hvert procentpoint vejer mere i lave tal end høje. Begge steder (4.3.1.b og c) er det tydeligt, hvor værdierne ligger tæt op ad hinanden i 'Andrea'. Måske er det resultatet af informanternes gruppeprojekter (jf. afsnit 4.1.5.2), hvor grupperne præsenterer deres arbejde for andre i klassen, som gør sig gældende her.

Figur 4.3.1.d Viser enkelte informanternes sprogproduktion og forholdet mellem antal ordformer og antal ord i alt.



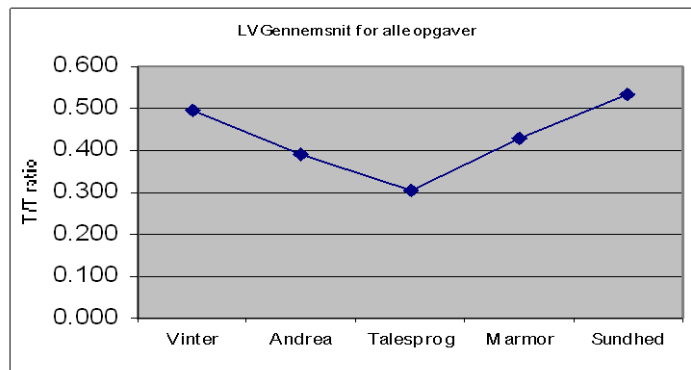
Figur 4.3.1.d viser hvor forskellige de enkelte informanter er når hele sprogproduktionen tælles sammen. Tre informanter ligger under 1.500 ord, med informant 1 som den laveste. Tre informanter ligger mellem 1.500 og 2.000 og de sidste tre går over de 2.000 ord.

Der er store ligheder med de to grafer 'ordformer' og 'antal ord', selv om de ikke ligger på samme målestok. Forskellen mellem højeste og laveste værdi er dog 186 ord. Værdierne varierer fra 377 til 563.

4.3.2 Leksikalsk variation. T/T ratio

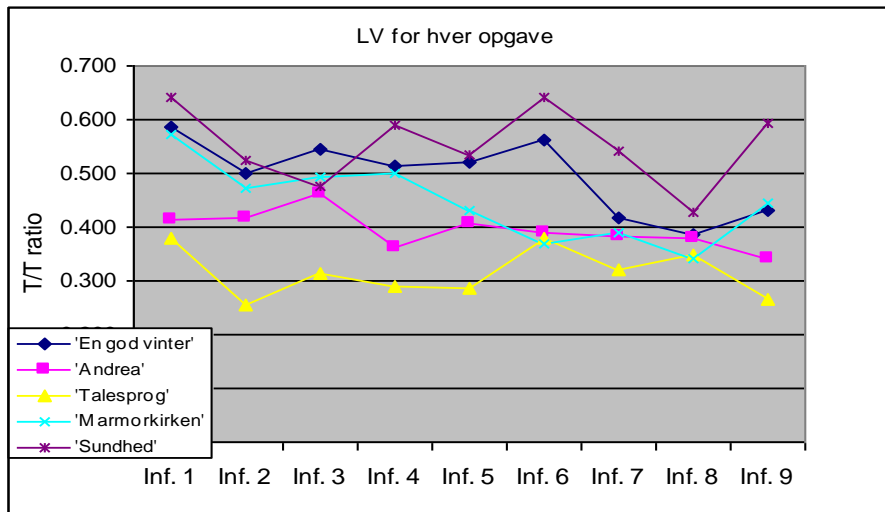
I dette afsnit vil jeg gennemgå informanternes LT (T/T ratio) for gruppen som helhed, for enkelte opgaver og for enkelte informanter.

Figur 4.3.2.a Viser leksikalsk variation – gennemsnit for hver af de fem opgaver.



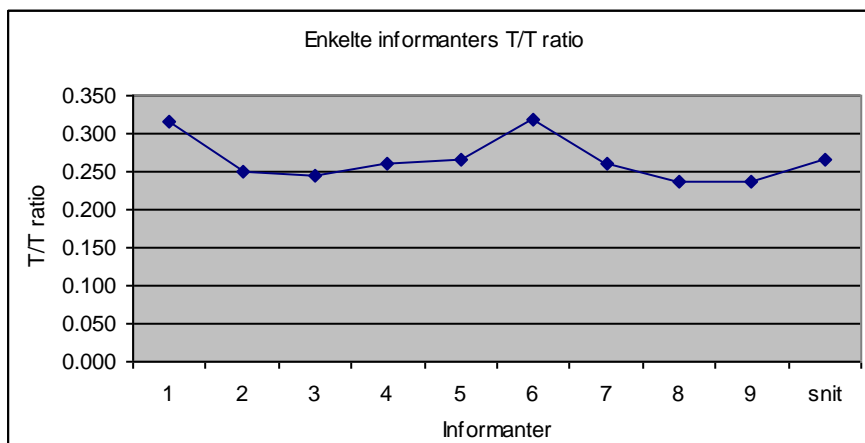
Figur 4.3.2.a viser gennemsnit i 'Leksikalsk variation' for alle opgaver. T/T ratioen er konsekvent i omvendt forhold til teksternes længde undtagen for 'Sundhed'. Det vil sige, at for fire af opgaverne gælder det, at jo kortere teksterne er, jo højere er T/T ratioen. Her er ordforrådet sorteret efter opgaver og ikke enkelte informanter. Grafen viser, hvordan forskellige opgavetyper fungerer indenfor ordforrådet som helhed. Som det fremgår af afsnit 4.3.1, ser det ud til, at opgavetyperne spiller mere afgørende rolle end tidsfaktoren i sprogproduktionen hos informanterne som gruppe, selvom der er individuelle variationer.

Figur 4.3.2.b Leksikalsk variation - alle opgaver – alle informanter.



Figur 4.3.2.b viser T/T ratio for de forskellige opgaver hos hver enkelt informant. Her kan man se de individuelle variationer. Informanter 8 og 9 er påfaldende forskellige, hos informant 8 har alle opgaver næsten den samme T/T ratio, men hos informant 9 er der kun to opgaver, der falder nogenlunde sammen og de resterende tre har stor spredning. Informanter 2 og 6 har tre opgaver, der falder sammen. Mønstret er meget uregelmæssigt hvilket er ensbetydende med at informanterne er meget forskellige indenfor de enkelte opgaver. Det er svært at få øje på en fællesnævner her. Der er ikke to opgaver, der viser parallelt mønster i diagrammet.

Figur 4.3.2.c T/T af hele sprogproduktionen hos hver enkelt informant + gennemsnit.



Figur 4.3.2.c viser det mønster, der fremkommer når alle linjerne fra figur 4.3.2.b bliver sat sammen til en. Her fremgår det tydeligt, at to informanter, 1 og 6, er dygtige til at bruge det ordforråd, de har til rådighed. De andre syv afviger ikke meget fra linjen, som viser T/T ratioen 0,250. Den tiende størrelse, som viser gennemsnits-ratioen, er taget med for at vise hvor meget informanter 1 og 6 gør for at forhøje gennemsnittet.

Tabel 4.3.2.a Oversigt over hele sprogproduktionen i tal, samt gennemsnitsværdi pr. informant.

| | Inf. 1 | Inf. 2 | Inf. 3 | Inf. 4 | Inf. 5 | Inf. 6 | Inf. 7 | Inf. 8 | Inf. 9 | Snit |
|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|
| Ordformer | 378 | 459 | 563 | 546 | 439 | 406 | 377 | 448 | 530 | 461 |
| Ord i alt | 1201 | 1829 | 2287 | 2103 | 1658 | 1274 | 1451 | 1881 | 2241 | 1769 |
| T/T ratio | 0,315 | 0,251 | 0,246 | 0,260 | 0,265 | 0,319 | 0,260 | 0,238 | 0,237 | 0,266 |

Tabel 4.3.2.a viser de hårde tal, der ligger til grundlag for figur 4.3.2.c. Der kan man tydeligt se, at informanter 1 og 6 ligger som nummer 2 og 3 i antal ordformer og nummer 1 og 2 - fra bunden. Men deres T/T ratio ligger nummer 2 og 1 - på toppen.

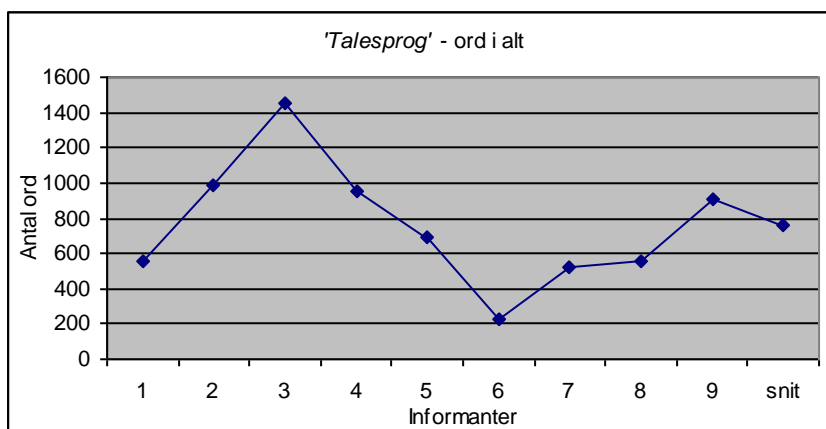
De resultater passer meget godt med det, forskere (Linnarud, Golden, Laufer & Nation jf. afsnit 3.5) har sagt om denne metode, nemlig at forskellig længde af tekster giver forskellige resultater.

Informant 3 har 45% flere *ordformer* end informant 7. Informant 7 har 82% af gennemsnitsantal *ordformer*. Informant 3 er 22% over gennemsnitsantal *ordformer*.

Informant 3 har 90% flere *ord i alt* end informant 1. Informant 1 har 68% af gennemsnitsantal *ord i alt*. Informant 3 er 77% over gennemsnitsantal *ord i alt*.

Informant 3 har den særstilling indenfor gruppen, at hun sammen med sin partner (ikke en af informanterne), afleverede en talesprogsopgave dobbelt så lang, som det forlangte. Ganske vist var der andre informanter, som overskred det, der blev bedt om i deres talesprogsproduktion, men ikke i så høj grad som informant 3. Denne store produktion forhøjer gennemsnittet og gør det sværere for dem, der holdt sig til det forlangte, eller ikke magtede det helt at stå mål med gennemsnitstallene. I denne opgave findes der de største individuelle forskelle i sprogproduktionen. Hvor meget så skyldes ambitioner eller interesse for faget eller at nogle af eleverne fandt opgaven morsom er svært at afgøre.

Figur 4.3.2.d 'Talesprog' Hver enkelt informants sprogproduktion.



Figur 4.3.2.d er taget med for at vise de store udsving i sprogproduktionen indenfor denne ene opgave. Informant 6 har 229 ord mens informant 3 har 1.456 ord (fordelt på to opgaver) Gennemsnittet er 763 ord. Se nærmere fig. 4.3.1.c og tabel 4.3.1.f.

4.3.3 Ni informanter – to opgaver

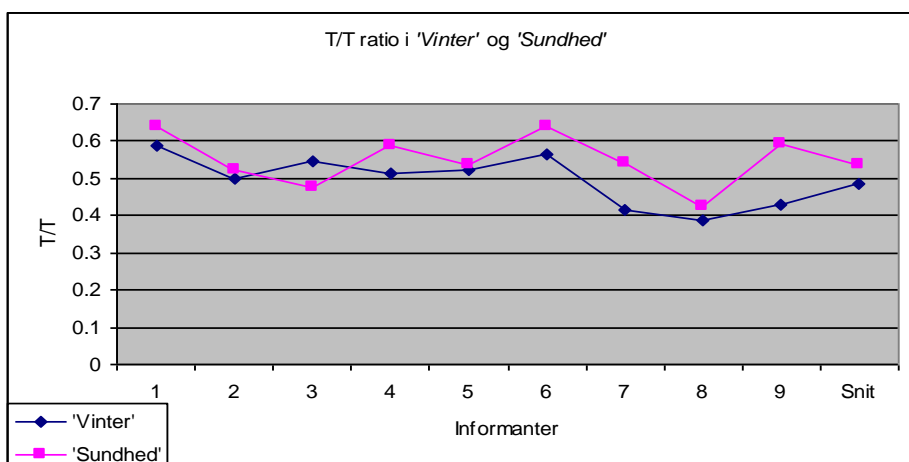
Da de to opgaver, der bygger på associationer, den første og den sidste, ser ud til at være de eneste to, der er muligheder for at sammenligne med hinanden, har jeg valgt at se dem lidt efter i sømmene.

Tabel 4.3.3.a T/T ratio i to opgaver 'En god vinter' (V) og 'Sundhed' (S)

| Informant | Ordformer/ord i alt | | T/T ratio | | Antal ord/sætninger | | Antal ord i sætning | |
|-------------|---------------------|---------|-----------|-------|---------------------|--------|---------------------|------|
| | V | S | V | S | V | S | V | S |
| Informant 1 | 51/87 | 64/100 | 0,586 | 0,640 | 87/5 | 100/11 | 17,4 | 9,1 |
| Informant 2 | 60/120 | 65/124 | 0,500 | 0,524 | 119/7 | 123/11 | 17,0 | 11,2 |
| Informant 3 | 90/165 | 122/257 | 0,545 | 0,475 | 165/7 | 258/13 | 23,6 | 19,8 |
| Informant 4 | 72/140 | 83/141 | 0,514 | 0,589 | 140/13 | 141/12 | 10,8 | 11,8 |
| Informant 5 | 71/136 | 57/107 | 0,522 | 0,535 | 136/13 | 101/14 | 10,5 | 7,2 |
| Informant 6 | 71/126 | 87/136 | 0,563 | 0,640 | 126/10 | 136/13 | 12,6 | 10,5 |
| Informant 7 | 62/149 | 78/144 | 0,416 | 0,542 | 149/8 | 143/9 | 19,8 | 15,9 |
| Informant 8 | 61/158 | 88/206 | 0,386 | 0,427 | 158/9 | 206/12 | 17,6 | 17,2 |
| Informant 9 | 91/211 | 105/177 | 0,431 | 0,593 | 211/16 | 173/19 | 13,2 | 9,1 |
| Snit | 70/144 | 83/155 | 0,486 | 0,535 | 143/10 | 153/13 | 14,3 | 11,8 |

Tabel 4.3.3.a viser T/T ratio for de to tekster ('En god vinter' og 'Sundhed') hos de ni informanter. Den første spalte viser informanternes *nummerering*. De andre spalte er delt i to, den til venstre angiver tal for 'En god vinter' mens den til højre (*kursiv*) angiver tal for 'Sundhed'. Anden og tredje 'dobbeltspalte' viser forholdet mellem antal *ordtyper* og antal *ord i alt* i hver tekst. Fjerde og femte 'dobbeltspalte' viser *gennemsnitslængde* af sætninger i hver tekst. Nederste linje indeholder gennemsnitstal for hver spalte. Af tabellerne fremgår, at alle informanter, undtagen én (nr. 4), skriver kortere sætninger i den sidste opgave. Disse oplysninger giver et fingerpeg om, at informanterne muligvis har lidt mere styr på formuleringen af det de skriver i den sidste opgave (jf. bilag 1:i-xix).

Figur 4.3.3.a Sammenligning af T/T ratio i de to opgaver.



Figur 4.3.3.a viser T/T ratioen i de to opgaver og her ses der tydelig forskel mellem de to opgaver. Det er kun informant 3, der ikke forøger sin T/T ratio. Informant 5 har næsten den samme T/T ratio, mens informant 7 og 9 scorer højest.

Tablet 4.3.3.b Ændring i T/T ratio over tid og ændring i antal ord over tid i % hos alle informanter – og gennemsnit for begge størrelser.

| | Inf. 1 | Inf. 2 | Inf. 3 | Inf. 4 | Inf. 5 | Inf. 6 | Inf. 7 | Inf. 8 | Inf. 9 | Snit |
|-----------------------------------|--------|--------|-------------|--------|--------------|--------|--------|-----------|-------------|------|
| Ændring i T/T over tid i % | 14,9 | 14,4 | -9,5 | 12,9 | 5,5 | 16,8 | 34,5 | 14,8 | 46,4 | 16,7 |
| Ændring i % af antal ord over tid | 13,6 | 4,2 | 56 | 1,4 | -25,7 | 7,9 | -3,4 | 31 | -15,6 | 7,7 |

Tabel 4.3.3.b viser, at seks ud af ni informanter skriver flere ord i den sidste tekst til trods for at de havde kortere tid til rådighed. Kun to informanter, 5 og 9, skriver betydeligt kortere tekster. En informant, 3, scorer lavere T/T ratio, men skriver samtidigt den længste tekst. Informanter 7 og 9 har den største fremgang i T/T ratio, samtidig med at de skriver en kortere tekst. Informant 5 har en ubetydelig fremgang, men skulle egentlig være den store højdespringer, hvis hun fulgte det samme mønster, som informanter 7 og 9, fordi hun ligger højt i forvejen – i den første opgave.

Det er fristende at tale om udvikling hos informanterne i løbet af de 13 uger, men der må tages nogle forbehold. Ved semestrets start havde eleverne ikke beskæftiget sig med dansk i tre måneder. En lille tekst, skrevet i løbet af 15 minutter, kan umuligt vise hele deres ordforråd. Den viser i grunden kun, hvad informanterne kunne præstere i skrivende stund. Den første opgave kan siges at være nemmere i den forstand, at informanterne kunne skrive om sig selv og deres forhold til begrebet vinter og de kunne derfor holde sig til noget, de kender ret godt. Den sidste opgave omfatter et sværere emneområde og mere abstrakte begreber. Informanterne vidste i forvejen, at denne opgave (dvs. at skrive om et af semestrets temaer) drejede sig om at bruge et så varieret ordforråd, som muligt indenfor det tema, de skulle dække. T/T ratio målingen viser, at de alle på nær en brugte mere varieret ordforråd i denne opgave og at de fleste skrev en længere tekst.

Målingen siger dog ikke noget om, hvor bredt informanternes ordforråd er, hvor mange ordfamilier de kender eller hvor primitivt eller avanceret deres ordforråd er.

Denne metode er ikke valid for tekster af forskellig længde og heller ikke for tekster, der er kortere end 200 ord (Laufer & Nation 1995). Derfor kan mine resultater ikke tillægges den store videnskabelige værdi, men skal først og fremmest betragtes som et forsøg på at vise informanternes leksikalske variation på to forskellige tidspunkter, informanternes indbyrdes forhold og den eventuelle udvikling, der er sket i tidsrummet mellem skrivningen af de to opgaver. Sammenligning af informantgruppens gennemsnit i T/T i de to opgaver viser, at forskellen er signifikant; $p < 0.05$.

Signifikansen bliver endnu mere markant i de følgende eksempler, hvor teksterne hos tre af informanterne bliver forkortede til 120 ord.

Ved at se nærmere på informanter 2, 4, og 7 (jf. tabel 4.3.3.a), der har **mindst** forskel i længde af første og sidste opgave, kan der konstateres en betydelig forhøjelse af T/T

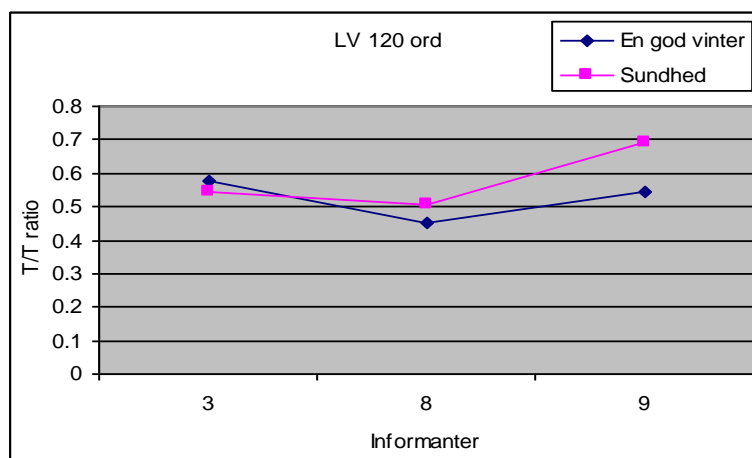
ratio, især hos informant 7, hele 30%. Hvad så med de andre, der har lavet længere tekster, kan der måles en større forhøjelse i deres T/T ratio, hvis deres tekster bliver forkortet til den samme længde? For at opnå et mere konkret sammenligningsgrundlag blev de **længste** tekster, hos informanter 3, 8 og 9, klippet af ved 120 ord.

Tabel 4.3.3.c T/T Ratio i de originale tekster versus de forkortede tekster – forskel i %

| Informant | 'En god vinter' | | | 'Sundhed' | | |
|-------------|-----------------|-----------------------|-------------|-----------------|-----------------------|-------------|
| | Original længde | Forkortet til 120 ord | Forskel i % | Original længde | Forkortet til 120 ord | Forskel i % |
| Informant 3 | 0,545 | 0,575 | 5,5% | 0,481 | 0,542 | 12,7% |
| Informant 8 | 0,398 | 0,450 | 13,1% | 0,437 | 0,508 | 16,2% |
| Informant 9 | 0,436 | 0,542 | 24,3% | 0,622 | 0,692 | 11,3% |

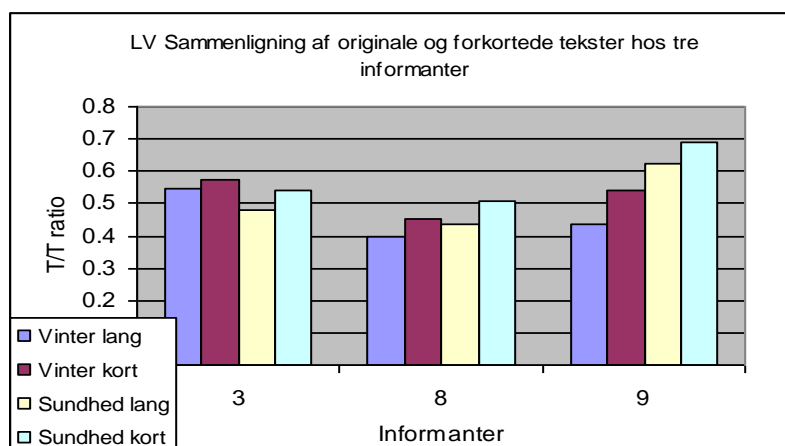
Tabel 4.3.3.c viser en højere T/T ratio, når alle tekster er lige lange, hvilket passer meget godt med at jo kortere teksten er, jo højere bliver T/T ratioen. Informanternes resultater kan nu sammenlignes på samme grundlag, ikke bare når man ser på opgaverne hver for sig, men også ved at konstatere, at forskellen på T/T ratioen mellem de to opgaver er markant større end ved måling af teksterne i den oprindelige længde.

Figur 4.3.3.b Tekster skåret ned til 120 ord hos tre informanter



Figur 4.3.3.b viser, at informant 3 ligger lavere i 'Sundhed', mens de andre gør det omvendte, dvs. har et højere forhold mellem ordformer og ord i alt i den sidste opgave.

Figur 4.3.3.c LV Sammenligning af originale og forkortede tekster hos tre informanter.



Figur 4.3.3.c viser, at informant 3 har den mindste forskel mellem den lange og korte udgave af den første opgave. Alle har nogenlunde den samme forskel på T/T ratio mellem den lange og den korte udgave af den sidste opgave (jf. tabel 4.3.3.c).

Selvom disse tre informanter har skrevet de længste tekster begge gange, er de dog meget forskellige med hensyn til T/T ratio-målingen. Der er en lille tilbagegang hos 3, en lille fremgang hos 8, mens 9 har den største fremgang af samtlige informanter.

4.3.3 Lige lange opgaver

Jeg valgte opgaver fra informanter 2, 6 og 9, fordi de forekommer mig at være de mest forskellige blandt informanterne. Informant 2 har middelhøj T/T ratio i opgaverne, forholdsvis korte tekster, undtagen '*Talesprog*'. Informant 6 har høj T/T ratio i alle opgaver, varierende længde i tekster (i forhold til andre), kort '*Talesprog*'. Informant 9 har de fleste ordformer og alle tekster er lange.

Deres respektive 5 opgaver blev skåret ned til 120 ord hver. Størrelsen 120 blev valgt, fordi de korteste opgaver netop var af den længde. '*Vinter*' hos informant 2 nåede lige de 120 ord og derfor er der ingen forskel i T/T ratio i tabellen for kort og lang '*vintertext*'.

Jeg valgte at tage de første 120 ord fra hver tekst. Det giver måske ikke det mest retfærdige billede af informantens formåen da det er muligt, at han/hun enten begynder sit skriverseri dårligt med en masse gentagelser og så kommer godt i gang – eller det kan være omvendt, det bedste kommer i starten og resten er gentagelser. En del af teksterne er

dog så korte, at de ikke byder på meget valg. Laufer & Nation (1995) anbefaler at skære ved 200 ord for at få pålidelige resultater, men af ovennævnte grunde var det ikke muligt, hvis alle opgaver skulle med.

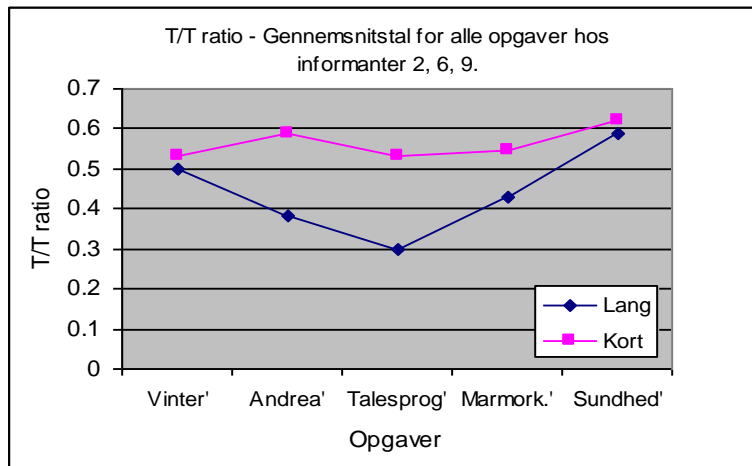
Tabel 4.3.4.a T/T ratio henholdsvis i de originale opgaver i den oprindelige længde og de forkortede tekster på 120 ord, samt gennemsnit for alle værdier.

| | | 'Vinter' | 'Andrea' | 'Talesprog' | 'Marmork' | 'Sundhed' | Gennemsnit |
|-------------|------|----------|----------|-------------|-----------|-----------|--------------|
| Informant 2 | Lang | 0,500 | 0,417 | 0,255 | 0,474 | 0,524 | 0,434 |
| | Kort | 0,500 | 0,617 | 0,517 | 0,542 | 0,533 | 0,542 |
| Informant 6 | Lang | 0,563 | 0,391 | 0,380 | 0,368 | 0,640 | 0,467 |
| | Kort | 0,558 | 0,558 | 0,517 | 0,500 | 0,658 | 0,558 |
| Informant 9 | Lang | 0,431 | 0,337 | 0,267 | 0,444 | 0,593 | 0,415 |
| | Kort | 0,542 | 0,592 | 0,558 | 0,592 | 0,675 | 0,592 |
| Gennemsnit | Lang | 0,498 | 0,382 | 0,301 | 0,429 | 0,586 | 0,439 |
| | Kort | 0,533 | 0,589 | 0,531 | 0,545 | 0,622 | 0,564 |

Tabel 4.3.4.a viser de forandringer, der bliver i T/T ratioen i alle opgaver for hver af de tre informanternes vedkommende, nr. 2, 6, og 8. I tabellen kan det tydeligt ses, at alle informanter har højere T/T ratio i de forkortede tekster end i de originale. Der er to undtagelser, begge to i opgaven 'Vinter'. Den ene hos informant 2, hvor den lange og den korte tekst er identiske, den anden hos informant 6, hvor ratioen er lavere i den korte udgave. De to tal understreger, at det ikke er lige meget, hvor man skærer af teksten når den skal forkortes.

Gennemsnitstallene har en stor interesse her. Vi har set på figur 4.3.2.a, at T/T ratioen hos alle informanter var lavest for 'Talesprog' og bliver højere i rækkefølgen 'Andrea', 'Marmorkirken', 'Vinter' og 'Sundhed'. Det samme gælder for informanter 2, 6 og jf. tabel 4.3.4.a. Gennemsnitstallene for de forkortede tekster viser et andet mønster; 'Talesprog' ligger stadigvæk lavest. 'Marmorkirken' i midten og 'Sundhed' højest, mens 'Vinter' og 'Andrea' bytter pladser. Det er dog ikke rækkefølgen, der siger det meste, men det at gennemsnitstallene for de forkortede tekster ligger meget tæt på hinanden.

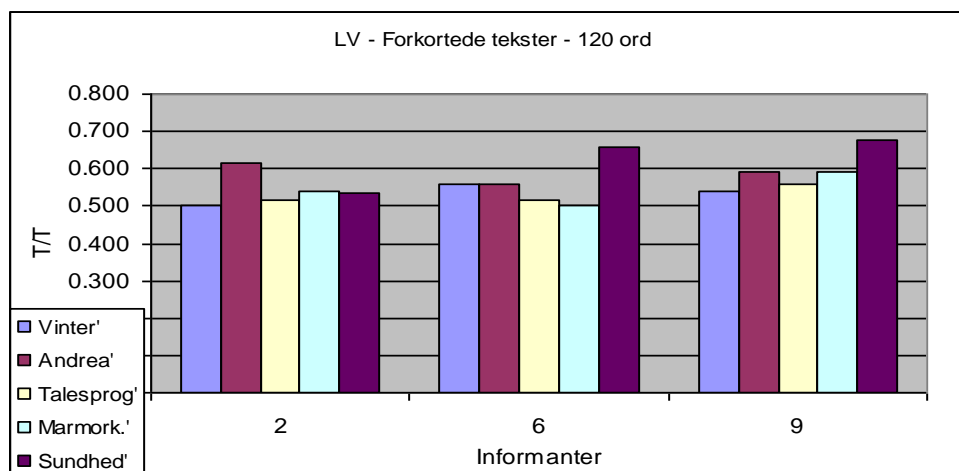
Figur 4.3.4.a Forskel i gennemsnitstal for lange og korte tekster hos tre informanter.



Figur 4.3.4.a viser kronologisk udvikling i leksikalsk variation, hvor alle tekster er skåret ned til 120 ord. De tre informanter tegner hver sit billede. Informant 2 viser et stort opsving i opgave to ('*Andrea*'), men ellers sker der ikke meget. Hos informant 6 går det nedad med opgave tre ('*Talesprog*') og fire ('*Marmorkirken*'), men den sidste opgave ('*Sundhed*') viser stor fremgang, der virker lidt ulogisk. Informant 9 viser det mønster, som man i forvejen ville forvente at se mere eller mindre hos alle informanter – en lille tilbagegang i opgave tre ('*Talesprog*'), men ellers støt fremgang. Forskere (fx Golden 2000:73) siger, at talesprog viser lavere T/T ratio end skriftsprog, men her er det næppe tilfældet. Det skyldes muligvis først og fremmest det, at informanterne ikke behersker et naturligt talesprog og derfor næsten undlader brug af tøveord og fyldeord. Her ville det have været interessant at kunne sammenligne informanternes talesprog med autentisk talesprog fra en indfødt.

Her kunne det måske også have været formålstjenligt at prøve T/T-måling af tekster med 200 ord eller måske 500 ord. I så fald ville første og sidste opgave ikke tælle med, informanterne ville højst være seks og noget af det, der følger i de næste afsnit ville blive udeladt.

Figur 4.3.4.b LV. Tre informanter – alle opgaver – forkortet til 120 ord.



Figur 4.3.4.b viser tre informanternes tekster i alle opgaver, forkortet til 120 ord. Hvis den bliver sammenlignet med fig. 4.3.2.b, hvor alle tekster vises i original længde, kan man se, at der er langt mindre forskel på T/T ratioen, når alle tekster er af samme længde. De tre informanter følger dog hver sit mønster; informant 2 når det bedste resultat i 'Andrea', mens andre opgaver er nogen lunde ens. Hos informant 6 går det ned ad i de første fire, for så at springe højt op i den sidste. Informant 9 har en lille afvigelse i opgave 2, men ellers er det opad fra først til sidst.

4.3.5 Konklusion

I afsnit 4.3 har jeg gjort rede for størrelse og inddeling af informanternes ordforråd. Jeg har anvendt en enkelt målingsmetode, (LV = leksikalsk variation) på alle opgaver fra alle informanter og derved fundet ud af, at informanterne er forskellige indbyrdes i deres sprogproduktion. Den store tendens, er at næsten alle gør fremskridt, hvis man kun ser på første ('Vinter') og sidste ('Sundhed') opgave. Det er noget man ikke kan se bort fra, selv om opgaverne måske er lidt for korte i forhold til de opgaver forskere har brugt ved udvikling og afprøvning af denne metode (jf. Nation og Laufer, afsnit 3.5 og 4.3.3).

Første og sidste opgave er i grunden også de eneste to opgaver, der kan sammenlignes hvor man ser på variablene tid og produkt. De tre andre opgaver var meget forskellige, både angående forberedelse og den tid der blev brugt til at løse dem. Det lader til, at den

største produktion fås ved at lade informanterne skrive i et bestemt antal minutter ud fra en ideassociation, hvor de selv kan bestemme over teksten, men ikke behøver at holde sig tæt til et bestemt indhold og dertil hørende ordforråd.

Når alle tekster er lige lange, bliver forskellen i T/T ratio mellem opgaverne langt mindre (0,500 – 0,675) end når alt bliver talt med (0,255-0,640), jf. figur 4.3.4.a.

Hvad kan det så bruges til? Denne metode er forholdsvis nem til at finde ud af, hvor mange ordformer eleverne bruger i forhold til hele produktionen. I undervisning, hvor elever afleverer opgaver gennem computer, tager det ikke lang tid at foretage en T/T-måling og på den måde kortlægge det kvantitative ordforråd. Eleverne kan på den måde blive mere bevidste om, hvor de står i forhold til deres kammerater og endnu mere, om de gør fremskridt i den henseende.

4.4 Afprøvning af metoder III

Leksikalsk frekvensprofil er en metode, der bliver brugt for at måle størrelsen af ordforrådet. Denne metode indebærer også en vurdering af ordforrådet. Ordforrådet bliver inddelt i frekvensbånd. Ideen er, at jo større del af ordforrådet en person kan hente fra de sjældnere frekvensbånd, jo større ordforråd har vedkommende. De hyppigste ord i sproget ligger på frekvensbånd 1-1000, dvs. omfatter de tusind mest brugte ord. I denne gruppe finder man hovedparten af de funktionsord, som ikke kan undværes i nogen tekst.

Udover det at kunne bruge de højst frekvente ord i sproget skal man helst kunne benytte sig af ord fra andre og mindre hyppige frekvensbånd. Hvis man vil kunne bruge ord der passer til forskellige anledninger, er det nødvendigt at beherske ord, som kun bruges i begrænset omfang. Derfor er det positivt og vidner om kvalitet i ordforråd, hvis man kan hente en betydelig del af sit ordforråd i de lavere (sjældnere) frekvensbånd.

I dette afsnit bliver der gjort rede for omfanget af hele ordforrådet og hvordan det inddeles på forskellige frekvensbånd. Derefter kommer en redegørelse for enkelte opgaver og der bliver gjort et forsøg på at kortlægge informanternes sprogkvalitet i henhold til de forskellige frekvenskategorier. Så kommer der en udførlig sammenligning af to opgaver, den første ('*Vinter*') og den sidste ('*Sundhed*') hos alle ni informanter. Til

sidst er der sammenligning mellem tre informanter, hvor jeg prøver at vurdere deres ordforråd i enkelte opgaver nærmere.

Til allersidst er der gjort et forsøg på sammenligne informanternes resultater med resultatet af en anden slags måling på den anden side (jf. 4.4.6 kerneord/ikke-kerneord).

4.4.1 Leksikalisk frekvensprofil (LFP)

De rammer, der bruges ved LFP måling i engelsk (jf. 3.6), kan ikke bruges i dansk. For der foreligger ikke en opdeling af det danske ordforråd i frekvensbånd på samme måde som i engelsk.

Jeg vil dog gøre et forsøg på at gå videre med ideen om inddeling i frekvensbånd på de præmisser som Dansk frekvensordbog lægger op til.

Dansk frekvensordbog (DFO) rummer ordforråd fra romaner, noveller, aviser og ugeblade fra årene 1987-1990, men den angiver kun forekomster af ortografiske ord. Frekvensen af de første 5.000 ord er angivet, men der tages ikke højde for homonymer og derfor kan fx et substantiv variere i frekvens alt efter om det står i ental, flertal, med eller uden artikel. I DFO er alle forekomster af 'lad' registreret som samme ortografiske ord selv om ordet betydningsmæssigt kan defineres på mindst 5 måder (DFO:XIII).

DFO giver frekvensrang af 5000 ord, og jeg har valgt i første omgang at inddele dem i 5 frekvensbånd på 1000 ord hvert. Det sjette frekvensbånd dækker de ord i DFO, der ikke har angivet frekvensrang og har derfor lavere frekvens end 5000. Det syvende frekvensbånd rummer de ord, der ikke findes i DFO, men dog er rigtige danske ord eller ord, der er lavet af informanterne, og er logiske og gennemsigtige (fx snehus og hjemmelektier). Den sidste kategori dækker andre hjemmelavede og 'udanske' ord eller ord, som tydeligvis har deres ophav i islandsk eller engelsk.

4.4.2 LFP af hele ordforrådet

I dette afsnit vil jeg først vise hvordan hele sprogproduktionen inddeles i kategorier, dernæst hvordan enkelte opgaver ser ud i LFP. Efter det går turen til alle informanter i to

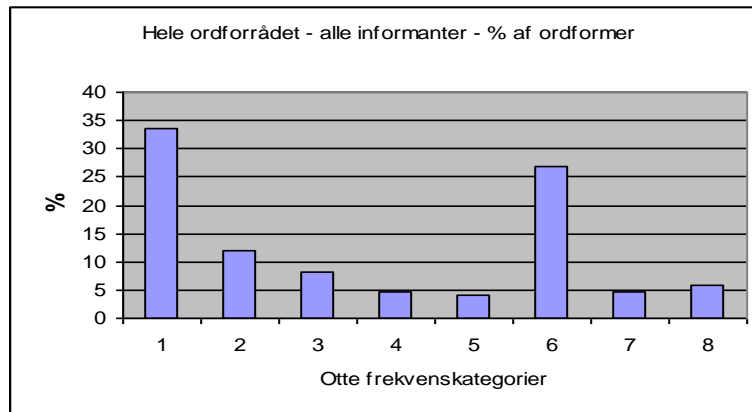
opgaver, hvor LFP bliver sammenlignet i alle kategorier i semestrets første og sidste opgave. Og til sidst tre informanternes LFP i alle opgaver.

Tabel 4.4.2.a Fordeling af ordformer i frekvenskategorier

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | Sjældnere end 5000 | Ord der ikke er med i DFO ²⁰ | Ord der falder udenfor kategorier ²¹ |
|----------------------------|--------|-----------|-----------|-----------|-----------|--------------------|---|---|
| Antal ordformer | 581 | 208 | 144 | 79 | 71 | 467 | 80 | 99 |
| Procentdel | 33,6 | 12,0 | 8,3 | 4,6 | 4,1 | 27,0 | 4,6 | 5,7 |
| Fire kategorier Procentdel | 33,6 | 29,0 | | | | 27,0 | 10,3 | |

Tabel 4.4.2.a viser, hvordan hele informanternes sprogproduktion i alle fem opgaver fordeles over de otte kategorier.

Figur 4.4.2.a Inddeling af hele ordforrådet i otte frekvenskategorier hos alle informanter – % af ordformer



Figur 4.4.2.a viser, hvordan informanternes samlede ordforråd i de fem opgaver ser ud på et søjlediagram.²²

²⁰ Ord, der ikke forekommer i Dansk frekvensordbog, men er dog enten rigtige danske ord eller kunne være det. (jfr. ”snehus” i stedet for snehytte eller snehule)

²¹ Ord, der ikke kan defineres indenfor ovenstående kategorier. Ord, der enten er islandske, engelske eller et mislykket og uigennemtsigtigt forsøg på at fremkalde et dansk ord. (jfr. ’efterdågen’ i stedet for eftermiddagen.)

²² Se bilag 2:xx-xxvi, hvor hele ordforrådet er skrevet op i frekvenskategorier.

Det forbyr ikke, at den største del af ordforrådet ligger indenfor det første frekvensbånd (de 1000 højst frekvente ord). Derimod er det overraskende, hvor mange ord der har lavere frekvens end 5000, hvorimod forholdsvis få ord ligger indenfor frekvensbånd to til fem. I tabel 4.4.2.a er frekvensbånd 2-4, dvs. for ord af frekvens 1001-5000 samlet til en størrelse. På samme måde er kategorier 7 og 8 lavet om til en størrelse. Resultatet er fire kategorier i stedet for otte. Det viser sig, at størrelserne for kategorier 2-5 udgør et tal der ligger i nærheden af kategori seks. På et diagram er det langt nemmere at sammenligne fire søjler end otte.

Hele ordforrådet i de fem opgaver er 1.729 forskellige ordformer. Godt en tredjedel af ordforrådet regnes til højfrekvente ord. I kategorier 1001 til 5000 bliver ordene færre og færre, men i kategori seks vokser antallet og bliver næsten lige så stort, som det samlede tal i de fire foregående. Her har vi et mønster der går igen, dog med små variationer i næste afsnit (4.4.3)

De to sidste kategorier udgør godt 10% af alle ordformer. Det må fremhæves, at de tekster der udgør ordforrådet i DFO nu er 13-17 år gamle. Mange nye ord er kommet ind i sproget og på tryk i mellemtiden. Man kan derfor godt regne med, at en del af de ord, der falder indenfor kategori syv, er rigtige danske ord, selvom de endnu ikke står i ordbøger. Hvis vi regner med at halvdelen af de 4,6%, som ikke er med i DFO, i virkeligheden tilhørte kategori seks, så ville de 27% vokse til 29%.

4.4.3 LFP af enkelte opgaver

I dette afsnit vil jeg regne ud LFP for hver enkelt opgave.

Tabel 4.4.3.a 'En god vinter' Antal ordformer 305 – i alle opgaver.

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | Sjældnere end 5000 | Ikke i DFO | Udenfor kategorier |
|-------------------------|--------|-----------|-----------|-----------|-----------|--------------------|------------|--------------------|
| Antal ordformer | 182 | 37 | 13 | 12 | 9 | 28 | 5 | 19 |
| Procentdel | 59,7 | 12,1 | 4,3 | 3,9 | 3,0 | 9,2 | 1,6 | 6,2 |
| 4 kategorier Procentdel | 59,7 | 23,3 | | | | 9,2 | 7,8 | |

Tabel 4.4.3.a viser måske det billede af ordforrådsinddelingen, som man i forvejen ville regne med. Knap 60% i højeste frekvenskategori, næsten en fjerdedel i de næste fire kategorier og knap 10% sjældnere end 5000. Når man ved, at denne opgave blev skrevet lige da skolen startede efter sommerferien, kommer man til at spekulere på, om denne sammensætning af ordforråd genspejler den viden informanterne havde med fra grundskolen eller at det er denne del af ordforrådet, der overlever en pause fra indlæringen.

Tabel 4.4.3.b 'Andrea elsker mig' Antal ordformer 754 – i alle opgaver.

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | Sjældnere end 5000 | Ikke i DFO | Udenfor kategorier |
|-------------------------|--------|-----------|-----------|-----------|-----------|--------------------|------------|--------------------|
| Antal ordformer | 370 | 94 | 52 | 19 | 29 | 153 | 19 | 18 |
| Procentdel | 49,1 | 12,5 | 6,9 | 2,5 | 3,8 | 20,3 | 2,5 | 2,4 |
| 4 kategorier Procentdel | 49,1 | 25,7 | | | | 20,3 | 4,9 | |

Sammenlignet med tabel 4.4.3.a ('Vinter') er stort set ca.10 procentpoint flyttet fra højeste frekvenskategori til kategorien sjældnere end 5000. På det tidspunkt var der gået 9 uger fra den første opgave, så man må antage, at informanterne både havde fået tid til at bearbejde en del af det input, de fik i mellemtiden, og nået at genopfriske noget af ældre viden. Her kan man bogstavelig talt se for sig romanens specifikke ordforråd forhøje procenterne i kategori seks. Det er dog muligt, at forberedelsestiden til opgaven har størst påvirkning i fordeling af ord i frekvenskategorier. Hvis det er tilfældet, harmonerer det med T/T målingen (jf. 4.3.1), hvor værdier fra alle informanter ligger forholdsvis tæt på hinanden.

Tabel 4.4.3.c 'Talesprog' Antal ordformer 824 – i alle opgaver.

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | Sjældnere end 5000 | Ikke i DFO | Udenfor kategorier |
|-------------------------|--------|-----------|-----------|-----------|-----------|--------------------|------------|--------------------|
| Antal ordformer | 363 | 93 | 63 | 39 | 24 | 161 | 41 | 39 |
| Procentdel | 44,1 | 11,3 | 7,6 | 4,7 | 2,9 | 19,5 | 5,0 | 4,7 |
| 4 kategorier Procentdel | 44,1 | 26,5 | | | | 19,5 | 9,7 | |

Sammenlignet med tabel 4.4.3.b (*'Andrea'*) er en bevægelse på ca. 5 procentpoint fra kategori ét til kategorier syv og otte. I forvejen ville man måske have regnet med, at informanterne ville gøre større brug af den mest frekvente kategori i talesproget. Derimod har informanterne højeste værdi sammenlagt i kategorier 2-5 i denne opgave. Den store forskel ligger i, at de har på den ene side fået fat i ord, der ikke er med i DFO og på den anden side hentet ord fra islandsk eller engelsk.

Tabel 4.4.3.d *'Marmorkirken'* Antal ordformer 387 – i alle opgaver.

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | Sjældnere end 5000 | Ikke med i DFO | Udenfor kategorier |
|-------------------------|--------|-----------|-----------|-----------|-----------|--------------------|----------------|--------------------|
| Antal ordformer | 247 | 34 | 16 | 7 | 9 | 57 | 6 | 11 |
| Procentdel | 63,8 | 8,8 | 4,1 | 1,8 | 2,3 | 14,7 | 1,6 | 2,8 |
| 4 kategorier Procentdel | 63,8 | 17,0 | | | | 14,7 | 4,4 | |

Her vender billedet brat i forhold til de foregående opgaver. Denne opgave giver den højeste værdi i kategori ét, hele 5 procentpoint mere end *'Vinter'*. Igen mener jeg, det er i selve opgaven forklaringen ligger. Gengivelse af en indviklet novelle, både hvad angår komposition og ordforråd, resulterer i en kort produktion i forhold til tid (jf. 4.3.1 og bilag 1) og brug af mere elementært ordforråd end i de andre opgaver. I opgaven kan informanterne vælge mellem to forskellige synsvinkler ved genfortælling af novellens indhold. De fleste vælger at bruge begge synsvinkler, én til hver del af novellen – i forhold til novellens komposition. Genfortællingen er meget grundig på visse punkter, men ikke nødvendigvis de vigtigste ud fra fortolkningsmæssigt synspunkt.

Tabel 4.4.3.e *Vigtige ord, som sjældent eller aldrig forekommer i 'Marmorkirken'.*

| Forekommer aldrig | Forekommer én gang | Forekommer to gange | Forekommer 5-6 gange |
|---|--------------------|--|----------------------|
| alibi offer dræbe underslæb kærlighed advokat kontor kriminalassistent | (kriminalpoliti) | motiv jalousi jurist (politimand) politiet | myrde 5 morder 6 |

Tabel 4.4.3.e viser ord, der bruges i novellen og kan karakteriseres som nøgleord, men bliver slet ikke eller meget lidt brugt i de skriftlige opgaver. De to ord i parentes er dog informanternes stedfortrædere for de udeblevne ord.

Tabel 4.4.3.f 'Sundhed' Antal ordformer 380 – i alle opgaver.

| Kategorier | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | 5001- | Ikke med i DFO | Udenfor kategorier |
|-------------------------|--------|-----------|-----------|-----------|-----------|-------|----------------|--------------------|
| Antal ordformer | 199 | 24 | 25 | 14 | 10 | 81 | 10 | 17 |
| Procentdel | 52,2 | 6,3 | 6,6 | 3,7 | 2,6 | 21,3 | 2,6 | 4,5 |
| 4 kategorier Procentdel | 52,2 | 19,2 | | | | 21,3 | 7,1 | |

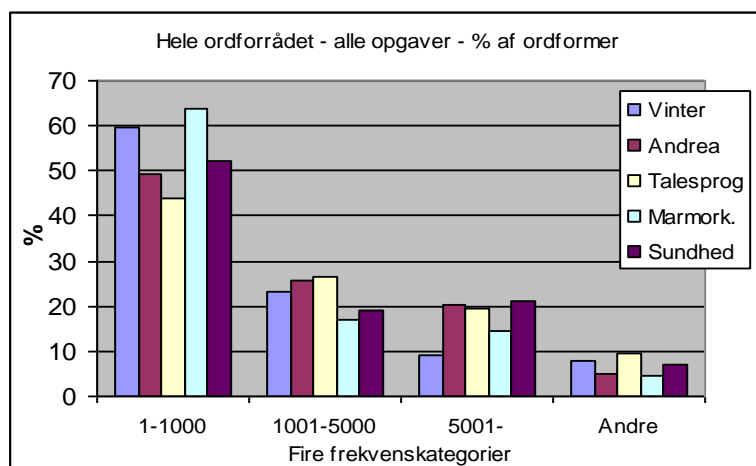
Når de to sidste opgaver ('Marmorkirken' og 'Sundhed') sammenlignes, kan man klart se, hvor stor en rolle selve opgavetyper spiller. Disse to opgaver er skrevet den samme dag, inden for de samme 60 minutter, men forskellen i brug af ordforråd er stor.

Den store tilførsel er fra kategori ét (1-1000) til kategori seks (5001-). Tilførslen kan næppe forklares med, at ordforrådet der ligger bag de to opgaver stammer fra forskellige frekvenskategorier. Bag opgaven 'Sundhed' ligger adskillige autentiske tekster om sygdomme, misbrug, hygiejne og sund levevis. Forklaringen må først og fremmest ligge i den frihed, som associationsopgaven giver. Hvis det er rigtigt, kan man konkludere som så, at det fremmer sprogproduktion og varieret brug af ordforråd, selv at bestemme forløbet af det man skriver. Som eksempler på ord fra kategorien 'frekvens mindre end 5000' er følgende ordliste (informanter i parentes):

| | | | |
|------------------|-------------------|------------------------|--------------------|
| bivirkninger (6) | kondomer (2) | narkotika (3) | stofmisbrugere (4) |
| effekter (5) | kostvaner (5) | reklamer (2) | *barnelømmelse (9) |
| epidemier (9) | menstruation (6) | restitution (1) | *skægvgæst (2) |
| gravid (2) | muskultræning (6) | sportsinteresseret (2) | *turbikloses (9) |

Eksemplerne er taget på må og få i opgaverne. Informanterne kender ordenes betydning og bruger dem rigtigt i konteksten. De tre sidste er taget med som eksempler på ord, som informanterne bruger med held, selvom de er stavet forkert. Her er der tale om delvis forståelse, som nærmer sig den præcise.

Fig. 4.4.3.a Inddeling af informanternes ordforråd i frekvenskategorier.



Figur 4.4.3.a belyser de variationer, der fremkommer i de forskellige opgaver, når ordforrådet inddeles i frekvenskategorier. Ved kun at kigge på første og sidste opgave ville man måske fristes til at tro, at den forskel, der viser sig mellem de to opgaver, kan være vokset fra opgave til opgave hele semestret igennem. Her spiller de forskellige opgavetyper klart ind og stiller ulige krav til informanterne. Der bør dog tages højde for, at andre ting kan spille ind og påvirke udfaldet af informanternes arbejde. En undersøgelse, der strækker over en længere periode, to eller tre semestre, ved udelukkende at bruge en type opgave, kunne give et klarere billede og mere konkrete resultater.

4.4.4 Sammenligning af frekvenskategorier i to opgaver

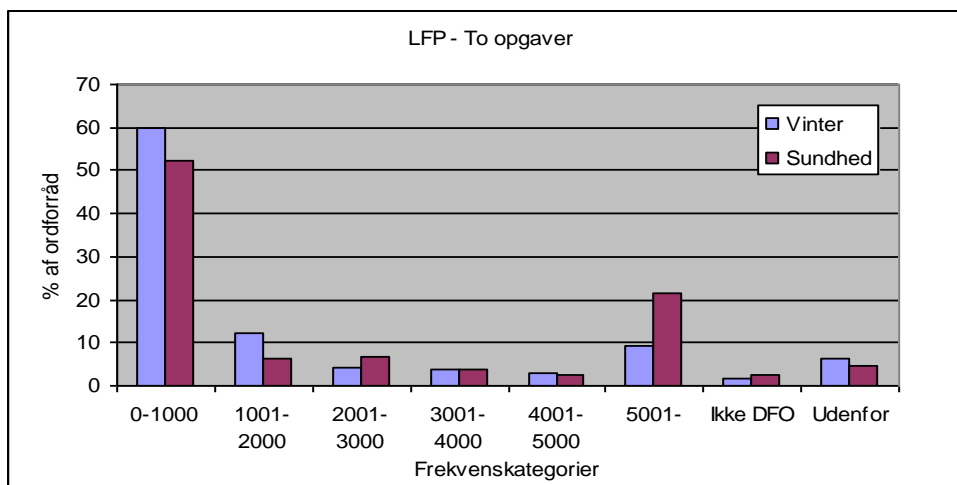
I dette afsnit vil jeg se nærmere på den første og den sidste opgave for at se, om der kan konstateres en udvikling. For det første fordi der er foregået længst tid mellem disse to opgaver. Begge to er bygget omkring associationer (den første uforberedt, den anden bygger på tekster fra læsebogen, jf. 4.1.5.1), og begge opgaver er skrevet i løbet af meget kort tid, som giver anledning til forholdsvis spontant sprogproduktion.

Tabel 4.4.4.a Ordforråd hos alle informanter i de to opgaver inddelt på forskellige frekvensbånd.

| Frekvenskategorier | | En god vinter Antal ord | Sundhed Antal ord | En god vinter % af ordforråd | Sundhed % af ordforråd |
|----------------------|--------------------|----------------------------|----------------------|---------------------------------|---------------------------|
| 1. | De første 1000 ord | 182 | 199 | 59,7 | 52,2 |
| 2. | 1001-2000 | 37 | 24 | 12,1 | 6,3 |
| 3. | 2001-3000 | 13 | 25 | 4,3 | 6,6 |
| 4. | 3001-4000 | 12 | 14 | 3,9 | 3,7 |
| 5. | 4001-5000 | 9 | 10 | 3,0 | 2,6 |
| 6. | 5001- | 28 | 81 | 9,2 | 21,3 |
| 7. | Ikke med i DFO | 5 | 10 | 1,6 | 2,6 |
| 8. | Udenfor | 19 | 17 | 6,2 | 4,5 |
| Ordeksemplarer i alt | | 305 | 380 | | |

Tabel 4.4.4.a gengiver nøjagtig inddeling af ord fra 'Vinter' og 'Sundhed' i frekvenskategorier og hvordan de tal forholder sig procentvis til hinanden. Som allerede er sagt (jf. 4.4.3), er de store forskelle at se i kategorier ét og seks. Der er et fald i kategori to, men de andre viser ikke store udsving. Disse to opgaver er korte og indeholder derfor forholdsvis få ord, der falder inden for de forskellige kategorier. Derfor kan forskel på forholdsvis få ord ændre hele billedet.

Figur 4.4.4.a Inddeling af ord i frekvenskategorier i to opgaver fra alle informanter.



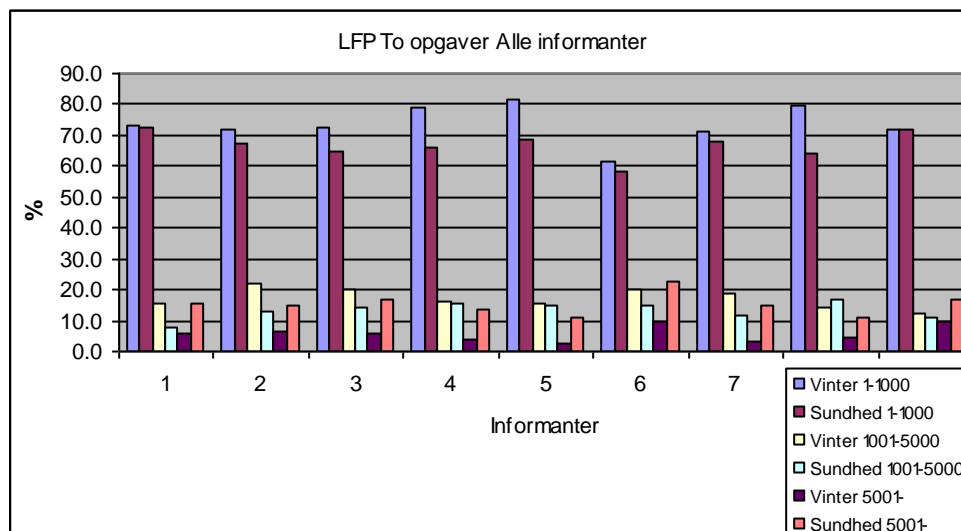
Figur 4.4.4.a viser godt det billede jeg havde forventet at se, dvs. en udvikling over en 13 ugers periode. Hvis ikke de andre opgaver var taget med, ville jeg se dette som et afgørende resultat af undersøgelsen. I dette tilfælde ville der kun være to målinger, og et søjlediagram som det på fig. 4.4.3.a ville slet ikke findes.

Tabel 4.4.4.b Enkelte informanternes ordforråd inddelt i frekvenskategorier. Tal for 'En god vinter' til venstre og for 'Sundhed' i kursiv til højre.

| Frekvenskategorier | Inf. 1 | | Inf. 2 | | Inf. 3 | | Inf. 4 | | Inf. 5 | | Inf. 6 | | Inf. 7 | | Inf. 8 | | Inf. 9 | |
|--------------------|--------|-----------|--------|-----------|--------|------------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|------------|
| | | | | | | | | | | | | | | | | | | |
| 1-1000 | 38 | <i>47</i> | 43 | <i>46</i> | 65 | <i>80</i> | 58 | <i>55</i> | 58 | <i>37</i> | 43 | <i>51</i> | 45 | <i>53</i> | 50 | <i>58</i> | 66 | <i>77</i> |
| 1001-2000 | 5 | <i>1</i> | 2 | <i>2</i> | 4 | <i>4</i> | 3 | <i>3</i> | 6 | <i>1</i> | 5 | <i>6</i> | 5 | <i>3</i> | 5 | <i>3</i> | 4 | <i>7</i> |
| 2001-3000 | 1 | <i>2</i> | 6 | <i>3</i> | 8 | <i>6</i> | 4 | <i>5</i> | 3 | <i>3</i> | 2 | <i>3</i> | 4 | <i>3</i> | 2 | <i>6</i> | 4 | <i>1</i> |
| 3001-4000 | 1 | <i>2</i> | 2 | <i>3</i> | 3 | <i>6</i> | 3 | <i>3</i> | 1 | <i>3</i> | 3 | <i>2</i> | 2 | <i>1</i> | 1 | <i>2</i> | 3 | <i>2</i> |
| 4001-5000 | 1 | <i>0</i> | 3 | <i>1</i> | 3 | <i>2</i> | 2 | <i>2</i> | 1 | <i>1</i> | 4 | <i>2</i> | 1 | <i>2</i> | 1 | <i>4</i> | 0 | <i>2</i> |
| 5001- | 3 | <i>11</i> | 4 | <i>10</i> | 5 | <i>21</i> | 3 | <i>11</i> | 2 | <i>6</i> | 7 | <i>20</i> | 2 | <i>12</i> | 3 | <i>10</i> | 9 | <i>18</i> |
| Ikke i DFO | 0 | <i>0</i> | 0 | <i>1</i> | 0 | <i>2</i> | 0 | <i>1</i> | 0 | <i>1</i> | 2 | <i>3</i> | 0 | <i>2</i> | 0 | <i>2</i> | 2 | <i>0</i> |
| Udenfor | 3 | <i>2</i> | 0 | <i>2</i> | 2 | <i>3</i> | 1 | <i>3</i> | 0 | <i>2</i> | 4 | <i>1</i> | 4 | <i>3</i> | 1 | <i>5</i> | 4 | <i>0</i> |
| Ordformer i alt | 52 | <i>65</i> | 60 | <i>68</i> | 90 | <i>124</i> | 74 | <i>83</i> | 71 | <i>54</i> | 70 | <i>88</i> | 63 | <i>79</i> | 63 | <i>90</i> | 92 | <i>107</i> |

Tabel 4.4.4.b gengiver værdier i alle frekvenskategorier for alle informanter i de to opgaver 'En god vinter' til venstre og 'Sundhed' i kursiv til højre. På den måde kan frekvenskategorier for de to opgaver sammenlignes på alle niveauer. Alle informanter har flere ordformer i den 'Sundhed' end 'En god vinter', undtagen informant 5. Alle har flere ord i kategorien 5001- i den sidste opgave, og de fleste har flere ord i den første kategori. Antal ord i andre kategorier er så få, at det er svært at se, om de har noget at fortælle. Det er først når tallene er omregnet til procenter – og de små kategorier taget sammen – at der dannes en struktur.

Figur 4.4.4.b Procentdel af ordformer inddelt efter frekvenskategorier for to opgaver 'En god vinter' og 'Sundhed' og alle informanter.



På figur 4.4.4.b er data fra tabel 4.4.4.b. De to første søjler hos hver informant viser værdi for de første 1000 ord, henholdsvis for 'En god vinter' og 'Sundhed'. De to søjler i midten giver en samlet størrelse for ord af frekvens 1001-5000. De sidste to søjler viser brug af ord af lavere frekvens end 5000, men dog med i DFO. De sidste to kategorier 'ikke i DFO' og 'udenfor' er udeladt.

Her fremgår det tydeligt, at de enkelte informanter udvikler sig forskelligt. Informant 6 er forholdsvis den største forbruger af lavfrekvente ord i begge opgaver. Informanter 1 og 9 viser ingen ændring indenfor frekvensbånd ét, mens alle andre går ned. Alle viser en markant forhøjelse i brug af de lavest frekvente ord.

4.4.5 Tre informanter – oversigt

I sidste afsnit var fokus rettet mod første og sidste opgave og sammenligning af informanterne. I dette afsnit har jeg valgt at se nærmere på de samme tre informanter (en pige og to drenge), som i afsnittet om lige lange tekster (4.3.4)

Informant 2

Tabel 4.4.5.a Informant 2 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd.

| Frekvens-kategorier /Opgaver | 'En god vinter' | 'Andrea' | 'Talesprog' | 'Marmorkirken' | 'Sundhed' |
|------------------------------|-----------------|----------|-------------|----------------|-----------|
| 1-1000 | 71,7 | 67,0 | 59,9 | 72,8 | 67,7 |
| 1001-5000 | 21,6 | 17,3 | 20,7 | 14,7 | 13,8 |
| 5001- | 6,7 | 14,0 | 12,7 | 11,1 | 13,8 |
| Udenfor DFO | 0,0 | 1,7 | 6,8 | 1,2 | 4,6 |

Hos denne informant er det opgaven 'Talesprog', der først springer i øjnene med en høj score i mindre frekvente ord. I kategorien 'udenfor DFO' har han den højeste score i 'Talesprog', både fordi han tager chancer (gætter) og inddrager forholdsvis mange egennavne (steder og personer) i samtalen. Hans samtale med en bekendt lader til at være logisk (fælles baggrund), de taler om ting der vedrører dem direkte. De andre opgaver (undtagen 'Vinter') går mere ud på at holde sammen på en fiktion (novelle + roman) eller færdiglavede sandheder (klichéer) uden for eller i periferien af hans erfaringsverden ('Sundhed'). Ord af frekvens mindre end 5000 topper straks med opgave nr. 2, evt. fordi opgaven kræver, at informanten gør brug af et bestemt ordforråd. Forberedelserne til opgaven har evt. gjort det muligt at gøre brug af forholdsvis mange sjældne ord i opgaven. Her gives der nogle eksempler på sjældne ord hos informant 2 (jf. bilag 5: ii).

bekendtskab
bibliotekar
journalistlev

liberale
morgenavis
optagelsesprøven

pædagogmedhjælper
terrorist
troværdig

Informant 6

Tabel 4.4.5.b Informant 6 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd.

| Frekvens-kategorier /Opgaver | 'En god vinter' | 'Andrea' | 'Talesprog' | 'Marmorkirken' | 'Sundhed' |
|------------------------------|-----------------|----------|-------------|----------------|-----------|
| 1-1000 | 61,8 | 57,1 | 62,1 | 75,5 | 58,6 |
| 1001-5000 | 19,7 | 24,1 | 17,1 | 14,8 | 16,1 |
| 5001- | 9,4 | 16,2 | 11,5 | 7,8 | 20,7 |
| Udenfor DFO | 8,4 | 2,5 | 9,1 | 2,0 | 4,6 |

Informant 6 har de laveste værdier i ord af højeste frekvens, undtagen i 'Marmorkirken', som hele vejen i gennem skiller sig ud. Det omvendte viser sig så i frekvenskategorien sjældnere end 5000, hvor han er lav i 'Marmorkirken', men ekstremt

høj i 'Sundhed' og 'Andrea'. Denne informant er ikke konsekvent i sin sprogproduktion. Nogle steder ser han ud til at kunne beherske ret indviklet ordforråd og at have tilegnet sig den del af teksternes ordforråd, som ikke findes i DFO. Hans 'Talesprog' er langt kortere end de andres, mest fordi han aldrig kom ud af rollen som interviewer, men også fordi han taler utydeligt og derfor mangler der nogle ord i hans replikker. Her gives der nogle eksempler på informantens replikker (jf. bilag 1:xv):

- | | |
|--|---|
| 2: Hvad er det de unge spiser i dag? Er det sundt or usundt? | 22: Det er ikke mange museer på Island. |
| 3: De ikke tager madpakke i skolen? | 23: Ja, bar og diskotek. |
| 4: Ja, men i gamle dage xxxxx man sig. | 24: Det er mange, mange diskotek på Island. |
| | 27: Badmuligheder, ja, den er meget god. |

I dialogens første halvdel stiller informant 6 kun spørgsmål, hvorimod hans partner siger mindst 3 gange så meget. I dialogens sidste halvdel prøver de at få en samtale op at stå, men de er uforberedte og springer fra det ene til det andet – uden forløb.

Informant 9

Tabel 4.4.5.c Informant 9 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd.

| Frekvens-kategorier /Opgaver | 'En god vinter' | 'Andrea' | 'Talesprog' | 'Marmorkirken' | 'Sundhed' |
|------------------------------|-----------------|----------|-------------|----------------|-----------|
| 1-1000 | 71,4 | 71,5 | 72,8 | 75,6 | 75,2 |
| 1001-5000 | 12,1 | 16,3 | 12,0 | 11,0 | 9,6 |
| 5001- | 9,9 | 10,0 | 10,3 | 11,8 | 15,2 |
| Udenfor DFO | 6,6 | 2,3 | 4,9 | 1,6 | 0,0 |

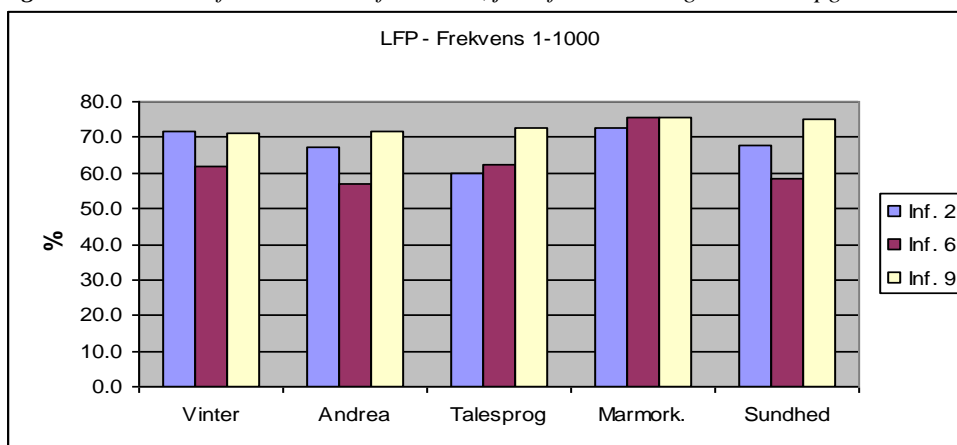
Denne informant karakteriserer sig som stabil. Der er ingen uventede sving i hendes præstationer. Det første frekvensbånd går lidt op i de sidste to opgaver. Der er stigning i frekvens sjældnere end 5000 hele vejen i gennem. Der er fald i ord udenfor DFO. Mønstret går ikke helt op i den sidste opgave, hvor værdien af de meste frekvente ord er lidt for høj, og værdien af den næste kategori lidt for lav. I det hele taget virker informant 9 som om hun har lagt stort arbejde i sine opgaver og gjort sikre fremskridt i sin tilegnelse af ordforråd (jf. bilag 1).

1. På Island er ikke altid god vinter fordi vejret her er meget slemt. En dag skiner solen meget men den næste dag er sner. ('En god vinter')

- Der er mange epidimer som har kommet. De største er HIV, turbiklose, barnelømmelse og nogle flere. Risikoen er meget (stor?) at få HIV. ('Sundhed')

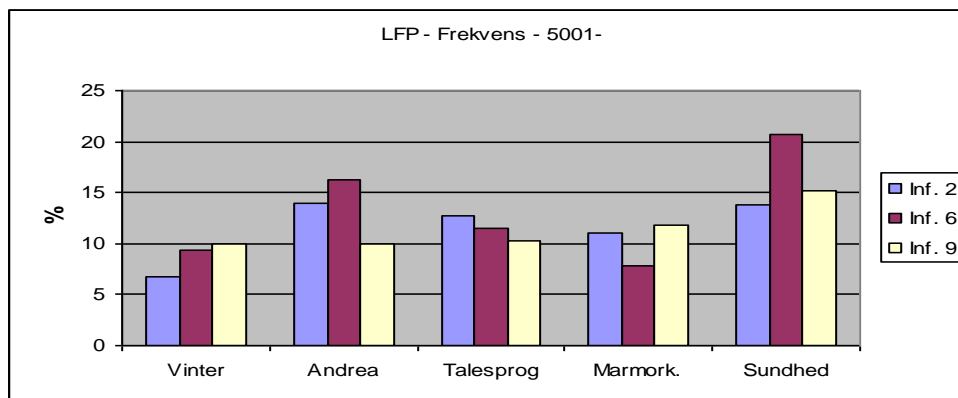
Disse to eksempler (jf. bilag 1: i, xix) viser stor forskel i ordforråd. Ganske vist mangler der lidt afpudsning i retskrivning af de sværeste ord, men på den anden side må der tages hensyn til, at teksterne er skrevet på henholdsvis 15 og 10 minutter. Informanten har muligvis været bevidst om, at nogle af ordene var forkert stavet – og kunne have rettet de største fejl ved hjælp af ordbog og lidt længere tid. Spørgsmål af denne art kan man aldrig få svar på. Det er dog klart, at informanten har været ambitiøs og prøvet at få så mange nye ord (temaord) med i teksten som muligt – og beviser, at hun i det skrivende øjeblik virkelig har været i nærheden af præcis forståelse af disse ord.

Figur 4.4.5.a Tre informanternes ordforråd i højeste frekvenskategori – alle opgaver.



Figur 4.4.5.a viser et diagram over frekvensbåndet hos de tre informanter i alle opgaver. Som før er sagt viser informant 9 mest stabilitet, de andre to afviger især i 'Talesprog' og 'Marmorkirken'. Forklaringen kunne være den, at nummer 2 og 6 er mere ensformige i deres talesprog. De benytter sig af færdiglavede og forholdsvis korte sætninger i højere grad end informant 9. Nummer 2 og 6 er højere i opgave 4 end i deres andre opgaver, mens informant 9 er konsekvent i forhold til sine andre opgaver.

Figur 4.4.5.b Tre informanternes ordforråd i frekvenskategori sjældnere end 5000 – alle opgaver.



Figur 4.4.5.b viser, at to af informanterne scorer højest i den sidste opgave, hvilket understreger, at de bruger flere sjældne ord sidst på semestret end i begyndelsen – informant to er en anelse højere i opgave 2 (jf. tabel 4.4.5.a). Det er bemærkelsesværdigt, at opgave tre ikke er den laveste, informanterne bruger ikke et mindre antal sjældne ord i talesproget. Årsagen er sikkert den, at 'Talesprog' ikke er et spontant talesprog, da informanterne har haft muligheden for at lave et manuskript, som de så kunne lære mere eller mindre uden ad. Man kan hverken se eller høre, at de benytter sig af manuskript under optagelserne.

1.

Inf-1: Men du kan måske komme med mig til Strøget. Jeg skal købe tøj og slik.

Inf-2: Jeg trænger bukser nu. Jeg skal købe bukser på i morron. Så trænger jeg altid slik.

(Bilag 1:vi)

2.

Inf-5: Jeg skal have en bærbar computer i skolen, så jeg har tænkt, at du ville måske betale den computer.

Inf-9: Nej, jeg tror, at du skal betale din computer selv, fordi du har valgt at gå i disse gymnasium.

Inf-5: Men jeg har ikke så mange lommepenge og alle mine venner vil gå i det gymnasium.

Inf-9: Men hvad har du gjort med alle dine penge? Du har arbejdet hele sommerferien.

(Bilag 1:xiii)

3.

Inf-7: Ja, men der skal være meget mange udgifter.

Inf-8: Ja, vi skal betale husholdning, mad og varme.

Inf-7: Ja, og selvfølgelig skat og så har vi mange flere xxxx.

(Bilag 1:xvi)

Disse tre eksempler viser, at det er lykkedes informanterne inddrage en del svære / lavfrekvente ord. Derimod er det svært at bevise, at opgaven 'Talesprog', ikke beror på spontant dagligdags talesprog. Selvfølgelig er det nemt nok at se, at det ikke er indfødte

danskere som taler, men det er umuligt at vise med eksempler, hvor informanternes spontane talefærdigheder holder op og den i forvejen konstruerede samtale begynder.

4.4.6 Kerneord

Da den model, jeg bruger til LFP måling, ikke er i overensstemmelse med den engelske, synes jeg det er hensigtsmæssigt at se nærmere på begrebet 'kerneord i dansk' (Ruus 1995).

Ordforrådet er inddelt i ordfamilier, dvs. grundformen (ubøjet opslagsform) er fællesnævner for andre ord i familien og der bliver skelnet mellem homonymer. Kerneordene i dansk er 1.117 og inddeles i 15 ordklasser, hvoraf substantiver, verber og adjektiver udgør 72%, som må være den andel disse ordklasser har i de 1.250.000 ord Ruus har udarbejdet 'kerneord i dansk' efter.

For at kunne afgøre, hvor stor andel af informanternes ord falder under begrebet 'kerneord', måtte jeg først sætte alle ord i grundform. Opgaven 'En god vinter', som har 305 forskellige ordformer (jf. tabel 4.3.1.a), har 241 forskellige ord i grundform, hvoraf 173 er 'kerneord'. Opgaven 'Sundhed', som har 381 forskellige ordformer (jf. tabel 4.3.1.a), har 280 forskellige ord i grundform, hvoraf 174 er 'kerneord' (jf. bilag 3:xxvii).

Tabel 4.4.6.a Ordforråd fra to opgaver 'Vinter' og 'Sundhed' inddelt i højfrekvente og mindre frekvente ord ved brug af to forskellige metoder.

| | Kerneord | Ikke-kerneord | Frekvens 1-1000 | Frekvens 1001- |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| 'En god vinter' | 71,4% (172 ord) | 28,6% (69 ord) | 73,3% (223 ord) | 26,7% (82ord) |
| 'Sundhed' | 62,1 (174 ord) | 37,9% (106 ord) | 65,3% (248 ord) | 34,7% (133 ord) |

I tabel 4.4.6.a kommer frem, at ved kun at bruge to kategorier til inddeling af ordforrådet, *kerneord* på den ene side og *ikke-kerneord* på den anden side, viser det sig at informanterne benytter sig af forholdsvis flere kerneord i den første opgave end den sidste. Sammenligner man det så med brug af de første tusind mest frekvente ord som den ene kategori, og alle mindre frekvente ord som den anden, viser det sig, at for 'En god vinter' falder resultaterne forbavsende godt sammen. Hvad angår inddelingen for 'Sundhed' er der fald i både brug af kerneord og de tusind mest frekvente ord. Brug af kerneord i 'Sundhed' falder lidt mere end antallet af de mest frekvente ord. En del af

forklaringen er muligvis, at kerneord og de tusind mest frekvente ord²³ ikke falder sammen (jf. tabel 4.5.1.b), fordi disse to begreber beror på arbejde, der er udført på forskellige måder (jf. 4.4.1). Denne forskel gør det svært at lave en sammenligning af ord indenfor de to kategorier. For nærmere præcision henvises der til bilag 3, der indeholder alle kerneord, informanterne har brugt i de to opgaver. Tendensen er klar, informanterne går over til at bruge forholdsvis flere ord, som ikke falder indenfor de mest højfrekvente ord eller kerneord.

Tabel 4.4.6.b Antal og andel kerneord indenfor tre ordklasser; substantiver, adjektiver og verber.

| | Substantiver kerneord | | Adjektiver kerneord | | Verber - kerneord | | Subst.+Adj.+Vb. % af alle kerneord |
|-----------------|-----------------------|-------|---------------------|-------|-------------------|-------|---------------------------------------|
| | antal | % | antal | % | antal | % | |
| 'En god vinter' | 35 | 50,0% | 26 | 78,8% | 43 | 82,6% | 60,1% |
| 'Sundhed' | 31 | 32,0% | 32 | 69,6% | 41 | 80,4% | 53,6% |

Tabel 4.4.6-b viser, hvordan brug af kerneord inden for de tre ordklasser; substantiver, adjektiver og verber, er gået ned hos informanterne fra 60,1% til 53,6% (dvs. 10,8%) fra første til sidste opgave. Af de enkelte ordklasser er forskellen størst inden for substantiver, hvor kerneordene går ned om 18 procentpoint (eller 36%). Det beviser, at det først og fremmest er substantiver, informanterne har tilegnet sig af de mere lavfrekvente ord. Der er også en pæn nedgang i adjektiverne, mens forskellen i verberne er meget lille. Måske kan man ræsonnere som så, at informanterne er ved at komme op på en normal standard for indfødte danskere i deres skriftsprog hvad angår andel af indholdsord over for funktionsord. Bliver denne måling lagt til grund, kan man tale om store fremskridt i retning mod et mere avanceret sprog.

4.4.7 Diskussion og konklusioner

Til trods for, at den hjemmestrikkede model af LFP viser et interessant billede af forskelle mellem den første og sidste opgave og endvidere tilkendegiver at informanterne har tilegnet sig en del nye ord i løbet af semestret, synes jeg det er for optimistisk at drage nogle afgørende konklusioner. Disse resultater kan eventuelt bruges i forbindelse med

²³Ifølge Dansk frekvensordbog

fremtidige undersøgelser på området. På nuværende tidspunkt er det for tidligt at lave statistiske regnestykker eller påstå noget om validitet og reliabilitet. Især er det relativt lille antal ord i frekvensbånd 2-5 påfaldende. Ved at lave lidt om på fremgangsmåden er der måske mulighed for, at denne metode kan gøres brugbar til måling af LFP i det danske sprog.

Efter at have afprøvet de forskellige metoder til måling af ordforråd, synes jeg bedst om at gå videre med LV (leksikalsk variation) og LFP (leksikalsk frekvens profil). Det er også de metoder, jeg har gjort mest ud af, da formålet med mit projekt er at kortlægge selve ordforrådet, uden at tage specielt hensyn til andre ting så som elevteksternes komposition. LR (leksikalsk raffinement) er dog en metode, jeg gerne vil afprøve – hvis og når det bliver muligt og nødvendige data foreligger.

Ordforråd er som et levende organ – nye ord indlæres hele tiden, nye ord bliver lavet, nogle ord går i glemmebogen. Derfor vil det være en spændende opgave at lave et korpus over det ordforråd (læsetekster og supplerende materiale) eleverne bliver præsenteret for i danskundervisningen. Det bør være forholdsvis nemt, ved hjælp af alle de elektroniske hjælpemidler der står til rådighed, at dække størstedelen. Talesproget ville nok være det sværeste, men dog ikke umuligt at få fat i nogle stikprøver for at se, om det afviger meget fra teksterne. Alt i alt vil det være muligt at få en god ide om hvor stort ordforrådet er.

4.5 Andre evalueringsmetoder

4.5.1 Receptive prøver

Hidtil har jeg fokuseret på informanternes produktive ordforråd. Forskere er enige om, at den produktive viden er (langt) mindre end den receptive (jf. 2.9). Hvor stor forskellen er er svært at afgøre og kommer an på forskellige ting.

Et forsøg på at måle informanternes ordforråd ved at bruge den model Nation (2001:416) præsenterer i sin bog. Seks ord er listet op og forklaringer gives til tre af dem. Opgaven består i at danne par af ord og forklaring. Hele prøven består af 4 dele, A, B, C og D (jf. bilag 5). De tre første dele af prøven er lavet efter Nations model, mens den sidste del er en udfyldningsøvelse, hvor der er fire valgmuligheder hver gang, og kun én rigtig.

Eksempel fra prøvens første del hvor informanterne skal kende ordene 'dreng', 'maske', 'stilling'.

| | |
|--|---|
| 1. dreng 2. hemmelighed 3. maske 4. regel 5. stilling 6. verden | __ ansættelse, embede __ noget der bruges til at skjule ansigtet __ et menneskebarn |
|--|---|

Informanterne havde ikke set denne opgaveform før, men de så ikke ud til at have problemer med at forstå, hvad den gik ud på. De fleste brugte kun 10-15 minutter på opgaven. Der bør ikke lægges for stor vægt på resultaterne, de bør snarere ses som et fingerpeg. Jeg vil dog gerne gennemgå de to første dele af prøven.

Test A

De 33 ord (21 substantiver, 6 verber, 6 adjektiver) i den første del af testen (A) er hentet fra en liste over de 1.117 ord der er blevet defineret som kerneord i dansk (Ruus 1995). Ordene er tilfældigt trukket ud af listen.

Denne liste var det eneste, jeg kendte til da prøven blev lavet, og som kunne tjene det formål at lave en eller anden form for frekvensgruppe eller kategori i retning af det man bruger i engelsk (jfr. Nation 2001). Denne del af testen skulle derfor dække de mest frekvente ord og være den del af prøven hvor informanterne skulle score højest. Umiddelbart kunne man tro, at disse ord stort set ville falde indenfor gruppen af de 1000 mest frekvente ord i dansk. Men det gør de dog ikke. Kun 8 af testens 33 ord er blandt de 1000 mest frekvente ord ifølge *Dansk frekvensordbog* (hvis eksistens jeg først opdagede nogle måneder senere).

Tabel 4.5.1.a Prøvens del A: 33 kerneord – inddelt efter frekvensbånd.

| Frekvensbånd | 1-1000 | 1001-2000 | 2001-3000 | 3001-4000 | 4001-5000 | 5001- |
|---|---|--|---|--|---|--|
| Antal point, ud af 8 mulige ²⁴ | 8 ansigt 8 dreng 8 kærlighed 7 lejlighed 8 mulighed 8 rolle 1 stilling 1 vende | 2 dejlig 0 dronning 8 ferie 4 fest 6 hurtig 6 resultat 5 størrelse | 8 artikel 3 bestemme 8 døgn 5 smag 8 voksen | 6 brænde 6 farlig 5 gemme 5 optræde | 8 forelske 7 maske 5 oplysning 7 populær | 0 fætter 2 klokke 7 nyhed 5 rund 8 terning |

²⁴ En af informanterne (5) var ikke med til denne prøve, så der er kun resultater fra de andre otte.

Tabel 4.5.1.b Antal points som hvert enkelt ord har fået idel A..

| | | | | | | | | | |
|--------------|----|---|---|---|---|---|---|---|---|
| Antal points | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| Antal ord | 11 | 4 | 7 | 6 | 4 | 1 | 2 | 2 | 2 |

Tabel 4.5.1.a viser hvordan de 33 kerneord falder i frekvensbånd. Tallene foran ordene angiver, hvor mange elever kender dem hvert og et. Denne fordeling af kerneord på alle frekvensbånd kan i nogle tilfælde forklares med varierende frekvens af ord inden for den samme ordfamilie eller forskellige lemmaer af det samme ord. Grundformen af et ord kan være et forholdsvis lavfrekvent ord, men en anden form eller former af ordet kan være højfrekvente. Det gælder fx *klokke/klokken* og *forelske/forelsket*. En anden forklaring er at Dansk frekvensordbog (DFO) og 'Danske kerneord' ikke er udarbejdet på samme grundlag.

Det er muligt at score 264 point, hvis alle informanter kender alle ord, men ifølge tabel 4.5.1.b får de 183 point. I forhold til det kan man sige, at informanterne kender knap 70% af kerneord i dansk

Tabel 4.5.1.b (og tabel 4.5.1.a) viser, at alle informanter kender en tredjedel (11) af ordene ud fra den betydning af ordene, som testen angiver. Den viser også at der er to ord som ingen af informanterne kender. Der er dog noget der tyder på, at testen ikke lever op til de krav der skulle stilles til den, fx lader det til at ingen af de otte informanter kender ordene 'dronning' og 'fætter', hvilket jeg synes er usandsynligt, især når det drejer sig om 'dronning'. Det er dog klart, at de hverken kender 'dronning' i betydningen 'regent' eller 'fætter' i betydningen 'slægtning'. En af informanterne forbinder 'regent' med 'historie' og en anden forbinder 'fætter' med 'billede', men ellers er der ikke andre forsøg med disse ord. Det kan ikke siges at være hensigtsmæssigt at teste informanter i et højfrekvent ord med et lavfrekvent ord, som der er blevet gjort i tilfældet '*dronning*'/'*regent*'. Det samme gælder '*fætter*'/'*slægtning*', hvor begge ord er lavfrekvente. Der er mange ting at tage hensyn til, hvis en prøve som denne skal være pålidelig og måle det den skal måle, jf. medfølgende eksempel.

Dette eksempel understreger hvor svært det kan være at lave prøver der måler det de skal måle.

| | | |
|--------------|---------------------------------|---------------------|
| 1. dronning | | 1. billede |
| 2. historie | __ <i>regent</i> | 2. <i>fætter</i> |
| 3. materiale | __ bruges om forskellige syn på | __ nummer i tøj |
| 4. resultat | __ fx mad, tøj, kunst | __ bolig |
| 5. smag | __ virkning, udfald | __ <i>slægtning</i> |
| 6. vægt | | 5. præmie |
| | | 6. størrelse |

Test B

Ordene i den anden del af testen (B) er hentet fra to tekster fra informanternes pensum. Resultaterne er dårligst i denne del af prøven og derfor ser det ud til at denne del af prøven er for svær.

Tabel 4.5.1.c Prøvens del B: 39 ord fra pensum – inddelt efter frekvensbånd

| Frekvensbånd | 1001-2000 | 2001-3000 | 3001-4000 | 5001- | Ikke i DFO | |
|-----------------------------|------------------------------|--|------------------------------|--|---|-----------------------------------|
| Antal point, ud af 8 mulige | 6 krig 2 kæmpe 4 viden | 8 <i>journalist</i> 0 miljø 4 <i>stoffer</i> | 6 effektiv 1 undervisning | 0 administration 7 ansøger 2 ballade 2 berette 2 <i>bivirkning</i> 3 <i>epidemi</i> 6 <i>fejlagtig</i> 2 fornærme 2 forurene 8 <i>gymnasium</i> 5 idealbillede 2 katastrofe 5 klima 4 konsekvens 1 kortvarig | 3 kønshormon 1 monument 0 norm 0 opfordre 4 optimisme 3 orientering 0 overhøre 6 springvand 5 spørgeskema 7 toldvæsenet 8 udvekslingsstudent 2 vaccine 0 voldgrav | 6 forlystelse 3 hormonpræparat |

Tabel 4.5.1.d Antal points som hvert enkelt ord har fået idel B.

| | | | | | | | | | |
|--------------|---|---|---|---|---|---|---|---|---|
| Antal points | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| Antal ord | 3 | 3 | 5 | 3 | 4 | 4 | 8 | 3 | 6 |

Tabel 4.5.1.c viser, at 31 ud af 39 er lavfrekvente ord, sjældnere end 5000. Det forklarer til dels, hvorfor informanternes resultater er dårligst her; lange, fremmedartede, sammensatte ord hvor betydningen er abstrakt eller specifik. Dels skal der tages højde for, at forklaringerne til ordene ikke altid er vellykkede og, at de måske nogle gange

forvirrer mere end de forklarer. Denne prøve giver derfor ingen udtømmende oplysninger om informanternes kendskab til ordene udover det 'rigtigt/forkert forhold' prøven måler. Det kan dog nævnes at 6 af disse ord er med på listen over informanternes produktive ordforråd (jf. bilag 2).

Tabel 4.5.1.d viser, hvor mange informanter kender de forskellige ord. Kun 3 af ordene er kendt af alle 8 informanter, mens ingen af informanterne kender 6 af ordene. Hele 8 ord er kun kendt af to informanter.

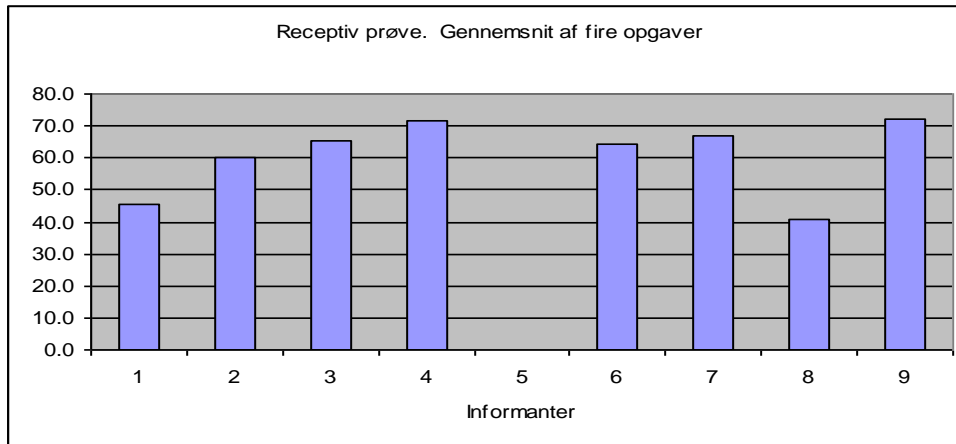
Der bliver ikke gjort nærmere rede for de sidste to dele af prøven, C og D, end de oplysninger der kommer frem i følgende tabeller og grafer. Prøven i helhed følger med som bilag 4.

Tabel 4.5.1.e Informanternes resultater i de enkelte dele af den receptive prøve.

| Prøve / Informanter | A (33 ord) | B (39 ord) | C (15 udtryk) | D (10 huller) | Gennemsnit |
|----------------------------|-------------------|-------------------|----------------------|----------------------|-------------------|
| 1 | 78,8 | 30,8 | 13,3 | 60,0 | 45,7 |
| 2 | 63,6 | 41,0 | 86,3 | 50,0 | 60,2 |
| 3 | 81,8 | 56,4 | 73,3 | 50,0 | 65,4 |
| 4 | 78,8 | 43,6 | 73,3 | 90,0 | 71,4 |
| 5 | - | - | - | - | - |
| 6 | 63,6 | 41,0 | 73,3 | 80,0 | 64,5 |
| 7 | 75,8 | 51,3 | 80,0 | 60,0 | 66,8 |
| 8 | 42,4 | 28,2 | 53,3 | 40,0 | 41,0 |
| 9 | 69,7 | 61,5 | 66,7 | 90,0 | 72,0 |
| Gennemsnit | 69,3 | 44,2 | 64,9 | 65,0 | 60,9 |

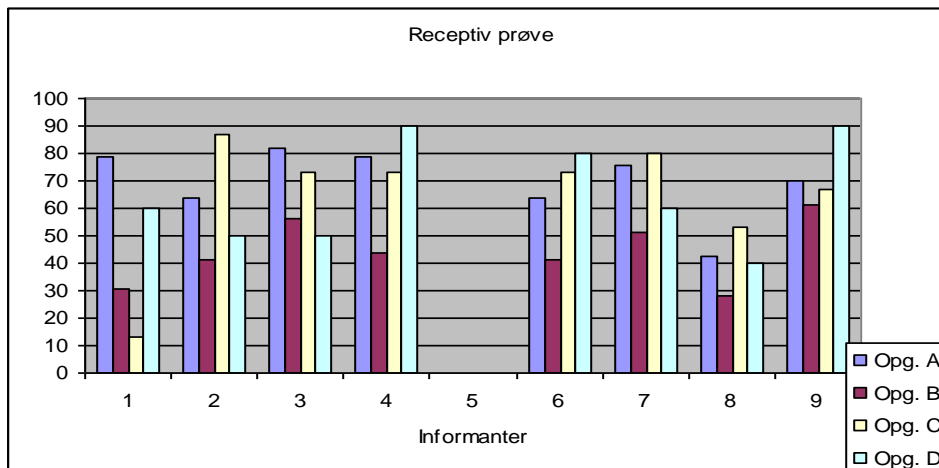
Tabel 4.5.1.e viser resultater fra prøvens fire dele for alle informanternes vedkommende, desuden gennemsnit for hver del af prøven for sig og gennemsnit for hver informant for sig. Det er tydeligt, at del B skiller sig ud med det laveste gennemsnit og dermed det laveste resultat hos alle undtagen to informanter, 1 og 3, men henholdsvis del C og D er lavest hos dem. Informant 1 viser de største udsving inden for hele prøven, især i A og i C, informant 8 er lav hele vejen i gennem, men ellers er der ikke tale om noget mønster i resultaterne.

Figur 4.5.1.a Gennemsnitsresultater af de fire receptive prøver.



Figur 4.5.1.a viser gennemsnitsresultater for alle fire prøver. Alle undtagen informanter 1 og 8 scorer over 50%, men kun to når over 70%. Informant 4 er den eneste der ville nå de 80% ved at smide det dårligste resultat ud, dvs. del B.

Figur 4.5.1.b Resultater af de fire receptive prøver – hver for sig.



På figur 4.5.1.b kan det tydeligt ses, at del B er langt den sværeste og del C varierer mest fra informant til informant, men alt i alt er resultaterne meget forskellige. Disse prøver kan ikke bruges i denne sammenhæng til stort andet end at vise hvilke fælder, der ligger og lur, når man skal lave en prøve til måling af receptiv sprogviden. Alligevel har det været en vis erfaring at have arbejdet med dem på denne måde, og denne erfaring kan forhåbentlig lægges til grundlag for et senere og mere effektivt arbejde.

Ved at se på tabel 4.3.3.c tegner der sig et lille mønster angående informanter 8 og 9. Informant 8 har de dårligste resultater i den receptive prøve og hendes T/T ratio er ligeledes forholdsvis lille hele vejen i gennem. Hun har den laveste score, både når målingen angår forkortede tekster og tekster af original længde. Informant 9 ligger altid på toppen eller i nærheden af toppen. Hun har det højeste gennemsnit i den receptive prøve, og hun viser en flot fremgang i T/T ratio over tid, jf. fig. 4.3.3.c.

Henriksen (1995a:14-15) mener, at mange ordforrådstester primært fokuserer på størrelsen af indlærerens ordforråd og ikke giver mulighed for delforståelse. Der er behov for at udvikle metoder, der kan afsløre hvor megen viden sprobrugeren har om det enkelte ord, dvs. dybdeforståelse. Det er vigtigt, at blive mere bevidst om indlærernes skjulte ressourcer og blive bedre til at udnytte deres store receptive potentialer, bl.a. delforståelse af ord.

Den receptive prøve jeg lagde for mine informanter viser, at informanterne behersker ordforrådet på forskellig vis, både indenfor de enkelte dele af prøven og i gennemsnit. Resultaterne ville sikkert være blevet anderledes, hvis informanterne havde haft mulighed for fx sige/gætte hvad ordene hedder på islandsk eller havde set dem i en anden kontekst/med andre/bedre forklaringer, eller hvis de havde fået forelagt en Wesche og Paribakht-prøve (jf. s. 15). Det vil give en større mulighed for at kortlægge informanternes viden ud fra teorier om delforståelse af ord. Jeg ville også få indsigt i den hypoteseafprøvning, der, ifølge teorien, uafbrudt foregår mens indlærere arbejder med fremmedsproget (jf. 2.5).

En af de mange metoder, der har været udviklet til måling af produktiv skrivning, er en parallel til den receptive test, *Vocabulary Levels Test*. Prøven er sammensat af ord fra WFL (word frequency levels) 2000 og 3000 og UWL (University Word List) 5000 og 10 000. De første bogstaver i prøveordet er givet, fx *I'm glad we had this opp_____ to talk* (Nation 2001).

De fandt en signifikant korrelation mellem total score i aktiv-level-test og det forholdsvis antal af ord i elevernes skrevne opgaver. Jo større ordforråd, jo færre ord fra de 1000 mest frekvente ord og flere fra UWL og de mere lavfrekvente niveauer.

4.5.2 Adjektiver

En varieret brug af adjektiver viser, at indlæreren er en god sprogbruger og blandt andet også at han/hun behersker et godt ordforråd. Henriksen & Haastrups (2000) undersøgelse angående tilegnelse af adjektiver hos elever i 4.-6. klasse giver et fingerpeg om, at adjektiver er svære at tilegne sig.

Tabel 4.5.2.a Andel adjektiver i alle informanternes ordforråd. Tre forskellige kriterier

| | i grundform | ordformer | i alt |
|-----------------|--------------------|------------------|--------------|
| Adjektiver | 161 | 235 | 883 |
| Alle ordklasser | foreligger ikke | 1.729 | 15.924 |
| Adj./Alle ord | | 13,6% | 5,5% |

Alle informanterne bruger i alt 161 adjektiver (grundform) i 235 forskellige ordformer, i alt har de brugt 883 gange et adj. ord eller 5,5 % af hele brugte ord. (Hele ordforrådet 15.924 ord jfr. afsnit 4.3.1)

Tabel 4.5.2.b Andel adjektiver i ordforrådet hos informanter 2, 6, og 9.

| Adjektiver | i grundform | ordformer | i alt |
|--------------------------|-------------|-----------|-------|
| Informanter 2, 6, 9 | 77 | 108 | 347 |
| Alle informanter | 161 | 235 | 883 |
| Tre informanternes andel | 47,8% | 42,6% | 39,3% |

De tre informanter bruger samlet knap halvdelen af alle adjektiver (i grundform) og omkring 40% af ordformer og alle adjektiver, der bliver brugt i teksterne, dvs. en større andel end deres repræsentation i gruppen måler. Hvis man kigger på dem hver for sig viser der sig en forskel. Informant 9 kender de fleste adjektiver eller 23,5% flere end informant 2. De bruger alle intetkøn, flertal og bestemt form, men informant 9 bruger 7 gange komparativ og superlativ, mens de andre gør det en gang hver.

Det kan påstås at informanterne kender forholdsvis mange adjektiver, da de ikke alle sammen bruger de samme adjektiver hele tiden. Så vidt jeg ved, er der ingen tilgængelige oplysninger om, hvorvidt disse forhold passer til tekster hos indfødte dansktalende mennesker.

Tabel 4.5.2.c Tre informanternes indbyrdes forhold i brug af adjektiver

| Adjektiver | i grundform | ordformer | i alt |
|-------------|-------------|------------|-------------|
| Informant 2 | 34 (21,1) | 41 | 122 |
| Informant 6 | 37 (23%) | 44 | 96 |
| Informant 9 | 42 (26%) | 55 (23,4%) | 129 (15,5%) |

I tabel 4.5.2.c kan det tydeligt ses, at informant 9 bruger flere adjektiver end de andre.

Tabel 4.5.2.d Brug af adjektiver i de forskellige tekster (forholdet grundform/ordformer af adjektiverne).

| | 'Vinter' | 'Andrea' | 'Talesprog' | 'Marmork' | 'Sundhed' |
|-------------|----------|----------|-------------|-----------|-----------|
| Informant 2 | 6/7 | 18/19 | 15/18 | 3/3 | 8/9 |
| Informant 6 | 5/8 | 18/21 | 15/15 | 4/4 | 8/8 |
| Informant 9 | 5/6 | 19/26 | 21/23 | 7/7 | 11/14 |

En ting er dog iøjnefaldende – *'Udsigt til Marmorkirken'* har langt de færreste adjektiver hos alle informanter (jf. tabel 4.5.2.d), måske skyldes det novellens udviklede indhold, der igen fører til, at informanterne har så svært ved at gengive den røde tråd og gøre rede for selve plottet, at der ikke rigtig bliver plads til adjektiver (eller at informanterne ikke formår at udvikle holdninger til selve indholdet i den grad, at det kræver brug af adjektiver ud over det minimale). Det hænger sammen med, at selv om informanterne skriver dobbelt så lange tekster om *'Udsigt til Marmorkirken'* bruger de omtrent lige mange ordformer (alle ordklasser) som i *'Sundhed'* (jf. 4.4.3.e) hvor der er en liste over det manglende ordforråd.

4.5.3 Udtryk og vendinger

Til sidst vil jeg prøve at se efter, om og i så fald, i hvor høj grad informanterne bruger faste vendinger. Da jeg ikke kan benytte den teknik, der skal til at gennemsnøge informanternes ordforråd for en nærmere bestemt række ord, må jeg bruge den gamle metode at gennemsnøge informanternes tekster på papir. Da dette er ment som et ekstra indslag i mit speciale og ikke en udtømmende undersøgelse mente jeg det var nok at gå tre af informanterne efter i sømmene.

Det kan siges med det samme, at de tre informanter hverken brugte ordsprog eller idiomer. Hvorvidt de brugte faste leksikaliserede udtryk, kommer an på definitioner:

Problemet ved at studere kollokationer er at bestemme på en konsekvent måde, hvad der skal klassificeres som en kollokation. Det er et problem, fordi de forekommer i variationer af generelle former og i varierende forbindelser/fællesskab i mellem de ord, der udgør kollokationen. I denne bog bliver begrebet kollokation brugt til at beskrive enhver generelt accepteret gruppering af ord i fraser eller sætninger.

(Nation, 2001:317)

At et udtryk er blevet leksikaliseret, vil sige, at en sproglig sammenstilling er blevet accepteret i sproget som en fast enhed, ofte med en helt fast betydning eller brugsfunktion. Der findes mange forskellige slags faste enheder i sproget, lige fra faste sammenstillinger på et par få ord til lange idiomatiske udtryk.

(Henriksen, 1998:25)

Henriksen giver også overblik over nogle typer faste enheder (jf. også 2.8 s. 32)

- kollokationer – faste ordsammenstillinger mellem fx *navneord og udsagnsord* eller *navneord og tillægsord*
- mindre sætningsdele med en helt fast struktur: kommenterer det sagte, skaber struktur i en samtale eller en tekst.
- faste udtryk, som har en helt bestemt pragmatisk funktion; signalerer en bestemt sproglig hensigt.

Jeg har valgt at inddele de udtryk som jeg fandt i fire kategorier og har brugt fed skrift hvor fx præpositionen er forkeret og hvor der mangler noget, har jeg valgt at bruge parentes. Nogle af de udtryk, som jeg har valgt at tage med, falder knap nok under definitionen 'faste sammenstillinger på et par ord', men er taget med for at belyse hvor slemt det egentlig står til med informanternes produktive beherskelse af faste leksikaliserede udtryk.

Informant 2

Udtryk med præpositioner

i fjor
i julen
i morgen

i skolen
på Strøget

på badeværelset
om ugen

om vinteren
på sporet af

Udtryk med være/blive

blive arrig
blive bekymret
blive forelsket i

blive gravid
blive lykkelig
blive myrdet

blive skudt
blive syg

er sammen med
være efterlyst

Udtryk med andre verber

| | | | |
|---------------------------------|-----------------|----------------|-------------|
| gå en tur | have det svært | sige sandheden | løbe på ski |
| gå i biografen | komme på besøg | sove godt | synes du |
| gå i Tivoli | ringe til nogen | spille fodbold | tage bussen |
| have råd <i>for</i> (til) noget | se på piger | spise mad | trænge til |

Forskellige udtryk

| | | | |
|--|--------------------|--------------|-----------------|
| hans forhold med (til) hvor gammel er du | hvor kommer du fra | en kop kaffe | (et) stykke tid |
|--|--------------------|--------------|-----------------|

Informant 6

Udtryk med præpositioner

| | | |
|--------------------------|-----------------------------|----------|
| for eksempel for fred | i sin fritid på grund af | i verden |
|--------------------------|-----------------------------|----------|

Udtryk med være/blive

| | | |
|-------------------------------------|---------------------------|--------------------------------|
| blive bekymret blive forelsket i | blive slank blive træt | blive vred være forelsket i |
|-------------------------------------|---------------------------|--------------------------------|

Udtryk med andre verber

| | | | |
|---|--|---|---------------------------------------|
| arbejde sammen dyrke sport få børn give et interview | have det dejligt holde sig til kigge ned kigge på | klare prøven kunne (godt) lide sige sandheden | spille guitar stole på synes om |
|---|--|---|---------------------------------------|

Informant 9

Udtryk med præpositioner

| | | | |
|------------------------------|--|------------------------------------|---------------------------|
| efter skoletid i hinanden | om vinteren <i>om</i> (i) weekenden | <i>omkring</i> (rundt i) Island | under måtten på avisen |
|------------------------------|--|------------------------------------|---------------------------|

Udtryk med være/blive

| | | |
|-------------------------------|----------------------------------|------------|
| blive chokeret blive dræbt | blive forelsket blive smittet | være sjovt |
|-------------------------------|----------------------------------|------------|

Udtryk med andre verber

| | | | |
|---|--|---|--|
| arbejde hårdt arbejde sammen banke på døren | <i>brække</i> (bryde) en regel føle sig godt | jeg tror kom og sæt dig ned komme til skade | ringe til nogen rydde op se op til |
|---|--|---|--|

betale regninger
blande sammen
bruge penge

gå i byen
gå på cafe
gå på ski (stå)
have et job

kunne lide
køre rundt
lyde spændende
passe huset

solen skinner
tage elevatoren
udvikle sig
vande blomster

Forskellige udtryk

det er koldt
hele tiden
hvor meget

godt vejr
ude på landet

Informanter 2, 6 og 9

Udtryk med præpositioner

efter skoletid
for eksempel
for fred
i verden
i fjor

i hinanden
i julen
i morgen
i skolen
om ugen

om (i) weekenden
om vinteren
omkring (rundt i)
Island

på avisen
på badeværelset
på grund af
på sporet af
på Strøget

Udtryk med være/blive

blive arrig
blive bekymret
blive chokeret
blive dræbt

blive forelsket i
blive gravid
blive lykkelig
blive myrdet

blive skudt
blive slank
blive smittet
blive syg
blive træt

blive vred
er sammen med
være efterlyst
være forelsket i
være sjovt

Udtryk med andre verber

arbejde hårdt
arbejde sammen
banke på døren
betale regninger
blande sammen
bruge penge
brække (bryde) en
regel
dyrke sport
føle sig godt
få børn
give et interview
gå en tur
gå i biografen

gå i byen
gå i Tivoli
gå på cafe
gå på ski (stå)
have det dejligt
have det svært
have et job
have råd **for** (til) noget
holde sig til
jeg tror
kigge ned
kigge på
klare prøven

kom og sæt dig ned
komme på besøg
komme til skade
kunne lide
køre rundt
lyde spændende
passe huset
ringe til nogen
rydde op
se op til
se på piger
sige sandheden
snakke om
solen skinner

sove godt
spille fodbold
spille guitar
spise mad
stole på
stå på ski
synes du
synes om
tage bussen
tage elevatoren
trænge til
udvikle (sig)
vande blomster

Forskellige udtryk

det er koldt
en kop kaffe
(et) stykke tid

forhold **med** (til)
godt vejr
hele tiden

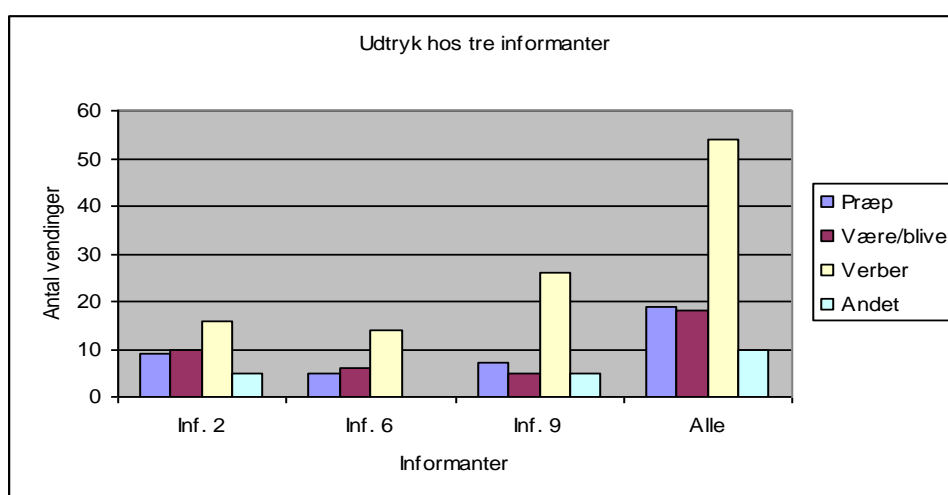
hvor gammel er du
hvor kommer du fra

hvor meget
ude på landet

Tabel 4.5.3.a Udtryk i tal efter grupper fordelt på informanter

| | Informant 2 | Informant 6 | Informant 9 | Alle tre informanter | Overlaping |
|---------------|-------------|-------------|-------------|----------------------|------------|
| Præpositioner | 9 | 5 | 7 | 19 | 2 |
| Være / blive | 10 | 6 | 5 | 18 | 3 |
| Verber | 16 | 14 | 26 | 54 | 2 |
| Forskelligt | 5 | 0 | 5 | 10 | 0 |

Figur 4.5.3.a Udtryk inddelt i fire grupper



Inddeling af udtryk i grupper opstod egentlig af sig selv da jeg syntes, at der var brug for en eller anden form for kategorisering af de udtryk, som informanterne kunne præstere. Udtryk med

- præpositioner (19 forskellige udtryk)
- være eller blive (18 forskellige udtryk)
- andre verber end være/blive (54 forskellige udtryk)
- andre ting end led a/b eller c indebærer (10 forskellige udtryk)

Ovenstående tabel og diagram viser, hvor mange udtryk informanterne 2, 6 og 9 bruger i deres opgaver. Det fremgår meget tydeligt af tabel 4.5.3.a, at hver informant har sit eget forråd af udtryk og at de overlapper hinanden i meget beskednen grad. Udtryk med 'andre verber' er alt i alt 56, men der er 54 forskellige, altså kun 2 der bliver brugt mere

end en gang. Denne gruppe udtryk er også den informanterne har flest af. Udtryk med 'være' eller 'blive', 18 forskellige udtryk og 3 af dem bliver brugt mere end en gang.

De udtryk informanterne bruger er alle korte, 2-3, højst 4 ord, alle gennemsigtige, dvs. ikke idiomer eller udtryk, som i sin helhed betyder andet, end hvad summen af ordene giver anledning til. Alt i alt må man sige, at informanternes brug af udtryk eller vendinger er meget elementær. Det er nok først og fremmest på dette område, man kan se at informanternes sprogproduktion er elevtekster.

5 Sammenfatning og konklusion

Her kommer en sammenfatning af de forskellige dele af undersøgelsen og til sidst en diskussion om, hvorvidt man kan tale om udvikling i ordforrådet i løbet af et semester. Eller om der kun er tale om at kortlægge situationen, som det nu lader sig gøre ud fra de ni informanter, som gruppe på den ene side, og som de forskellige individer de nu er på den anden.

Spørgsmålet, der blev stillet i begyndelsen: *Hvordan forholder det sig med udviklingen af elevernes sprogproduktion (ordforrådstilegnelse) i dansk i gymnasietidens første semester?*

- Hvor stort ordforråd har eleverne i begyndelsen af semestret?

Det samlede ordforråd i den første opgave var 1.292 ord, 305 ordformer. Gennemsnitslængden af hver opgave var 144 ord. Informanterne havde et kvarter til rådighed. Dette betyder dog ikke, at informanterne har kunnet nedfælde hele deres ordforråd i dansk på papir i løbet af et kvarter. Er det meget eller lidt, at ni elever producerer tekster med 305 forskellige ord på 15 minutter?. Kunne de have skrevet længere tekster med flere forskellige ord på islandsk - eller engelsk? Svar på dette spørgsmål må vente til en anden gang, en anden undersøgelse.

De 305 ordformer tilhører 241 ordfamilier (jf. 4.4.6), 71,4% er kerneord, 73,3% er højfrekvente ord (frekvens 1-1000), og T/T ratioen er 0,486 (jf. 4.3.3).

- Hvor stort ordforråd har eleverne i slutningen af semestret?

Det samlede ordforråd i semestrets sidste opgave var 2.136 ord, 387 ordformer. Gennemsnitslængden var 155 ord og informanterne havde 10 minutter til rådighed.

De 387 ordformer tilhører 280 ordfamilier (jf. 4.4.6), 62,1% er kerneord, 65,3% er højfrekvente ord (frekvens 1-1000), og T/T ratioen er 0,535 (jf. 4.3.3).

Ved hjælp af disse to 'punktmålinger' med 13 ugers mellemrum, er det muligt at konstatere en udvikling i alle størrelser.

- mindre tid
- flere ordformer (og ord i alt)
- flere ordfamilier
- nedgang i brug af kerneord
- nedgang i brug af højfrekvente ord
- højere T/T ratio

Hvad siger så denne udvikling i grunden? Den er signifikant (jf. 4.3.3), hvilket betyder at forskel i målinger af de to opgaver ikke er tilfældig. Det giver dog ikke et svar på spørgsmålet, om denne udvikling er stor eller lille. Et fingerpeg er dog, at forskere ikke er enige hvorvidt det kan lade sig gøre at konstatere en udvikling i ordforrådstilegnelse i løbet af så kort tid (jf. Henriksen og Haastrup 4.2.1 og Laufer s.21).

Sprogproduktionen i de fem opgaver er 15.925 ord, 1.729 ordformer. Fire af opgaverne skulle genspejle pensummets ordforråd hen ad vejen.

- en roman ('Andrea')
- to temaer – som grundlag for '*Talesprog*' De to opgaver udgør knap 70% af bruttoproduktionen
- fire noveller – gengive indhold fra en af dem ('*Marmorkirken*')
- tre temaer – hurtigskrivning omkring et af dem ('*Sundhed*')

Udviklingen i de målinger, der er blevet foretaget, viser ikke en lige linje fra første til sidste opgave for informanterne som helhed (T/T jf. fig. 4.3.1.b og c og fig. 4.4.3.a) og heller ikke individuelt (jf. fig. 4.3.2.b og fig. 4.4.4.b og afsnit 4.4.5.) Hvor mange ordfamilier det drejer sig om har jeg ikke et nøjagtigt tal på, men jeg vil anslå dem til at være et sted mellem 1.000 og 1.200. I forhold til Anglins resultater (jf. 2.1) har informantgruppen mindre end en tredjedel af det ordforråd en 6-årig modersmålstalende har. Nation taler om 2.000 højfrekvente ord (foruden teknisk og akademisk ordforråd) som tilstrækkeligt for den som stiler mod en akademisk uddannelse (jf. 2.1). Hvis tallet 2.000 er et fornuftigt mål at stræbe efter, er informanterne kun halvvejs. Heller ikke her kan man finde et rigtigt ståsted for de størrelser Anglin og Nation taler om må være receptiv viden.

Sammenligning mellem L2, produktiv viden (dansk) hos mine 16-årige informanter og receptiv viden hos de L1 talende 6-årige må være urealistisk. Ved at se nærmere på de 2000 højfrekvente ordfamilier (plus andet ordforråd), som receptiv viden hos Nation, er det måske muligt at pejle frem til en slags sammenligningsgrundlag. Man kan fx antage at produktiv viden hos informanterne omfatter 1.000 ordfamilier. Hvis receptiv viden er dobbelt så stor som den produktive, bliver det til 2.000 ordfamilier.

Den receptive prøve, del A, siger at informanterne kender knap 70% af kerneord i dansk, og del B siger at de kender knap 45% af pensumrelateret ordforråd. Informanterne brugte 15% af testordene i deres opgaver. Ved at regne lidt videre kan man konkludere som så, at forholdet i dette eksempel viser tre gange så stor receptiv viden som produktiv

viden. Så er det muligt at ræsonnere frem til, at informanterne muligvis kender op til 3.000 ordfamilier.

Ganske vist udgør informanterne en gruppe på ni, men hvis noget af det, jeg her har sagt, holder stik, er det ikke helt hen i luften at antage at en del af dem har kendskab nok til at studere i Danmark i dansk sprogmiljø, især hvis der tages hensyn til at de skal have mindst et semester mere i dansk.

- Hvad slags ordforråd er der tale om; kerneord, specielt ordforråd, akademisk ordforråd? Frekvensen af de ord der bliver brugt?

Kerneord blev kun målt i to opgaver (jf. 4.4.6), men de udgør ca. 67% af ordforrådet i de opgaver. Andelen kan nok siges at være over de 60% for de andre opgaver også.

Frekvens af hele sprogproduktionen følger som bilag 2 og er gengivet i tabel 4.4.2.a og figur 4.4.2.a. Disse tal viser, hvordan produktionens 1.729 ordformer fordeles over frekvenskategorier. Figur 4.4.3.a og 4.4.4.a viser hvordan ordforråd fra enkelte opgaver fordeles på kategorier. Derimod gives frekvenskategorier for hver enkelt informant i to opgaver (jf. 4.4.4.b). Billedet forandres efter synsvinklen, men her er fokus på enkelte opgaver og det individuelle, er det interessante i denne sammenhæng.

Den højfrekvente del (1-1000) svinger fra 44,1% op til 63,8% i de enkelte opgaver, den lavfrekvente del (5000-) svinger fra 9,2-21,3%, midten fra 14,7-21,3%.

Laufer mener, at Leksikalsk frekvensprofil forholder sig nogenlunde ens mellem samme slags skriveopgaver, lavet af samme elever med få dages mellemrum (jf. 3.6) og at udviklingen bliver den, at ordforråd af frekvens mindre end 2000 bliver større (jf. s. 39).

Det, at forskellen er så stor fra opgave til opgave, kan skyldes at opgavetyperne er for forskellige. Det kan også være, at Meara har ret, når han påstår, at LFP er et dårligt måleapparat (jf. 3.7), som kun går ud på at tælle ord fra kategorier 3 og 4 (jf. 2.3 s. 15), dvs. den del af ordforrådet, eleverne kan bruge produktivt.

Ved at se på de enkelte informanters resultater i første og sidste opgave (jf. fig. 4.4.4.b), er der fem informanter, som viser ingen eller små forskelle i LFP mellem de to opgaver, mens fire viser større forskelle. Igen dukker spørgsmålet om ensartethed af opgaver versus dårlige måleegenskaber op.

Specielt ordforråd er forholdsvis lavfrekvente ord, som er forbundet et specifikt emne eller tema. Af de opgaver, som informanterne skrev, kommer *'Sundhed'* nok nærmest det at kræve et specielt ordforråd. *'Marmorkirken'* kræver også en del specielle ord, hvor mange glimrer ved fravær i teksterne (jf. tabel 4.4.3.e).

Generelt kan man sige, at den største del af det specielle ordforråd hører til i kategorien af ord af mindre frekvens end 5000. Denne kategori udgør 27% af informanternes samlede ordforråd. Tre af opgaverne tangerer de 20%, men kun én af informanterne når der op i *'Sundhed'* og ingen i *'En god vinter'*. (*'Andrea'* og *'Talesprog'* er ikke inddelt i frekvenskategorier pr. informant).

Akademiske ord er der nok ikke mange af i teksterne. Mig bekendt findes der ikke en liste over akademiske ord i dansk. En stor del af det akademiske ordforråd stammer sikkert fra latin og har derfor mange fælles træk med den engelske.

- Kan der spores en udvikling i ordforrådet?

Der er uden tvivl tale om en udvikling mellem første og sidste opgave. Informanterne bruger flere ordformer (højere T/T ratio), skriver flere ord, har flere lavfrekvente ord i den sidste opgave end den første (jf. 4.3.3 og 4.4.4). Udover det, der allerede er blevet sagt, kan man tale om en udvikling i opgaverne *'En god vinter'* – *'Andrea'* – *'Talesprog'* (jf. fig. 4.3.3.a). Der er samtidig tale om *korte - mellemlange - lange tekster*, foruden *kort – længere – længst* tid til rådighed. *'Marmorkirken'* virker så som et stort skridt tilbage, nærmest som en optakt til en ny udviklingsproces. I denne opremsning er der for mange variabler for at kunne fremsætte en absolut påstand om udvikling i ordforrådet. Min erfaring kunne dog bruges som grundlag, hvis og når en ny undersøgelse af denne art finder sted.

- Er der tale om udvikling i brug af kollokationer?

Alle tekster fra tre af informanterne blev gennemløst og det kan påstås, at informanterne befinder sig på et elementært stadium, hvad angår brug af kollokationer og derfor er der ikke tale om nogen udvikling, i hvert fald kan den ikke måles (jf. 4.5.3). Hvad der angår receptiv viden om kollokationer og faste leksikaliserede udtryk, er der kun resultatet af den receptive prøve, del C (jf. bilag 5:xxxi), som kan lægges som grundlag for en diskussion. I gennemsnit forstår informanterne 65% af udtrykkene, men hvis vi ser bort fra det laveste resultat, bliver forståelsen 72%. Set i dette lys, er den receptive viden

enorm i forhold til den produktive. Det er klart, at her har sproglærere/dansklærere i Island en indfaldsvinkel til nye prioriteringer i undervisningen.

- Er der forskel på elevernes ordforråd, når de udtrykker sig mundtligt og skriftligt.

I så fald, hvilken forskel er der tale om?

Der kan ikke spores nogen afgørende forskel på mundtlige og skriftlige udtryk i denne undersøgelse. Grunden må være den, at den mundtlige opgave ikke beror på spontan tale (jf. 4.3.4, 4.4.3, 4.4.5), men grundig forberedelse og kunne derfor lige så godt have været en skriftlig opgave. *'Talesprog'* er den længste opgave, et pararbejde, lavet udenfor undervisningstimerne, men dog bygget over to temaer med forskellige tekster; hvoraf en betydelig del af det ene tema bestod af taleopgaver i timerne. Informanterne brugte ikke manuskript under selve optagelserne, det kan både høres og ses med det samme. For de flestes vedkommende var det tydeligt, at de havde noget bestemt de ville/skulle tale om og derfor bliver alle målinger lige så høje som i de skrevne opgaver. Udeblivelsen af faste udtryk bliver særlig påfaldende i denne opgave. Informanterne oversætter direkte fra islandsk, når de hilser og spørger til hinandens velbefindende, og andet i den stil. Det er i grunden utrolig meget de kan producere af talt sprog og langt det meste ville de danskere forstå uden vanskeligheder.

- Hvilke strategier bruger eleverne når de ikke kender det rigtige ord for det de vil udtrykke

Nogle af informanterne prøver sikkert at omformulere det, de vil sige, når de mangler det rigtige ord. Den slags er svært at se, medmindre de streger ud og prøver på nyt.

Der er tale om hypoteseafprøvning; navneord, ental, bestemt, fælleskøn ender på –en; udsagnsord har –(e)r i nutid; og endelsen –lig svarer til –legur. på islandsk. Eksempler: *munen* munurinn/forskellen, *hugser* hugsar/tænker, *smukkelig* fallegur/smuk eller de prøver engelske/engelskagtige ord (windoen glugginn/vinduet, *liften* lyftan/elevatoren) (jf. bilag 2:xxvi).

I denne undersøgelse falder 5,7% af de ord informanterne bruger udenfor de kriterier, som gør det muligt at godkende dem som danske ord. Her findes ord som *føjdig, *multistioner, *faintedes. Jeg kan ikke engang ud fra konteksten gætte mig frem til hvad de skulle betyde. Ord som *nyårskveld er nemmere at regne ud. Her er et eksempel på delvis forståelse af et ord. Informanten har sikkert tænkt på ordet 'nytårsaften' og husket,

at det ikke lød som det islandske 'gamlárskvöld'; resultatet er blevet en kompromis mellem det danske og det islandske ord. For at få et bedre indblik i strategier vil det være nødvendigt at høre informanterne forklare hvad de gør og hvorfor (jf. Haastrup s. 21). En anden metode er at gøre det til en del af undervisningen/sprogtilægnelsen, at eleverne gør sig klart hvad de vil udtrykke og prøver sig frem til en acceptabel løsning.

- Hvilken rolle spiller opgavernes forskellighed?

Opgavernes forskellighed spiller en langt større rolle end jeg havde troet muligt. En opgave, som er bundet til en bestemt roman eller en novelle, virker hæmmende på sprogproduktionen. En opgave, hvor informanten står frit stillet overfor et emne, giver en størst produktion i forhold til tid (jf. 4.3.1).

- Hvilke metoder er anvendelige til måling af ordforråd i elevtekster?

Jeg mener, at der kan måles fremskridt i løbet af et semester, men det ville være en fordel at have længere tid til rådighed. Elever i 1g har et lille ordforråd i dansk i forhold til en modersmålstalende, men det er vores opgave som lærere at sørge for, at det ordforråd er anvendeligt i flere sammenhænge – for turister, studerende, på arbejdsmarkedet, i hjemmet.

Når jeg står ved vejs ende i dette speciale, kan jeg roligt sige, at denne proces har været en opdagelsesrejse fra først til sidst. Jeg har fået svar på mange spørgsmål, men undervejs er der opstået langt flere nye spørgsmål. Nogle af de nye spørgsmål går ud på at finde metoder til at fremme indlæringsprocessen hos eleverne. Der er også brug for enkle, afgørende metoder til måling og evaluering, der tager til begrænsede områder indenfor fremmedsprogtilægnelsen.

6 Litteratur

- Aitchison, J. (1994). *Words in the Mind. An Introduction to the Mental Lexicon*. 2nd edition. Blackwell, Oxford UK & Cambridge USA.
- Anglin, J. M. (1993). 'Vocabulary Development: a Morphological Analysis'. *Monographs of the Society for Research in Child Development*, Serial no. 238, 58, 10.
- Bergenholtz, H. (1992). *Dansk Frekvensordbog*. GEC Gads Forlag, København.
- Carter, R. og M. McCarthy. (1988). *Vocabulary and Language Teaching. Applied Linguistics and Language Study*. Red: C.N. Candlin. Longman, London & New York.
- Ellis, N. (1994). 'Implicit and Explicit Language Learning – an Overview'. *Implicit and Explicit Learning of Languages: 1-32*. Red: N. Ellis. Academic Press, London.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford University Press, Oxford.
- Færch, C. og G. Kasper. (1983). 'Procedural Knowledge as a Component of Foreign Language Learners' Communicative Competence'. *Kommunikation im (Sprach-) Unterricht*. Red: Boete, H. & Herrlitz W. Utrecht.
- Færch, C., K. Haastруп, og R. Phillipson, (1984). *Learning Language and Language Learning*. Gyldendals Sprogbibliotek, København.
- Fries, C. C. (1945). *Teaching and Learning English as a Foreign Language*. University of Michigan Press, Ann Arbor.
- Gimbel, J. (1995). 'Bakker og udale'. *Sprogforum nr. 3*, 1995:28-34.
- Golden, A. (2000). *Ordforråd, ordbruk og ordlæring i et andrespråksperspektiv*. Ad Notam Gyldendal, København.

- Hafdís Ingvarsdóttir og Kirsten Friðriksdóttir. (1997). *Panorama*. Mál og menning, Reykjavík.
- Henriksen, B. og K. Haastrup. (2000). 'Describing Foreign Language Learners' Productive Vocabulary use and Development'. *Language Use, Language Acquisition and Language History: (Mostly) Empirical Studies in Honour of Rüdiger Zimmermann*. Red: I. Plag & K.P. Schneider. WVT Wissenschaftlicher Verlag Trier, Trier.
- Henriksen, B. (1995a). 'Hvad vil det sige at kunne et ord? Ordforståelse og ordbeherskelse'. *Sprogforum nr. 3:12-18*.
- Henriksen, B. (1995b). 'Hvordan husker man ord? Fokus på forskellige opgavetyper'. *Sprogforum nr. 3: 45-55*.
- Henriksen, B. (1998). 'At lære i helheder'. *Sprogforum nr. 10: 24-31*.
- Henriksen, B. (1999). 'Three Dimensions of Vocabulary Development'. *Studies in Second Language Acquisition, 21 (2): 303-317*.
- Henriksen, B. (1999a). 'Ordforråd og ordforrådsindlæring'. *Studier i dansk som andetsprog*. Red. A. Holmen og K. Lund. Akademisk Forlag, København.
- Henriksen, B. (1999b) 'Three Dimensions of Vocabulary Development.' *Studies in second language acquisition 21 (2): 303-317*
- Holmen, A. (1988). 'Sprogindlæring og undervisningsmetoder'. *Kvan. Et tidsskrift for læreruddannelsen og folkeskolen: 79-90*. 8. årgang. Februar 1988.
- Haastrup, K. (1994) 'On Word Processing and Vocabulary Learning'. *Kognitive Linguistik und Fremdesprachenerwerb: 129-147*. Red: W. Börner & K. Vogel. Gunter Narr Verlag, Tübingen.
- Jörgen Pind. (1997). *Sálfræði ritmáls og talmáls*. Háskólaútgáfan, Reykjavík.

- Laufer, B. og Nation, P. (1995). 'Vocabulary Size and Use: Lexical Richness in L2 Written Production'. *Applied Linguistics* 16 (3): 307-322.
- Laufer, B. (1994). 'The Lexical Profile of Second Language Writing: Does it Change over Time?' *RELC Journal* 25 (2): 21-33.
- Laufer, B. (1998). 'The Development of Passive and Active Vocabulary: Same or Different?' *Applied Linguistics* 19: 255-271.
- Linnarud, M. (1986). *Lexis in Composition. A Performance Analysis of Swedish Learners' Written English*. Lunds Studies in English 74. CWK Gleerup, Liber Förlag, Malmö.
- Lund, K. (1997). *Lærer alle dansk på samme måde? - En længdeundersøgelse af voksnes tilegnelse af dansk som andetsprog*. Special -Pædagogisk forlag.
- McLaughlin, B. (1987). *Theories of Second Language Learning*. Edward Arnold. A division of Hodder & Stoughton London, New York, Melbourne, Auckland.
- Meara, P. (1999). 'The Vocabulary Knowledge Framework'. *Vocabulary Acquisition Research Group*. Virtual Library. <http://www.swan.ac.uk/cals/calsres/vlibrary/pm96d.htm>
Posted on 19th November 1996, revised November 1999.
- Meara, P. og H. Bell. (2001). 'P_Lex: A Simple and Effective Way of Describing the Lexical Characteristics of Short L2 Texts'. *Prospect Vol. 16, No. 3 December 2001*.
- Meara, P. og Fitzpatrick, P. (2000). 'Lex30: an Improved Method of Assessing Productive Vocabulary in an L2'. *System* 28, 1: 19-30.
- Meara, P. og H. Bell. (2001). 'P_Lex: A Simple and Effective Way of Describing the Lexical Characteristics of Short L2 texts'. *Prospect* 16, 3: 5-18.
- Meara, P.M. og R. Sanchez. (1993). 'Matrix Models of Vocabulary Acquisition: an Empirical Assessment'. *CREAL Symposium on Vocabulary Research*. Ottawa.

- Melka, F. (1997). 'Receptive vs. Productive Aspects of Vocabulary'. *Vocabulary, Description, Acquisition and Pedagogy*: 84-102. Red: N. Schmitt & M. McCarthy. Cambridge University Press, New York.
- Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. English Language Institute, University of Wellington.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press, Cambridge.
- Nation, P. og J. Coday. (1988). 'Vocabulary and Reading'. *Vocabulary and Language Teaching*: 97-110. Red: Carter, R. & M. McCarthy. Longman. London and New York.
- Nattinger, J. (1988). 'Some Current Trends in Vocabulary Teaching'. *Vocabulary and Language Teaching*: 62-82. Red: Carter, R. & M. McCarthy. Longman, London and New York.
- Nattinger, J.R. og J.S. DeCarrico. (1992). *Lexical Phrases and Language Teaching*. Oxford University Press, Oxford.
- Pawley, A. og F.H. Syder. (1983). 'Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency'. *Language and Communication*: 191-226. Red: J.C. Richards & R.W. Schmidt. Longman, London.
- Richards, I.A. (1943). *Basic English and its uses*. New York: Norton.
- Richards, J. (1976). 'The Role of Vocabulary Teaching'. *TESOL Quarterly* 10 (1): 77-89.
- Richards, J.C. og T.S. Rodgers. (2001). *Approaches and Methods in Language Teaching*. 2nd Edition . Cambridge University Press.
- Rivers, W. M. (1968). *Teaching Foreign Language Skills*. University and Chicago Press, Chicago.
- Ruus, H. (1995). *Danske kerneord*. Museum Tusulanums.

Swain, M. (1995) 'Three Functions of Output in Second Language Learning'. *Principles and Practice in Applied Linguistics*. Red: G. Cook og B. Seidlhofer. Oxford University Press, Oxford.

Twaddell, F. (1972). 'Linguistics and the Language Teacher'. *Readings on English as a Second Language*: 268-276. Red: K. Croft.

Wesche, B. og S. Paribakht. (1996). 'Assessing Vocabulary Knowledge: Depth Versus Breadth'. *Canadian Modern Language Review* 53, 1: 13-40.

Wilkins, D.A. (1972). *Linguistics and Language Teaching*. Edward Arnold, London.

7 Oversigt over tabeller og figurer

Tabeller

| | | |
|----------------|--|----|
| Tabel 2.3.a | Hvad vil det sige at kunne et ord (Nation 2001:27) | 17 |
| Tabel 4.2.2.a. | 'En god vinter' og 'Sundhed': Antal ord hos alle elever og gennemsnit af sætningslængde i de to opgaver. | 49 |
| Tabel 4.2.3.a. | Leksikalsk tæthed hos alle ni informanter i 'En god vinter' og 'Sundhed'. | 50 |
| Tabel 4.2.4.a. | Procentdel og antal originale ord hos alle ni informanter i to opgaver ('En god vinter' og 'Sundhed'). | 52 |
| Tabel 4.3.1.a | Antal ord, ordformer, andel i helheden og T/T ratio i alle opgaver | 56 |
| Tabel 4.3.1.b | Sammenligning af opgavers længde i forhold til tidsfaktoren. | 58 |
| Tabel 4.3.1.c | Sammenligning af opgavers længde i forhold til tidsfaktoren | 60 |
| Tabel 4.3.1.d | Forhold mellem tid og produktion af ordformer og ord i alt. | 60 |
| Tabel 4.3.1.e. | Det højeste og laveste antal ordformer i hver enkelt opgave. I midten ligger gennemsnitstal for ordformer i alle opgaver. | 61 |
| Tabel 4.3.1.f | Det højeste og laveste antal ord i alt i hver enkelt opgave. I midten ligger gennemsnitstal for ord i alt i alle opgaver. | 62 |
| Tabel 4.3.2.a | Oversigt over hele sprogproduktionen i tal, samt gennemsnitsværdi pr. informant. | 66 |
| Tabel 4.3.3.a | T/T ratio i to opgaver 'En god vinter' og 'Sundhed' | 68 |
| Tabel 4.3.3.b | Ændring i T/T ratio over tid og ændring i antal ord over tid i % hos alle informanter – og gennemsnit for begge størrelser. | 69 |
| Tabel 4.3.3.c | T/T Ratio i de originale tekster versus de forkortede tekster – forskel i % | 70 |
| Tabel 4.3.4.a | T/T ratio henholdsvis i de originale opgaver i den oprindelige længde og de forkortede tekster på 120 ord, samt gennemsnit for alle værdier. | 72 |
| Tabel 4.4.2a | Fordeling af ordformer i frekvenskategorier. | 77 |
| Tabel 4.4.3.a | 'En god vinter' Antal ordformer 305 – i alle opgaver. | 79 |
| Tabel 4.4.3.b | 'Andrea elsker mig' Antal ordformer 754 – i alle opgaver. | 79 |
| Tabel 4.4.3.c | 'Talesprog' Antal ordformer 824 – i alle opgaver. | 80 |
| Tabel 4.4.3.d | 'Marmorkirken' Antal ordformer 387 – i alle opgaver. | 80 |

| | | |
|---------------|---|-----|
| Tabel 4.4.3.e | Vigtige ord, som sjældent eller aldrig forekommer i 'Marmorkirken'. | 81 |
| Tabel 4.4.3.f | 'Sundhed'. Antal ordformer 380 – i alle opgaver. | 81 |
| Tabel 4.4.4.a | Ordforråd hos alle informanter i de to opgaver inddelt på forskellige frekvensbånd. | 83 |
| Tabel 4.4.4.b | Enkelte informanters ordforråd inddelt i frekvenskategorier. Tal for 'En god vinter' til venstre og for 'Sundhed' i kursiv til højre. | 84 |
| Tabel 4.4.5.a | Informant 2 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd. | 86 |
| Tabel 4.4.5.b | Informant 6 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd. | 86 |
| Tabel 4.4.5.c | Informant 9 - Ordforråd fra enkelte opgaver inddelt i frekvensbånd. | 87 |
| Tabel 4.4.6.a | Ordforråd fra to opgaver 'Vinter' og 'Sundhed' inddelt i højfrekvente og mindre frekvente ord ved brug af to forskellige metoder. | 90 |
| Tabel 4.4.6.b | Antal og andel kerneord indenfor tre ordklasser; substantiver, adjektiver og verber | 91 |
| Tabel 4.5.1.a | Prøvens del A: 33 kerneord – inddelt efter frekvensbånd | 93 |
| Tabel 4.5.1.b | Antal points som hvert enkelt ord har fået idel A.. | 94 |
| Tabel 4.5.1.c | Prøvens del B: 39 ord fra pensum – inddelt efter frekvensbånd | 95 |
| Tabel 4.5.1.d | Antal points som hvert enkelt ord har fået idel B. | 95 |
| Tabel 4.5.1.e | Informanternes resultater i de enkelte dele af den receptive prøve. | 96 |
| Tabel 4.5.2.a | Andel adjektiver i alle informanters ordforråd. | 99 |
| Tabel 4.5.2.b | Andel adjektiver i ordforrådet hos informanter 2, 6, og 9. | 100 |
| Tabel 4.5.2.c | Tre informanters indbyrdes forhold i brug af adjektiver | 100 |
| Tabel 4.5.2.d | Brug af adjektiver i de forskellige tekster (forholdet grundform/ordformer af adjektiverne). | 100 |
| Tabel 4.5.3.a | Udtryk i tal efter grupper fordelt på informanter | 104 |

Oversigt over figurer

| | | |
|---------------|--|----|
| Figur 2.3.a | Viser en forenklet ide om, hvordan ord kan 'flyde' frem og tilbage mellem forskellige stadier, samtidigt med at først receptiv viden, og så produktiv viden bliver større. | 20 |
| Figur 2.5.a | Indlæringsmodel Færch et al. (1984:200). Modellen viser hvordan input og viden bliver genstand for hypotesedannelse, og relationer overfor udfaldet af hypoteseafprøvningen. | 23 |
| Figur 2.7.a | I denne model er de to processer føjet til hypotesedannelsen. og –afprøvningen som elementer i indlæringen (Holmen, 1988). | 27 |
| Figur 2.7.b | De forskellige vidensformer og deres samspil (Lund, 1997:119). | 26 |
| Figur 2.8.a | Samspil mellem pragmatisk komponent, ordforråd og syntaks (Nattinger og DeCarrico, 1992:16). | 28 |
| Figur 4.2.3.a | Leksikalsk tæthed hos ni informanter i to opgaver ('En god vinter' og 'Sundhed'). | 50 |
| Figur 4.2.4.a | Viser procentdel originale ord i de to tekster 'Vinter' og 'Sundhed'. | 52 |
| Figur 4.2.4.b | Viser antal originale ord i de to tekster 'Vinter' og 'Sundhed'. | 53 |
| Figur 4.3.1.a | Hele sprogproduktionen inddelt efter opgaver i antal ord. | 57 |
| Figur 4.3.1.b | Antal ordformer – yderste værdier og gennemsnit for hver opgave. | 61 |
| Figur 4.3.1.c | Viser højeste, laveste og gennemsnitsantal ord i hver opgave. | 63 |
| Figur 4.3.1.d | Viser enkelte informanters sprogproduktion og forholdet mellem antal ordformer og antal ord i alt. | 64 |
| Figur 4.3.2.a | Viser leksikalsk variation – gennemsnit for hver af de fem opgaver. | 64 |
| Figur 4.3.2.b | Leksikalsk variation - alle opgaver – alle informanter. | 65 |
| Figur 4.3.2.c | T/T af hele sprogproduktionen hos hver enkelt informant +gennemsnit. | 66 |
| Figur 4.3.2.d | 'Talesprog' Hver enkelt informants sprogproduktion. | 67 |
| Figur 4.3.3.a | Sammenligning af T/T ratio i de to opgaver. | 68 |
| Figur 4.3.3.b | Tekster skåret ned til 120 ord hos tre informanter. | 70 |
| Figur 4.3.3.c | LV Sammenligning af originale og forkortede tekster hos tre informanter. | 71 |

| | |
|---|-----|
| Figur 4.3.4.a Forskel i gennemsnitstal for lange og korte tekster hos tre informanter. | 73 |
| Figur 4.3.4.b LV Tre informanter – alle opgaver – forkortet til 120 ord. | 74 |
| Figur 4.4.2.a Inddeling af hele ordforrådet i otte frekvenskategorier hos alle informanter – % af ordformer. | 78 |
| Fig. 4.4.3.a Inddeling af informanternes ordforråd i frekvenskategorier. | 82 |
| Figur 4.4.4.a Inddeling af ord i frekvenskategorier i to opgaver fra alle informanter. | 83 |
| Figur 4.4.4.b Procentdel af ordformer inddelt efter frekvenskategorier for to opgaver 'En god vinter' og 'Sundhed' og alle informanter. | 85 |
| Figur 4.4.5.a Tre informanters ordforråd i højeste frekvenskategori – alle opgaver. | 88 |
| Figur 4.4.5.b Tre informanters ordforråd i frekvenskategori sjældnere end 5000 – alle opgaver. | 88 |
| Figur 4.5.1.a Gennemsnitsresultater af de fire receptive prøver. | 97 |
| Figur 4.5.1.b Resultater af de fire receptive prøver – hver for sig. | 97 |
| Figur 4.5.3.a Udtryk inddelt i fire grupper | 105 |

8 Bilag

| | |
|---|-----------|
| 1. Bilag. Elevtekster | i-xix |
| 2. Bilag til afsnit 4.4: Ordforråd i frekvensbånd. | xx-xxvi |
| 3. Bilag til afsnit 4.4.6: Brug af kerneord | xxvii |
| 4. Bilag til afsnit 4.4.6: Brug af adjektiver. | xxviii |
| 5. Bilag til afsnit 4.5: Prøve i receptivt ordforråd. | xxix-xxxi |