Gunna Funder Hansen

Ph.d.-afhandling

Det arabiske skriftsystems betydning for læsefærdighed i arabisk som fremmedsprog

Center for Mellemøststudier
Syddansk Universitet
2004

INDHOLDSFORTEGNELSE

FIGUROVERSIGT 3

APPENDIXOVERSIGT 7

FORORD 9

1. INDLEDNING 11
   1.1 Problemformulering og begrebsafklaring 12
      Det arabiske skriftsystem 13
      Fremmede elementer 15
      Visuel ordgenkendelse 16
      Danske lørnere 18
   1.2 Eksisterende forskning på området 18
      Strategier 19
      Automatiserede kompetencer 20
      Pædagogik 22
      Sammenfatning 23
   1.3 Fremgangsmåde 24

2. TEORETISK UDGANGSPUNKT 27
   2.1 Kognitive læseteorier i historisk perspektiv 28
   2.2 Ordgenkendelse 30
      Dualruteteorien 31
      Konnektionisme 32
      Teoretisk udgangspunkt for afhandlingen 36
   2.3 Læsning på tværs af ortografier 37
   2.4 Læsning på fremmedsprog 40
   2.5 Læsning med et nyt alfabet 42
   2.6 Sammenfatning 45

3. DE SEMITISKE SKRIFTSYSTEMER 49
   3.1 Lidt om skriftsystemer generelt 49
   3.2 Det arabiske skriftsystem 51
      Skrift 51
      Fonologi 53
      Morfologisk struktur 55
   3.3 Diglossi 58
   3.4 Sammenligningsgrundlag for arabisk og hebraisk 62
      Skrift og fonologi 64
      Morfologisk struktur 65
      Diglossi 68
      Sammenfatning 68
   3.5 Den semitiske morfologis mentale repræsentation 69

4. GRAFISK DIMENSION: ARABISK SKRIFT 73
4.1 Læseretning 74
4.2 Bogstavarkitektur 76
Semiotiske modeller og bogstavgenkendelsesteorier 77
Tilegnelse af en ny skrift 81
Tilegnelse af arabisk skrift 84
4.3 Sammenfatning 87

5. ORTOGRAFISK DIMENSION: DET KONSONANTISKE SKRIFTSYSTEM 89
5.1 Læsning på semitiske sprog som L1 89
Tværsproglige studier af øjenbevægelser og læsehastighed 89
Vokalernes betydning for ordbestemmelse, ordgenkendelse og læseforståelse. 90
Rødders og mønstres rolle for ordgenkendelse 97
Studier af dyslexi 101
Betydningen af diglossi 103
Antydninger af en læseprocesteori for semitiske sprog 106
5.2 Læsning på semitiske sprog som L2 112
5.3 Sammenfatning 114

6. DANSKE LØRNERES LÆSNING PÅ ARABISK 117
6.1 Hypoteser og metode 117
Den grafiske dimension – Test A 117
Den ortografiske dimension – Test B 122
Den ortografiske dimension – Test C 125
Spørgeskemaundersøgelse 126
Analyseforløb 126
6.2 Beskrivelse af respondenterne 127
6.3 Testresultater og analyser 132
De danske lørneres indtryk og erfaringer med læsning på arabisk 132
Statistiske kontrolanalyser 137
Test A 140
Test B 151
Test C 158
Kvalitativ analyse af omkodningsfejlene i test A 159
6.4 Sammenfatning 167

7. KONKLUSIONER OG PÆDAGOGISKE PERSPEKTIVER 173
7.1 Kommunikativ undervisning i læsefærdighed på arabisk 174
7.2 Pædagogiske overvejelser og anbefalinger 178

LITTERATURLISTE 189

RESUMÉ & ENGLISH SUMMARY 211

APPENDIX

FIGUROVERSIGT

Figur 1.1 Skriftsystem, skrift og ortografi 15
Figur 1.2 Kategorisering af kompetencer, der aktiveres ved læsning 16
Figur 1.3 Centrale processer og stadier ved ordgenkendelse 18
Figur 1.4 Visuelle og fonologiske fejl fordelt efter niveau ved læsning af enkeltord og sætninger hos Khaldieh 21
Figur 1.5 Eksempler på *glyphs* til indlæring af hebraisk og arabisk hos Bar-Lev 23

Figur 2.1 Hierarki over nøgler til læseforståelse 29
Figur 2.2 Ordgenkendelsesprocessen ifølge dualruteteorien 31
Figur 2.3 Konnektionistisk model for genkendelse af isolerede ord 34
Figur 2.4 Tilpasset konnektionistisk ordgenkendelsesmodel 37
Figur 2.5 Komplet læseprocesmodel med udspecificering af primær og sekundær genkendelse i ordgenkendelsesfasen 45

Figur 3.1 Symbolsystemer 49
Figur 3.2 Skriftsystemernes udvikling 49
Figur 3.3 Arabisk skrift 52
Figur 3.4 Eksempler på grafisk og fonetisk slægtskab i det arabiske skriftsystem 54
Figur 3.5 Arabiske konsonantfonemer med kendt og mindre kendt udtale 54
Figur 3.6 Verbalformer 56
Figur 3.7 Vigtigste verbalafledte nomenformer 56
Figur 3.8 Nomenformer med relation til verbalsystemet 57
Figur 3.9 Eksempler på orddannelse på arabisk med tre rødder og tre mønstre 57
Figur 3.10 Eksempel på homografi på arabisk 58
Figur 3.11 Arabiske kasusendelser 59
Figur 3.12 Model af talesprogsvariation på arabisk 60
Figur 3.13 De i dag talte semitiske sprog 63
Figur 3.14 Hebraisk skrift 63
Figur 3.15 Arabisk og hebraisk skrift uden og med vokalisering 64
Figur 3.16 Hebraiske verbalformer 66
Figur 3.17 De mest konsekvent anvendte verbalafledte nomenformer 66
Figur 3.18 Nomenformer med relation til verbalsystemet 66
Figur 3.19 Nogle arabiske og hebraiske ord afledt af roden كتب og roden ספר 67

Figur 4.1 Eksempel på komponentmatrix og forvekslingsmatrix for de latinske majuskler 79
Figur 4.2 Boumas gruppering af bogstavernes generelle former baseret på forveksling af bogstaver 80
Figur 4.3 Brooks artificielle skriftsystemer 83
Figur 4.4 De latinske majusklers grafiske komposition 86
Figur 4.5 De latinske minusklers grafiske komposition 86
Figur 4.6 De arabiske finalers grafiske komposition 86

Figur 4.7 De arabiske medialers grafiske komposition 86
Figur 4.8 Gruppering af de arabiske bogstavers generelle former baseret på identiske basisformer 87
Figur 4.9 Uddrag af spørgeskemaundersøgelse vedrørende danske L2-lørneres opfattelse af deres læsetilegnelsesproces på arabisk 88

Figur 5.1 Korrelationsmatrix over signifikante forskelle mellem læsning af isolerede ord og tekststykker på vokaliseret og uvokaliseret arabisk og hebraisk 95
Figur 5.2 Reaktionstider for ordbestemmelse og ordgenkendelse samt fejl ved ordbestemmelse hos Frost, Forster & Deutsch og Abu-Rabia & Awwad . 99
Figur 5.3 Forskel i reaktionstider i ms. og % hos Frost, Forster & Deutsch og Abu-Rabia & Awwad 99
Figur 5.4 Model af den valgte læseteoretiske ramme 106
Figur 5.5 Gennemsnitlige fejlprocenter og tider for højtlæsning af hebraiske ord bestående af henholdsvis 1 og 3 morfemer 113
Figur 5.6 Gennemsnitlig tid for læsning og antal læsefejl ved højtlæsning af tekststykke 113

Figur 6.1 Model for oplæsning af pseudoord 118
Figur 6.2 Model for læsning efter forståelse 118
Figur 6.3 Pseudoord konstrueret til brug for test C samt deres ortografiske struktur og mulige valide mønstre 125
Figur 6.4 Fag med arabisk sprog som emne og/eller undervisningsmedium på bachelordelen af kandidatuddannelsen i arabisk ved KU og Cand. negot. uddannelsen i arabisk ved SDU 128
Figur 6.5 Oversigt over respondenter i datasættet 130
Figur 6.6 Køns- og aldersfordeling hos respondenterne på hvert enkelt niveau 131
Figur 6.7 Respondenternes holdning til udsagn om, at det er svært at læse 132
Figur 6.8 Respondenternes holdning til udsagn vedrørende undervisningen 133
Figur 6.9 Respondenternes holdning til udsagn vedrørende specifikke vanskeligheder 134
Figur 6.10 Respondenternes holdning til udsagn vedrørende læsehastighed 135
Figur 6.11 Respondenternes holdning til udsagn om vokalernes betydning 135
Figur 6.12 Respondenternes holdning til udsagn vedrørende læsestrategier 136
Figur 6.13 Udvalgte testresultater for respondenterne på niveau 1 fordelt efter studiested 138
Figur 6.14 Korrelationsdiagrammer over hastighed og korrekthed i test A og B 139
Figur 6.15 Tid for læsning af arabiske pseudoord uden fremmede fonemer 141
Figur 6.16 Tid for læsning af arabiske pseudoord med fremmede fonemer 141
Figur 6.17 Tid for læsning af pseudoord med latinsk skrift 141
Figur 6.18 Tid for læsning af pseudoord på tværs af testtype, niveau 1 142
Figur 6.19 Tid for læsning af pseudoord på tværs af testtype, niveau 2 142
Figur 6.20 Tid for læsning af pseudoord på tværs af testtype, niveau 3 142
Figur 6.21 Tid for læsning af pseudoord på tværs af testtype, niveau 4 143
Figur 6.22 Bogstavfejl ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII) 144

Figur 6.23 Realfejl ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII) 144
Figur 6.24 Fejl i diakritiske tegn ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII) 144
Figur 6.25 Omkodningsfejl ved læsning af pseudoord med latinsk skrift 146
Figur 6.26 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 1 147
Figur 6.27 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 2 147
Figur 6.28 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 3 147
Figur 6.29 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 4 148
Figur 6.30 Tid for læsning af arabiske tekststykker både med og uden vokalisering 151
Figur 6.31 Tid for læsning af arabiske tekststykker med vokalisering (BI) og uden vokalisering (BII) 152
Figur 6.32 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 1 152
Figur 6.33 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 2 153
Figur 6.34 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 3 153
Figur 6.35 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 4 153
Figur 6.36 Kontrolgruppens gennemsnitlige læsehastighed ved læsning af de to typer tekstmateriale i test B 154
Figur 6.37 Læseforståelse ved læsning af arabiske tekststykker med vokaler (BI) og uden vokaler (BII) 155
Figur 6.38 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 1 156
Figur 6.39 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 2 156
Figur 6.40 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 3 156
Figur 6.41 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 4 157
Figur 6.42 Antal korrekte mønstre appliceret på uvokaliserede pseudoord 158
Figur 6.43 Oversigt over substitutionsfejl hos de indfødte sprogbrugere 160
Figur 6.44 Display over samtlige forvekslingsfejl hos de danske lørnere 163
Figur 6.45 Display over forvekslinger af bogstaver, der ikke har samme grafiske base 163
Figur 6.46 Display over forvekslinger af bogstaver, der hverken har åbenlyst grafisk sammenfald i base eller samme antal prikker 164
Figur 6.47 Grafiske og fonetiske kategorier, inden for hvilke bogstavsubstitution forekommer hos danske lørnere 164
Figur 6.48 Forveksling af bogstaver der udelukkende har grafisk eller fonetisk lighed, og som forekommer mere end én gang 165
Figur 6.49 Forveksling af bogstaver der både har grafisk og fonetisk lighed, og som forekommer mere end én gang 166
Figur 6.50 Gennemsnitlige omkodningshastigheder på alle niveauer og ved alle testtyper i test A 168
Figur 6.51 Gennemsnitligt antal realfejl på alle niveauer og ved alle testtyper i test A 168
Figur 6.52 Gennemsnitligt antal bogstavfejl ved læsning af arabiske pseudoord i test A 168

Figur 6.53 Gennemsnitlige læsehastigheder på alle niveauer og ved begge testtyper i test B 170
Figur 6.54 Gennemsnitligt antal korrekte besvarelser i forståelsestestene på alle niveauer og ved begge testtyper i test B 170
Figur 6.55 Gennemsnitligt antal korrekte mønstre appliceret på uvokaliserede pseudoord i test C 171

Figur 7.1 Eksempler på prioritering af læsetræning i to lærebogssystemer til undervisning i arabisk som L2 175
Figur 7.2 Pædagogiske traditioner i fremmedsprogsundervisningen 176
Figur 7.3 Elementer der er afgørende for motivation til at læse på L2. Efter Day & Bamford 180
Figur 7.4 Elementer der er afgørende for motivation til at læse på L2. Efter Mori 180

APPENDIXOVERSIGT

APPENDIX 1: TESTMATERIALE
1.1 Instruktionsark på dansk
1.2 Instruktionsark på arabisk
1.3 Test AI
1.4 Test AII
1.5 Test AIII
1.6 Test BI, niveau 1, tekst A
1.7 Test BII, niveau 1, tekst A
1.8 Svarark på dansk, niveau 1, tekst A
1.9 Svarark på arabisk, niveau 1, tekst A
1.10 Test BI, niveau 1, tekst B
1.11 Test BII, niveau 1, tekst B
1.12 Svarark på dansk, niveau 1, tekst B
1.13 Svarark på arabisk, niveau 1, tekst B
1.14 Test BI, niveau 2, tekst A
1.15 Test BII, niveau 2, tekst A
1.16 Svarark på dansk, niveau 2, tekst A
1.17 Svarark på arabisk, niveau 2, tekst A
1.18 Test BI, niveau 2, tekst B
1.19 Test BII, niveau 2, tekst B
1.20 Svarark på dansk, niveau 2, tekst B
1.21 Svarark på arabisk, niveau 2, tekst B
1.22 Test C
1.23 Spørgeskema på dansk
1.24 Spørgeskema på arabisk

APPENDIX 2: ANALYSEARK
2.1 Fejlregistreringsskema til test A
2.2 Resultatregistreringsskema til test A, B og C

APPENDIX 3: STATISTISK MATERIALE
3.1 Datasæt
3.2 Statistiske analyser
3.3 Kommandoer til programmering af grafer

APPENDIX 4: ØVRIGE DATA
4.1 Samtlige registrerede substitutionsfejl i test A
4.2 Besvarelser i spørgeskemaets åbne svarkategori
4.3 Lydfiler (kun vedlagt eksemplarer til bedømmelse)

Forord

Jeg vil gerne takke alle de mange personer, der på den ene eller anden måde har ydet et bidrag til denne afhandling. Især vil jeg takke de 81 studerende, der indvilgede i at aflægge læsetest til brug for den empiriske undersøgelse. Også tak til Salim Abu-Rabia fra Haifa Universitet, der tog sig tid til at drøfte testdesignet med mig under mit forskningsophold i Israel, og til Sonja Wehberg fra Center for Sprogtilegnelse ved Syddansk Universitet, som fungerede som konsulent på statistikken. En særlig tak til Børge Spang-Thomsen, der med stor omhu har læst og kommenteret manuskriptet. Min vejleder, Helle Lykke Nielsen, skal også have tak for god og konstruktiv kritik. Endelig vil jeg takke min søde mand, Abdennbi, der uden at kny flyttede familien til Israel i 6 måneder, gennem hele forløbet tog det store slæb med hjem og børn, og som – altid velvilligt – lagde øre til mine monologer om arbejdets fremskridt eller mangel på samme.

1. Indledning

Arabisk har traditionelt været betragtet som et eksotisk fremmedsprog i Danmark og resten af Vesten, og i universitetssammenhæng har det haft status af et filologisk forsknings- og undervisningsområde på linje med latin og oldgræsk. Derfor har arabiskundervisning i universitetsregi normalt været domineret af litteraturstudier med vægt på læsefærdighed baseret på grammatikstudier og oversættelse, mens skrive-, lytte- og talefærdighed først og fremmest blev trænet med det formål at styrke elementer i læseindlæringen.

Filologien har fortsat en dominerende position i de arabiske forsknings- og undervisningsmiljøer, men den såkaldte globaliseringsproces har inden for de seneste par årtier været anledning til, at arabisk i stadig højere grad efterspørges som et moderne fremmedsprog. Således har øget turisme i, samhandel med og indvandring fra de arabiske lande gjort kommunikationsfærdigheder i moderne arabisk efterspurgt inden for mange forskellige erhverv, ligesom det politiske fokus på islamisk relateret terrorisme i kølvandet af terrorangrebene 11. september 2001 har medført efterspørgsel på arabiske sprogfærdigheder hos efterretningstjenester og udviklingsorganisationer i den vestlige verden. Derfor har flere universitetsmiljøer efterhånden omstillet eller suppleret den filologiske tilgang med kommunikativt orienteret undervisning i moderne standardarabisk. Sammenholdt med at efterspørgslen på arabiskundervisning i samme periode er steget markant – ved Syddansk Universitet er arabisk i dag det næststørste sprogfag efter engelsk – er behovet for en moderne fremmedsprogspædagogisk metode for undervisning i arabisk blevet påtrængende, og i visse undervisningsmiljøer, især i USA, har undervisningsformen har da også været igennem en løbende reformproces siden 1980'erne. I Danmark har Syddansk Universitet siden 1992 udbudt undervisning i kommunikativt arabisk, først på kombinationsuddannelsen Cand. negot. og siden også som sidefag på både dagstudier og Åben Uddannelse. Også ved Københavns Universitet har man etableret undervisning i moderne standardarabisk som et element i det ellers filologisk funderede arabiskstudium, mens man ved Aarhus Universitet har bibeholdt en entydig filologisk profil.

Naturligt nok har den pædagogiske tilpasning, der har fundet sted, været inspireret af den kommunikative tilgang, der har præget fremmedsprogspædagogikken generelt siden 1970'erne. Dette har i de arabisk-lingvistiske forskningsmiljøer givet anledning til diskussion om problemstillinger som hvilken variant af sproget, der skal undervises i, hvorvidt tale- eller skriftsprog skal introduceres først, og hvorvidt der skal undervises eksplicit i grammatik. Imidlertid mangler man en egentlig forskningsbaseret tilgang til disse problematikker – ikke mindst fordi de faglige miljøer inden for arabisk er små og spredte og desuden har haft svært ved at omstille sig til den nye situation. Derfor er den samlede kommunikativt orienterede fremmedsprogspædagogiske forskning vedrørende arabisk som målsprog særdeles begrænset. Som resultat har de pædagogiske metoder hidtil i meget høj grad måttet forlade sig på den generelle fremmedsprogspædagogik, der er stærkt domineret af forskning i europæiske sprog – især engelsk – som målsprog.

Denne tilgang er problematisk, når det gælder undervisning i læsefærdighed på arabisk. Ganske vist hviler de moderne pædagogiske metoder, der i den tvær-europæiske kontekst danner udgangspunkt for undervisningen i fremmedsprogslæsning, på formodningen om at erhvervede læsefærdigheder fra modersmålet (L1) kan bruges ved læsning på fremmedsproget (L2). Imidlertid peger en del tværsproglig læseforskning på, at ortografiske forskelle

mellem L1 og L2 er relevante for læseprocessen og dermed for læsefærdighed, hvilket indicerer, at læseprocesser – og dermed også læsetilegnelsen – forløber forskelligt ved brug af forskellige skriftsystemer. Alligevel har læsning på arabisk som fremmedsprog hidtil kun været levnet særdeles beskeden forskningsmæssig opmærksomhed, og selv de nyeste materialer til undervisning i læsefærdighed på arabisk savner et moderne pædagogisk svar på dette problem. De nyskabende elementer begrænser sig således til større fokus på bestemte læsestrategier, der var centrale for den kommunikative tilgang i fremmedsprogs-undervisningen i 1980'erne, men som i dag udfordres fra flere sider. Samtidig adskiller de elementer, der skal tilgodese specifikt arabisk-relaterede indlæringsbehov, sig ikke afgørende fra den filologiske tradition: Alfabetet præsenteres i et indledende forløb og får derefter ingen særbehandling, og den semitiske morfologiske struktur behandles ikke fundamentalt anderledes end tidligere.[1]

Formålet med dette forskningsprojekt er at tage hul på dette reelt uopdyrkede forskningsområde og derved tilføre fremmedsprogspædagogikken for arabiskundervisningen et moderne læseteoretisk perspektiv.

1.1 Problemformulering og begrebsafklaring

Mere specifikt er målet med denne afhandling at definere nogle centrale problemområder på et forskningsbaseret grundlag, og på den baggrund anvise pædagogiske retningslinier, der kan opkvalificere læseundervisningen i arabisk.

Afhandlingen er altså afgrænset til at omhandle aspekter af tilegnelsen af arabisk, der

- vedrører læseprocesser
- bevirker, at tilegnelsen adskiller sig fra tilegnelse af fremmedsprog, der skrives med det latinske alfabet

Med hensyn til det første punkt skal det præciseres, at læsning i denne sammenhæng betragtes som en kommunikativ kompetence. Det kommunikative aspekt indebærer, at blandt de mange mulige mål, der kunne være for at engagere sig i læsning (fx indlæring, kritik eller indhentning af specifik information), er målet for læsning i denne sammenhæng generel tekstforståelse, og selve læsningen er en umiddelbar, flydende proces. Således betragtes analytiske mellemled som eksempelvis ordbogsopslag ikke som et led i læse-processen. Dette indebærer ikke en værdisætning af de forskellige tilgange til tekstmateriale, men er blot udtryk for en afgrænsning af læsning defineret som et løbende, uafbrudt forløb, hvor sproganalytiske overvejelser som undervejs fjerner læserens opmærksomhed fra selve teksten, udgør brud på processen. Tilgange, der vægter analytiske processer kan betegnes som strategier, hvilket kan defineres som færdigheder, der er under et individs bevidste kontrol. I modsætning hertil står automatiserede processer. I praksis er en konsekvent skelnen mellem strategier og automatiserede kompetencer i relation til læsning ikke mulig, idet læseaspekter, der i teorien betegnes som strategiske, fx at springe over et ukendt ord eller genlæse en passage for at sikre korrekt forståelse, hos den kompetente læser i vid udstrækning er automatiserede processer. Automatiserede kompetencer tilegnes ofte som strategier, der efterhånden automatiseres, og der er således tale om en glidende overgang.

[1] De fremførte påstande uddybes og dokumenteres i kapitel 7

Anledningen til alligevel at introducere denne skelnen er, at L2-læseteori traditionelt fokuserer på en eksplicit strategisk tilgang til tekst, hvilket ikke er tilfældet her.

Med hensyn til tilegnelsesaspektet er udgangspunktet for projektet tesen om, at det arabiske skriftsystem indeholder elementer, der er så fremmede for lørnere med et latinskskrevet L1, at de kræver særlig opmærksomhed i læseundervisningen i arabisk. Dette fokus på de skriftmæssige forskelle indebærer, at selvom læsning involverer mange forskellige aspekter og aktiverer helt generelle sproglige kompetencer, bliver de tekstmæssige sider af læseprocessen, dvs. ordgenkendelse[2], projektets læseteoretiske omdrejningspunkt.

Således vil jeg undersøge:

Hvilken betydning har de elementer i det arabiske skriftsystem, der er fremmede for danske lørnere, for deres visuelle ordgenkendelse på arabisk?

Undersøgelsen skal udmunde i nogle konkrete anbefalinger af, hvordan undervisningen i læsefærdighed på arabisk kan målrettes de pædagogiske behov, som de særegne dimensioner ved skriftsystemet afstedkommer.

I det følgende uddybes og defineres de centrale begreber 'det arabiske skriftsystem', 'fremmede elementer', 'visuel ordgenkendelse' og 'danske lørnere'. I tilfælde, hvor der ikke findes nogen dansk tradition for de anvendte begreber, tilstræbes det at etablere en dansk terminologi.[3]

Det arabiske skriftsystem

Hvad angår problemformuleringens centrale begreb 'skriftsystem' finder man i den engelsksprogede faglitteratur, der er relevant for dette emne, en inkonsekvent og usystematisk brug af begreberne *script, orthography* og *writing system*, og en nærmere definition af begreberne – og deres danske udgaver – er derfor nødvendig:

På dansk er der ikke etableret nogen egentlig fagspecifik terminologi på området – 'skrift' bruges normalt enten som modsætningen til tale eller som synonym for skrifttype, fx gotisk skrift eller kursiveret skrift, og 'ortografi' bruges udelukkende i en ordret oversættelse fra græsk om retskrivning. 'Skriftsystem' findes ikke i nogen af de danske ordbøger og leksika, der er konsulteret, men anvendes som fagudtryk af den centrale danske læseforsker Elbro (2001), der definerer det som et glottografisk symbolsystem – et grafisk symbolsystem, der gengiver det talte sprog – som modsætning til semasiografier, som er grafiske symbolsystemer, der er sproguafhængige. Dette falder sammen med den franske skriftforsker Sirats definition af '*l'écriture*':

> *" L'écriture est le système de signes graphiques utilisés dans les diverses civilisations. Elle est liée à la langue orale qui l'a précédée dans le temps."* (Sirat 1990)

I den engelsksprogede faglitteratur anvendes begreberne anderledes og ofte med en langt bredere betydning. To af de mest centrale forskere i læsning på tværs af skriftsystemer, Katz

---

[2] Begrebet ordgenkendelse defineres nærmere nedenfor.

[3] Visse engelske begreber, der er meget veletablerede i dansk fagterminologi, som fx *schema*, *top-down-* og *bottom-up*-læsemodeller, fastholdes dog.

& Frost, definerer *script* som '*a set of symbols*' (Katz & Frost 1992:68), og den mest internationalt refererede tyske læseforsker, Scheerer skriver:

> *"A script is defined by the set of characters it uses and by the mapping principle on which it relies. […] when considered with reference to languages using them, scripts become writing systems."* (Scheerer 1986:263)

I *The Blackwell Encyclopedia of Writing Systems* (Coulmas 1996:454) defineres *script* i overensstemmelse hermed som '*the graphic form of a writing system*'. Men skriftforskeren Coulmas skriver desuden:

> *"The same writing system may be written in a variety of scripts. For example, the Roman, Cyrillic, Greek, Russian and runic scripts are different graphic instantiations of the same writing system, the alphabet."* (Coulmas 1996:454)

Her er *writing system* lig med det, som Scheerer kalder '*the mapping principle*' og ikke en grafisk repræsentation. Denne implicitte definition på *writing system* stemmer da heller ikke overens med Coulmas' egen definition på netop dette begreb:

> *"Writing system: A set of visible or tactile signs used to represent units of language in a systematic way, with the purpose of recording messages which can be retrieved by everyone who knows the language in question and the rules by virtue of which the units are encoded in the writing system."* (Coulmas 1996:560).

Her synes *writing system* at være det samme som *script.* Som udgangspunkt for denne afhandling sammenholdes Elbros, Katz' og Scheerers definitioner (Scheerer 1986:263-264; Katz & Frost 1992:68; Elbro 2001:56) til følgende: Et skriftsystem er et system for grafisk repræsentation af et sprog, og skriften er det grafiske symbolsæt, skriftsystemet anvender. Det latinske alfabet, det arabiske alfabet, de japanske syllabiske systemer og det kinesiske tegnsystem er altså forskellige eksempler på skriftsystemer, hvor skriften er betegnelsen for deres grafiske udtryk.

I en stor del af faglitteraturen bruges *orthography* tilsyneladende som metabetegnelse for det hele og altså ikke bare bredere, men faktisk helt anderledes, end ortografi bruges på dansk. Kun få har bestræbt sig på at give en definition. En undtagelse er tyskeren Scheerer (1986), som harcelerer over den slappe brug af begrebet i den psykologiske faglitteratur og kalder på en opstramning, hvor begrebet udelukkende bruges om det normative aspekt, nemlig retskrivning (Scheerer 1986) – faktisk ligesom det bruges på dansk. Men noget tyder på, at synspunktet er europæisk forankret og ikke kan finde opbakning i den amerikansk dominerede fagdisciplin. De følgende 15 års litteratur viser i hvert fald med al tydelighed, at opsangen ikke har haft nogen effekt. Faktisk bruges begrebet nogle steder snævert om skriftsprogs visuelle udtryk. Som de amerikanske læseforskere Willows & Geva påpeger:

> *"[...] it is fairly common in the growing literature on orthographic processing for researchers to refer to orthographic processing as 'visual/orthographic' as though these two terms were essentially synonymous"* (Willows & Geva 1995:356)

Ligeledes har kognitions- og læseforskerne Carr & Posner ifølge den amerikanske forsker i ordgenkendelse, Foorman defineret *orthography* som "*the visual organization of the writing system*" (Foorman 1994:334), hvilket kommer i åbenlys karambolage med Scheerers definition af *script* og Coulmas' definition af *writing system*. I den nyeste amerikanske litteratur bruges *orthography* meget bredt om alle de sprogspecifikke aspekter ved skriften. Ortografien er her måden, et sprog skrives på, og både retstavning, skrift og skriftsystem er således en del af ortografien (fx Seidenberg 1992:85). Det kunne være fristende at forsvare

den europæiske tradition for at tage det græske betydningsindhold alvorligt og kun bruge begrebet om retskrivning, men det vil for det første gøre formidlingen af forskningen på området langt mere besværlig og desuden skabe usikkerhed omkring andre begreber, der er helt centrale i læseteorien. På den anden side synes det uhensigtsmæssigt at benytte begrebet i den ovenfor omtalte meget brede betydning. Derfor defineres ortografi hér ud fra den kernebetydning, begrebet har i litteraturen, og på en måde, der samtidig giver det en komplementær frem for overlappende funktion i forhold til de andre begreber, således, at mens skriften repræsenterer det 'ydre', repræsenterer ortografien det 'indre': Ortografien er de sprogspecifikke karakteristika ved den grafiske repræsentation af et sprog, dvs. både retskrivningen og strukturen, eksempelvis for koreansk *hanguls* vedkommende systemet, hvor tegnene illustrerer artikulationen af de fonemer, de repræsenterer.

SKRIFTSYSTEM	
SKRIFT	**ORTOGRAFI**
Skriftsystemets grafiske dimension sproguafhængig	Skriftsystemets sprogspecifikke dimension

Figur 1.1 Skriftsystem, skrift og ortografi

Sammenfattende er skriftsystemet – som illustreret i figur 1.1 – systemet for grafisk repræsentation af et sprog, hvor skriften er betegnelsen for den grafiske dimension, der er uafhængig af hvilket sprog, der er tale om, mens ortografien repræsenterer den sprog-specifikke dimension, som i princippet er uafhængig af skriften. Skriften og ortografien er altså tilsammen skriftsystemet. Det arabiske skriftsystem omfatter således både arabisk skrift i form af det arabiske alfabet med øvrige grafiske tegn og den særegne semitiske morfologiske struktur. Der gives en egentlig præsentation af det arabiske skriftsystem i kapitel 3, hvor det desuden sammenlignes med det hebraiske skriftsystem, som en del af den anvendte litteratur bygger på, ligesom den semitiske, rodbaserede morfologi og dens påståede mentale repræsentation beskrives nærmere.

Fremmede elementer

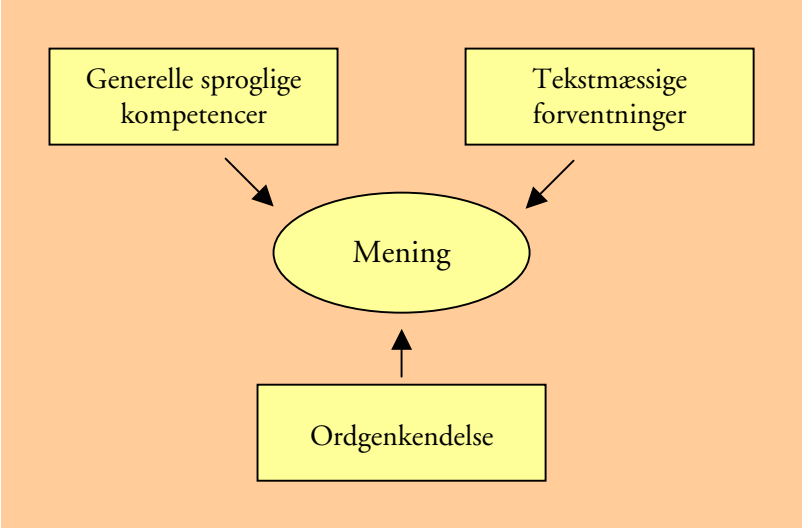
Problemformuleringens afgrænsning til fokus på de 'fremmede elementer' indebærer, at projektet koncentrerer sig om de tre visuelle dimensioner, der gør det arabiske skriftsystem anderledes for studerende, der er vant til det latinske alfabet:

- Læseretningen går fra højre mod venstre, hvilket som bekendt er modsat af læseretningen på sprog skrevet med det latinske alfabet.
- Bogstaverne er grafisk forskellige fra de latinske.
- Skriftsystemet er konsonantisk, hvilket kommer til udtryk ved at kun konsonanter og lange vokaler markeres som bogstaver, mens korte vokaler udelades eller markeres som diakritiske tegn.

Mens de første to dimensioner er entydigt grafiske aspekter, er den tredje dimension mere kompleks. Ud over at den grafisk manifesterer sig ved kortere ord, repræsenterer den reducerede fonologiske information og den morfologiske struktur, som ligger til grund for udtryksformen, endnu en fremmed dimension, der kan tænkes at have betydning for den visuelle ordgenkendelse.

Visuel ordgenkendelse

Ordgenkendelse er en sproglig perceptionsproces, og at den er visuel betyder at den er forbundet med læsning. Visuel ordgenkendelse er en af flere former for kompetencer, der aktiveres, når man læser. I figur 1.2 ses en grov kategorisering af disse. Ud over ordgenkendelse aktiveres læserens generelle sproglige kompetencer og forventninger til teksten.



Figur 1.2 Kategorisering af kompetencer, der aktiveres ved læsning

Ordgenkendelse kan i modsætning til de øvrige kompetencer ses som en afgrænset delproces ved læsning. I ordgenkendelsesteorierne opererer man med centrale begreber som *decoding*, *recoding, encoding* og *word recognition*, men der findes i den engelsksprogede faglitteratur ikke nogen entydig praksis for, hvordan de defineres. Nogle, fx Austad et al., Ehri, Calfee & Drum, Samuels & Kamil (gengivet i Kulbrandstad 1996:75-99) har defineret *decoding* (afkodning) som 'den tekniske side af læsningen', dvs. syntesen af de enkelte tegn til ord. Andre, fx Goodman (1982), der var en af 1970'ernes mest indflydelsesrige læseforskere, mener at syntaktisk, semantisk og pragmatisk viden aktiveres i afkodningsprocessen. Goodman (1982:128) skelner mellem *recoding* og *decoding*, hvor *recoding* (omkodning) er processen, hvor grafem overføres til fonem, mens *decoding* er processen, hvor grafemisk eller fonemisk input oversættes til mening. Elbro (2001) synes at have adopteret Goodmans distinktion:

> *"Ved afkodningen bevæger man sig fra den skriftligt kodede meddelelse til en identifikation af ordene – med alt hvad det indebærer af genkendelse af ordenes generelle betydning, lyd og grammatiske egenskaber. [...] Afkodningen er altså ikke*

> *bare en omkodning (eller oversættelse) til et andet kodesystem, fx til sproglyd. Afkodningen leder til en egentlig genkendelse af ordene."* (Elbro 2001:28)

Afkodning synes således hos Elbro at være det samme som ordgenkendelse (*word recognition*), der i et dansk hovedværk af Dalby et. al. (1989) forklares således:

> *"Genkendelse af et ord vil sige, at læseren ved synet af ordet forbinder det med noget bekendt, en betydning og/eller et navn. [...] Genkendelse af et ord indebærer, at synet af et ord 'vækker' ordet i læserens ordforråd (indre leksikon)."* (Dalby et al. 1989:101)

Just & Carpenter (1987:9,326) refererer i deres meget anvendte lærebog i læsepsykologi til *decoding* som knyttet til begynderlæsning og *encoding* som øvede læseres transformation af input til mentale symboler, mens *word recognition* er en kombination af de to delprocesser *encoding* og *lexical access* (adgangen til ordets betydning i det indre leksikon):

> *"The term decoding primarily focuses on the process of extracting information from the orthographic code. The term's focus is appropriate for early reading because written language is a code that beginning readers must crack. By contrast, the term encoding focuses on the process of transforming the information into mental symbols. In fact, both children and adults execute both processes, although skilled readers may perform both functions with a single process."* (Just & Carpenter 1987:326)

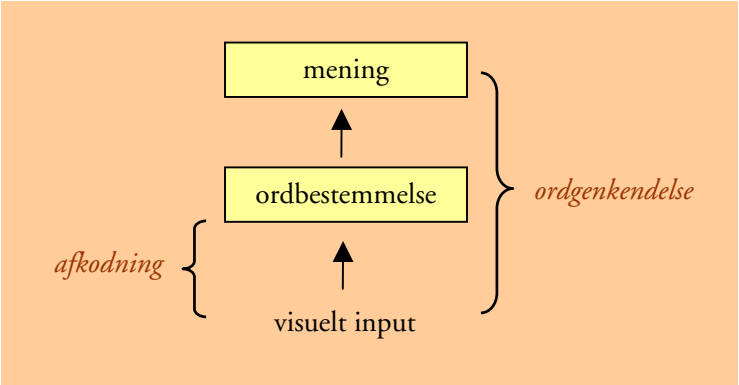
> *"Word recognition…[involves]…a combination of two processes: (1) encoding the visual pattern of a printed word, and (2) accessing its meaning in the mental dictionary or lexicon (lexical access)."* (Just & Carpenter 1987:9)

Samtidig betegner Just & Carpenter den rent fonologisk medierede omkodning, dvs. selve oversættelsen af tegn til lyd som *speech recoding* (Just & Carpenter 1987:93,338).

Just & Carpenter opererer altså med forholdsvis mange forskellige, ret uklart definerede, kategorier, som tilsyneladende i et vist omfang overlapper hinanden. Goodman og Elbro holder sig omvendt lidt snævert til distinktionen mellem den altfavnende afkodning og den specificeret lyd-medierede omkodning, hvilket udelukker skelnen mellem afkodning og ordgenkendelse. Det er i denne sammenhæng uhensigtsmæssigt, da netop disse forskellige perceptionsstadier er udgangspunkt for en betydelig del af den forskning i læsning på semitiske sprog, der præsenteres i kapitel 5. Her udnyttes i vidt omfang forskellen mellem ordbestemmelse (*lexical decision*), dvs. genkendelse af en bogstavstreng som et gyldigt ord, og ordgenkendelse, som jo forudsætter *lexical access*. Samtidig er distinktionen relevant, fordi forskellen synes så reel – vi kender vel alle tilfælde, hvor vi afkoder et ord uden at vide præcist hvad det betyder, ligesom vi af og til den anden vej søger et ord i 'det mentale leksikon', hvis betydning vi meget klart kender, men som vi blot ikke lige her og nu kan komme i tanker om. I figur 1.3 illustreres disse delprocesser med en syntese af Goodman/Elbros og Just & Carpenters terminologi, som fremover vil blive brugt i denne afhandling. Bemærk, at selvom vejen fra visuelt input til leksikon **kan** foregå som to separate processer, gør de det ikke nødvendigvis, da ordbestemmelse ofte indebærer forståelse af ordets betydning, hvorved afkodning *de facto* **er** ordgenkendelse.[4]

---

[4] Modellen udbygges i kapitel 2, og her vil de to niveauer præsenteres som ét, idet ordbestemmelse som separat niveau ikke har nogen relevans før gennemgangen af læsning på semitiske sprog i kapitel 5.



Figur 1.3 Centrale processer og stadier ved ordgenkendelse

Hvad der mere nøjagtigt karakteriserer den kognitive proces ved ordgenkendelse diskuteres nærmere i kapitel 2, hvor afhandlingens læseteoretiske ramme defineres og sættes i relation til læsning på fremmedsprog, herunder fremmedsprog, der grafisk repræsenteres af et andet skriftsystem end det der anvendes på modersmålet. I den forbindelse skal det nævnes, at der – på trods af, at min tese i sidste instans indebærer, at læseprocessen måske forløber anderledes ved læsning på arabisk end på fx dansk, engelsk og fransk – tages udgangspunkt i det eksisterende læseteoretiske grundlag, som ensidigt er grundlagt på empiri fra europæiske sprog. Denne konflikt imødegås ved, at de gængse teorier efterfølgende sættes i relation til de sprogspecifikke aspekter.

Danske lørnere

Ordet 'lørner' er en fordanskning af det engelske *learner* og er et forholdsvis veletableret begreb i dansk fremmedsprogsteoretisk terminologi. I dette forskningsprojekt undersøges danske arabiskstuderendes læsning på arabisk. Der er imidlertid flere grunde til at omtale forsøgspersonerne som lørnere frem for studerende. For det første har enkelte af deltagerne afsluttet deres studium, hvorfor de i teknisk forstand ikke længere er studerende. Derimod er de fortsat lørnere. For det andet er begrebet 'lørner' mere hensigtsmæssigt i denne sammenhæng, fordi det mere eksplicit relaterer til indlæringsprocessen end 'studerende', der nærmere henviser til aktiviteten at studere.

At forsøgspersonerne er danske lørnerne indebærer i denne sammenhæng, at de har dansk som modersmål og ikke har stiftet bekendtskab med sprog, der skrives med det arabiske alfabet, før de påbegyndte deres arabiskuddannelse på universitetet. Afgrænsningen skyldes at alle forsøgspersonerne er danske, men projektets konklusioner vil i princippet kunne generaliseres til alle lørnere med et modersmål, der skrives med det latinske alfabet.

1.2 Eksisterende forskning på området

Mens der efterhånden er et lille udbud af studier i indlæring af de produktive færdigheder på arabisk, findes der kun ganske få forskningspublikationer, der beskæftiger sig med læsning på arabisk som L2. Da forskningsgrundlaget er så beskedent og derfor ikke i sig selv kan repræsentere en ramme for afhandlingen, udgør denne gennemgang ikke en forskningsoversigt i traditionel forstand. En mere dybtgående forskningsoversigt, som inddrager forskning fra flere forskellige fagdiscipliner, som på forskellig vis kan belyse

problemstillingen foreligger i kapitel 4 og 5, som omhandler hhv. den grafiske og den ortografiske dimension. Formålet med nedenstående oversigt er således udelukkende at positionere nærværende afhandling i forhold til eksisterende forskning om læsning på arabisk som L2. I det følgende præsenteres al tilgængelig forskning om emnet. Den falder inden for tre kategorier:

- Studier, der vedrører læsestrategier
- Studier, der vedrører automatiserede læsekompetencer
- Studier, der vedrører læsepædagogik

Den følgende gennemgang er struktureret efter disse kategorier. De studier, der falder inden for de sidste to, beskæftiger sig alle med læseprocessens lavere niveauer som ordgenkendelse, ortografi og fonologi, der også er nogle af de delemner, denne afhandling omhandler.

Strategier

Shakir (1996) sidestiller læseproblemer hos arabiskstuderende med korrekt oversættelse. Han præsenterer en række typiske lingvistiske og tekstuelle problemer i en konkret oversættelsesopgave, fx at især konjunktioner og substantiver ofte oversættes fejlagtigt, fordi deres mening behandles isoleret og ikke som en del af en kontekst-bundet struktur. Shakir forsøger med undersøgelsen at dokumentere, at samtale om teksten fremmer oversættelsesresultatet. Efter en fælles diskussion om teksten mellem læreren og de 12 deltagende studerende, hvor både overordnede aspekter (fx tekstforfatterens og de studerendes holdning til emnet og opdeling af teksten efter delemner) og mere specifikke tekstuelle aspekter (fx kontekstens betydning for ordenes semantiske indhold), viste en efterfølgende nyoversættelse da også langt færre fejl hos de studerende. Resultatet er ikke overraskende, da det er den samme tekst, der oversættes igen efter klassediskussionen.

I en undersøgelse, der omhandler samme problematik mere udførligt, undersøger Fragman (1999) læseforståelse på arabisk hos en gruppe gymnasieelever i Israel. Ligesom hos Shakir afspejler studiet, at oversættelse normalt anvendes som eneste element i læseundervisning på arabisk, men til forskel fra Shakir forsøger Fragman at rette blikket mod en metode, der helt eliminerer oversættelsesopgaverne. I løbet af en periode på syv uger instruerede han de 20 deltagende lørnere i at referere de læste tekster i stedet for som normalt at oversætte dem. Derefter testede han deres læsning af to ukendte tekster, hvor den ene skulle oversættes og den anden refereres på hebraisk (de studerendes L1). Fragman fandt, at de studerende blev bedre til at besvare overordnede spørgsmål om teksten, men også at de modsat hans forventninger besvarede grammatiske spørgsmål og mere specifikke spørgsmål om enkelte dele af teksten omtrent lige så korrekt, når de lavede referat af teksterne, som når de oversatte dem. Endvidere viste et sideløbende, kvalitativt studium, at lørnerne følte, at de blev bedre til fx at skelne mellem tekstens overordnede budskab og underordnede, argumenterende elementer. De følte også, at deres generelle tekstforståelse var bedre, når de refererede, end når de oversatte, at de gjorde bedre brug af ordbogen (i stedet for at slå hvert ord op, brugte de den til at kontrollere betydningen af nøgleord, de var i tvivl om), og nogle studerende følte endog, at de også var blevet bedre til at håndtere tekster på modersmålet, fordi de kunne overføre de tilegnede læsestrategier til læsning på hebraisk.

Der findes andre studier fra samme periode, der beskæftiger sig med læser-strategier uden at relatere til oversættelsesmetoden. Disse afspejler tydeligt den tidlige kommunikative undervisningsmetodes indtog i arabiskundervisningen, og fokuserer næsten udelukkende på overordnede processer som hypotesedannelse og andre læserstrategier: Aweiss (1993) undersøger læsning på arabisk hos fem amerikanske arabiskstuderende på begynderniveau. Data består af de deltagende studerendes kommentarer i forbindelse med læsning, hvor de blev bedt om at 'tænke højt'. De studerendes udsagn under læsningen afspejler, at de danner hypoteser om tekstens indhold og aktiverer baggrundsviden om de emner, teksten omhandler, og at de to strategier anvendes nogenlunde lige meget. Dertil kommer udsagn i form af spørgsmål, som ikke vedrører hypotesedannelse, citater fra teksten, der be- eller afkræfter hypoteser, validerende udsagn, når en hypotese bekræftes og metakommentarer, fx "det her er for svært, det giver ingen mening". Aweiss anbefaler på den baggrund, at instruktion i de anvendte læsestrategier indgår i den pædagogiske praksis for L2-læseundervisning.

I et lignende studium af Khaldieh (1999) undersøges læser-strategier hos 15 studerende opdelt i kategorierne svage L2-læsere, gode L2-læsere og L1-læsere ved hjælp af samme introspektive metode som hos Aweiss. I dette studium registreres hypotesedannelse på grundlag af skimning, overskrifter og billeder samt meningskonstruktion på grundlag af fx tekststruktur, baggrundsviden og forud- og bagudlæsning. Khaldieh registrerede dog også andre, mere tekstfunderede læsestrategier, som fx brug af grammatisk viden, leksikalsk analyse og analyse af enkeltsætningers betydning, og fandt at disse strategier blev langt hyppigere anvendt hos L2-læsere end hos L1-læsere, og at gode L2-læsere brugte dem mere end svage L2-læsere.

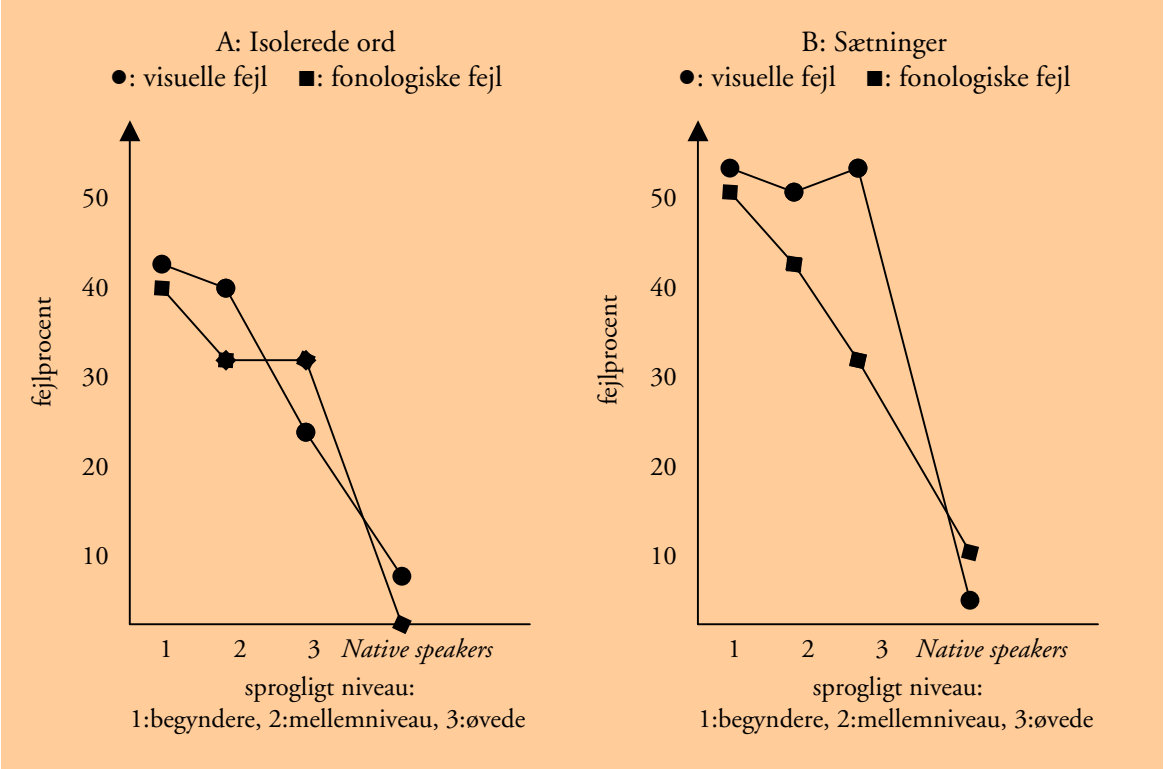
Automatiserede kompetencer

Hana (1972)[5] gennemførte en analyse af læsefejl hos tolv amerikanske arabiskstuderende på begynderniveau. Artiklen omfatter to delstudier. Det første præsenteres som et forsøg på at etablere en ramme for kvantitativ analyse af læsefejl på arabisk, men er i praksis mest af alt en kvalitativ analyse af de deltagende studerendes fejl. På baggrund af en diagnostisk metode anvendt på engelske børn med læsevanskeligheder (Monroes system) opstilles fejlkategorierne 'vokalfejl', 'konsonantfejl', 'tilføjelse af lyd', 'udeladelse af lyd' og 'repetition af lyd'. Fejltypen 'ombytning af lyde', som anvendes i Monroes originale analysemetode, er udeladt, da den slet ikke forekommer i datamaterialet. Studiet viser, at 'tilføjelse af lyd' er den hyppigste fejl, mens 'udeladelse af lyd' og 'gentagelse af lyd' er de mindst hyppige. Det mest interessante resultat er dog, at de 12 deltagende studerende har meget forskellige svagheder. Forfatteren konkluderer, at Monroes system er et godt værktøj til at diagnosticere de studerendes individuelle svagheder. I det opfølgende studium testes de studerende igen, efter at otte af dem har modtaget specialundervisning rettet mod at udbedre de individuelle svagheder, som var blevet diagnosticeret. Her konstateres det, at de studerende, der modtog specialundervisning, har halveret antallet af læsefejl og dermed viser signifikant bedre resultater end kontrolgruppen. Forfatteren konkluderer desuden, at studerende, der falder igennem i normal klasseundervisning, kan opnå gode resultater ved hjælp af specialundervisning tilpasset deres individuelle behov. Hanas studium er

[5] Denne artikel er et sammendrag af to tidligere publicerede artikler: Hanna (1964) og (1966). Da indholdet er fuldstændig det samme, vil disse ikke blive behandlet yderligere. At forfatterens navn staves anderledes, skyldes formentlig inkonsekvent translitteration.

interessant, fordi det afspejler datidens interesse for grafemers og fonemers rolle for ordgenkendelse, som nærmest forsvandt med den kommunikative tilgangs indtog. Imidlertid er den anvendte fejldiagnosticering på visse punkter problematisk, fordi den kommunikative dimension er helt fraværende. Fejlregistreringen afspejler datidens fokus på korrekt oplæsning, hvilket fx medfører at fejlkategorien 'gentagelse af lyd' også omfatter tilfælde, hvor testpersonen retter sine egne fejl. Sådanne fejl spiller selvsagt ikke samme rolle for ordgenkendelse ved læsning efter mening, som gentagelsesfejl, hvor en konsonant eller vokal fordobles og ordet måske som resultat heraf ikke genkendes eller omkodes forkert.

I et nyere studium af Khaldieh (1996) undersøges, om arabisk fonologi og skrift udgør et problem for ordgenkendelse hos amerikanske arabiskstuderende. For at gøre dette testede han 36 studerende på fire forskellige færdighedsniveauer for ordgenkendelsesfejl ved brug af grafiske og fonologiske 'distraktorer'. Fx lod han dem læse et testark med fire ord, hvor et af dem var عين [ʕain] (øje), fulgt af et kontrolark med de samme ord, men hvor عين var erstattet af ordene غين [rain] (navnet på et bogstav) som visuel distraktor og أين [aina] (hvor) som fonologisk distraktor. Testpersonerne skulle så identificere, hvilke ord de havde set på det første testark. På sætningsniveau skulle testpersonerne læse sætninger med samme form for distraktorer og blev derpå bedt om at afgøre, om sætningerne var semantisk korrekte eller ej. Som det ses i figur 1.4 fandt Khaldieh, at de studerende lavede færre fejl, jo højere færdighedsniveau de havde i arabisk generelt, men han kunne også konstatere, at selv studerende på højt niveau lavede langt flere fejl end kontrolgruppen af indfødte sprogbrugere i begge test, og han konkluderede, at L2-arabiskstuderende med latinsk skrevet L1 synes at tilegne sig det arabiske skriftsystem og arabisk fonologi forholdsvis langsomt. Dette studium er helt centralt for den grafiske dimension, der behandles i denne



Figur 1.4 Visuelle og fonologiske fejl fordelt efter niveau ved læsning af enkeltord og sætninger hos Khaldieh (1996)

afhandling, men hans testmetode er problematisk, fordi det i mange tilfælde ikke helt klart kan godtgøres, hvorvidt testpersonernes fejl er visuelle eller fonologiske, eller i nogle tilfælde simpelt hen skyldes manglende ordforråd.

I endnu et studium af Khaldieh (2001) undersøges, om ordforråd og kendskab til *iʕraab*[6] – dvs. systemet af syntaktisk bestemte kasusendelser – har betydning for læseforståelse. Undersøgelsen inddrog 46 engelske lørnere på to forskellige færdighedsniveauer, som blev bedt om at læse en tekst så mange gange de ønskede og umiddelbart derefter på modersmålet nedskrive, hvad de havde forstået. Derpå blev deres ordforråd og kendskab til *iʕraab* testet. Mens der ikke kunne påvises nogen sammenhæng mellem hverken læseforståelse og kendskab til *iʕraab* eller mellem kendskab til *iʕraab* og ordforråd, var respondenternes læseforståelse direkte proportional med deres testscore i ordforrådstesten.

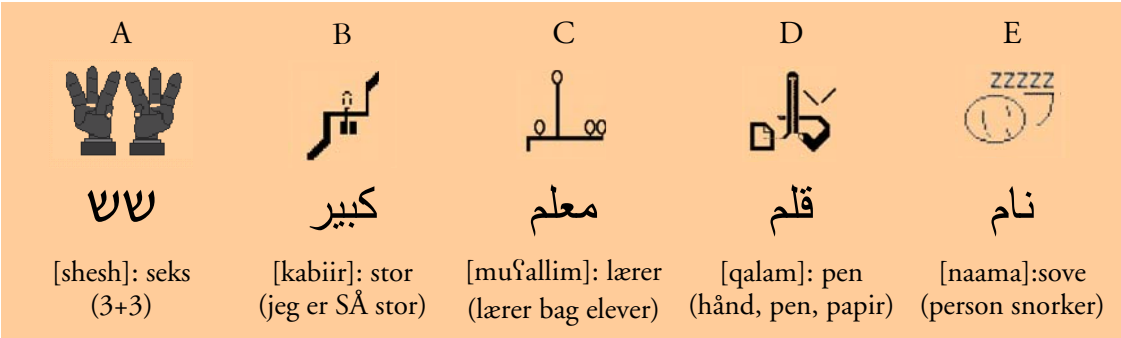
Pædagogik

Ud over de ovennævnte studier findes to publikationer, der først og fremmest er undervisningsmetodiske input til, hvordan tilegnelsen af det nye alfabet kan lettes, men som ikke er forskningsmæssigt underbyggede:

Magrath (1988) argumenterer for, at undervisning i ikke-latinske alfabeter bør baseres på en kommunikativ, '*whole word vocabulary in context*'-model, adopteret fra en metode anvendt til alfabetisering på engelsk som fremmedsprog. Argumentationen baseres udelukkende på engelsk L2-metodik og ikke på nogen konkret forskning relateret til de særlige karakteristika ved arabisk. Samtidig begrænser analysen af metodens brugbarhed i arabisk sammenhæng sig til nogle eksempler på, hvilke ord, der kunne udgøre gode *sight words*. Her nævnes som eksempler elementære ord som personlige pronominer og hilsner, som begynderen meget tidligt har i sit mundtlige ordforråd. Essensen er altså, at bogstaver ikke skal læres selvstændigt, men i en kontekst af meningsfulde enheder.

Bar-Lev (1991) har samme teoretiske udgangspunkt, men til forskel fra Magrath sigter han mod at udvikle en metode, der netop tager udgangspunkt i det arabiske skriftsystems særlige karakteristika. Metoden, som han kalder *glyphs*, er oprindeligt udviklet af Bar-Lev til undervisning i hebraisk skrift (Bar-Lev 1983,1989,1990), men er her tilpasset arabisk. En central pointe hos Bar-Lev er, at den studerende bruger ordbilledet som basis for genkendelse af det arabiske ord uden at kende ordets bogstaver, som først derefter læres enkeltvis, og ligesom Magrath arbejder Bar-Lev med *sight words*, som skal sikre, at bogstaverne læres i en meningsfuld kontekst, men Bar-Lev udstyrer desuden hvert *sight word* med en *glyph* – dvs. et billede, der visuelt minder om ordet og samtidig skaber associationer til dets mening (se figur 1.5). I figurens eksempel A, som er et eksempel fra hebraisk, ser ideen meget god ud, fordi de seks fingre så oplagt minder om både ordets visuelle udtryk og semantiske indhold. Eksempel B og C kan muligvis også give de korrekte associationer, men hovedparten af de ca. 50 *glyphs*, der er udviklet til arabisk, mangler forbindelse til enten det semantiske indhold, som i eksempel D, hvor det er vanskeligt at gennemskue, at det forestiller en hånd, der med en blyant skriver på papir, eller hvor den visuelle lighed med ordet er svær at se, som i eksempel E. Når den oplagte associationsværdi

[6] Kendskab til *iʕraab* kan naturligvis anvendes strategisk, men behandles i dette studium primært som en automatiseret kompetence. Måling af respondenternes eksplicitte viden om *iʕraab* indgår kun som kontrol.



Figur 1.5 Eksempler på *glyphs* til indlæring af hebraisk (A) og arabisk (B-E) hos Bar-Lev (1983, 1989, 1990, 1991)

som i eksempel D og E udebliver, synes det ikke umiddelbart overbevisende, at introduktionen af sådanne billeder skulle give de studerende bedre og/eller mere automatisk ordgenkendelse end almindelig eksponering for tekst. Og skulle det virkelig være tilfældet, synes det overordentlig problematisk, at eksempelvis billedet, der skal associere til ordet معلم (lærer), faktisk har større grafisk lighed med ordet قلم (pen). Bar Lev præsenterer ikke i sine publikationer om metoden (Bar-Lev 1983,1989,1990,1991) nogen empirisk dokumentation for dens brugbarhed, og det synes derfor ikke umiddelbart sandsynliggjort, at normale lørnere uden læsehandicap overhovedet har brug for denne form for støtte til at skabe forbindelse mellem visuelt udtryk og semantisk indhold. Tværtimod synes metoden blot at medføre et ekstra niveau af fremmed grafisk information, der skal tilegnes.

Det skal desuden bemærkes, at ud over at både Magraths og Bar-Levs publikationer mest forekommer at være baseret på løse påstande om deres respektive metoders brugbarhed i arabisk sammenhæng, har den teori, som metoder af denne art bygger på, undergået markante ændringer siden artiklerne blev skrevet.[7]

Sammenfatning

Mens de studier, der vedrører læsestrategier kun har begrænset berøringsflade med emnet for denne afhandling, er både studierne vedrørende kompetencer og pædagogik relevante i denne sammenhæng. Desværre berører de kun skriftsystemets grafiske dimension, og mens Hana (1972) og Khaldieh (1996) udelukkende behandler problematikken empirisk, behandler Magrath (1988) og Bar-Lev (1991) som nævnt kun det undervisningsmetodiske aspekt. Ingen af dem søger teoretisk at forklare de vanskeligheder, der er forbundet med begynderstuderendes tilegnelse af arabisk skrift, som de dog alle anerkender eksistensen af. Hvad angår den ortografiske dimension – dvs. det konsonantiske skriftsystem – findes der slet ingen studier, der behandler dens indflydelse på L2-læsetilegnelse på arabisk.

Denne afhandling behandler altså blandt andet den problematik som Hana (1972) og Khaldieh (1996) berører, men adskiller sig på flere punkter: For det første søges problematikken belyst teoretisk. For det andet er udgangspunktet for fejldiagnosticeringen i afhandlingens empiriske del, at målet for læseundervisningen er kommunikativt, og den kvalitative analyse af læsefejlene omfatter en langt større population end Hanas studium. For det tredje er studiet metodisk tilrettelagt, så de svagheder, der som nævnt præger

[7] Dette teori- og metodeskift beskrives nærmere i kapitel 2 og diskuteres i pædagogisk sammenhæng i kapitel 7.

Khaldiehs studium, elimineres. Desuden inddrages den ortografiske dimension, som ikke tidligere er behandlet, og hele empirien relateres til en forholdsvis nyudviklet teoretisk ramme for ordgenkendelsesprocesser, som tilsyneladende endnu er ladt ubemærket i L2-læseforskning, og slet ikke tidligere er sat i relation til læsning på semitiske sprog. Endelig diskuteres pædagogisk praksis i forhold til både teori og empiri, hvilket bringer dette studium langt videre end både Magrath og Bar-Lev, der som nævnt ikke underbygger deres metoder empirisk og desuden funderer deres metodik på en læseteoretisk ramme, som den moderne læseforskning efterhånden har forkastet.

1.3 Fremgangsmåde

Som det fremgår af ovenstående, er forskningen i læsefærdighed på arabisk som L2 særdeles sparsom. Endvidere findes der ikke nogen egentlig teoretisk ramme for afhandlingens specifikke emne, idet den generelle og fremmedsprogspædagogiske læseteori udelukkende omhandler læsning på indoeuropæiske sprog skrevet med det latinske alfabet, mens teoretiske studier, der anerkender læseprocesmæssige forskelle ved læsning af forskellige skriftsystemer, udelukkende behandler læsning på L1. Ud over den overordnede læseteoretiske ramme består de teoretiske dele af denne afhandling derfor først og fremmes af en præsentation af en række forskningsområder, der fra forskellige vinkler kan belyse emnet. Projektets empiriske del omfatter tre deltests, hvor effekten af henholdsvis bogstavarkitektur, vokalisering og morfologisk struktur testes. Effekten af læseretning, som jo også vedrører den grafiske dimension, testes ikke, idet den er vanskelig at undersøge under de testforhold, der i øvrigt anvendes. Afhandlingen er således struktureret som følger:

I **kapitel 2** præsenteres det basale læseteoretiske grundlag. Kapitlet har til formål at forklare begreber og teorier, der ligger til grund for analysen i kapitel 4 og kapitel 5 om læsning på semitiske sprog, og som danner grundlag for empirien i kapitel 6. Først gives en overordnet gennemgang af de centrale læseteorier i historisk perspektiv. Den historiske dimension skal tjene til at begrunde valget af teoretisk ramme for projektet. Efter den overordnede introduktion behandles mere specifikt de mest anerkendte ordgenkendelsesteorier, mens læseforståelse kun inddrages på ordgenkendelsesniveau og i relation til læsning på L2. Til ordgenkendelsesniveauet regnes kontekseffekt, som er et centralt omdrejningspunkt for den forskning i læsning på semitiske sprog, der præsenteres i kapitel 5. Derpå gives en redegørelse for forskning vedrørende læsning på tværs af ortografier, og endelig relateres teorierne til læsning på L2, herunder L2, der skrives med et andet skriftsystem end L1. Endelig sammenfattes elementerne til en læseprocesmodel, der er tilpasset de forskellige stadier i ordgenkendelsesprocessen, som er centrale for afhandlingens to fokusområder.

I **kapitel 3** illustreres de særlige karakteristika ved arabisk. I kapitlet præsenteres det arabiske skriftsystem, de arabiske fonemer og det arabiske sprogs morfologiske struktur. Denne præsentation er tiltænkt læsere, der ikke i forvejen har kendskab til arabisk. Endvidere inddrages hebraisk, der som semitisk sprog har mange fællestræk med arabisk og har været genstand for langt mere forskning.[8] Det hebraiske skriftsystem præsenteres, og sammenligningsgrundlaget mellem de to semitiske skriftsystemer analyseres med henblik på at afgrænse, i hvor høj grad forskningsresultater fra hebraisk er repræsentative for arabisk og

[8] Grunden hertil er sandsynligvis, at der i vestlige forskningsmiljøer findes flere personer med kendskab til hebraisk (som er andetsprog for mange jøder) end forskere, der kan arabisk. Hertil kommer, at læseforskningen er langt mere etableret i Israel end i den arabiske verden.

på den baggrund kan inddrages senere i afhandlingen. Endelig præsenteres en række forskningsresultater, der belyser gyldigheden af den traditionelle betragtning af de semitiske sprogs morfologiske struktur som særegen i kraft af opdelingen af ord i rødder og mønstre, der opfattes som separate morfemer. Dette spørgsmål vedrører ikke blot den faktuelle lingvistiske beskrivelse af arabisk og hebraisk, men vil senere i afhandlingen vise sig relevant for den teoretiske beskrivelse af læseprocesser på semitiske sprog.

De følgende to kapitler omhandler hver én af de førnævnte fremmede dimensioner ved det arabiske skriftsystem:

I **kapitel** 4 behandles læseretningen og den fremmede bogstavarkitektur samt dennes indflydelse på visuel ordgenkendelse ved læsning af trykt skrift. Håndskrift og ordgenkendelse af håndskrevet tekstmateriale er en relevant dimension, men da den for det første gør problemstillingen langt mere kompleks og for det andet er mindre relevant i forhold til fremmedsprogslæsning, der især retter sig mod trykte tekster, lades den ude af betragtning. Med henblik på at dokumentere, at fremmed bogstavarkitektur komplicerer bogstavgenkendelse og på det teoretiske plan udgør en afgørende barriere for danske arabiskstuderendes læsetilegnelse på arabisk, præsenteres relevante semiotiske og perceptionspsykologiske modeller og bogstavgenkendelsesteorier for det latinske alfabet sammenlignes med det arabiske. Kapitlet er et teoretisk oplæg til test A i kapitel 6.

Kapitel 5 omhandler det konsonantiske skriftsystems betydning for læsetilegnelsen på arabisk hos danske arabiskstuderende. I mangel på direkte relateret forskning præsenteres eksisterende forskningsresultater vedrørende læsning på arabisk og hebraisk som L1, og disse forsøges anvendt til at formulere nogle retningslinier for en selvstændig læseprocesteori for semitiske sprog. Kapitlet er et teoretisk oplæg til test B og C i kapitel 6.

I **Kapitel 6** præsenteres projektets empiriske del, som omfatter tre deltest: Test A vedrører den grafiske dimension. Effekten af den fremmede bogstavarkitektur måles vha. læsetest, hvor den sproglige dimension udelades ved brug af pseudoord, og hvor der kontroltestes for effekt af fremmede fonemer. Der testes på arabiskstuderende på tre forskellige færdighedsniveauer med en kontrolgruppe af indfødte sprogbrugere. Testresultaterne analyseres både kvantitativt og kvalitativt, således at både en overordnet sammenligning af progressionen mellem niveauerne og en nærmere analyse af typiske læseproblemer afdækkes. Test B og C vedrører den ortografiske dimension. I test B undersøges effekten af vokalisering på læsehastighed og –forståelse, og i test C undersøges i hvor høj grad det arabiske sprogs morfologiske struktur kan bruges som kompensationsstrategi til afkodning af uvokaliseret tekst. Der testes på de samme færdighedsniveauer og med samme kontrolgruppe som i test A.

I **kapitel** 7 opsamles de teoretiske og empiriske elementer i en samlet konklusion, som sættes i relation til de undervisningsmetoder, der anvendes i dag. På baggrund af forskningsprojektets resultater fremsættes nogle bud på, hvordan en specialiseret fremmedsprogspædagogisk indsats i læseundervisningen på arabisk for arabiskstuderende kan tilpasses de særlige udfordringer, det anderledes skriftsystem giver voksne lørnere, der ønsker at tilegne sig læsefærdigheder på arabisk.

Gennem hele afhandlingen translittereres efter standarden i det internationale fonetiske alfabet.

2. Teoretisk udgangspunkt

Med henblik på at præcisere den teoretiske ramme for afhandlingen og for at definere en række begreber, der er centrale for forskningsprojektet som helhed, gives i dette kapitel en kort præsentation af de mest relevante teorier om læsning. Først gennemgås de kognitive læseteorier i historisk perspektiv, hvilket skal tjene til at begrunde valget af den teori, der bygges videre på. Derpå behandles det moderne teorigrundlag for delprocessen ordgenkendelse, mens læseforståelsesaspektet behandles efterfølgende i relation til læsning på fremmedsprog. Denne vægtning afspejler, at ordgenkendelse jvf. afsnit 1.1 er afhandlingens centrale teoretiske omdrejningspunkt. Forståelse er selvsagt relevant, men indgår som et aspekt i alle områder af fremmedsprogstilegnelsen, og behandles derfor som en generel, implicit faktor, og i empirien indgår læseforståelse som en indikator for ordgenkendelse. Efter denne beskrivelse af den generelle læseteori gennemgås teorier, der specifikt behandler læsning på tværs af ortografier.

Disse afsnit, der udgør størstedelen af kapitlet, beskriver udelukkende teorier og forskningsresultater, der vedrører kompetente læseres[1] læseprocesser på deres L1. Baggrunden herfor er for det første, at den eksisterende forskning i læsning på L2 er meget beskeden, hvorfor den ikke selvstændigt kan gøre det ud for et brugbart teoretisk grundlag. For det andet er de resultater der foreligger, vanskelige at teoretisere udfra, fordi de samme forhold ikke nødvendigvis gør sig gældende for alle sprog – og antallet af mulige kombinationer af L1 og L2 er selvsagt betragteligt. Samtidig kan L2-læsere have vidt forskellige profiler, hvad angår fx alder, forudsætninger, behov og sprogfærdigheder på det aktuelle L2. Alle disse forhold gør det særdeles vanskeligt at generalisere ud fra enkeltstående forskningsresultater. I modsætning hertil er kompetent L1-læsning et bredt dyrket forskningsfelt, der – uanset at indlæringsprocessen og tilegnelseshastigheden givetvis er forskellig for læsefærdigheder på L1 og L2 – kan afdække hvordan kompetencen er etableret, og hvordan en hensigtsmæssig læseudvikling bedst kan støttes, ikke mindst fordi målet for L2-læsning i princippet er en læsekompetence, der er sammenlignelig med kompetent L1-læsning. Så selvom mange studerende aldrig bliver egentlig kompetente L2-læsere, og selvom L2-læseundervisning også bør afspejle dette forhold, må indsigt i kompetent L1-læsning på nuværende tidspunkt betragtes som det bedste tilgængelige teoretiske udgangspunkt for udvikling af undervisningsstrategier, der kan hjælpe L2-læsere i den rigtige retning.

Efter denne gennemgang præsenteres det beskedne forskningsgrundlag for læsning på fremmedsprog, og endelig behandles problematikken omkring læsning med et nyt alfabet. Hvad angår disse to afsnit skal det præciseres, at L2-læsning betragtes som en proces, hvor der umiddelbart ekstraheres mening fra tekst uden analytiske mellemled som fx morfologiske og syntaktiske overvejelser eller ordbogsopslag, jvf. afsnit 1.1.

---

[1] Begrebet 'kompetente læsere' kan sidestilles med det i L1-læseteorien meget anvendte begreb 'funktionelle læsere'. For begge begreber gælder det, at det ikke er muligt at definere dem objektivt, idet måling af, i hvor høj grad læsning er kompetent eller funktionelt må baseres på en normativ vurdering.

2.1 Kognitive læseteorier i historisk perspektiv

De moderne læseteorier er udviklet i et samspil mellem de psykologiske, neurologiske og lingvistiske fagfelter. Psykologerne beskæftigede sig allerede fra midten af det 19. århundrede med de visuelle processer – primært øjenbevægelser – der finder sted ved læsning (Rayner & Pollatsek 1989:5-6), og omkring begyndelsen af det 20. århundrede blev denne forskning anset for at være helt central for den eksperimentelle psykologi, da læsning mentes at afspejle generelle psykologiske processer. Psykologer satte således dagsordenen for de tidligste teorier om læsning, men mellem 1913 og ca. 1960 – hvor psykologien i USA, hvor læseteorierne især udfoldede sig, var domineret af behaviorismen – blev forskningsfeltet mere eller mindre overladt til lingvisterne. De fik dermed stor indflydelse på de egentlige læseprocesteorier, der blev udviklet i perioden. Lingvisterne satte teksten i centrum ved at introducere fonembegrebet med henblik på at anskueliggøre betydningen af, at børns forståelse af forbindelsen mellem fonem og grafem er essentiel for god læseindlæring (Kulbrandstad 1996:51). I forbindelse med kognitionsforskningens opståen i kølvandet på den store omvæltning inden for psykologien fra behaviorisme til rationalisme og i takt med udviklingen af specielle undersøgelsesteknikker, begyndte neurologer fra 1950'erne at dokumentere, hvilke dele af hjernen, der operationaliseres ved læsning (Dalby et al. 1983:204), især gennem studier af patienter, der udviste dyslexi som følge af hjerneskade. Samtidig var et værk, der i væsentlig grad bidrog til behaviorismens fald – Chomskys *Syntactic Structures* (1957) – netop en lingvistisk funderet eksplicitering af rationalismen, der fremhævede sammenhængen mellem sprog og tænkning (Ravn 1992; Schøsler 1998). I takt med at de kognitive (herunder sproglige) processer således igen blev et kerneområde i psykologien, blev læsning – set som en kognitiv og sproglig aktivitet frem for en isoleret visuel perceptuel proces – igen et tværfagligt interesseområde (Tzeng 1981:4; Rayner & Pollatsek 1989:6-7).

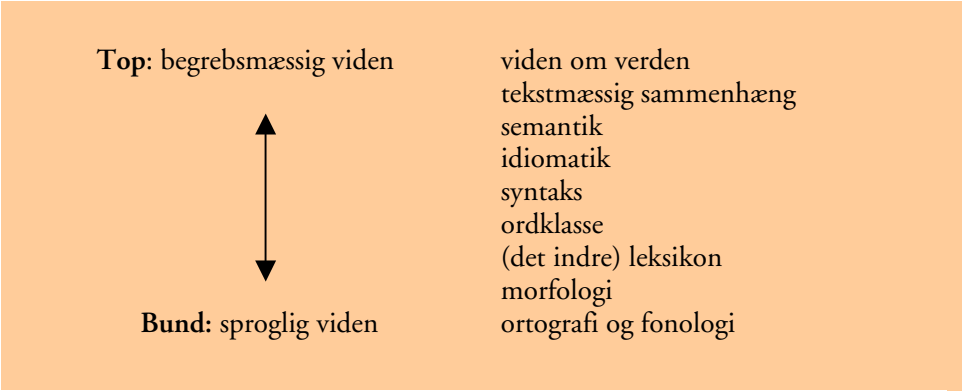
I dag er der blandt psykologer, neurologer og lingvister enighed om, at læseprocessen er en interaktiv proces mellem en række forskellige kognitive færdigheder. Det kognitive perspektiv indebærer for læseforskningen ifølge Bernhardt følgende:

> *"taking a cognitive perspective means examining the reading process as an intrapersonal problem solving task that takes place within the brain's knowledge structures".*(gengivet i Urquhart & Weir 1998:8)

Man kunne også sige, at målet for den kognitive læseforskning er at beskrive hvordan input, dvs. de grafiske tegn, omsættes til output, dvs. mening eller 'ny viden' i læserens bevidsthed. Udgangspunktet for denne tilgang er, at mennesket har en begrænset 'arbejdshukommelse': Et blik på en side fuld af tekst giver ikke umiddelbart mening. Den må **læses** først. Essensen i læseprocessen er således, at læseren opdeler teksten i mindre, overskuelige enheder for at skabe mening. Den kognitive læseforskning har derfor primært koncentreret sig om to spørgsmål:

- Hvilke kognitive processer aktiverer vi, når vi omsætter grafiske tegn til mening?
- Og hvor små enheder opdeler vi tekst i for at kunne overkomme denne opgave?

Ifølge de tidlige læseprocesmodeller fra begyndelsen af 1960'erne opløser læseren teksten i dens mindste enheder, nemlig bogstaver, for at afkode den. Læsningen foregår som en lineær, fonologisk afkodningsproces, begyndende med det første ords første bogstav. Når det første ord er afkodet, og ordgenkendelse har fundet sted, går man videre til næste ords



Figur 2.1 Hierarki over nøgler til læseforståelse (Haastrup 1989:11)

første bogstav. Processen fortsætter på samme måde til alle ordene i sætningen er blevet behandlet. Først derefter appliceres syntaktiske og semantiske regler for at give sætningen mening (se fx Gough 1972).

Denne og lignende modeller kaldes *bottom-up* modeller med henvisning til figur 2.1, som viser et hierarki over nøgler til læseforståelse: Læsningen begynder i hierarkiets bund med tekstens mindste enheder, og først når ét niveau er 'på plads' bevæger man sig et trin op i hierarkiet for at trække på stadig bredere kategorier (Gough 1972; Urquhart & Weir 1998:40). Denne lineære forståelse af læseprocessen synes at hænge tæt sammen med kognitionsforskningens teknologisk inspirerede rationalisme. Den nyopfundne computer blev set som et teknologisk eksempel på, hvordan hjernen modtager input, behandler det, og omdanner det til intelligent output (Ravn 1992). Kognitive processer blev altså forstået som menneskelig databehandling, således at læseprocessen forløber som et computer-program – altså en trinvis opbygget forskrift for hvad der skal ske, for at et givet input bliver transformeret til korrekt output.

Bottom-up modellerne blev imidlertid fra midten af 1960'erne mødt af kritik fra flere fronter. For det første blev de såkaldte *top-down* modeller introduceret. Her er læserens forventninger til teksten afgørende. Læseren danner hypoteser om tekstens indhold og bruger teksten til at bekræfte eller forkaste disse hypoteser. For Goodman (1968; 1976) relaterer disse hypotesetest sig i høj grad til betydningen af de enkelte ord. Hos andre relateres de til hele teksten med reference til *schema*: læserens forventninger til teksten, dannet på baggrund af fx viden om emnet, den kulturelle kontekst og den information om teksten, som læseren har fået før læsningen påbegyndes (Rayner & Pollatsek 1989:305-311). Såvel *top-down* som *bottom-up* teoretikerne kunne fremvise empiriske data til støtte for, at hhv. de lavere og de højere niveauer i hierarkiet var vigtige elementer i læseprocessen. Efterhånden forsvandt forsvarene for de rendyrkede versioner af de to teorier – dvs. læsning forstået som en stort set autonom, input-domineret proces hhv. en kognitivt domineret 'gætteleg', hvor læserens baggrundsviden sætter rammerne for, hvilke potentielle ord, der overhovedet er mulighed for at behandle som input – og empiriske data fra begge lejre blev brugt til at understøtte en interaktiv model, som siden ca. 1980 har været bredt anerkendt. Den interaktive model er stort set en syntese af *bottom-up* og *top-down* modellerne, og dens hovedpointe er, at de forskellige former for tekstrelevant viden aktiveres samtidig (Urquhart & Weir 1998:45). Ved læsning operationaliseres input således under indflydelse af såvel læserens sproglige som begrebsmæssige viden, dvs. hele spektret af baggrundsviden fra

ortografi over morfologi, syntaks og semantik til generel viden om verden (jvf. figur 2.1). Stanovich (1980) har desuden beskrevet læseprocessen som 'interaktiv-kompensatorisk', hvor det kompensatoriske går på, at en læser kan kompensere for svagheder i ét vidensfelt med styrker i et andet. Har læseren fx dårlig ortografisk viden, kan han kompensere herfor med stærk syntaktisk viden. Ifølge Stanovich læser den ideelle læser interaktivt uden behov for at kompensere (Stanovich 1980:47-48).

Desuden begyndte en række psykologer at påpege, at de fonologiske aspekter, som lingvisterne havde introduceret, måske var blevet for dominerende. Undersøgelser viste nemlig, at ord genkendes hurtigere end enkeltbogstaver, og at selv pseudoord kan genkendes med samme hastighed som bogstaver, hvilket tyder på at bogstaver i ordgenkendelsesfasen behandles parallelt frem for lineært (Urquhart & Weir 1998:41). Psykologerne mente, at dette burde være anledning til, at man igen begyndte at vise læsningens visuelle processer mere opmærksomhed (Henderson 1984:1-9; Berninger 1994:1; Foorman 1994:321).

På den ene side blev der således sat fokus på betydningen af de højere niveauer i 'læsehierarkiet' (se figur 2.1), og læseres individuelle verbale, lingvistiske og metalingvistiske kompetencers betydning for læseprocessen blev præciseret og empirisk dokumenteret (Vellutino & Scanlon 1986; Rayner & Pollatsek 1989). På den anden side kom der stor opmærksomhed på de generelle kognitive processer, der på de lavere niveauer i læsehierarkiet forbinder det visuelle input med mening. Mens praktisk talt alle læseteoretikere i dag støtter en interaktiv læsemodel, er læsehierarkiet nærmest blevet afløst af en opfattelse af, at læseprocessen består af afkodning og forståelse som to adskillelige delkomponenter (Elbro 2001:28-30), og forskningsområdet er karakteriseret ved at være skarpt opdelt i dem, der beskæftiger sig med ordgenkendelsesprocesser, og dem, der beskæftiger sig med sprogforståelse.

I det følgende præsenteres i hovedtræk centrale teorier om ordgenkendelse med hovedvægt på de elementer, der er relevante for afhandlingens empiriske del.

2.2 Ordgenkendelse

De tidligste studier beskæftigede sig med øjenbevægelser under læsning. Man fandt, at modsat det umiddelbare indtryk, man som læser får af, at øjnene 'glider' hen ad siden, laver øjnene serier af saccader (dvs. spring), afbrudt af fikseringer af ca. et kvart sekunds varighed (Rayner & Pollatsek 1989:5-6). Moderne undersøgelser, der kan udføres med langt større præcision, har vist, at man fikserer på næsten hvert eneste ord, og at kun meget korte, forudsigelige ord (som konjunktioner, præpositioner og artikler (Adams 1990:100)) springes over. Man fikserer i længere tid på lange og lavfrekvente[2] ord og på ord i en sætnings slutning. Ca. hver tiende saccade er en regression, dvs. et spring tilbage i teksten. Regressioner forekommer hyppigst hos begynderlæsere og svage læsere. Fikseringstiden er – alt andet lige – udtryk for, hvor meget kognitiv kapacitet læseren må bruge på at behandle materialet (Rayner 1997; Elbro 2001:22-24). Ved en fiksering opfatter man kun en meget begrænset del af teksten. Man har et centralt opmærksomhedsfelt, der ligger asymmetrisk omkring fikseringspunktet, sådan at man opfatter mest materiale i læseretningen. Ved læsning af det latinske alfabet opfatter man klart 6-8 bogstaver til højre og 2-4 bogstaver til

[2] Frekvens er her et udtryk for, hvor hyppigt ordet forekommer i de almindeligste tekstgenrer.

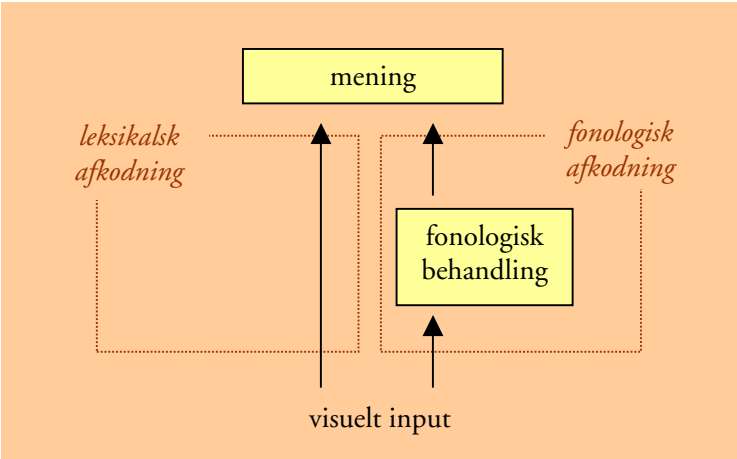
venstre for fikseringspunktet. Herudover har man et perifert opmærksomhedsfelt, hvor man fx kan opfatte det følgende ords længde eller liniens slutning (Elbro 2001:25).

Dualruteteorien

Studier i øjenbevægelser har altså dokumenteret, at man fokuserer på næsten hvert enkelt ord, når man læser. Hvad angår selve behandlingen af de enkelte ord, har hovedparten af læseforskningen i de seneste to årtier faktisk drejet sig om netop at afklare, hvordan afkodning af ord foregår (Geva & Willows 1994:359). Diskussionen har især handlet om, i hvor høj grad fonologiske hhv. leksikalske 'veje' til leksikon operationaliseres.

Udgangspunktet for disse studier, der metodemæssigt er funderet i psykologien, var til at begynde med lingvistiske beskrivelser af korrespondancen mellem bogstav og lyd i engelske ord, som opdeltes i kategorierne 'regelmæssige ord' og 'uregelmæssige ord', hvor kun ord, hvis udtale følger den oftest forekommende fonem-grafem korrespondance, anses for regelmæssige (fx Wijk 1966; Venezky 1970). Psykologerne mente gennem undersøgelser af læsehastighed af hhv. regelmæssige og uregelmæssige ord at kunne dokumentere, at de blev behandlet læseprocesmæssigt forskelligt (Henderson 1984:3). Det passede dårligt med den gængse opfattelse af afkodningsprocessen, ifølge hvilken den regelbaserede omkodning fra grafem til fonem kunne benyttes på alle slags input. Man forklarede uoverensstemmelsen med, at der må findes en alternativ, visuelt baseret vej til leksikon, som læsere kan benytte ved læsning af uregelmæssige ord. Ifølge denne dualruteteori *(dual route theory)* er der altså to veje til ordgenkendelse: Den fonologiske vej, der går fra visuelt input over lyd til mening (også kaldet fonolektisk adgang) og den leksikalske vej (eller den ortografiske vej), der går direkte fra visuelt input til mening (også kaldet logolektisk adgang). Mens regelmæssige ord kan udtales på baggrund af begge procedurer, må uregelmæssige ord være et produkt af den leksikalske vej, mens læsning af pseudoord må være et produkt af den regelbaserede, fonolektiske afkodning, siden de ikke er genkendelige, leksikalske enheder (Henderson 1982:111-169; 1984:3).

Dualruteteorien støttes af eksperimentel forskning, der viser, at uregelmæssige ord – angiveligt fordi de afkodes direkte, uden involvering af det fonetiske niveau – oplæses hurtigere end pseudoord, og at regelmæssige ord – angiveligt fordi de kan afkodes ad begge



Figur 2.2 Ordgenkendelsesprocessen ifølge dualruteteorien

veje samtidig – oplæses hurtigere end uregelmæssige ord, men at forskellen er lille for højfrekvensord, mens den er stor for lavfrekvensord – angiveligt fordi ord afkodes direkte, når de er set tilstrækkeligt mange gange (Rayner & Pollatsek 1989:87-88). Desuden støttes teorien af neurologisk forskning: Man har konstateret tilfælde af fonologisk dyslexi hos hjerneskadede, hvor patienten kan læse, men har mistet evnen til at læse pseudoord, dvs. kun den fonologiske vej er ødelagt. Tilsvarende har man konstateret tilfælde af overfladedyslexi, hvor patienten kan læse både regelmæssige ord og pseudoord, men fejllæser uregelmæssige ord og homofoner, dvs. hvor kun den leksikalske vej til leksikon er ødelagt (Rayner & Pollatsek 1989:393-435; Coltheart & Coltheart 1997; Miceli et al. 1997; Rapp et al. 1997).

Konnektionisme

De seneste år er dualruteteorien blevet udfordret fra flere sider. Konkret fremsættes kritikken først og fremmest som et opgør med grundlaget for den empiriske forskning, der støtter teorien. De mest anvendte former for tests er baseret på læsning af pseudoord (som skulle garantere fonologisk afkodning, da de ikke findes i leksikon) og ordbestemmelses-opgaver,[3] eksempelvis som valg mellem to homofoner – fx: hvilket er et ord: 'rejn' eller 'regn' (hvilket skulle garantere leksikalsk afkodning, da de fonetisk er ens, og afgørelsen af, hvilket af dem, der er et gyldigt ord, må være baseret på leksikalsk afkodning af ordet). Imidlertid er der blevet rejst tvivl om, hvorvidt de fonologiske og leksikalske processer i virkeligheden kan isoleres fra hinanden og fra de øvrige kognitive processer, der aktiveres ved læsning (fx Seidenberg 1992; Berninger 1994; Geva & Willows 1994; Vellutino et al. 1995). Samtidig er der i nogen grad blevet sat spørgsmålstegn ved den neurologiske dokumentation, idet isoleret overfladedyslexi eller fonologisk dyslexi normalt kun forekommer ved erhvervet dyslexi (hjerneskade), mens udviklingsdyslektikere ('almindelige' ordblinde) har vanskeligt ved at afkode både fonologisk og leksikalsk. Dette underminerer ifølge kritikerne dualruteteoriens troværdighed (fx Foorman 1995:397). På den baggrund udfordres dualruteteorien i dens traditionelle form nu af en mere fleksibel ordgenkendelsesmodel, hvor de fonologiske og de ortografiske processer integreres (Seidenberg 1992; Foorman 1994,1995; Vellutino et al. 1995), ligesom andre former for sproglige kompetencer inddrages (Norris 1990; Sharkey 1990).

Set i en større teoretisk sammenhæng hænger kritikken sammen med en generel udvikling i moderne kognitionsforskning, hvor nativismen,[4] inden for lingvistikken repræsenteret af den generative lingvistik[5], siden midten af 1980'erne er blevet udfordret af en ny retning, konnektionismen (*connectionism* eller *parallel distributed processing*)(Ravn 1992; Ramsey 1999; Smolensky 1999). Ligesom rationalismen, herunder den generative lingvistik, havde sammenhæng med udviklingen af computere til databehandling, er konnektionismen tæt forbundet med forskning i kunstig intelligens. Konnektionismen tilstræber ligesom den generative lingvistik at internalisere sproget (i modsætning til den traditionelle strukturalismes eksternalisering af og fokus på sproget som system, opfattes sproget som udtryk for, hvordan kognitive processer generelt fungerer). Men mens Chomsky opfattede

---

[3] Opgaver, hvor forsøgspersoner skal tilkendegive, om bogstavstrenge, de bliver præsenteret for, er gyldige ord eller ej.

[4] Psykologisk/filosofisk term for den rationalistisk funderede betragtning, at mennesket bærer en naturlig, medfødt viden – i modsætning til empiricismen, der betragter al viden som erfaringsbaseret (Ramsey 1999).

[5] Dvs. Chomskys transformationsgrammatik og deraf afledte teorier.

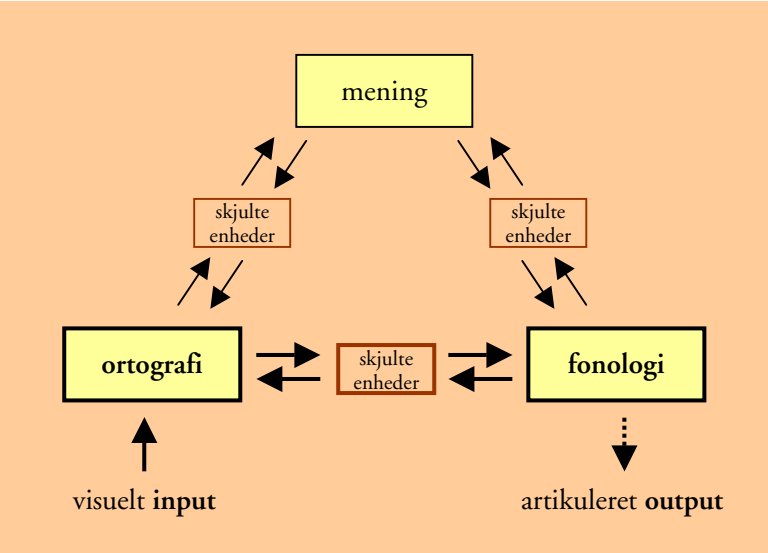
sprogkompetence som et medfødt sæt af regler i form af syntaktiske strukturer placeret i et selvstændigt modul i hjernen, betragter konnektionisterne sproget som en erfaringsbaseret kompetence, der fungerer som en del af den almene kognitionsevne (Ravn 1992; McClelland 1999; Ramsey 1999; Smolensky 1999; Garson 2002).[6] Denne forskel er på mange måde et spejlbillede af forskellen mellem den klassiske computer og de nye såkaldte neurocomputere, der forsøger at imitere det biologiske neurale netværk. Mens den klassiske computer fungerer ved hjælp af et program, der i praksis er en liste af instrukser for, hvordan 'videnselementer' i en nøje fastlagt rækkefølge kombineres for at frembringe det rette output, er neurocomputeren et netværk af enheder (jvf. hjernens synapser), typisk grupperet i input-enheder og outputenheder og derimellem et eller flere lag af 'skjulte' enheder. Enhederne er forbundet i et netværk af konnektioner, og informations-behandlingen sker som en spredning af information fra input-enhederne rundt i netværket, hvor mønstret af informationsstrømme afhænger af de enkelte konnektioners vægte. Informationen (i praksis elektrisk strøm) afvejes således mod den information, der i form af vægtenes forskellige størrelser allerede befinder sig i netværket. Netværket 'optrænes' ved at hver enkelt vægt justeres, når output er forkert. Herved erhverver netværket 'erfaring', som det kan bruge til at generalisere på input, det aldrig har mødt under sin 'optræning' (Ravn 1992; McClelland 1999; Garson 2002).

Konnektionismen er blevet anvendt som et teoretisk grundlag for en ny ordgenkendelses-model, der altså afviser den duale ordbehandlingsmekanisme til fordel for en enkelt. Desuden afvises forestillingen om et 'mentalt leksikon'. Der er således ikke noget sted at 'slå ordene op', og der findes dermed ingen leksikalsk vej til ordgenkendelse. I stedet behandler og sammenkobler man ortografiske, fonetiske og semantiske koder. Seidenberg, som er en af hovedkræfterne inden for konnektionisme i relation til ordgenkendelse, beskriver forskellen således:

> *"According to this theory, codes are not accessed, they are computed; semantic activation accrues over time, and there can be partial activation from both orthographic and phonological sources. So, for example, whereas in the standard dual-route model, 'phonological mediation' required deriving the complete phonological code for a word and using it to search lexical memory, in the present framework there can be partial activation of phonology from orthography, or of meaning from phonology. Thus, the meaning of a word is built up by means of activation from both routes, […] rather than accessed by means of whatever route wins the race."* (Seidenberg 1992:105)

Der er altså fortsat tale om behandling af fonologisk og ortografisk materiale, men frem for at løbe om kap ad to separate veje til listen af kendte ord, samles det i en suppegryde, hvor det sammen med andre former for tekstrelevante erfaringer koger sammen til mening – bogstavkombinationer giver hints om kendte fonologiske mønstre, fonologiske sammenstillinger giver semantiske associationer, etc. At blive en god læser er således et

[6] Nogle fremhæver dog konnektionismen som en bro mellem nativismen (jvf note 4) og empiricismen, idet teorien på samme tid accepterer, at vores perception afhænger af et medfødt sæt af simple sensorer og at viden er erfaringsbaseret (fx Adams 1990:201-204).



Figur 2.3 Konnektionistisk model for genkendelse af isolerede ord (efter Seidenberg 1989:526)

spørgsmål om, at man, på baggrund af gentagelsesfrekvens og konsistens i sammenhængen mellem ortografiske og fonetiske enheder, efterhånden finjusterer vægtningen af konnektionerne mellem de forskellige enheder. Grafem-fonem korrespondance er altså stadig essentiel – ikke som isolerede regler, men som karakteristiske stavningsmønstre, der efterhånden genkendes, når man har mødt dem mange gange.

Dette forklarer både ordfrekvenseffekt – at man genkender højfrekvente ord hurtigere end lavfrekvente ord (Monsell et al. 1989), regelmæssighedseffekt – at ord med regelmæssig stavning og udtale genkendes hurtigere end uregelmæssige ord (Metsala et al. 1998) – og samspillet mellem de to fænomener, også kaldet 'nabo-frekvenseffekt' (Grainger 1992) – at ordfrekvens for 'naboer', dvs. ord, der ligner målordet i kraft af at de har fælles bogstavsammenstillinger, influerer på genkendelseshastigheden: Hvis ordet og dets naboer er regelmæssige (fx 'gav', 'sav' og 'lav') er nabofrekvenseffekten positiv, og et højfrekvent ord øger hastigheden for ordgenkendelse af mindre frekvente naboer. Men for en uregelmæssig nabo (fx 'hav') kan effekten være negativ. Dette er især tilfældet for lavfrekvente ord, hvor påvirkningen fra de højfrekvente naboer gør genkendelse langsommere, mens frekvenseffekten hos øvede læsere udvisker den negative effekt på højfrekvente ord (Massaro et al. 1979; Seidenberg & McClelland 1989; Grainger 1992; Johnson 1992). I træningsforsøg med computere, hvor man har simuleret den optrukne del af ordgenkendelsesmodellen i figur 2.3, har man faktisk registreret bl.a. frekvens-, regelmæssigheds- og nabofrekvenseffekter, der er meget lig de resultater, man har opnået med levende forsøgspersonerpersoner (Seidenberg & McClelland 1989).[7]

Et andet interessant og overbevisende aspekt ved teorien er, at den forklarer læseprocessens kompleksitet og betragter den endnu bredere end de traditionelle interaktive modeller: På

[7] Bemærk desuden, at den konnektionistiske ordgenkendelsesmodel i figur 2.3 meget enkelt redegør for højtlæsning som et fænomen, der kan involvere alle elementer af ordgenkendelsesprocessen, men ikke nødvendigvis involverer mening (til forskel fra tidligere læsemodeller, hvor dette aspekt ignoreres eller illustreres mindre overbevisende, se fx Rayner & Pollatsek 1989:92&461-473).

læsehierarkiets øvre plan favner den en bred fortolkning af *schema*-teorien, idet alle former for viden påvirker processen, mens den på det nedre plan inddrager de mindste del-komponenter. Så selvom det kan dokumenteres, at øvede læsere har hurtig ordgenkendelse, der ikke synes at involvere en fonologisk omkodning, og at de kan udnytte ords holistiske form i ordgenkendelsesprocessen, betyder det ikke, at ordene genkendes som helheder. Og selvom kontekst påvirker afkodning, betyder det ikke, at læsning består af løbende, kontekstbaserede hypotesetests. Tværtimod står det klart, at øvede læsere behandler nærmest hvert eneste bogstav visuelt – men ikke isoleret:

> *"Even while the individual letters of the text are the basic perceptual data of reading, they are not perceived one by one, independently of each other. Instead, their efficient and productive perception depends additionally on ready knowledge of words – their spellings, meanings, and pronunciations – and on consideration of the context in which they occur. In the mind of the skillful reader, each such type of knowledge is represented by constellations of elementary units, connected in specific, learned relation to each other: Simple patterns are represented by interrelated clusters of units, more complex patterns of clusters of clusters of units, and so on such that the whole of any percept or idea is defined, at core, by the particular relations that hold among its parts."* (Adams 1990:14-15)

Mening opstår således gennem kobling af segmenter på mange niveauer: Bogstaver er repræsentationer af indbyrdes forbundne visuelle delkomponenter, ligesom et ord er en repræsentation af indbyrdes forbundne bogstavkombinationer. På samme måde korresponderer ordets udtale med et kompleks af fonemer, ligesom dets betydning relaterer til indbyrdes forbundne meningselementer (Adams 1990:15). Ordgenkendelse (og læseforståelse) er altså et flerdimensionalt puslespil af erfaringsbaserede videnselementer, der bringes i spil og forbindes med hinanden.

Samtidig favner teorien både *bottom-up* og *top-down*-teorierne på en anden måde end de traditionelle interaktive modeller: Hvis for eksempel man fodrer et netværk med bogstavkomponenter og lærer det at klassificere ord, pseudoord og bogstaver, så lærer det at identificere bogstaver i ord bedre end bogstaver i pseudoord. Modellen fungerer altså *bottom-up* (fodres med de mindste enheder) men udviser en klar *top-down*-effekt (applicerer sit kendskab til ord) (Norris 1990). Dette skyldes at systemet har et flow af information tilbage i netværket for at korrigere de enkelt vægte under indlæringen. Denne effekt forklarer fænomener, hvis eksistens man tidligere nok har kunnet konstatere, men som man aldrig på tilfredsstillende vis har kunnet inkorporere i de traditionelle læsemodeller, eksempelvis kontekstefekt (se fx oversigt i Tabossi 1991). Tidligere har man således aldrig kunnet godtgøre, om kontekst influerer på ordgenkendelse eller først bliver aktiveret i en senere forståelsesproces, hvor sætningens fulde meningsindhold behandles, ligesom det har været uklart, hvorvidt et flertydigt ord aktiverer én eller flere meninger. I den konnektionistiske læsemodel er kontekstefekten konstant, idet al information påvirker vægtene i netværket – i traditionel terminologi kunne man sige, at kontekst løbende bliver en del af *schema*, dvs. den inkorporeres kontinuerligt i de forventninger har til teksten på baggrund af vores viden om alverdens ting (Seidenberg 1990:59; Whitney & Warning 1991). Ligeledes elimineres problematikken om, hvorvidt én eller flere meninger aktiveres ved læsning af flertydige ord (se fx oversigt i Seidenberg 1982:490-496 og Simpson, 1984), idet der ikke som i dualruteteorien opereres med et leksikon. I stedet er ordenes semantiske indhold fleksible sammensmeltninger af mange former for information, lagret i netværkets konnektioner (Seidenberg 1992:58).

Bagsiden af medaljen er, at modellen i sin altfavnende universalitet er meget uspecifik. Et væsentligt kritikpunkt er, at den udpræget interaktive forståelse af netværkets funktion underkender, at det synes dokumenteret, at komplicerede processer faktisk behandles i en sekventiel form, ligesom de gamle modeller antager (fx Norris 1990).[8] Eksempelvis kan der ved læsning af flertydige ord konstateres en tidsafstand på nogle hundrede millisekunder mellem ordbestemmelse og ordgenkendelse (Whitney & Warning 1991:205), jvf. figur 1.3. En betydelig del af den terminologi, der er etableret på basis af de traditionelle læsemodeller, er derfor stadig i anvendelse – også hos læseteoretikere, der støtter konnektionismen. Fx bruges begreber som ortografisk og fonologisk afkodning fortsat – blot mere til at beskrive, at der er tale om behandling der involverer ortografisk eller fonologisk materiale, end for at sige, at afkodningen er isoleret ortografisk eller fonologisk (fx Seidenberg 1992). Så selvom dualruteteorien i sig selv muligvis står for fald, kan de nævnte begreber altså godt anvendes på det analytiske plan.

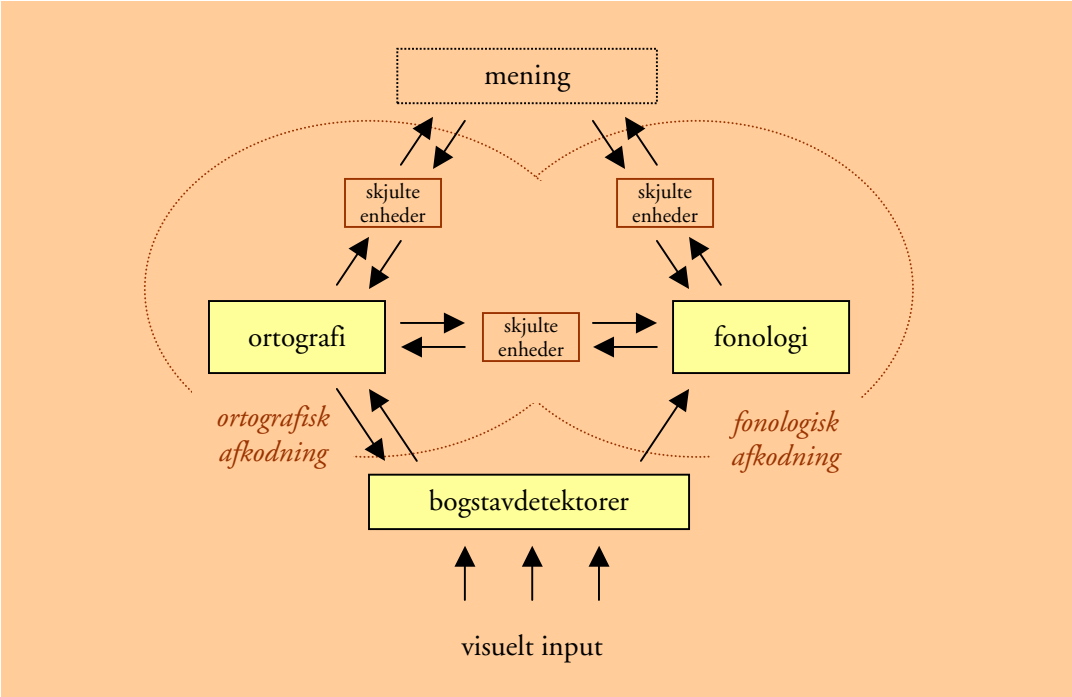
Teoretisk udgangspunkt for afhandlingen

En sådan analytisk syntese af dualruteteoriens terminologi og konnektionismens kompleksitet synes hensigtsmæssig som udgangspunkt for de videre analyser i denne afhandling. Men de gængse konnektionistiske modeller ignorerer, at der kan være tale om delprocesser som fx initial bogstavafkodning (Richman & Simon 1989) og anvender desuden en i denne sammenhæng uklar terminologi, idet skriften betragtes som en del af ortografien, jvf. begrebsafklaringen i afsnit 1.1. Begge dele er uhensigtsmæssige i denne sammenhæng, hvor skriften og dermed genkendelse af enkeltbogstaver udgør et problemfelt i sig selv. En tilpasset model, der kompenserer for disse svagheder ses illustreret i figur 2.4. I modsætning til hos Seidenberg & McClelland opretholdes her et element af sekvens-opdelt progression, idet bogstavgenkendelse ses som et selvstændigt led i processen.[9] Som det ses, involverer ordgenkendelse ikke nødvendigvis fonologisk behandling – så langt ligner modellen dualruteteorien – men i modsætning til dualruteteoriens lineære struktur med separate veje til et eksisterende leksikon, er output her et produkt af en proces, som foregår interaktivt i et net af enheder, der sender signaler frem og tilbage mellem hinanden. Bemærk, at bogstavdetektorer, ortografi og fonologi ikke skal betragtes som statiske elementer, men som 'undernetværk', der ligeledes behandler input i lag af skjulte enheder. Udover at modellen i figur 2.4 således er tilpasset denne afhandlings terminologi, giver den altså bedre end Seidenberg & McClellands et udgangspunkt for en analyse af læseprocesser, der ikke er optimale, fordi skriften for L2-læseren er uvant.

---

[8] Et centralt kritikpunkt mod konnektionismen generelt er, at netværkene kræver megen træning for at fungere. De er forholdsvis dårlige til at generalisere på baggrund af faste regler (til forskel fra mennesker og den traditionelle computer) og de kan ikke kan lære af en enkelt slem fejl, som mennesker kan. På et mere abstrakt niveau kritiseres konnektionismen for at være en teknisk forsimpling, idet basale aspekter af menneskelig kognition lades ude af betragtning. Humanister peger på tro, ønsker, planer, etik og moral, og biologer på neurontransmittere og hormoner som eksempler. Modargumentet er, at disse mangler blot afspejler, at de netværk man på nuværende tidspunkt kan konstruere, er forsimplinger, der i en biologisk version vil kunne rumme det hele (Ramsey 1999; Garson 2002).

[9] At ordgenkendelsesprocessen faktisk er sekvens-opdelt med genkendelse af bogstavformer som en initial proces er forholdsvis veldokumenteret i neurologisk forskning (Miozzo & Caramazza 1998).



Figur 2.4 Tilpasset konnektionistisk ordgenkendelsesmodel

2.3 Læsning på tværs af ortografier

Relanceringen af det visuelle inputs betydning for læseprocessen og diskussionen om dualruteteoriens holdbarhed har haft den vigtige sidegevinst, at man blev opmærksom på, at læseprocessen ikke nødvendigvis er universel, men muligvis forløber forskelligt ved læsning af forskellige ortografier. (Hung & Tzeng 1981; Henderson 1984:4; Just & Carpenter 1987:305-). Det var således først i midten af 1980'erne, at læseforskere rigtigt begyndte at interessere sig for læsning på tværs af ortografier. Ved hjælp af neurologiske undersøgelser af patienter med erhvervet dyslexi begyndte man at undersøge hvilke elementer af læseprocessen, der er kognitivt universelle hhv. afhængige af ortografi, ligesom man med udgangspunkt i dualruteteorien forsøgte at determinere i hvilket blandingsforhold de to veje til leksikon manifesterer sig ved læsning af forskellige ortografier (Besner et al. 1984; Coltheart 1984; Morton & Sasanuma 1984; Sasanuma 1984; Just & Carpenter 1987:314).

Et udgangspunkt for en stor del af forskningen i forskelle i læseproces mellem ortografier har været hypotesen om ortografisk kompleksitet (*the orthographic depth hypothesis*) (Feldman & Turvey 1983; Frost et al. 1987; Katz & Frost 1992). Hypotesen bygger på, at alfabetiske ortografier adskiller sig fra hinanden i kraft af, at ikke alle har samme grad af korrespondance mellem grafem og fonem. En gennemskuelig ortografi (*shallow orthography*) har en simpel og konsistent én til én –korrespondance mellem grafem og fonem, mens en kompleks ortografi (*deep orthography*), på trods af tilknytning til det alfabetiske princip, har en række mere komplicerede sammenhænge mellem bogstaver og lyd. På et kontinuum mellem gennemskuelig og kompleks er fx. serbokroatisk ortografi gennemskuelig, da hvert bogstav kun repræsenterer ét fonem, og hvert fonem kun

repræsenteres af ét bogstav, mens fx engelsk ortografi er mere kompleks, da der er en lang række regulære fonetiske forskelle mellem ord med samme bogstavkombination (fx *heal-health*) og ens udtale for ord med uens bogstavkombinationer (fx *peel-deal*). Ifølge hypotesen om ortografisk kompleksitet støtter gennemskuelige ortografier en fonetisk medieret ordgenkendelsesproces, mens komplekse ortografier lægger op til ortografisk afkodning. Afkodningsstrategien afhænger altså ifølge hypotesen af, hvor meget fonetisk materiale ortografien rummer. Hypotesen om ortografisk kompleksitet var som udgangspunkt baseret på dualruteteoriens gyldighed. Det aktuelle oprør mod dualruteteorien stiller en del af den forskning, der støtter hypotesen, i en usikker position, for dualruteteoriens kritikere påpeger jo netop, at fonologisk hhv. ortografisk behandling ikke kan isoleres. Det truer imidlertid ikke umiddelbart hypotesen i sig selv (Katz & Frost 1992:78; Berninger 1994:13), for, som Berninger skriver:

> *"… reading any orthography probably requires orthographic-phonological connections, although specific orthographies may differ in the units of orthography and units of phonology involved in forging connections."* (Berninger 1994:13)

Samtidig kan det anføres, at selvom det ikke kan godtgøres, at hhv. fonologisk og ortografisk afkodning kan isoleres, så kan man udmærket argumentere for, at i den kobling, som finder sted ved læsning, er andelen af fonologisk materiale i overvægt ved læsning af gennemskuelige ortografier (som det er tilfældet for pseudoord), mens andelen af ortografisk materiale omvendt er mest markant ved læsning af komplekse ortografier (som det er tilfældet for uregelmæssigt stavede ord).

The European Commission has just announced an agreement whereby English will be the official language of the EU rather than German, which was the other possibility. As part of the negotiations, Her Majesty's Government conceded that English spelling had some room for improvement and has accepted a 5 year phase-in plan that would be known as 'EuroEnglish':

In the first year, 's' will replace the soft 'c'... Sertainly, this will make the sivil servants jump with joy. The hard 'c' will be dropped in favor of the 'k'. This should klear up konfusion and keyboards kan have one less letter.

There will be growing publik enthusiasm in the sekond year, when the troublesome 'ph' will be replased with the 'f'. This will make words like 'fotograf' 20% shorter.

In the 3rd year, publik akseptanse of the new spelling kan be expekted to reach the stage where more komplikated changes are possible. Governments will enkourage the removal of double letters, which have always ben a deterent to akurate speling. Also, al wil agre that the horible mes of the silent 'e''s in the language is disgraseful, and they should go away.

By the 4th year, peopl wil be reseptiv to steps such as replasing 'th' with 'z' and 'w' with 'v'. During ze fifz year, ze unesesary 'o' kan be dropd from vords kontaning 'ou' and similar changes vud of kors be aplid to ozer kombinations of leters.

After zis fifz yer, ve vil hav a reli sensibl riten styl. Zer vil be no mor trubls or difikultis and evrivun vil find it ezi tu understand ech ozer.

- ZE DREM VIL FINALI KUM TRU!!

EuroEnglish – Fra en kompleks til en gennemskuelig ortografi ? (*http://www.thehumorsource.com*)

Konsekvensen i serbokroatisk ortografi har affødt studier, der sammenlignede læsning på serbokroatisk og engelsk, og sådanne studier dominerede da også feltet i begyndelsen (fx Katz & Feldman 1981; Feldman & Turvey 1983; Katz & Feldman 1983; Turvey et al. 1984), men efterhånden så nogle teoretikere potentialet i at bevæge sig uden for den snævre kreds af europæiske sprog, og man begyndte også at interessere sig for helt anderledes skriftsystemer.[10] Interessen koncentrerede sig især om hebraisk og japansk, der på hver deres måde kunne bidrage med nye, interessante dimensioner: Hebraisk eksisterer i to fonografiske ortografier, hvor den ene (vokaliseret) er meget gennemskuelig, mens den anden (uvokaliseret) er stærkt kompleks, mens japansk anvender både et logografisk (*kanji*) og to syllabiske (*hirgana* & *katagana*) skriftsystemer, som gør det muligt at skrive de samme ord både logografisk og fonografisk. Disse studier viste forskellige responsmønstre for forskellige skriftsystemer (Sasanuma 1984; Turvey et al. 1984), ligesom kliniske studier viste, at logografisk og syllabisk afasi skyldtes forskellige former for hjernelæsioner, og at japanske patienter kunne miste evnen til at afkode den ene ortografi men bibeholde evnen til at afkode den anden (Sasanuma 1984).

Disse undersøgelser syntes at underbygge dualruteteorien og dermed antagelsen om, at ortografisk og fonologisk afkodning er separate processer, og ræsonnementet vedrørende sammenhængen mellem ortografiers grafem-fonem-korrespondance og afkodningsstrategi blev udvidet til at omfatte afkodningsprocessen på det enkelte sprog generelt. På den baggrund blev afkodning ved læsning af logografiske og meget komplekse alfabetiske ortografier anset som nødvendigvis leksikalsk, mens ordgenkendelse ved læsning af helt gennemskuelige, fonografiske ortografier mentes at være udelukkende fonologisk. Man mente således at kunne godtgøre, at fonologisk afkodning er den eneste mulige strategi på serbokroatisk (Turvey et al. 1984:88), at leksikalsk afkodning er den eneste mulige på kinesisk (Coltheart 1984:71), mens begge strategier må i brug på japansk, hvor *kanji* afkodes leksikalsk, mens *kana* (dvs. *hirgana* & *katagana*) må overføres til en fonologisk kode, før ordgenkendelse er mulig (Morton & Sasanuma 1984; Sasanuma 1984).

I takt med den teoretiske udviskning af grænsen mellem dualruteteoriens to veje, opfattes leksikalsk og fonologisk afkodning også i denne forbindelse efterhånden mere som to strategier, der sideløbende er i spil og støtter hinanden. Nyere undersøgelser underbygger dette ved at dokumentere, at leksikalsk ordgenkendelse faktisk spiller en rolle ved læsning af eksempelvis serbokroatisk (Katz & Frost 1992:72), mens fonologisk kodning spiller en rolle for ordgenkendelse på kinesisk (Henderson 1982:192; Hung et al. 1994; Jackson et al. 1994; Hanley & Huang 1997). Imidlertid synes det teoretiske opbrud at have skabt en fundamental usikkerhed om, hvad der egentlig karakteriserer ortografisk behandling. Således skriver Geva & Willows eksempelvis så sent som i 1994:

> *"[Recent research] highlight[s] the importance of examining carefully in different writing systems what is meant by orthographic knowledge, and the contribution of underlying cognitive and linguistic factors to its development in different orthographic systems. It is clear that theoretical claims regarding the universal role of orthographic and phonological processing in reading and spelling, based on*

[10] Dette afspejles i en variant af hypotesen om ortografisk kompleksitet, kaldet *script dependence hypothesis*, (fx hos Gholamain & Geva 1999; Geva & Siegel 2000).

> *learning to read and spell in English and other Roman-based alphabets, need to be examined carefully."* (Geva & Willows 1994:365)

Helt centralt i dette arbejde er studier af simultan tosproget læseindlæring, som giver mulighed for at vurdere tværortografiske forskelle i læsetilegnelsen hos de samme forsøgspersoner. Sådanne undersøgelser har vist, at japanske begynderlæsere læser *kanji* bedre end de læser *kana*, tilsyneladende fordi *kanji*-symboler associeres med noget meningsfuldt (Henderson 1982:192). Hvad angår alfabetiske skriftsystemer, tyder flere undersøgelser på, at børn hurtigere lærer at afkode ved læsning af gennemskuelige ortografier (vokaliseret hebraisk og vokaliseret persisk) end komplekse ortografier (engelsk) (Geva et al. 1993; Geva & Wade-Woolley 1998; Gholamain & Geva 1999; Geva & Siegel 2000), også selvom den gennemskuelige ortografi er andetsprogets og langt dårligere udviklet end modersmålets. Desuden er de almindeligste kategorier af afkodningsfejl for sådanne tosprogede begynderlæsere ortografi-specifikke (Geva & Siegel 2000).

Disse forskningsresultater indicerer, at forskellige ortografier i nogle henseender trækker på forskellige kompetencer: Hos engelsk-hebraisk tosprogede børn synes fonologisk opmærksomhed at være vigtigere for læsefærdigheden ved læsning på engelsk end på hebraisk, mens hurtig ordgenkendelse synes at være vigtigere ved læsning på uvokaliseret hebraisk end på engelsk (Geva & Wade-Woolley 1998). Forskning i tosprogethed støtter altså både hypotesen om, at læsefærdighed på alle sprog afgøres af individuelle kognitive faktorer (*the central processing hypothesis)* og hypotesen om ortografisk kompleksitet, ifølge hvilken læseprocessen forløber forskelligt for forskellige ortografier (Geva & Wade-Woolley 1998; Gholamain & Geva 1999; Geva & Siegel 2000). Dette er på alle måder i overensstemmelse med konnektionismen: De generelle færdigheder og evner (t*he central processing*) er givne og på teoretisk plan et udtryk for hvor mange lag af skjulte enheder den enkelte neurocomputer rummer – hos mennesket er de biologisk betingede og svarer til antallet af synapser i hjernen – mens den ortografiske kompleksitet gennem træning påvirker vægtene i konnektionerne og dermed bliver ortografi-specifikke færdigheder.

2.4 Læsning på fremmedsprog

Frem for at udvikle en selvstændig teori for læsning på L2 er teoretikere inden for fremmedsprogspædagogik som udgangspunkt gået ud fra, at læsning på L2 aktiverer samme former for kognitive processer som læsning på L1 (Carr & Levy 1990). Som nævnt er der imidlertid efterhånden blevet sat spørgsmålstegn ved universaliteten i L1-læseteorierne, og ikke mindst på den baggrund er generaliseringerne til L2-læsning også blevet udfordret, hvorfor man er begyndt at undersøge, om forskelle i ortografisk, leksikalsk og morfo-syntaktisk kompleksitet mellem L1 og L2 påvirker læseprocessen.

Langt hovedparten af forskningen i tværsproglig læsning har imidlertid været koncentreret om studier i tosprogethed som eksemplificeret ovenfor. Dvs. tilfælde, hvor de mundtlige sproglige færdigheder på L2 er på et L1-lignende niveau, ofte som resultat af, at L2 er indlært allerede tidligt i barndommen. Meget af denne forskning i tosprogethed har drejet sig om, hvorvidt der sker en overførsel (*transfer*) af læsestrategier fra det ene sprog til det andet. Det har vist sig, at læsefærdighed på L1 er en god indikator for læsefærdighed på L2, hvilket jo støtter *the central processing hypothesis*. Læsefærdighed opfattes på den baggrund generelt som en funktion af individuelle kognitive og lingvistiske evner, såsom hukommelse, proceshastighed, ortografisk viden og fonologisk opmærksomhed (Geva & Wade-Woolley 1998; Gholamain & Geva 1999; Geva & Siegel 2000). Ifølge Cummins'

interdependenshypotese (*interdependence hypothesis*) påvirker læsning på L1 og L2 hinanden gensidigt, således at læsefærdighed på L2 afhænger af læsefærdighed på L1, mens læsning på L2, ud over at have indflydelse på læsefærdigheden på L1, har en positiv effekt på generel sproglig bevidsthed og andre akademiske kompetencer (Cummins 1984; Verhoeven 1991).

Hvad angår tværsproglig læsning i perioden før fremmedsproget beherskes så godt, at det kan kaldes et andetsprog, er forskningen sparsom. Nogle har dog konstateret, at L2-læsere ikke er i stand til at overføre gode læsefærdigheder fra L1 til L2, før de har opnået et vist niveau af kompetence på L2: Ifølge Alderson (1984) og Cummins (1991) findes der en 'sproglig tærskel', som overskrides, når L2-læsere har 'flydende ordgenkendelse' og tilstrækkeligt kendskab til L2 til at kunne bruge læsefærdigheder på L1 til at styrke deres L2 (*linguistic threshold hypothesis)*. Koda (1992) opererer med en *limited-capacity theory*, ifølge hvilken der er grænser for, hvor mange mentale ressourcer, vi kan aktivere på én gang i forbindelse med en så kognitivt krævende aktivitet som læsning. Kun hvis vores behandling af de ord, vi læser, er automatisk, kan de kognitive og metakognitive processer, som forekommer ved normal læsning, finde sted.

Grundlæggende påpeger denne form for teoretisering blot, at manglende sproglige kompetencer belaster læseprocessen, fordi sprogproblemet beslaglægger megen kognitiv kapacitet, og det præciseres ikke, i hvor høj grad problemet med at etablere gode læsefærdigheder på L2 er et specifikt læseproblem, og i hvor høj grad det skyldes mangel på generelle sproglige kompetencer. Samtidig har kun få søgt empirisk at undersøge, hvad der mere specifikt karakteriserer den læseproces, der – selvom den ikke er optimal – finder sted, når man begynder at lære at læse på et fremmedsprog. MacNamara et al. (1968) viste, at læsere gjorde bedre brug af semantiske og syntaktiske redundanser på L1 end på L2, ligesom Oller & Tullis (1973) viste, at L2-læsere havde øjenfiksering af længere varighed L1-læsere, på trods af at læseforståelsen og antallet af fikseringer ikke var signifikant anderledes. Også studier af nyere dato viser, at læseprocessen på L2 er mere tidskrævende end på L1 (Favreau et al. 1980; Segalowitz et al. 1991). Disse studier støtter altså ovennævnte hypoteser om, at man ved læsning på L2 udviser mindre kognitivt overskud end ved læsning på L1. Favreau & Segalowitz (1982) kunne desuden dokumentere, at den nævnte hastighedsforskel ikke blot afspejler en generelt lavere proceshastighed på L2 inden for alle de sproglige færdigheder, men også skyldes en mindre automatiseret ordgenkendelse – dvs. et specifikt læserelateret problem – fordi tendensen også fandtes hos tosprogede med et ellers fuldstændig flydende L2. Den lavere hastighed måtte således være et resultat af den senere og mere begrænsede eksponering for tekst på L2. Dette synspunkt støttes af Bossers (1991), som viste, at i tilfælde, hvor akademiske færdigheder er tilegnet på både L1 og L2, eller udelukkende på L2, kan læsefærdigheden på L2 godt være bedre end på L1.

På trods af dette er et gennemgående træk ved den eksisterende litteratur om læsning på L2 en stærk tendens til at fokusere på generelle sprogkompetencer og processer, der finder sted i den øverste del af læsehierarkiet (Randall & Meara 1988; Koda 1992; Koda 1997; Urquhart & Weir 1998:50), jvf. figur 2.1. Den interaktivt-kompensatoriske læseteori er beredvilligt blevet adopteret af fremmedsprogsforskningen, hvor det er en bredt accepteret hypotese, at baggrundsviden ved læsning på fremmedsprog kan kompensere for mangelfulde sproglige færdigheder på det pågældende sprog (Urquhart & Weir 1998:45). Dette på trods af, at L2-læsere typisk tyer til en meget bundstyret læsning, når opgaven er for svær, hvilket i litteraturen betegnes som en uhensigtsmæssig strategi, mens aktivering af de øvre niveauer antages at være det optimale (fx Haastrup 1993: 12-13).

I løbet af de seneste år, hvor læsehierarkiets lavere niveauer har fået større opmærksomhed, er ordgenkendelse ved læsning på fremmedsprog også blevet genstand for større interesse. Man har eksempelvis vist, at ordgenkendelsesstrategier fra L1 overføres og anvendes ved læsning på L2, og at L2-læsning derfor kompliceres af store forskelle mellem ortografierne på de to sprog (Koda 1997; Muljani et al. 1998). Desuden er især effekten af, at ordforrådet (i dualruteteoretisk terminologi beskrevet som 'det indre leksikon') på fremmedsproget er begrænset, blevet genstand for interesse. Og meget tyder på, at netop ordforrådet er en afgørende determinant for læsefærdighed på L2 (se fx Coady 1997; Laufer 1997; Paribakht & Wesche 1997; Yang 1997). Laufer skriver således:

> *"[It has been consistemtly demonstrated that] by far the greatest obstacle to good [L2] reading is insufficient number of words in the learner's lexicon. Lexis was found to be the best predictor of succes in reading, better than syntax or general reading ability."* (Laufer 1997:31)

Dette er blevet udnyttet til at forsøge at definere den 'sproglige tærskel', og det anslås, at 95% af ord i en tekst – beregnet til at svare til 3000 ordfamilier eller 5000 ord – er et minimum for at tidligere erhvervede læsestrategier kan bringes i anvendelse (Laufer 1997:31), mens 98% af ordene skal være kendte, for at afslappet 'lystlæsning' er mulig (Hirsh & Nation 1992).

Der er således inden for L2-læseteorien en begyndende erkendelse af, at automatiseret ordgenkendelse spiller en central rolle for læsefærdighed, og at ordforråd i den forbindelse er den måske vigtigste komponent at udvikle. Til gengæld er læsehierarkiets allerlaveste niveauer – skriftsystemet og fonologien – fortsat næsten helt uberørte områder inden for forskning i læsning på L2 (se dog Koda 1992; Jackson et al. 1994; Akamatsu 1998; Koda 1998; Akamatsu 1999; Jackson et al. 1999, som alle behandler forskelle mellem alfabetiske og logografiske ortografier). I lyset af at automatisering etableres på baggrund af frekvens, synes det imidlertid indlysende, at netop et nyt skriftsystem udgør en ikke ubetydelig barriere for L2-læseren.

2.5 Læsning med et nyt alfabet [11]

Ikke overraskende er de mest studerede tilfælde af læsning af en uvant skrift børns begynderlæsning. Og udviklingen inden for læseforskningen har naturligt nok påvirket den tilbagevendende diskussion om, hvordan man mest hensigtsmæssigt lærer børn at læse: Lingvisternes indflydelse på de tidlige læseprocesteorier gjorde lydmetoden fremherskende. Dens udgangspunkt er, at læsning er tolkning af det alfabetiske system, dvs. fonologisk afkodning, hvorfor læsetilegnelse kræver træning i forholdet mellem grafem og fonem. *Top-down* modellerne og parallelruteteoriens fokus på den direkte, leksikalske afkodning har inspireret til mere holistiske fremgangsmåder, herunder ordbilledmetoden og helheds-læsning, der har som udgangspunkt, at læsning er en naturlig sproglig aktivitet, hvorfor læsetilegnelse handler om at benytte sproglige færdigheder, der er kendt fra talesproget, til

[11] I dette afsnit behandles udelukkende alfabetiske skriftsystemer, idet de yderligere komplikationer, tilegnelse af andre former for skriftsystemer måtte indebære, ikke er relevante i denne sammenhæng. Det er dog relevant at nævne, at undersøgelser af japanske lørneres L2-læsning på engelsk (Akamatsu 1998) og amerikanske lørneres L2-læsning på japansk (Koda 1992) synes at vise, at der er en sammenhæng mellem god bogstav-/tegngenkendelse og god læseforståelse på L2.

at nå frem til mening, gerne ved at gætte ud fra konteksten, mens der på ordgenkendelses-niveau lægges vægt på genkendelse af hele ords form (Just & Carpenter 1987:329; Frost 1998:25). Eksperimentel forskning har inden for de senere år givet de holistiske metoder betydelig modstand, da man har dokumenteret, at fonologisk opmærksomhed er afgørende for hurtig læsetilegnelse, og at elever med svage fonologiske forudsætninger falder bagud, når de ikke trænes i sammenhængen mellem grafem og fonem. Egentlige sammenligninger mellem de to metoders resultater falder ifølge Adams (1990:31) også ud til lydmetodens fordel. Nogle mener dog, at fremgangsmåden ikke kan stå alene og inkorporerer derfor elementer fra de andre nævnte metoder (fx Frost 1998:25). Andre påpeger, at eksistensen af børn, der lærer at læse af sig selv inden skolestart, burde inspirere til at lade skrift indgå lige så naturligt som tale i voksnes kommunikation til børn længe før skolealderen, gennem en uformel præsentation af skrevne ord uden eksplicit fokus på forbindelsen mellem grafem og fonem (fx Kjertmann 2002). Alligevel kan man sige, at der blandt læseforskere generelt er etableret en bred konsensus om, at evne til at afkode er essentiel for at kunne udvikle effektiv læseforståelse (Just & Carpenter 1987:326-), og at automatisk ordgenkendelse, der overflødiggør omkodning til lyd, kommer med øvelsen, når ordene i takt med omfanget af teksteksponering bliver højfrekvente for den enkelte elev (Adams 1990:98).

Der er imidlertid markante forskelle mellem begynderlæsere og læsere, der kan læse på deres L1, men lærer et nyt skriftsystem. At lære at læse forbindes først og fremmest med at 'knække koden' – dvs. opnå forståelse for det fonematiske princip – hvilket fremmedsprogs-studerende normalt har gjort mange år tidligere. Omvendt har begynderlæseren til forskel fra den studerende, der lærer et nyt sprogs skriftsystem, stor fortrolighed med det talte sprog, som det skrevne repræsenterer. Det de har tilfælles, er som udgangspunkt manglende kendskab til selve skriften. Når der er etableret hurtig bogstavgenkendelse (og hos begynderlæseren forståelse for det fonematiske princip) er afkodning mulig. Men for at kunne læse effektivt er det ikke tilstrækkeligt, at der finder afkodning sted. For at læseren kan have mentalt overskud til at koncentrere sig om tekstens meningsindhold, må afkodningen være automatisk:

> *"[...] gode læseres ordafkodning foregår automatisk. Det vil sige tre ting: for det første optager afkodningen ikke opmærksomheden, for det andet foregår den obligatorisk (når man ser et ord, kan man ikke lade være med at læse det) [...], og for det tredje kan man ikke få indblik i ordafkodningen ved at vende blikket 'indad' og forsøge at lægge mærke til, hvad man gør, når man læser."* (Elbro 2001:41)
>
> *"Automatisering opnås gennem øvelse. Så enkelt og så krævende er det".* (Elbro 2001:42)

Målet for både begynder- og fremmedsprogslæseren er altså at opnå tilstrækkelig øvelse til at kunne flytte koncentrationen fra afkodningen til indholdet. Dette kræver først og fremmest hurtig og sikker ordgenkendelse. I modsætning til dualruteteoriens forestilling om en direkte genkendelse af ordet i sin helhed – eller ordets **grafiske** identitet – følger det af konnektionismen og en mængde empiriske studier (Adams 1990:97), at automatisk ordgenkendelse hviler på genkendelse af ordets **ortografiske** identitet, dvs. den bestemte bogstavrækkefølge, der udgør ordet:

> *"An [...] aspect of the writing system that the child must learn about concerns the distribution of letter patterns [...]. Only some combinations of letters are possible, and the combinations differ in frequency. [...] Constraints on the form of written*

> *words may play an important role in the recognition process. The reader must discriminate the input string from other vocabulary items, a task that might be facilitated by knowledge of the letter combinations that are permissible or realized."* (Seidenberg & McClelland 1989:524)

Én strategi hos begynderlæsere med mangelfuld bogstav- eller ordgenkendelse er at gætte ud fra kontekst og i meget høj grad støtte sig til ordets samlede grafiske udtryk. Ofte bygger genkendelse på det første bogstav og et markant fikspunkt i midten af ordet, fx et højt bogstav (Adams 1990:112; Elbro 2001:39). Denne strategi giver selvsagt problemer, efterhånden som læsebogens ordforråd udvides (Elbro 2001:39). En anden strategi er at overfokusere på de enkelte bogstaver (Adams 1990:112), hvilket resulterer i en meget langsommelig omkodning, der på sprog som engelsk og dansk desuden kompliceres af uregelmæssighederne i grafem-fonem-korrespondancen (Elbro 2001:118). Begge strategier ses også hos L2-læsere (Hauge 1993:25).

For at blive en god læser kræves det altså, at man behandler ord på bogstavniveau, men samtidig at bogstaver ikke alene genkendes uafhængigt af hinanden, men først og fremmest i grupper af frekvente bogstavsammenstillinger. Faktisk har man vist, at genkendelse af i hvert fald engelske enstavelsesord forløber som en rullende tripel-bogstavanalyse, så fx ordet *'make'* behandles som overlappende sæt af tre bogstaver, dvs. *_ma, mak, ake, ke_* , hvor _ repræsenterer mellemrummet før eller efter ordet (Seidenberg & McClelland 1989). Ved indlæring af nære fremmedsprog er dette sandsynligvis en fordel, idet nabofrekvenseffekten til en vis grad kan overføres, selvom selve ordene på L2 er lavfrekvente. Men ved indlæring af et nyt skriftsystem er den L2-studerende lige så meget begynder som L1-begynderlæseren, og har tilmed langt dårligere forudsætninger for at gætte og skabe mening ud fra kontekst. Dette kan tvinge den studerende til i meget høj grad at læne sig op ad skriftens fonologiske information.

Studier af, hvordan voksne, alfabetiserede personer tilegner sig et artificielt skriftsystem[12] viser således – i lighed med de før omtalte undersøgelser af læsning på tværs af ortografier – at hvis skriftsystemet indeholder fonologisk materiale (i alle undersøgelserne med direkte relation til det latinske alfabet og forsøgspersonernes L1, engelsk), så afkoder forsøgspersonerne fonologisk. Sammenligninger af læsning af artificielle alfabetiske og logografiske skriftsystemer viser, at den fonologiske afkodning gør ordgenkendelse betragteligt langsommere (Baron 1977, 1978; Brooks, 1977, 1979), og forsinkelsen ved fonologisk afkodning svinder først efter en ret omfattende læsetræning med det nye alfabet (Baron 1977; Brooks 1977).[13] Ved indlæring af et nyt alfabet overfører forsøgspersonerne altså deres kendskab til det fonematiske princip og støtter sig til grafem-fonem-korrespondancen, på trods af at det – i hvert fald på kortere sigt – er en tidsrøvende strategi. Hvis de i stedet fokuserede på at lære hele ord som logografer, ville de opnå langt hurtigere ordgenkendelse. Den fonologiske afkodning bliver først efter omfattende træning effektiv, og tilsyneladende

---

[12] I princippet er selvfølgelig alle skriftsystemer artificielle, men betegnelsen anvendes gerne om skriftsystemer, der til forskningsbrug er er opfundet af forsøgslederen.

[13] Mærkeligt nok har grafem-fonem korrespondance imidlertid en positiv effekt på afkodningshastighed, hvis den er implicit – dvs. maskeret således at forsøgspersonerne ikke opdager den. Brooks & Miller (1979) fandt således, at ved læsning af ord skrevet med et artificielt skriftsystem, der havde direkte korrespondance til det engelske alfabet men skulle læses fra højre mod venstre (hvilket ingen af forsøgspersonerne gennemskuede), havde forsøgspersonerne hurtigere ordgenkendelse, end hvis der ingen relation var til alfabetet, eller hvis relationen var eksplicit i den forstand, at forsøgspersonerne var blevet informeret om den.

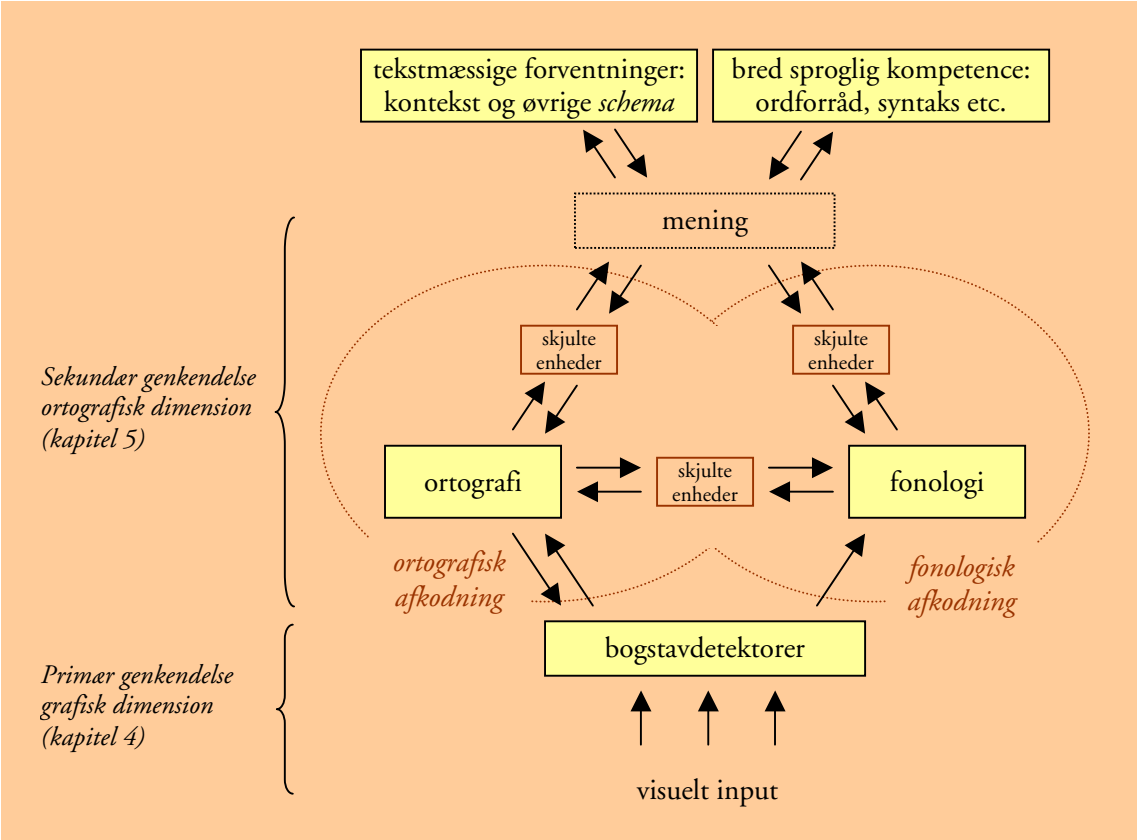
i høj grad i kraft af, at der etableres automatisk ordgenkendelse. I hvert fald falder afkodningshastigheden betragteligt, når nye ord introduceres (Brooks & Miller 1979).

Hos L2-studerende, der lærer et nyt skriftsystem, synes fonologisk omkodning af de enkelte segmentsymboler altså at være en almindelig strategi på begynderstadiet, hvor genkendelse af symboler og dermed også ord endnu ikke er hurtig og automatiseret. Dette kan være særlig problematisk, hvis sproget har meget fremmedartede fonemer (Koda 1998) – hvilket ofte er tilfældet for fjerne fremmedsprog, der skrives med et ikke-latinsk skriftsystem – da mangelfuld genkendelse og evt. forveksling af fonemer kan forstyrre etableringen af en mental forbindelse mellem ordbillede, lydbillede og mening.

2.6 Sammenfatning

Det læseteoretiske udgangspunkt for denne afhandling er en konnektionistisk læseproces-model, hvor ordgenkendelsesniveauet er sekvens-opdelt i bogstavanalyse (primær genkendelse) og den egentlige afkodning, hvor ortografisk og fonologisk behandling af det grafiske materiale skaber mening (sekundær genkendelse). Ved læsning af tekst og ikke blot enkeltord vil modellen som illustreret i figur 2.5 naturligvis også involvere videnselementer fra læsehierarkiets højere niveauer.

Hvad angår læsning på L2 kan man sammenfattende sige, at vores viden om læseprocesser på L2 i den fase, hvor L2 endnu ikke beherskes flydende, er meget begrænset. Traditionelt



Figur 2.5 Komplet læseprocesmodel med udspecificering af primær og sekundær genkendelse i ordgenkendelsesfasen.

har læseteorier udviklet på baggrund af undersøgelser om læsning på L1 langt hen ad vejen været generaliseret til også at gælde for læsning på L2. Dog er der efterhånden dannet konsensus om, at læsning på L2 adskiller sig afgørende fra læsning på L1 i kraft af, at ordforrådet på L2 er begrænset. Hertil kommer, at forskningsresultater vedrørende læsning på L1 gennem de seneste år har fået vigtige konsekvenser for L2-læseteorien: Teorien om den fonologiske repræsentations rolle for ordgenkendelse ved læsning på L1 er efterhånden veletableret, og det må på den baggrund betragtes som sandsynligt, at L2-specifikke lyde hæmmer læsning på L2. Ligeledes er der efterhånden overbevisende empirisk dokumentation for, at læseprocesser er forskellige ved læsning af forskellige ortografier, hvilket indebærer, at læsning på L2 nødvendigvis må adskille sig fra læsning på L1, især hvis de to sprog skriftligt repræsenteres af forskellige skriftsystemer. Hvis eksisterende læsefærdigheder og dermed også læsestrategier overføres fra L1 til L2, har det sandsynligvis vigtige konsekvenser for udviklingen af læsefærdighed på L2.

L2-læseforskningen har med få undtagelser (fx Koda 1996; Muljani et al. 1998) negligeret konnektionismen, men samspillet mellem de ovenfor nævnte aspekter falder udmærket ind under modellen i figur 2.5, og her bliver konnektionismens universalitet igen dens force: Modellen omfatter de elementer, der er relevante for tekstlæsning på både L1 og L2, og eksemplificerer hvordan afkodning, sprogforståelse og forventninger til teksten kan betragtes som selvstændige kompetencer. Læseprocessen fremstår således som en entydigt individuel proces, hvor output, nemlig meningen med teksten, er et produkt af læserens helt personlige forudsætninger. På ordgenkendelsesniveau er processen betinget af læserens tidligere erfaringer med den grafiske, ortografiske og fonologiske information, der aktiveres ved læsning af enkeltord på det aktuelle sprog, og som har sat sig spor i konnektionerne. På forståelsesniveau influerer på samme måde generelle sproglige kompetencer og læserens forventninger til teksten. Til de sproglige kompetencer hører ordforråd, syntaks og idiomatik, som hos L1-læseren normalt er langt bedre udviklet end hos L2-læseren. Læserens forventninger til teksten er skabt af erfaringer med både tekst og verden i øvrigt. De omfatter ikke blot områder som kendskab til tekstgenre og læserens kulturelle udgangspunkt, men justeres også løbende i henhold til kontekst. Sproglige kompetencer og forventninger til teksten skal altså ikke ses som en ramme for, hvor meget man kan forstå af teksten, men derimod som elementer, der i samspil med ordgenkendelsen skaber mening hos læseren. Og mening er ikke blot et produkt af processen, men indgår i et samspil med de anførte videnselementer og virker tilbage på dem, således at der sker en kontinuerlig udvikling af netværket.

Den konnektionistiske læseprocesmodel er altså også i relation til L2-læsning en god forklaringsmodel, som desuden teoretisk underbygger, at 'øvelse gør mester'. Eftersom den enkeltes præstation er afhængig af, hvor meget erfaring, der er lagret i konnektionernes vægte, er konsekvensen, at ordgenkendelse på L2 – ud over at afhænge af læserens individuelle kognitive forudsætninger og ligheder mellem L1's og L2's skriftsystemer – først og fremmest hviler på omfanget af eksponering for tekstmateriale og af de relevante ords frekvens på det pågældende sprog (Muljani et al. 1998). En betingelse for, at lørneren erhverver sig erfaring og dermed bliver en bedre læser, er imidlertid, at læsning finder sted, og dette er ikke tilfældet, hvis teksten er så svær, at læseprocessen bryder sammen, fordi meningen udebliver – enten fordi den mening, læseren danner, i så høj grad er baseret på gætteri, at den ingen forbindelse har til teksten – eller fordi den studerende må bruge så meget kognitiv kapacitet på at dechifrere tekstsegmenter, at korttidshukommelsen overbebyrdes, hvilket normalt også vil medføre, at motivationen for overhovedet at

beskæftige sig med teksten forsvinder. I begge tilfælde står lørneren tilbage uden at have erhvervet nye, brugbare læseerfaringer.

Konklusionen på dette burde være, at litteraturens meget udbredte brug af begreber som uhensigtsmæssige læsestrategier' og '*inadequate readers'* (se fx Hauge 1993; Haastrup 1993) helt afskaffes til fordel for 'uhensigtsmæssige undervisningsmaterialer' og *'inadequate texts*'.

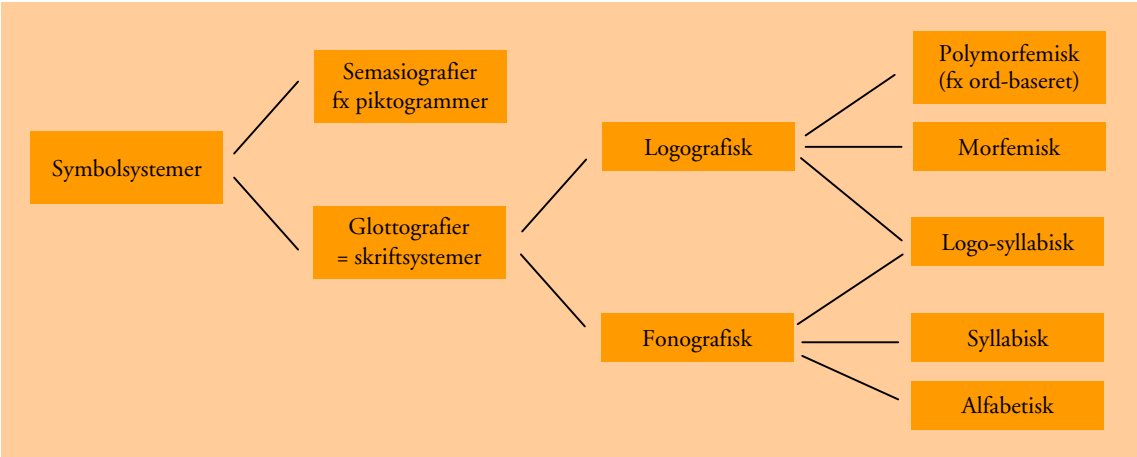
Et mål med denne afhandling er netop at give nogle bud på, hvad der så kendetegner hensigtsmæssige undervisningsmaterialer og '*adequate texts*' for undervisning i læsning på arabisk som fremmedsprog. Dette vil blive forsøgt i kapitel 7 – på baggrund af de følgende kapitler: Efter en beskrivelse af, hvad der karakteriserer de semitiske skriftsystemer, gennemgås ovenstående generelle aspekter af læsning, ordgenkendelse og læsning på fremmedsprog og af fremmed skrift specifikt i relation til læsning på arabisk som L2. Som markeret i figur 2.5 behandles i kapitel 4 den primære genkendelse. I mangel på direkte relateret litteratur sammenfattes her studier fra forskellige forskningsretninger for at anskueliggøre, hvilke vanskeligheder der mere nøjagtigt karakteriserer etableringen af bogstavgenkendelse, når det arabiske alfabet indlæres. I kapitel 5 behandles den sekundære genkendelse, hvor den arabiske ortografis indflydelse på ordgenkendelsesprocessen analyseres. Også her er den eksisterende litteratur særdeles beskeden, hvorfor det er nødvendigt at inddrage forskningspublikationer vedrørende læsning på hebraisk som L1 for på det grundlag at tegne et billede af, hvad der mere præcist karakteriserer ordgenkendelse på semitiske sprog. Sammenligneligheden mellem hebraisk og arabisk gennemgås i kapitel 3.

3. De semitiske skriftsystemer

I dette kapitel gives en kort redegørelse for skriftsystemernes oprindelse og sammenhæng for at placere de latinske og semitiske varianter i en bredere sammenhæng. Derefter gennemgås det arabiske skriftsystems særlige karakteristika, og diglossi-fænomenet i arabisk beskrives. Endvidere gives en analyse af grundlaget for den generalisering af de moderne semitiske skriftsystemer, som dele af de følgende kapitler bygger på. Hvor der ikke er angivet andre kilder, er oplysningerne baseret på Wright (1896), Haywood (2000), Glinert (1991), (1991) אלון og (1997) דהאן. Endelig diskuteres hypotesen om de semitiske ortografiers tre-konsonantiske, rodbaserede morfologiske system, som er relevant for den senere analyse af skriftsystemets mentale repræsentation i forbindelse med læsning.

3.1 Lidt om skriftsystemer generelt

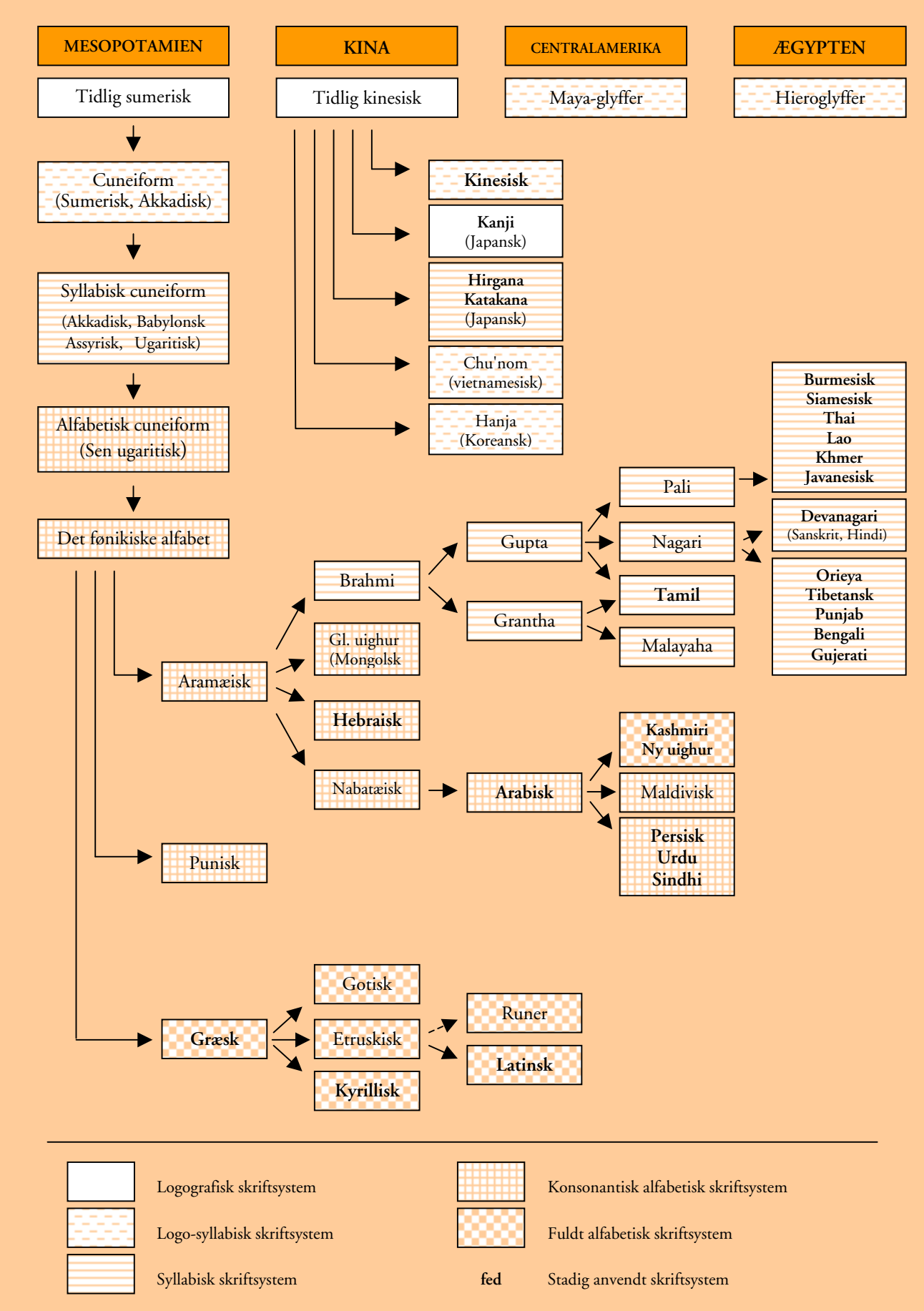
Skriftsystemer gengiver det talte sprog, og dermed adskiller de sig fra andre grafiske symbolsystemer som fx piktogrammer. Det der kendetegnede de første rigtige skriftsystemer var således, at de udnyttede talesprogets syntaks, så rækkefølgen af ordene svarede til sætningsopbygningen i talesproget.



Figur 3.1 Symbolsystemer (frit efter Elbro 2001)

Opfindelsen af et skriftsystem synes at være en uhyre sjælden foreteelse. Faktisk hævder nogle palæografer (fx Gelb 1952; Powell 1981), at alle kendte skriftsprog kan ledes tilbage til det sumeriske skriftsystem, og at opfindelsen derfor kun er gjort én gang. Hypotesen må imidlertid betragtes som utilstrækkeligt underbygget, og der er generelt bred enighed om, at sumerisk og ægyptisk hieroglyfskrift (begge fra ca. 3.000 fvt.), kinesisk (fra ca. 2.000 fvt.) og maya (fra ca. 300 fvt.) er opfundet uafhængigt af hinanden (Jensen 1970; Sampson 1985).

Sumerisk er dog ophav til langt hovedparten af verdens skriftsystemer. Til at begynde med var skriften logografisk, og de enkelte tegn repræsenterede hele ord. Problemet med logografisk skrift er, at den kun kan bruges til at beskrive ting, som både forfatter og læser kender til. Man kan eksempelvis ikke umiddelbart skrive navne på ting og personer, som



Figur 3.2 Skriftsystemernes udvikling. (Baseret på Jensen 1970; Sampson 1985; Coulmas 1996; Ager 2001). Figuren er ikke udtømmende.

læseren ikke kender. Sumererne kom over problemerne ved at supplere den logografiske skrift med fonografiske elementer, idet de begyndte at anvende tegn som lydsymboler. For eksempel havde de et tegn for ordet 'orakel', som hed *me*, og det tegn begyndte de også at anvende for flertalsendelsen med samme eller næsten samme lyd, nemlig *–me* og *–mesh* (Sampson 1985:54). Præcis den samme udvikling er sket i de tre andre tidlige skriftsystemer. På ægyptisk findes fænomenet i selv de tidligste skrifter, hvor personnavne er skrevet i en 'navnering', der indicerer, at tegnene skal fortolkes som lydværdier (Elbro 2001:58).

Denne kombination af en dominerende logografi og en supplerende (i praksis altid syllabisk) fonografi kaldes logo-syllabisk skrift. Akkaderne tilpassede systemet i retning af en egentlig syllabisk skrift, hvor hvert tegn står for en stavelse, og de videreudviklede kileskriften, hvor indskrivningen foregik ved, at pinde med kileformede ender blev trykket ned i blødt ler. Det er ingen tilfældighed, at det var akkaderne og ikke sumererne, der opfandt den syllabiske skrift, for betingelsen for, at en syllabisk skrift fungerer hensigtsmæssigt, er, at det sprog, skriften skal repræsentere, har forholdsvis få stavelser. Det havde akkadisk, der var et semitisk sprog, i modsætning til sumerisk, der er agglutinerende og tilsyneladende ikke deler slægtskab med andre kendte sprog (Jensen 1970; Sampson 1985; Coulmas 1996).

Det var også et semitisk sprog, formentlig en sen variant af ugaritisk, der omkring 1700 fvt. var grundlaget for udviklingen af det alfabetiske system, hvor der i princippet kun er ét bogstav for hvert af talesprogets fonemer. Bogstaverne blev dannet ved akrofoni, hvilket vil sige, at man lod dem stå for den første lyd i et ord. Således hed A *alef*, som betyder okse, B hed *bet*, hvilket betyder hus, G hed *gimel*, som betyder kamel, etc. De semitiske alfabeter var dengang som nu konsonantiske, og man havde ingen tegn for vokaler. Det var først efter at grækerne omkring 800 fvt. importerede det fønikiske alfabet, at vokalerne fik bogstaver, og som det kan ses af figur 3.2 er vokaler på skrift lige ved at være et isoleret europæisk fænomen (Jensen 1970; Sampson 1985; Coulmas 1996; Elbro 2001:60).

Vores eget alfabet kan altså historisk føres tilbage til de semitiske skriftsystemer, og det arabiske og det latinske alfabet har begge det fønikiske alfabet som ophav og er altså – om end lidt langt ude – i familie.

3.2 Det arabiske skriftsystem

I det følgende præsenteres det arabiske skriftsystem, dvs. såvel selve skriften og de fonemer, den repræsenterer, som det arabiske sprogs morfologiske struktur og andre relevante aspekter af arabisk ortografi.

Skrift

Arabisk skrives fra højre mod venstre. Der anvendes et alfabet på 28 bogstaver, hvoraf de fleste tager lidt forskellig form, afhængigt af om de står alene, sammenkædes med et bogstav til venstre (initial version), sammenkædes med et bogstav til højre (final version), eller

FINAL	MEDIAL	INITIAL	
ض	ضـ	ضـ	ض DAD
ط	طـ	طـ	ط TA
ظ	ظـ	ظـ	ظ DHA
ع	عـ	عـ	ع 'AIN
غ	غـ	غـ	غ RAIN
ف	فـ	فـ	ف FA
ق	قـ	قـ	ق QAF
ك	كـ	كـ	ك KAF
ل	لـ	لـ	ل LAM
م	مـ	مـ	م MIM
ن	نـ	نـ	ن NUN
ه	هـ	هـ	ه HA
و	و	و	و WAW
ي	يـ	يـ	ي YA

FINAL	MEDIAL	INITIAL	
ا	ا	ا	ا ALIF
ب	بـ	بـ	ب BA
ت	تـ	تـ	ت TA
ث	ثـ	ثـ	ث THA
ج	جـ	جـ	ج JIM
ح	حـ	حـ	ح HA
خ	خـ	خـ	خ KHA
د	د	د	د DAL
ذ	ذ	ذ	ذ DHAL
ر	ر	ر	ر RA
ز	ز	ز	ز ZAY
س	سـ	سـ	س SIN
ش	شـ	شـ	ش SHIN
ص	صـ	صـ	ص SAD

Det arabiske alfabet

	A	I	U
KORT	َ	ِ	ُ
LANG	ا/ى	ي	و

Det arabiske vokaliseringssystem

	AN	IN	UN	-				(A)T
tegn	ً	ٍ	ٌ	ْ	ـَ	ـّ	ء	ة
betydning	kasus-markører			ingen vokal	fordobl. af kons.	fordobl. af alif	glottal klusil	feminin-markør

Øvrige arabiske diakritiske tegn og bogstaver

Figur 3.3 Arabisk skrift

sammenkædes med bogstaver til både højre og venstre (medial version). Seks af bogstaverne sammenkædes aldrig med bogstaver der står til venstre for dem, hvorfor den mediale og finale version for disse bogstaver er ens. Men da hovedparten af bogstaverne altid sammenkædes i et ord, fremstår det arabiske alfabet i alle former for typologi som kursiv (dvs. løbende, sammenhængende). Som ved brug af det latinske alfabet markeres ord med mellemrum til begge sider. Det samlede sæt af bogstaver ses i figur 3.3.

Alle alfabetets bogstaver er konsonanter. Herudover har man udviklet et 'hjælpesystem', der tilfører skriften yderligere fonologisk information – først og fremmest vokalisering. For det første har tre af bogstaverne, *alif*, *waw* og *ja* (ا , و og ي), fået en dobbeltfunktion, så de i nogle tilfælde repræsenterer en lang vokal. Hertil kommer en række diakritiske tegn, som viser korte vokaler, kasusendelser og konsonantfordobling. En oversigt over disse ses nederst i figur 3.3.[1] De korte vokaler og de øvrige diakritiske tegn består af streger og 'krøller', der placeres over eller under bogstaverne, og de har i forhold til bogstaverne et meget beskedent grafisk udtryk. I praksis udelades de helt, bortset fra i visse former for religiøse tekster og poesi og i litteratur, der henvender sig til børn og begynderlæsere. I alle arabiske lande introduceres begynderlæsere først til vokaliseret tekst. Fra 3. klasse introduceres løbende tekst uden vokaler, og i slutningen af 6. klasse er stort set intet af den tekst, børn møder i skolen, vokaliseret. Dog kan enkeltstående diakritiske tegn anvendes i alle former for tekster, hvor meningen ellers ville være tvetydig, og hvor denne tvetydighed ikke umiddelbart afklares af konteksten.

Arabisk skrift er præget af stor variation. Arabisk kalligrafi er både en kunstart og et reelt aspekt af den typografiske variation, og forskellige geografiske områder har ret forskellige skrifttraditioner. Dog følger moderne skriftmateriale som dagblade, tidsskrifter og bøger normalt en jævn, læsevenlig standard (fx fontene Akhbar og Mudir MT, som her er anvendt i figurerne), hvor de enkelte bogstavers udformning ikke varierer mere end bogstaverne normalt gør det i tekster trykt med det latinske alfabet.

Fonologi

Som det ses i figur 3.3 har man på arabisk tre vokaler ([a],[u] og [i]) i en kort og en lang version. Alle tre vokalers udtale har stor variationsbredde og udtalen er desuden meget dialektafhængig. Konsonanterne repræsenteres meget regelret med hver sit bogstav, i modsætning til hvad der er tilfældet i dansk eller engelsk retskrivning, hvor ét bogstav kan repræsentere flere fonemer (fx 'c' både [s] og [k]), samme fonem kan repræsenteres af flere bogstaver (fx [s] både af 's' og 'c'), en bogstavgruppe man repræsentere ét fornem (fx engelsk 'th' for [θ]), og en gruppe fonemer kan repræsenteres af ét bogstav (fx 'x' for /ks/). Det fonetiske slægtskab mellem forskellige bogstaver kommer i visse tilfælde til udtryk rent grafisk som illustreret i figur 3.4.

---

[1] *Hamza* (glottal klusil) og femininmarkøren *Taa marbuuta* betragtes som bogstaver, men listes normalt ikke som en del af alfabetet. Ud over de i figuren viste diakritiske tegn findes to særlige markører, der dog selv i vokaliseret tekst ofte udelades: én til markering af, at lang vokal udtales, selvom den ikke markeres med *alif* (hvilket kun er tilfældet for ganske få, men meget almindelige ord (الله, رحمن, لكن, هذه, هذا), samt en *hamza*-variation, der bruges når vokaler, der bærer *hamza*, absorberes af vokaloverførsel fra et foregående ord, hvorved også udtalen af *hamza* falder væk (*hamza al-wasl*).

س	s	ش	ʃ (sh)
ت	t	ث	θ (th)
د	d	ذ	ð (blødt d)

Figur 3.4 Eksempler på grafisk og fonetisk slægtskab i det arabiske skriftsystem

ب	b som i bukser	ك	k som i krølle
ت	t som i tulipan	ل	l som i lomme
د	d somi dukke	م	m som i medalje
ذ	ð som i grød	ن	n som i nisse
س	s som i sild	ه	h som i hus
غ	ʁ som i rive	ي	j som i julemand
ف	f som i fantasi	ء	ʔ som i Møll'er

A. Fonemer kendt fra dansk

ث	θ som i engelsk thing	ص	$s^{ˤ}$ (dybt / mørkt s)
ش	ʃ som i engelsk shine	ض	$d^{ˤ}$ (dybt / mørkt d)
ج	ʒ/dʒ som fransk/engelsk j	ط	$t^{ˤ}$ (dybt / mørkt t)
خ	x som i tysk ach	ظ	$ð^{ˤ}$ (dybt / mørkt ð)
ر	r som spansk r (let rullet)	ق	q (uvulær klusil)
ز	z som fransk z (stemt)	ح	ħ (ustemt faryngal spirant)
و	w som engelsk well	ع	ʕ (stemt faryngal spirant)

B. Fonemer kendt fra nære L2 C. Fremmede fonemer

Figur 3.5 Arabiske konsonantfonemer med kendt og mindre kendt udtale (jvf. Det Internationale Fonetiske Alfabet)

Som det ses i figur 3.5, kender danske lørnere 14 af konsonanterne fra dansk, mens yderligere syv optræder i nære fremmedsprog som engelsk, tysk og fransk. Fonemet [ʔ], en ustemt glottal klusil, kan uden problemer erstattes med det danske stød (udtalen af [l] i 'Møller', modsat den stødløse udtale af [l] i 'møller'). Mens den rette beskrivelse af stødets rolle i dansk er omdiskuteret, er det arabiske *hamza* tydeligvis et konsonantfonem, og det er derfor helt konsekvent, at det har en repræsentant i skriften. De øvrige fonemer er mere fremmedartede. Fire af dem ligner meget fonemer, som kendes i forvejen, men de udtales lidt anderledes. De kaldes traditionelt emfatiske, hvad der siger meget lidt om deres udtale. De udtales velariseret (med hævet bagtunge), og det hulrum der derved fremkommer foran tungens hævning, har indflydelse på de omgivende vokaler, der får en mørk klang (Et dansk eksempel på modsætningen mellem lys-mørk vokal er er den mellem de to 'a' i 'and' og 'anderledes'). De sidste tre konsonanter udtales alle langt tilbage. [q] er en ustemt, uvulær klusil (der dannes lukke mellem det bageste af tungen og drøvelen), som den, der skrives 'q' i grønlandsk. [ħ] er en ustemt spirant (frikativ, hæmmelyd), som er faryngal (den udtales ved en indsnævring af svælget). Fonemet [ʕ], repræsenteret ved bogstavet '*ʕajn*', er en stemt faryngal spirant.

Morfologisk struktur

Arabisk morfologi beskrives af både arabiske og vestlige lingvister som et tre-konsonantisk, rodbaseret system, der er unikt for de semitiske sprog. Hovedparten af alle arabiske ord er bygget op af to morfologiske enheder: en rod bestående af tre konsonanter, der bærer en grundbetydning (normalt i form af et verbum i perfektum, 3. person ental, mask. sing.), samt et begrænset antal sæt af præ-, in- og suffixer, kaldet mønstre.

Som det ses i figur 3.6, tilfører disse præ- in- og suffixer for verbernes vedkommende information om fx kausalt eller refleksivt aspekt. En rod indgår i mange tilfælde kun i ord fra få af de ti grupper, og systemet er langtfra konsekvent. I nogle tilfælde er det beskrevne forhold mellem grupperne forvansket eller slet ikke til stede, og i andre tilfælde kan det være vanskeligt at spore et fælles semantisk indhold i ord med samme rod. Sådanne uregelmæssigheder kan naturligvis skyldes sproghistoriske forandringer, og systemet er ikke mere udvandet, end at det ofte vil være muligt at anvende det til at ræsonnere sig frem til meningen med et ukendt ord på baggrund af ordets form og rodens betydningsindhold.

Inden for de øvrige ordklasser findes et lignende system. Substantiver kan enten være uafledte eller afledt af verber i en af de ti anførte former, og adjektiver kan være afledt af enten verber eller uafledte substantiver. I figur 3.7 ses de mest almindelige verbalafledte nomenformer. Disse har alle semantisk relation til verbalform I. Derudover har المصدر (al-masˠdar), اسم الفاعل (ism al-faaʕil) og اسم المفعول (ism al-mafʕuul) også andre former, der relaterer til de øvrige verbalgrupper. En oversigt over alle formerne ses i figur 3.8. Der vises i figur 3.8 ingen eksempler, men princippet er det samme som i figur 3.7, blot relateres der til det semantiske indhold af verbet i den enkelte de ti verbalgrupper, der ses i figur 3.6.

Der findes visse afvigelser fra de viste nominale former, især i form af vokalvariationer, og som nævnt findes også nominer, der ikke er afledt af verbalsystemet. Disse er oftest ord for dyr, madvarer, kropsdele, planter og andre naturfænomener. I deres grundform består også disse tit af tre konsonanter. Adjektiver afledt af de originale substantiver dannes med et suffix (-ي).

	form		forklaring	translittereret eksempel
I	فَعَلَ	faʕala	grundform: 'gøre'	
II	فَعَّلَ	faʕʕala	transitiv, ofte kausativ af form I: 'få nogen til at gøre'	darrasa: undervise, rod: drs form I: darasa: studere
III	فَاعَلَ	faaʕala	transitiv for person-objekt: 'gøre ved nogen'	raasala: korrespondere, rod: rsl form I: rasala: sende
IV	أَفْعَلَ	afʕala	transitiv, ofte kausativ af form I: 'få nogen til at gøre'	aðʷhaka: få til at le, rod: ðʷhk form I: ðʷahika: le
V	تَفَعَّلَ	tafaʕʕala	refleksiv af form II: 'gøre ved sig selv'	tatʷawwara: udvikle sig, rod: tʷwr form II: tʷawwara: udvikle
VI	تَفَاعَلَ	tafaaʕala	reciprok af andre former: 'gøre ved/mellem hinanden'	tadaaxala: overlappe, rod: dxl form III: daaxala: sætte/putte ind
VII	اِنْفَعَلَ	infaʕala	passiv af form I: ' blive gjort ved'	inqa tʷaʕa: blive afskåret, rod: qtʷʕ form I: qatʷaʕa: skære/klippe
VIII	اِفْتَعَلَ	iftaʕala	refleksiv af form I: 'gøre ved/for sig selv'	iʒtamaʕa: mødes, rod: ʒmʕ form I: ʒamaʕa: samle
IX	اِفْعَلَّ	ifʕalla	særlig form til at udtrykke farver og fysiske defekter	-
X	اِسْتَفْعَلَ	istafʕala	'søge/ønske/få – form' af andre former refleksiv af form IV: 'gøre ved sig selv'	istaxdama: bruge, rod: xdm form I: xadama: servicere istamtiʕa: nyde, rod: mtʕ form IV: amtʕ: give nydelse

Figur 3.6 Verbalformer. De forskellige former dannet på baggrund af roden فعل anvendes i arabisk terminologi som grammatiske betegnelser for disse former. I den translittererede version af 'form' er rodbogstaverne grå og mønstrene markeret med fed. I grundformen kan vokalen på medialen variere.

grammatisk betegnelse		form		forklaring	translittereret eksempel
المصدر	al-masʷdar	فَعْل	faʕl	selve handlingen	dars: undervisningstime rod: drs, darasa: studere
اسم الفاعل	ism al-faaʕil	فاعِل	faaʕil	den der handler	haakim: dommer, rod: hkm hakama: bedømme, dømme
اسم المفعول	ism al-mafʕuul	مَفْعول	mafʕuul	det der behandles eller skabes vha. handlingen	maftuuh: åben, rod fth fataha: åbne
اسم المكان	ism al-makaan	مَفْعَل	mafʕal	sted for handlingen	madxal: indgang, rod: dxl daxala: gå ind
اسم الآلة	ism al-aala	مِفْعَال	mifʕaal	instrument for handlingen	miftaah: nøgle, rod: fth fataha: åbne

Figur 3.7 Vigtigste verbalafledte nomenformer

	المصدر [al-masˤdar]		اسم الفاعل [ism al- faaʕil]		اسم المفعول [ism al- mafʕuul]	
I	فَعْل	faʕl	فاعِل	faaʕil	مَفْعول	mafʕuul
II	تَفْعيل	tafʕiil	مُفَعِّل	mufaʕʕil	مُفَعَّل	mufaʕʕal
III	مُفاعَلة	mufaaʕala	مُفاعِل	mufaaʕil	مُفاعَل	mufaaʕal
IV	إفْعال	ifʕaal	مُفْعِل	mufʕil	مُفْعَل	mufʕal
V	تَفَعُّل	tafaʕʕul	مُتَفَعِّل	mutafaʕʕil	مُتَفَعَّل	mutafaʕʕal
VI	تَفاعُل	tafaaʕul	مُتَفاعِل	mutafaaʕil	مُتَفاعَل	mutafaaʕal
VII	اِنْفِعال	infiʕaal	مُنْفَعِل	munfaʕil	مُنْفَعَل	munfaʕal
VIII	اِفْتِعال	iftiʕaal	مُفْتَعِل	muftaʕil	مُفْتَعَل	muftaʕal
IX	اِفْعِلال	ifʕilaal	مُفْعَلّ	mufʕall	مُفْعَلّ	mufʕall
X	اِسْتِفْعال	istifʕaal	مُسْتَفْعِل	mustafʕil	مُسْتَفْعَل	mustafʕal

Figur 3.8 Nomenformer med relation til verbalsystemet. Grundformen for المصدر [al-masˤdar] varierer og antager i praksis mange andre former end denne, fx فَعَل, فُعُول, فَعَالة , فيعال eller

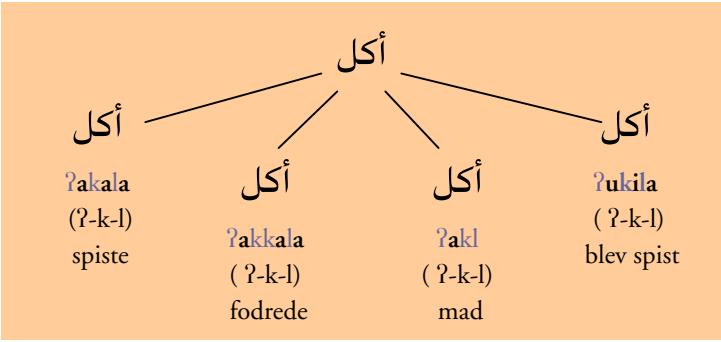
Som det fremgår, er orddannelse på arabisk altså afgrænset af forholdsvis faste rammer for, hvordan et ord kan se ud, og selvom ovenstående gennemgang ikke er udtømmende – der findes som ovenfor beskrevet yderligere ordformer og desuden et begrænset antal af låneord fra europæiske sprog, som helt bryder strukturen – er resultatet, at de samme mønstre går igen i et meget stort antal ord, hvor det alene er variation i konsonantroden, der udgør den fonologiske og semantiske forskel. Dette er anskueliggjort i figur 3.9, hvor systemet med udgangspunkt i tre forskellige rødder er eksemplificeret på tværs af skemaerne i figur 3.6, 3.7 og 3.8. Da mønstrene ikke kun hæfter sig foran eller efter roden, som det også ses på mange andre sprog, men også placerer sig imellem rodbogstaverne, betegnes den arabiske morfologiske struktur ofte som ikke-lineær. Og det ikke-lineære element præger også

الجِذْر rod	فَعَلَ [faʕala] verbum i grundform	اسم الفاعِل [faaʕil] 'den der handler'	اسم المكان: مَفْعَل [mafʕal] 'sted for handlingen'
ك ت ب [k-t-b]	كَتَبَ [kataba] (han skrev)	كاتِب [kaatib] (forfatter)	مَكْتَب [maktab] (kontor)
ش ه د [ʃ-h-d]	شَهِدَ [ʃahida] (han overværede)	شاهِد [ʃaahid] (vidne/tilskuer)	مَشْهَد [maʃhad] (udsigt)
ط ب خ [tˤ-b-x]	طَبَخَ [tˤabaxa] (han lavede mad)	طابِخ [tˤaabix] (kok)	مَطْبَخ [matˤbax] (køkken)

Figur 3.9 Eksempler på orddannelse på arabisk med tre rødder og tre mønstre

bøjningssystemerne inden for de forskellige ordklasser. Således adskiller de verbale bøjningsformer sig ved, at der opereres med både præ- og suffixer. Lidt forenklet kan man sige, at præfixer i imperfektum bruges til at udtrykke køns- og personbøjning, mens suffixer bruges til at udtrykke tal (singularis, dualis og pluralis). I perfektum bruges kun suffixer til at udtrykke bøjning, til imperativ bruges et fast præfix i alle former (hvor den korte vokal dog varierer fra ord til ord), mens person, køn og tal udtrykkes vha. suffixer. Futurum udtrykkes med imperfektum og et fast præfix. Også substantivernes pluralisform dannes i nogle tilfælde efter en ikke-lineær struktur, idet der anvendes både præ-, in- og suffixer. Pluralisformer, der tager præ- og infixer har ikke lige så faste og regelmæssige strukturer som verbalbøjningerne.

De personlige pronomener er i akkusativ og genitiv er klitiske og bliver altså – undtagen i nominativ – vedhæng på hhv. verber og substantiver eller præpositioner frem for at stå som selvstændige ord (fx احبك [uħibbu**ki**] 'jeg elsker dig', كتابك [kitaabu**ki**] 'din bog', منك [min**ki**] 'fra dig'. Hertil kommer en række præpositioner, den bestemte artikel og andre partikler, som også klitiserer og slet ikke forekommer som selvstændige ord.



Figur 3.10 Eksempel på homografi på arabisk.

Alt i alt synes såvel visuel som auditiv afkodning af arabiske ord i mange tilfælde at medføre en ikke altid lineær opsplitning i forskellige former for morfemer, hvor roden kan betragtes som nøglen til at udskille hvilke former for morfemer, der er i spil på både ordklasse-, bøjnings- og semantisk niveau. Da en del af variationen mellem de forskellige mønstre findes i de korte vokaler og konsonantfordobling, der som nævnt ikke markeres i moderne skriftligt materiale, kan den i grafisk form være usynlig. Derfor er arabisk i skriftlig form præget af et meget stort antal homografer. Et eksempel på dette ses i figur 3.10, hvor de fire ord, som det ses tager sig fuldstændig ens ud, selvom de både udtales forskelligt og betyder noget forskelligt.

3.3 Diglossi

De markante forskelle mellem arabisk skrift- og talesprog vedrører som et sociolingvistisk aspekt ikke skriftsystemet, men introduceres alligevel her, idet det ligesom de grafiske og fonologiske dimensioner i praksis har stor indflydelse på de psykolingvistiske processer, der knytter sig til læsning.

Den klassiske definition på diglossi blev fremsat af Ferguson (1959) som den situation, hvor to eller flere varianter af et sprog bruges under forskellige betingelser. I Fergusons terminologi består diglossien i en 'højsprogsvariant' (H) og en 'lavsprogsvariant' (L), hvor H er én standardiseret sprogvariant, der anvendes ved skriftlig fremstilling og ved talesprog i formelle sammenhænge, mens L er en mængde af lokale dialekter – med en i forhold til H simplificeret grammatisk struktur – der bruges ved uformel sprogbrug (Ferguson 1959:325-328). I praksis er der dog gerne tale om en mere kompleks sproglig variation. På arabisk karakteriseres dette ofte lidt simplificeret som tre varianter: Den egentlige højsprogsvariant – klassisk arabisk – bruges kun i religiøse og stærkt formelle sammenhænge samt i klassisk litteratur. Til den stadig formelle, men moderne orienterede formidling, der finder sted i medierne, den offentlige administration, den politiske diskurs og i private såvel som officielle breve, bruges en variant, der især i sit ordforråd adskiller sig fra klassisk arabisk, og som i vestlig terminologi betegnes moderne standardarabisk (MSA) (Ennaji 2001:50). De komplekse strukturer fra det klassiske sprog opretholdes i princippet, men i skriftlig form udelades de diakritiske vokaliseringstegn, og også i mundtlig form simplificerses MSA gerne ved at man taler med reduceret eller helt uden kasusvokalisering (se figur 3.11). Dette hænger naturligt sammen med, at kasusendelser appliceres i relation til et forholdsvist komplekst syntaktisk system, som gør at kun et fåtal af selv højtuddannede arabere er i stand til umiddelbart at vokalisere korrekt i skriftlig såvel som i mundtlig produktion (Parkinson 1994). Da brugen af klassisk arabisk er reduceret til især religiøse sammenhænge og ikke anvendes i interaktive kommunikationssammenhænge, bruges diglossi-begrebet i dag oftest sådan, at der ses bort fra den klassiske variant, så højsprogsvarianten الفصحة [al-fusʸha] er MSA, mens lavsprogsvarianten اللهجة [al-lahʒa] er en fællesbetegnelse for de lokale dialekter.

	Ubestemt form		Bestemt form	
Nominativ	ـٌ	[-un]	ـُ	[-u]
Akkusativ	ـً	[-an]	ـَ	[-a]
Genitiv	ـٍ	[-in]	ـِ	[-i]

Figur 3.11 Arabiske kasusendelser. Påføres nominer i vokaliseret skrift og meget formelt talesprog.

Nogle har defineret en fjerde sproglig variant betegnet *educated spoken Arabic* eller *formal spoken Arabic* (FSA) (El-Hassan 1978; Meiseles 1980; Mitchell 1980,1986; Ryding 1991). Der er altså tale om en formel talesprogsvariant, hvilket i den traditionelle diglossi-diskurs umiddelbart er en selvmodsigelse. FSA forener ifølge disse forfattere det prestigiøse fra H-varianten med det uhøjtidelige og strukturelt simplificerede fra L-varianten, hvorved der skabes en variant, der imødekommer kommunikative behov ved inter-dialektal kommunikation mellem veluddannede arabere i semi-formelle sammenhænge og ved andre sociale lejligheder, hvor H er for stilistisk højtravende og L er for uformel (Ibrahim 1986:121; Ryding 1991:212). FSA trækker i meget høj grad på ordforrådet fra MSA, men er samtidig karakteriseret af samme simplificerede grammatiske struktur og fonologi som talesprogsvarianterne:

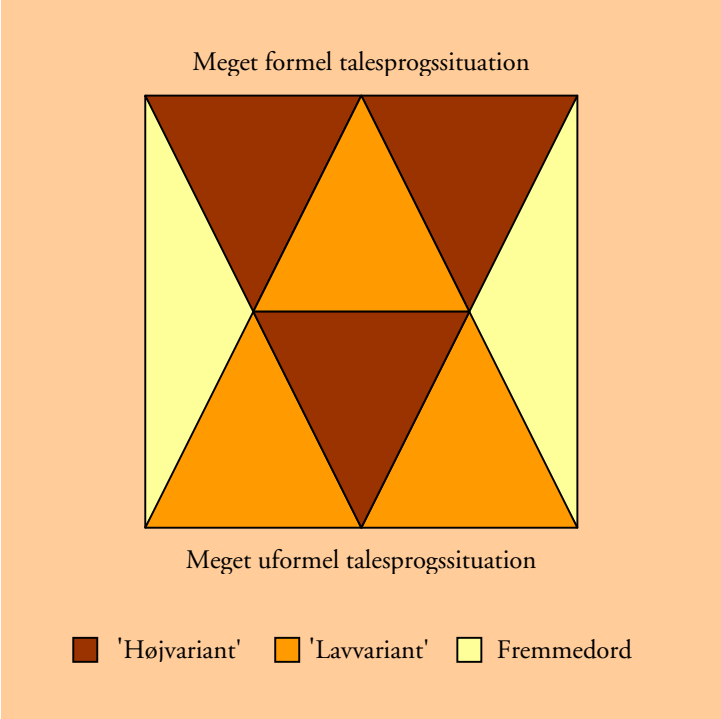
> *"FSA can be described as essentially solid at the core, surrounded by fuzzy areas of fluctuating language behavior [...], presenting a set of language features that vary on a regional and/or stylistic basis..."* (Ryding 1991:214)

FSA anvendes ifølge El-Hassan (1978) og Ryding (1991:214) i Egypten og det levantiske område, og ifølge beskrivelser af Ennaji (2001:60) synes en lignende mellemvariant at have etableret sig i Marokko:

> *"[une] variété de l'arabe qui commence à envahir le terrain linguistique au Maroc est l'arabe marocain moderne [...] C'est une variété pan-marocaine formelle, intermédiaire entre l'arabe dialectal et l'arabe standard. L'arabe marocain moderne est réservé au domaine formel ou technique car l'utilisation de l'arabe dialectal dans ce domaine est parfois impossible puisqu'il manque de jargon scientifique ou formel nécessaire. L'arabe dialectal est en fait la variété de l'échange quotidien [...] [et il] tire son prestige du fait qu'il est associé aux locuteurs d'un niveau scolaire assez élevé."* (Ennaji 2001:60)

Ligesom for FSA er det karakteristisk for denne moderne marokkanske sprogvariant, at samtidig med at der gøres brug af de simplere strukturer fra talesprogsvarianten, indgår der på det leksikalske niveau mange elementer fra standardarabisk. Desuden indgår der ikke låneord fra fransk, hvilket ellers er meget udbredt i det uformelle marokkanske talesprog (Ennaji 2001:60-63)

Imidlertid kan sprogvariationen på arabisk reelt ikke kategoriseres så skarpt, men kan bedre beskrives som et kontinuum fra det såkaldt klassiske sprog, anvendt i religiøse tekster og traditionelle former for poesi, over MSA, der er det moderne skriftsprog anvendt i aviser og nutidig litteratur og FSA, der som nævnt er en formel talesprogsvariant, til den helt



Figur 3.12 Model af talesprogsvariation på arabisk. Frit efter Badawi (بدوي ١٩٧٣).

uformelle dialekt, der anvendes i familien og mellem lavt- eller uuddannede personer – og omtrent hele dette spektrum er i spil i den veluddannede arabers sociolingvistiske adfærd. Dette spektrum er illustreret i figur 3.12, hvoraf det fremgår, at sprogbrugeren – frem for at skifte mellem forskellige kategorier – kan placere sig hvor som helst i spektret for at tilpasse sig de sociale kontekster, kommunikationen foregår i.

For at illustrere, hvad det er for sproglige variationer, der karakteriserer dette spektrum, præsenteres nedenfor nogle fremtrædende forskelle mellem formelt og uformelt sprog. Oversigten bygger på Saiegh-Haddad (2003:433) og Ryding (1991:216):

Fonetiske forskelle: De fleste arabiske talesprogsvarianter har et let simplificeret sæt af fonemer i forhold til MSA. For eksempel er den uvulære klusil [q] (ق) i ægyptisk dialekt sjælden og erstattes oftest af glottal klusil [ʔ], og på syrisk, libanesisk og palæstinensisk dialekt erstattes den upustede, dentale spirant [ð] (ذ) af den stemte, aveloære hvislelyd [z], ligesom den ustemte, dentale spirant [θ] (ث) i næsten alle talesprogsvarianter bortfalder til fordel for klusiven [t]. Hertil kommer store forskelle i fonotaktiske begrænsninger mellem MSA og de uformelle talesprogsvarianter. På MSA er initiale konsonantgrupper ikke mulige, men i de fleste talesprogsvarianter er sådanne grupper meget almindelige (fx. تُراب [turaab] bliver til تْراب [traab]), og især i de vestlige dialekter som marokkansk, algerisk og tunesisk, hvor korte vokaler falder væk, forekommer fænomenet ofte. Omvendt er finale konsonantgrupper, som er blandt de mest almindelige i MSA-ordstrukturer, meget sjældne i de centrale dialekter (fx. سَهْل [sahl] bliver til سَهَل [sahal]).

Leksikalske forskelle: En del af leksikon er i princippet fælles for MSA og talesprogs-varianterne, når der ses bort fra de ovennævnte fonetiske afvigelser. Mange, især højfrekvente, verber har dog særlige dialektale former, og nogle præpositioner, pronominer, adverbialer, negationer og andre partikler erstattes af andre og/eller bruges anderledes og gerne mindre grammatisk restriktivt end på MSA. Hvad angår substantiver, har nogle talesprogsvarianter et stort antal låneord fra kolonimagternes sprog, dvs. engelsk og fransk. Disse bruges også om begreber, der i MSA har en betegnelse baseret på arabisk morfologi (fx bruges أوتوبيس [autobiis] i stedet for حافلة [ħaafila], تلفون [telefoon] i stedet for هاتف [haatif] og i Maghreb-landene كوسينة [kusiina] (fr. 'cuisine') stedet for مطبخ [matbax]).

Syntaktiske forskelle: Bøjninger i køn, tal og tid er i alle uformelle talesprogsvarianter stærkt forenklede. Eksempler er udeladelse af kasusendelser, ændring af nominativ-endelser til den ikke-nominative form i dualis og pluralis, sammenlægning af dualis og pluralis og af den feminine og maskuline pluralis ved verber og pronominer, reducering af bøjningsendelsen i 2. person femininum, singularis og 2. og 3. person pluralis i imperfektum og reduktion af det relative pronomen til én form. Endvidere benyttes i nogle talesprogsvarianter præpositioner i forbindelse med pronomen-suffixer ved brug af indirekte objekt og/eller possessive pronominer, hvorved suffixerne skilles fra selve målordene (fx på marokkansk talesprog الكتاب دِالي [al-kitaab diaal**ii**] frem for كتابي [kitaab**ii**] ('min bog')), og ordstillingen er i talesprog SVO til forskel fra den traditionelle VSO-ordstilling på MSA .

Der er altså tale om ret markante forskelle mellem det arabiske skriftsprog og de uformelle talesprogsvarianter. Sådanne forskelle findes på alle sprog, og også mange danske dialekter afviger både leksikalsk og syntaktisk fra rigsdansk. Den primære forskel på den danske og den arabiske sprogsituation er, at skriftsproget på dansk i høj grad tilpasses den talesproglige

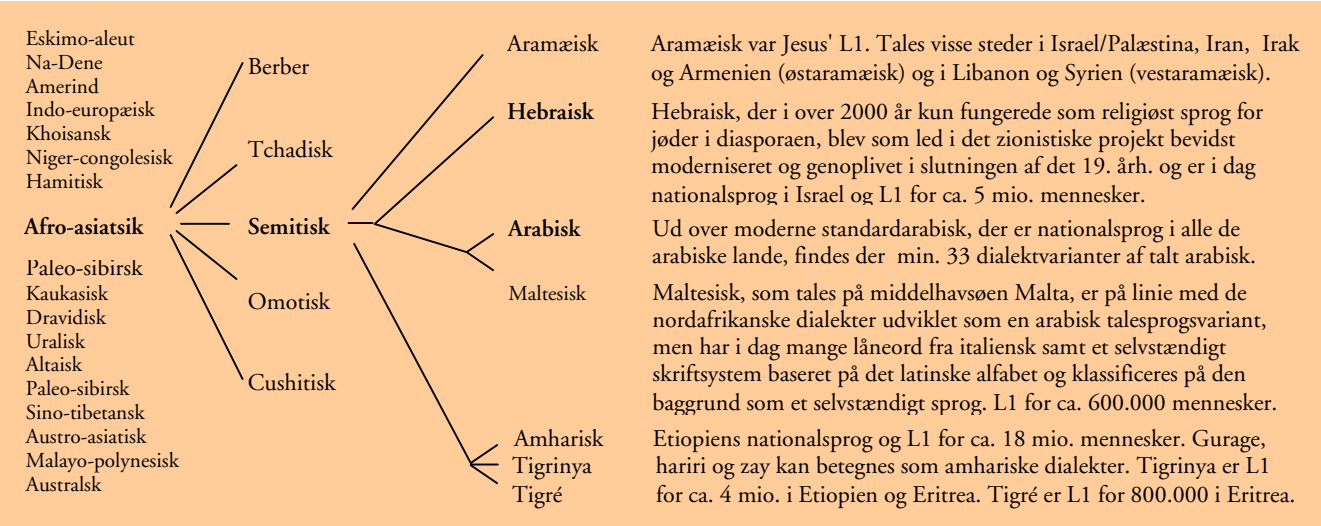
udvikling, og at de danske dialekter er uddøende, så talesprogsnormen i dag er stort set den samme i hele sprogområdet. At situationen er en anden i den arabiske verden, skyldes at religiøse, politiske og sociale forhold har skabt en inerti i de processer, der 'normalt' ligger til grund for sproghistoriske forandringer: Det arabiske sprog har traditionelt været – og er stadig – tæt forbundet med islam, idet Koranen opfattes som åbenbaret på arabisk. Som Maamouri skriver:

> *"Arabs viewed the Arabic language as sacred and insisted on safeguarding its original purity and unity from any variation, which could lead to dangerous religious deviation. The lack of receptive attitude towards language change barred all attempts at reforming the language in the direction of those who spoke it. It also led to the prevailing traditional ideology which has validated and preserved until now the cultural and historical uniqueness of Arabic by manifesting a highly pronounced sensitivity for purism and a low level of tolerance (and even some disdain) towards mistakes and error of common language use."* (Maamouri 1998:21)

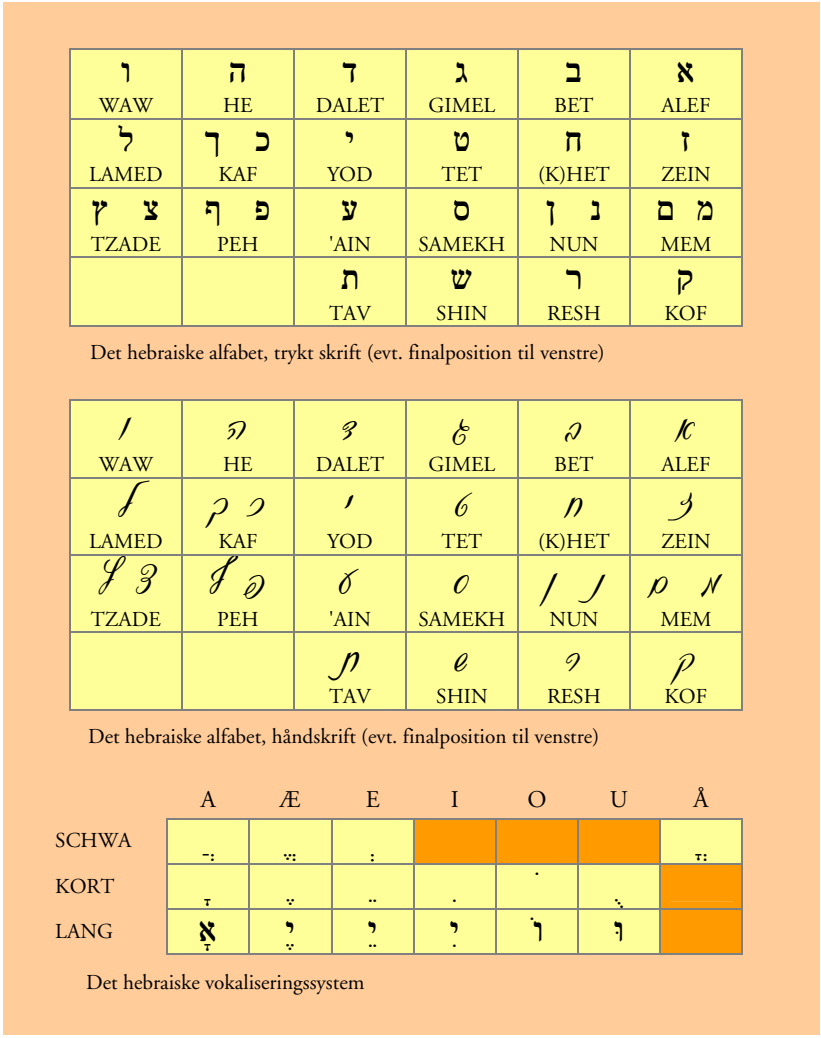
Samtidig har det arabiske sprog efter afkoloniseringen af de arabiske lande været et værdiladet element i den pan-arabiske ideologi, hvor det fælles sprog har været et centralt argument for den arabiske enhedstanke. Men sideløbende med, at sproget formelt er blevet vedligeholdt i sin traditionelle form og fremhævet som et forenende element for de arabiske nationer, har en lav alfabetiseringsgrad i alle de arabiske lande været med til at forstærke de lokale sprogforandringer, der har skabt stor sproglig afstand landene imellem (Maamouri 1998).

3.4 Grundlaget for sammenligning mellem arabisk og hebraisk

Arabisk og hebraisk er fra et dansk synspunkt marginale sprog, men internationalt har de en central position som de dominerende sprog i Mellemøsten og som sprog for hver sin verdensreligion. De tilhører begge den semitiske gren af den afroasiatiske sprogfamilie, hvis primære karakteristikum er det tre-konsonantiske rodsystem. Flere af de nulevende semitiske sprog – herunder arabisk – har mange talesprogsvarianter, og der er ikke konsensus om, hvilke af dem, der bør klassificeres som selvstændige sprog. I figur 3.13 gives en overordnet, 'grovsorteret' oversigt. Mens aramæisk skrift er uddød, maltesisk har adopteret det europæiske skriftsystem og der til de etiopiske sprog som amharisk, tigrinya og tigré er udviklet den syllabiske skrift *geez*, er de originale semitiske skriftsystemers særegne konsonant-dominerede karakteristika stort set bevaret i den skriftlige repræsentation af arabisk og hebraisk. Dette gør det oplagt at drage nogle paralleller mellem de to sprog i deres mundtlige såvel som deres skriftlige udtryk – ikke mindst fordi generaliseringer fra det ene sprog til det andet i nogen grad kan kompensere for den meget begrænsede viden vi har om kognitive processer, der involverer henholdsvis arabisk og hebraisk. Med henblik på at udnytte denne mulighed for at bruge forskningsresultater om hebraisk til at beskrive læseprocesser på arabisk, gives i det følgende en overordnet kontrastiv analyse af de to sprog, hvilket skal anskueliggøre i hvor høj grad generaliseringerne kan bære.



Figur 3.13 De i dag talte semitiske sprog. (Baseret på Pagh 1998; Spolsky, 1999; Boeree, 2000 og Ethnologue 2001)



ו WAW	ה HE	ד DALET	ג GIMEL	ב BET	א ALEF
ל LAMED	כ ך KAF	י YOD	ט TET	ח (K)HET	ז ZEIN
צ ץ TZADE	פ ף PEH	ע 'AIN	ס SAMEKH	נ ן NUN	מ ם MEM
		ת TAV	ש SHIN	ר RESH	ק KOF

Det hebraiske alfabet, trykt skrift (evt. finalposition til venstre)

WAW	HE	DALET	GIMEL	BET	ALEF
LAMED	KAF	YOD	TET	(K)HET	ZEIN
TZADE	PEH	'AIN	SAMEKH	NUN	MEM
		TAV	SHIN	RESH	KOF

Det hebraiske alfabet, håndskrift (evt. finalposition til venstre)

	A	Æ	E	I	O	U	Å
SCHWA	ְ	ֱ	ֲ				ֳ
KORT	ַ	ֶ	ֶ	ִ	ֹ	ֻ	
LANG	אָ	ֵ	ֵ	ִי	וֹ	וּ	

Det hebraiske vokaliseringssystem

Figur 3.14 Hebraisk skrift.

Skrift og fonologi

Både arabisk og hebraisk skrives fra højre mod venstre, men de to skriftsystemer tager sig ellers meget forskelligt ud. Mens arabisk skrift som beskrevet i afsnit 3.2 er kursiv, med bogstaver, der især består af buer og 'krøller' med forskelligt antal prikker over eller under, er hebraiske bogstaver som det ses i figur 3.14 nærmest bygget op omkring en kvadratisk skitse, og de sammenbindes aldrig. Hovedparten af de hebraiske bogstaver har samme form uanset placering i ordet; kun fem af dem har en særlig form i finalposition. Til gengæld afviger håndskrift så markant fra trykt skrift, at man reelt opererer med to forskellige alfabeter. De to sprog havde oprindeligt omtrent de samme fonemer, men moderne hebraisk har færre fonemer end arabisk. Imidlertid anvender skriften de oprindelige tegn, hvorfor nogle af fonemerne i dag kan er repræsenteret af flere forskellige bogstaver, hvorimod tilsvarigheden mellem fonem og grafem på arabisk er bevaret. Således bliver på moderne hebraisk fonemerne [k], [h] og [ʔ]dels skrevet med bogstaverne, כ, ה og א, dels med ק, ח og ע, der tidligere repræsenterede fonemerne [q], [ħ] og [ʕ].[2]

På hebraisk har man ligesom på arabisk et hjælpe-vokaliseringssystem, hvor *alef*, *waw* og *jod* (א, ו og י) – lige som arabisk *alif*, *waw* og *jaa* (ا, و og ي) – bruges som hjælpetegn for vokaler. På arabisk angiver de lange vokaler. På moderne hebraisk har forskellen mellem korte og lange vokaler fortonet sig, hvilket har haft indflydelse på både bogstavers og diakritiske tegns tilsvarighed i udtalen. Hvad angår uvokaliseret skrift, opretholder sproglige traditionalister den bibelske retskrivning, hvor brugen af disse bogstaver er meget begrænset, mens sproglige modernister i høj grad anvender især *waw* og *jod* som hjælpetegn for at undgå homografi og lette læsningen. Så selvom der i 1970 blev fastsat én officiel retskrivningsstandard, er der i praksis ingen fast norm – heller ikke i ordbøger og grammatikker. Hvad angår de diakritiske vokaliseringstegn, opretholdes den traditionelle, bibelske retskrivningstradition, når vokalisering forekommer. Oprindeligt har vokaltegnene repræsenteret hvert sit fonem, mens ét vokalfonem i dag vises med forskellige tegn i henhold til traditionen, der går tilbage til tidligere sprogtrin. Desuden har man i hebraisk

A. Uden vokalisering

القراءة باللغة العربية קריאה בעברית

B. Med vokalisering

القِراءَةُ بِاللُّغَةِ العَرَبِيَّةِ קְרִיאָה בְּעִבְרִית

Figur 3.15 Arabisk og hebraisk skrift uden og med vokalisering

[2] Dog er den traditionelle udtale fortsat i brug hos grupper af sefardiske jøder, der tidligere var – og i nogen grad stadig er – arabisksprogede.

diakritiske tegn, der adskiller konsonantiske fonemer. Således angiver en prik i bogstavets midte, at det drejer sig om en klusil (lukkelyd) og ikke om den tilsvarende spirant (hæmmelyd), som bogstavet betegner uden prik. Labiodentalen [v] betegnes altså med bogstavet 'ב', bilabialen [b] derimod med 'בּ'. Tilsvarende betegner 'פ' [f] men 'פּ' [p], og (כ) betegner [x], men 'כּ' [k]. Desuden forsynes *shin* 'ש' med en prik til venstre 'שׂ', når den betegner den alveolære hvislelyd [s], men med en prik til højre 'שׁ' for at betegne den palatale hvislelyd [ʃ]. I ikke-vokaliseret tekst benyttes disse diakritiske tegn ikke.

I figur 3.15 ses arabisk og hebraisk skrift i uvokaliseret og vokaliseret form. Det er vanskeligt at se skrifternes tætte historiske slægtskab: Mens arabiske bogstaver er sammenkædede, domineret af prikker og ligesom latinsk skrift har bogstaver, der strækker sig både over og under grundlinjen, er hebraiske bogstaver alle enkeltstående, og da hovedparten af bogstaverne følger den karakteristiske kvadratiske form, fremstår bogstavernes som mere grafisk homogene end de arabiske. De korte vokaler tager sig også meget forskelligt ud, men har dog det tilfælles, at de fremstår som sekundære i forhold til selve bogstaverne og dermed synes lette at se bort fra i begge skrifter.

Morfologisk struktur

Hebraisk morfologi ligner den arabiske, blot er der kun 7 verbalgrupper, og mønstrene er anderledes end på arabisk. De enkelte mønstre og deres relation til det arabiske system ses i figur 3.16.

Mens alle hebraiske verber ligesom de arabiske kan indpasses i et verbalsystem, er der på hebraisk en langt større fleksibilitet hvad angår nominer i forhold til på arabisk, hvilket skyldes, at et meget stort antal låneord er adopteret i moderne hebraisk. Et traditionelt system for nominer svarende til det arabiske findes dog, skønt mange ord falder forholdsvis langt fra de originale former. De mest anvendte former ses i figur 3.17. I lighed med hvad der er tilfældet i det arabiske system, har פָּעִילָה [pˤila] former med relation til fem af de syv verbalgrupper. Disse ses i figur 3.18. Ud over disse har man på hebraisk ligesom på arabisk uafledte substantiver og adjektiver afledt af disse. Både hvad angår form og semantisk indhold svarer de originale hebraiske substantiver mere eller mindre til de arabiske. Substantiviske adjektiver dannes ved tilføjelse af er suffix (ִי -) hvilket svarer til den tilsvarende arabiske form.

Den fleksibilitet, der præger især de arabiske nomenformer, er endnu mere udtalt på hebraisk, som i modsætning til arabisk har været igennem en bevidst moderniseringsproces og som ovenfor anført desuden har været åben for et stort antal låneord. Den angivne systematik er altså langt fra en udtømmende beskrivelse af moderne hebraisk morfologi.

Hvad angår de øvrige aspekter, der skaber en ikke-lineær struktur, svarer hebraisk imidlertid langt hen ad vejen til arabisk. Verbernes bøjningssystem er ganske vist mindre komplekst. I imperfektum findes kun fire former, nemlig de to køn i hhv. pluralis og singularis, hvor der på arabisk er 12, fordi der også personbøjes i 1. 2. og 3. person og desuden også opereres med dualis (der er kun 12, fordi der ikke skelnes mellem køn i 1. person og i dualis). I perfektum og imperativ svarer bøjningen på hebraisk næsten fuldstændig til den arabiske, bortset fra at man ikke opererer med dualis, og futurum dannes som den arabiske

	form		forklaring	svarer til på arabisk		
I	פָּעַל	paʕal	aktiv form – grundform: 'gøre'	فَعَلَ	faʕala	(I)
II	נִפְעַל	niffal	passiv form – passiv af form i: 'blive gjort'	اِنْفَعَلَ	infaʕala	(VII)
III	הִפְעִיל	hiffil	aktiv form – transitiv af form i: 'få noget/nogen til at gøre'	أَفْعَلَ	affala	(IV)
IV	הוּפְעַל	hupʕal	passiv form – passiv af form iii: 'blive gjort ved'			
V	פִּיעֵל	piʕæl	aktiv form - transitiv af form i: 'gøre ved nogen'	فَعَّلَ	faʕʕala	(II)
VI	פּוּעַל	puʕal	passiv form – passiv af form v: 'blive gjort ved'	فَاعَلَ	faaʕala	(III)
VII	הִתְפַּעֵל	hitpaʕæl	aktiv form – refleksiv eller reciprok af form v: 'gøre ved sig selv' 'gøre ved/mellem hinanden'	تفَعَّلَ تَفَاعَلَ اِسْتَفْعَلَ	tafaʕʕala tafaaʕala istaffala	(V) (VI) (X)

Figur 3.16 Hebraiske verbalformer

form		forklaring	svarer til på arabisk	
פְּעִילָה	pʕila	selve handlingen	المصدر	[al-masʕdar]
פַּעָל	paʕal	den der handler	اسم الفاعل	[ism al-faaʕil]
מַפְעֵל	maffæl	instrument for handlingen	اسم الآلة	[ism al-aala]
פעלון	piʕlon	sted eller ting der skabes af handlingen		
פָּעוּל	paʕlit	adjektiv for handlingen	اسم المفعول	[ism al-mafʕuul]

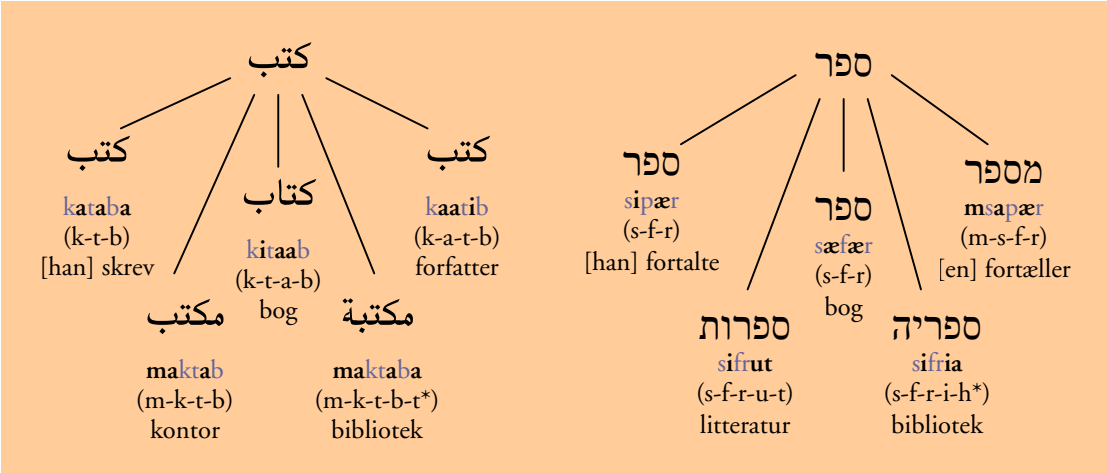
Figur 3.17 De mest konsekvent anvendte verbalafledte nomenformer. פְּעִילָה [pʕila] har former, der relaterer til andre verbalgrupper, se figur 3.18

gruppe	form		svarer til på arabisk	
I	פְּעִילָה	pʕila	فَعْل	faʕl
II	הִיפָּעְלוּת	hipaʕlut	اِنْفِعال	infiʕaal
III	הַפְעָלָה	haffalat	إِفْعال	iffaal
V	פִּיעוּל	piʕul	تَفْعيل	taffiil
VII	הִתְפַּעְלוּת	hitpaʕlut	تَفَعُّل، تَفَاعُل، اِسْتِفْعال	tafaʕʕul, tafaaʕul, istiffaal

Figur 3.18 Nomenformer med relation til verbalsystemet.

imperfektum ved hjælp af både præ- og suffixer. Herudover findes på hebraisk en infinitiv, der dannes ved hjælp af et præfix og forskellige infixer afhængigt af verbalgruppe. Mange hebraiske verber tager en objekt-markør i form af en præposition, der tilføjes ved brug af direkte objekt i ubestemt form. Dette fænomen findes ikke på arabisk. I modsætning til arabisk er substantivernes bøjning på hebraisk som hovedregel lineær, men som på arabisk klitiserer den bestemte artikel og visse præpositioner, ligesom de personlige pronominer i akkusativ og genitiv. Genitiv-formen, der altså svarer til det possessive pronomen, er dog stilistisk lidt højtravende, og i mange tilfælde erstattes den af en moderne form, hvor den som i visse arabiske talesprogsvarianter kombineres med en præposition og dermed antager et mere selvstændigt udtryk.

Generelt er den morfologiske struktur på arabisk og hebraisk således den samme: I princippet dannes ord med udgangspunkt i en tre-konsonantisk rod, der bærer på en grundbetydning, og mønstre i form af præ- in- og suffixer tilfører information om fx kausalt, refleksivt eller passivt aspekt. I figur 3.19 ses dette eksemplificeret med to helt forskellige rødder, der danner ord med samme eller lignende semantisk indhold. (Bemærk, at roden كتب [k-t-b]også eksisterer på hebraisk: קתב [k-t-b], hvor den betyder det samme, men ordene 'bog' og 'bibliotek', der findes i begge figurens ordtræer, dannes altså på arabisk og hebraisk af forskellige rødder – på arabisk refererer ordet bog til 'noget der er skrevet', mens det på hebraisk refererer til 'noget der bliver fortalt'.) Brugen af artikler, præpositioner og nogle pronominer cementerer på begge sprog den ikke-lineære struktur, og selvom moderne hebraisk på grund af den større åbenhed for låneord fra europæiske sprog og den meget bevidste sproglige moderniseringsproces, der har tilført en mere fleksibel brug af det morfologiske system, i mange tilfælde fremstår mindre konsekvent ikke-lineært end arabisk, er princippet i orddannelsen den samme på de to sprog. Den manglende vokalisering resulterer på begge sprog i et stort antal homografer. Umiddelbart synes homografi at være lidt mere udtalt på arabisk end på hebraisk, og Abu-Rabia (1998:109) angiver da også, at næsten hvert tredje ord på arabisk kan have flere forskellige betydninger, mens Shimron & Sivan (1994:10) talte andelen af homografer i de tekststykker, de anvendte til deres hebraiske læsetests, til 23%. Der findes imidlertid ingen egentlige komparative analyser, der kan dokumentere en forskel i den henseende. Der er dog den forskel mellem de to sprogs skrifsystemer, at hebraiske konsonanter, som vist, i nogle tilfælde noteres homografisk.



Figur 3.19 Nogle arabiske og hebraiske ord afledt af roden كتب (k-t-b) og roden ספר (s-p-r).
* denne arabiske t-variant og det hebraiske *het* udtales her som /a/.

Diglossi

Også på hebraisk ses diglossi-fænomenet i udtalt grad, men til forskel fra arabisk har hebraisk været igennem en bevidst, strategisk moderniseringsproces, der har haft til formål at genoplive sproget, der i ca. 2000 år var uddødt som aktivt talesprog. I dag anvendes den klassiske skriftsprogsstil næsten udelukkende i religiøse sammenhænge, men på grund af religionens centrale position i det israelske samfund og blandt mange jøder uden for Israel er den traditionelle sprogstil imidlertid stadig et centralt element i hebraisksprogede undervisningsmiljøer overalt i verden, hvorved det klassiske sprog ligesom på arabisk fortsat holdes ved lige. Mange af de sproglige forskelle mellem klassisk og moderne hebraisk sprogstil afviger ikke afgørende fra det, der ovenfor beskrives som forskellene mellem MSA og de arabiske talesprogsvarianter. Den moderne sprogstil findes også i uvokaliseret skriftligt materiale, og de stilistiske niveauer der findes i moderne hebraiske tekster, svarer dermed nogenlunde til hvad der findes på de fleste andre moderne sprog, fx i Europa. Dog opstår der konflikter på det leksikalske niveau, idet de to stilniveauer ofte støder sammen ved brug af vokaliseret tekst. Direkte misvisende bliver det, når bogstavet *bet* 'ב' i et ord som מבלבל ('forvirret') forsynes med spiranttegnet (en prik indeni) og altså indikerer læsningen [mevulbal] – en udtale, som er forsvundet, blandt andet i kraft af, at de diakritiske tegn normalt ikke anvendes. Af de ovenfor anførte grunde er selv veluddannede israelere er normalt ikke i stand til at gennemføre korrekt vokalisering af tekst (Ravid, D. 1996:1226).

I praksis er 'korrekt' vokalisering altså en sprogbrug, der på både arabisk og hebraisk opretholdes i religiøse og undervisningsmæssige sammenhænge og i de nationale medier – på trods af at denne formelt korrekte vokalisering praktiseres i henhold til nogle komplekse sprogregler, som kun sprogeksperter mestrer, og som ignoreres af almindelige sprogbrugere. På arabisk består det komplekse først og fremmest i systemet af kasusendelser, mens det på hebraisk først og fremmest består i det i forhold til talesproget langt mere udbyggede vokaliseringssystem.

Sammenfatning

Det arabiske og det hebraiske skriftsystem har som vist en række ligheder, men også nogle afgørende forskelle. Skriftsystemernes grafiske dimension er meget forskellige, men i begge skriftsystemer udgør de korte vokaler et sekundært niveau, der synes let at 'sortere fra' ved læsning.

Med hensyn til fonologi har de to sprog traditionelt haft omtrent de samme fonemer, men sprogudviklingen har medført, at der i moderne hebraisk kun er fem vokalfonemer, hvorfor samme fonem i skriften ofte repræsenteres af to forskellige bogstaver. I modsætning hertil har man på arabisk bibeholdt de traditionelle fonemer, og hvert bogstav repræsenterer dermed kun ét fonem. Samtidig har man på uvokaliseret hebraisk enkelte homografiske konsonanter, hvorfor ét bogstav kan repræsentere flere forskellige fonemer. Dette aspekt udgør muligvis en vigtig forskel i relation til læsning på semitiske sprog som L2. Det er imidlertid ikke muligt konsekvent at betegne det ene sprogs grafem-fonem-system som objektivt vanskeligere end det andet: Med hensyn til fonetisk variation er arabisk tydeligvis mere komplekst end hebraisk, men det omvendte gør sig gældende med hensyn til konsekvens i grafem-fonem-korrespondance.

Med hensyn til ortografien synes det derimod uproblematisk at antage, at de to sprog er sammenlignelige med hensyn til den morfologiske strukturs eventuelle indflydelse på kognitive processer i forbindelse med læsning. På moderne hebraisk er brugen af de vokaltegn, der oprindeligt betegnede lange vokaler, ikke er fastlagt, hvorved man i ganske mange tilfælde kan finde flere forskellige stavemåder for samme ord. Dette er udelukket på arabisk, når ses bort fra enkelte låneord, hvor brugen af lange vokaler kan udvise en vis variation. Imidlertid konstrueres ord ud fra nøjagtig de samme principper, og ordenes ikke-lineære struktur, der så markant afviger fra den morfologiske struktur på indoeuropæiske sprog, findes på begge sprog. På begge sprog resulterer denne struktur i et stort antal homografer. Selvom der muligvis er en forskel i antallet af disse, synes det rimeligt at antage, at homografien i sammenlignelig grad kan influere på, hvordan tekst afkodes på de to sprog under læsning.

Med hensyn til de sociolingvistiske aspekter er der afgørende forskelle. Selvom der findes meget forskellige sproglige stilarter på begge sprog, har det på hebraisk mindre at sige, fordi forskellen i moderne tekstmateriale kun kommer til udtryk i de diakritiske tegn, som jo normalt slet ikke anvendes. I modsætning hertil adskiller alle former for tekst på arabisk sig afgørende fra almindelig mundtlig sprogbrug. Først i forbindelse med alfabetiserings-processen igennem skoleårene tilegnes standardsproget, og det sprog arabiske børn møder, når de påbegynder deres skolegang, er meget forskelligt fra det sprog, de kender som deres modersmål. Gennem skoleforløbet udvikles de sproglige kompetencer, der gør det muligt at agere adækvat i mere og mindre formelle situationer, hvor elementer fra mere og mindre formelle stilarter blandes forskelligt. På samme måde tilegner veluddannede arabisksprogede personer med kontakt til den øvrige arabiske verden sig ofte strukturer fra andre talesprogsvarianter, således at også mindre formel sproglig interaktion på tværs af disse varianter kan foregå forholdsvis uproblematisk. For L2-lørnere indebærer dette, at de i praksis har et ret komplekst system af forskellige former for sprogvarianter at forholde sig til – et skriftligt standardsprog og et stort antal talesprogsvarianter – og for den ambitiøse lørner naturligvis også de forskellige stilniveauer, L1-sprogbrugeren behersker og fleksibelt kodeskifter imellem.

På den baggrund konkluderes det, at det er forsvarligt at betragte forskningsresultater fra hebraisk, der isoleret vedrører den morfologiske strukturs indflydelse på læseprocessen, som repræsentative for arabisk, for så vidt, at der i tilfælde hvor datamaterialet inddrager brug af homografiske konsonanter og ord med flere stavemåder, tages højde for dette. Desuden vil det i nogle tilfælde være nødvendigt at tage højde for forskellen mellem skrift- og talesprogsvarianterne, fx i forbindelse med forskningsresultater, der inddrager læsning af lavfrekvente ord, der udelukkende findes på MSA og ikke i forsøgspersonernes talesprog.

3.5 Den semitiske morfologis mentale repræsentation

Forekomsten af rødder og mønstre i semitisk lingvistik er traditionelt blevet betragtet som en ikke-linær morfologisk struktur, hvor mønstre såvel som rødder repræsenterer separate morfemer. I 1980'erne og begyndelsen af 1990'erne blev dette udfordret af forskere inden for både arabisk og hebraisk, der fremførte, at roden ikke kan betragtes som et selvstændigt morfem, men nærmere har karakter af en art 'paradigmatisk relation' mellem ord. Som argumentation blev fx anført, at både røddernes og mønstrenes relation til ordenes betydning er for forskellig fra tilfælde til tilfælde til at det giver mening at operere med disse begreber (Mahfoudi 2002). – Altså en altovervejende lingvistisk baseret kritik. Omvendt

peger visse sproglige aspekter modsat på en stærk rod-mønster-morfologi, fx behandlingen af nogle låneord. Importen sker gerne ved, at vokalerne ekstraheres fra ordet, hvorpå konsonanterne behandles som en rod og indsættes i de relevante mønstre (Shimron, Joseph 2002:11). Et eksempel er ordet فلم [film], som i pluralis tager den meget almindelige arabiske pluralisform med både præ- og infix: أفلام [aflaam]. Mange låneord har imidlertid mere end tre konsonanter. I sådanne tilfælde behandles de nogle gange som fire-konsonantiske rødder (Bolozky 2002:142; Shimron, Joseph 2002:11), som findes i et begrænset antal på både arabisk og hebraisk. Fire-konsonantiske rødder følger formerne فعّل [faʕʕala] eller تفعّل [tafaʕʕala] på arabisk og tilsvarende פיעל [piʕæl] eller התפעל [hitpaʕæl] på hebraisk, hvor den midterste konsonant oprindeligt var fordoblet som på arabisk. Herved opstår der fire 'konsonant-pladser'. Dette udnyttes fx ved ordet תלפון [tælæfoon], som tilpasses verbalsystemet til תילפן [tilfæn] 'han telefonerede'. Der kan dog også gives talrige eksempler på, at orddannelse både på baggrund af traditionelle ord og låneord fraviger den morfologiske struktur. Eksempelvis er verbet מִחשב [mihʃev] 'computerisere' etymologisk relateret til roden חשב [h-ʃ-b] som relaterer til meningsindholdet 'at tælle', men verbet er afledt af det allerede eksisterende substantiv מַחשב [mahʃev] 'computer' inklusiv præfixet מ (Bolozky 2002:132). Sådanne eksempler har fået flere lingvister til at afvise forestillingen om en særlig semitisk rod-mønster-morfologi (fx Bat-El 2002; Benmamoun 2002; Heath 2002). Heath skriver:

> *"Once we acknowledge that all of the productive Arabic ablaut processes are directional (i.e. have input stems and output stems), there is no good reason to decompose underived stems into roots, templates, and vocalic melodies, and Arabic lexical structure can be reconciled with those of the great majority of other languages."* (Heath 2002:129)

Imidlertid foreligger der en række nyere psykolingvistiske studier, der støtter hypotesen om roden som en selvstændig enhed i semitiske sprog (Obler 1989; Bentin & Feldman 1990; Feldman & Bentin 1994; Berg & Abd-El-Jawad 1996; Berent, I. & Shimron 1997; Ephratt 1997; Frost et al. 1997; Deutsch et al. 1998; Frost et al. 2000; Prunet et al. 2000; Berent, Iris et al. 2001; Frisch & Zawaydeh 2001; Berent, Iris & Shimron 2002; Berman 2002; Borer 2002; Mahfoudi 2002; Ravid, Dorit 2002; Shimron, Joseph 2002). Et eksempel er en række undersøgelser af metatesefejl[3] (fx Obler 1989; Berg & Abd-El-Jawad 1996; Prunet et al. 2000), herunder et studium af fransk-arabisk tosprogede med afasi, hvis metatese på arabisk kun berørte konsonanter, der var en del af rødderne, mens mønstrene – dvs. vokalerne og præ- in- og suffixernes konsonanter – aldrig berørtes. På fransk forekom metatese hos de samme patienter sjældnere, men i form af indbyrdes ombytning af både konsonanter, vokaler og stavelser (Prunet et al. 2000). Et andet eksempel er test af morfologisk repetition[4], som viser, at associationer der vedrører rødder, bevares længere i bevidstheden end associationer, der vedrører semantisk information (Bentin & Feldman

---

[3] Metatese er omstilling af sproglyd, fx historisk (Andreas bliver til Anders) eller som her i form af talefejl som følge af senhjerneskade.

[4] Morfologisk repetition (*repetition priming*) er en testprocedure udviklet af Forster & Davis (1984), hvor et associationsord med en bestemt ordform (fx '*nature*') præsenteres i ganske kort tid (50-60 ms.) før det egentlige målord, hvorved man – hvis målordet har samme form (fx '*mature*') – pga. den etablerede form-association kan registrere hurtigere ordbestemmelse, end hvis associationsord og målord ikke har samme form. Da forsøgspersoner ikke synes at være bevidst opmærksomme på relationen mellem associations- og målord, formodes resultater af metoden at afspejle en ubevidst association (Forster & Davis 1984; Frost et al. 1997:832).

1990; Feldman & Bentin 1994). I et sådant studium præsenteredes hebraisksprogede forsøgspersoner for målord og semantisk og morfologisk relaterede associationsord, fx målordet מִגדל [migdal] 'tårn', med associationsordene צריח [tzeriah] 'kanontårn' (semantisk), גִדול [gidul] 'tumor' (morfologisk) og גַדול [gadol] 'stor' (både semantisk og morfologisk). Hvis et associationsord blev vist lige efter målordet, blev det uanset relationen hurtigere genkendt som et ord. Hvis der derimod blev vist 15 ord uden relation til målordet mellem målordet og associationsordet, var der kun hurtigere ordbestemmelse for associationsord, der havde morfologisk relation til målordet. I ingen af testsituationerne gav associationsordet med både semantisk og morfologisk relation til målordet hurtigere ordbestemmelse, end hvor der kun var morfologisk relation. Disse resultater dokumenterer ifølge forfatterne, at roden har større associationsværdi mellem ord end semantisk indhold (Bentin & Feldman 1990). Andre eksempler er studier af forsøgspersoners reaktioner på hhv. initial og final dublering af rødders konsonanter. Disse undersøgelser bygger på fonotaktiske begrænsninger, der hænger sammen med, at rødder med dubleret konsonant finalt (fx bdd) på både hebraisk og arabisk er ret almindelige, mens rødder med dubleret konsonant initialt (fx bbd) er særdeles sjældne (Berent, I. & Shimron 1997; Berent, Iris et al. 2001; Frisch & Zawaydeh 2001). Studierne viser eksempelvis, at i en ordbestemmelses-opgave forkastes pseudoord konstrueret af rødder med dubleret konsonant initialt signifikant hurtigere end pseudoord konstrueret af rødder med dubleret konsonant finalt, og at denne følsomhed for røddernes sammensætning ses, uanset hvor i ordet rodens konsonanter forekommer (Berent, I. & Shimron 1997; Berent, Iris et al. 2001). Det synes desuden dokumenteret, at fænomenet ikke blot kan forklares med den statistiske forekomst af de enkelte strukturer (Berent, Iris & Shimron 2002).

Samtidig findes der studier, der viser, at arabisk- og hebraisksprogede personer kan ekstrahere mening fra rødder og selv konstruere ord baseret på rod-mønster-strukturen, og dette gælder vel at mærke også for børn, der endnu ikke har lært at læse og dermed ikke er blevet formelt indført i den anførte tradition for at beskrive morfologien som rod-baseret (fx Clark & Berman 1984; Berman 2002; Ravid, Dorit 2002). Eksempelvis viser et studium, hvor hebraisksprogede børn i alderen 3-7 år blev præsenteret for et verbum i forbindelse med et spørgsmål, hvor fx brugen af et instrument blev beskrevet, at børnene i mangel på et ord for et sådant instrument applicerede mønsteret for 'instrument for handlingen' på roden og på denne måde skabte ord, der var morfologisk korrekte og semantisk forståelige, men ikke desto mindre ikke-eksisterende i det hebraiske leksikon. I en tilsvarende forståelsestest blev de samme børn bedt om at beskrive det semantiske indhold af pseudoord konstrueret efter samme retningslinjer. Sådanne ord blev forstået i 59% af tilfældene blandt de 3-årige og i 87% af tilfældene blandt de 5-årige, mens kontrolgruppen af voksne forsøgspersoner klarede opgaven med en forståelse af de konstruerende ord på 100% (Clark & Berman 1984). Samtidig dokumenterer et studium af hebraisksprogede førskolebørns sproglige udvikling, at et toårigt barn producerer talefejl, hvor det mikser rødder og mønstre til ord, der ikke eksisterer, men analytisk følger den morfologiske struktur både hvad angår mønstrets fonetiske karakteristika og det semantiske indhold, som rod og mønster i princippet repræsenterer, nøjagtig som i studiet ovenfor (Borer 2002). Endvidere fandt Badry (1983), at marokkanske skolebørns kendskab til rødder – i sammenligning med en række andre former for sproglige ressourcer – var den faktor, der havde størst positiv indflydelse på deres produktion af verber i skriftlig fremstilling. Og Berman fremfører lige så kategorisk som Heath ovenfor gør for det modsatte standpunkt:

> *"[...] for speakers of Hebrew, the consonantal root has definite psychological reality as both the structural and the semantic basis for new-word formation."* (Berman 2002:245)

Anskuer man problematikken i historisk perspektiv, kan man argumentere for, at rod-mønster-morfologien hidrører fra det proto-semitiske sprog, der er ophav til de nulevende semitiske sprog, og at hvert enkelt sprog siden har udviklet sig i sin egen retning. Derfor optræder disse sproglige fænomener og nye orddannelser, der bryder den rodbaserede struktur. Omvendt kan man sige, at mange af de lingvistiske eksempler der gives for at afvise forestilling om en rod-mønster-morfologi – især på hebraisk – bygger på låneord fra andre sprog, hvor strukturer fra oprindelsessproget følger med på bekostning af de regler, der knytter sig til den semitiske morfologiske struktur og bøjningsmønstrene, hvilket ikke synes at afvige fra hvad der sker ved brug af låneord på andre sprog – fx dansk, hvor engelske ords bøjningsformer af og til følger med, hvilket ikke umiddelbart gør de traditionelle strukturer på dansk mindre relevante.

Under alle omstændigheder synes den psykolingvistiske forskning på området at indicere, at den traditionelle fremstilling af semitisk morfologi som rod-mønster-baseret faktisk afspejler, at rødder og mønstre har separate kognitive repræsentationer hos indfødte arabisk- og hebraisksprogede sprogbrugere.

Hvis dette er tilfældet – hvilke implikationer har det så for L2-lørnere af et semitisk sprog, der har et indoeuropæisk sprog som L1? Lørnere kan meget vel være opmærksomme på rod-mønster-strukturen, uden derfor at kunne bruge den til så meget. I almindelige kommunikationssituationer levnes der sjældent tid til at tænke over rod-mønster-kombinationer – og et ord er jo stadig et ord. Er det overhovedet muligt for lørnere at opbygge en helt ny måde kognitivt at behandle sprog på i stedet for at gøre brug af de sprogbehandlingsmekanismer, de allerede råder over? Man kan forestille sig, at lørnere på begynderniveau forsøger at passe sproget ind i de sprogstrukturer, de kender fra deres L1 og nære fremmedsprog, mens de, efterhånden som de tilegner sig sproget på højere niveau, vil begynde at opbygge mentale ordbehandlingsmekanismer, der ligner indfødte sprogbrugeres. Således kan man måske tænke sig, at der hos danske lørnere meget tidligt skabes en mental forbindelse mellem fx ordet دخل [daxala] 'gå ind' og ordet مدخل [madxal] 'indgang', mens der først senere skabes forbindelse mellem disse ord og ordet تداخل [tadaaxala] 'overlappe', fordi lørneren ikke umiddelbart kan gennemskue den semantiske forbindelse mellem dem og ikke i udgangspunktet besidder det mentale ordbehandlingssystem, der adskiller roden fra mønsteret. Da læsning er en proces, der involverer sådanne kognitive sprogbehandlings-mekanismer, vil nærværende afhandling berøre en flig af denne problematik, men som ovenfor anførte forskningsresultater vedrørende semitiske sprog som L1 illustrerer, er det et stærkt komplekst problemfelt, der kræver studier, der rækker langt videre end rammerne for dette forskningsprojekt.

4. Grafisk dimension: Arabisk skrift

Som det fremgår af afsnit 2.5, berører litteraturen om L2-læsning ikke problematikken om tilegnelse af et nyt skriftsystem, på trods af at flere forskningsresultater vedrørende tilegnelse af artificielle skriftsystemer indicerer, at tilegnelse af læsefærdighed med et nyt alfabet i sig selv – uafhængigt af det sprog, skriften repræsenterer – er en forholdsvis kompleks opgave. På den baggrund synes det indlysende, at det arabiske skriftsystems grafiske dimension er en afgørende barriere for læseprocessen for danske lørnere – i hvert fald på begynderniveau. Ifølge de traditionelle L2-læseteorier besidder de studerende ganske vist allerede forskellige former for læserelevante kompetencer – fx generel viden om verden og om den tekst-mæssige sammenhæng – der kan aktiveres i en *top-down* styret læsning. Men med en analogi til den generelle L2-tilegnelsesteoris tærskelhypotese, må der nødvendigvis være en nedre tærskel, for hvis skriften i sig selv er så fremmed for lørneren, at den mest af alt ligner kruseduller, giver det ingen mening at tale om læsning overhovedet.

Lørnere inden for de nære fremmedsprog[1], der jo er udgangspunkt for langt hovedparten af den forskning og dermed også de L2-læseteorier, der foreligger, kan overføre en god del af deres erfaringer med bogstaver og bogstavkombinationer fra modersmålet. Men studerende, der ønsker at lære arabisk, står altså over for helt anderledes udfordringer. Samtidig med, at deres ordforråd er begrænset, er de nemlig i arabisk sammenhæng nye læsere. Ganske vist kan de overføre deres kendskab til det fonematiske princip, dvs. sammenhængen mellem grafem og fonem, men bortset fra det er de som udgangspunkt mere eller mindre på bar bund – de har ingen erfaring med arabisk skriftligt materiale, så de har ingen mulighed for at støtte sig til højfrekvente ord, og læsetilegnelsen er derfor på mange måder at sammenligne med begynderlæserens. Og som begynderlæser forlader man sig i meget høj grad på fonologisk omkodning, og fordi det er umuligt at overskue meget mere end et enkelt bogstav ad gangen, staver man sig langsomt igennem – en proces der lægger beslag på så meget kognitiv kapacitet, at det er vanskeligt samtidig at forholde sig til tekstens indhold. Hertil kommer, at alting foregår 'omvendt', idet læseretningen går fra højre mod venstre, hvilket kolliderer med den retningssans, man har tillært sig som læser på L1.

Dette kapitel omhandler disse forholds rolle ved tilegnelse af læsefærdighed på arabisk. I det første afsnit gennemgås neurologiske og perceptionspsykologiske data, der berører læseretning. I andet afsnit behandles bogstavarkitekturen fra flere forskellige faglige vinkler, som tilsammen kan belyse, hvilke implikationer det arabiske alfabet har for lørnere, hvis L1 skrives med det latinske skriftsystem. Endelig sammenfattes de forskellige aspekter i en analyse af, hvilken indflydelse disse grafiske forskelle kan have på tilegnelse af læsefærdighed på arabisk som L2.

---

[1] Betegnelsen 'nære fremmedsprog' bruges her og i det følgende om indoeuropæiske sprog, der skrives med det latinske alfabet. I tilfælde hvor disse sprog nævnes som L1, betegnes de for nemheds skyld som 'europæiske sprog'.

4.1 Læseretning

Som beskrevet i afsnit 2.2 har man ved øjenfiksering på horisontalt organiseret tekstmateriale[2] et opmærksomhedsfelt, der ligger asymmetrisk omkring selve fikseringspunktet, så man opfatter mest materiale i læseretningen – dvs. ved læseretning fra venstre mod højre, opfatter man flere bogstaver til højre for fikseringspunktet end til venstre, mens det ved læseretning fra højre mod venstre er omvendt (Pollatsek et al. 1981; Roman 1982; Roman & Pavard 1987).

Umiddelbart synes læseretning ikke at spille nogen betydelig rolle. Imidlertid foreligger der en del neurolingvistisk dokumentation for, at netop placeringen af tekstmateriale i forhold til fikseringspunktet influerer på ordbehandling. Dette skyldes angiveligt, at kun materiale, der præsenteres centralt i synsfeltet, bliver tilgængeligt i begge hjernehalvdele samtidig, mens information set i venstre halvdel af synsfeltet (i begge øjne) sendes først til højre hjernehalvdel, mens information set i højre halvdel af synsfeltet først sendes til venstre hjernehalvdel (Rayner & Pollatsek 1989:21).[3] Der findes altså muligvis en kvalitativ forskel på læseretning: Mens billeder og ansigter behandles bedst i højre hjernehalvdel, finder sprogbehandling fortrinsvis sted i venstre hjernehalvdel, (Rayner & Pollatsek 1989:22), og man har da også kunnet dokumentere en generel højre synsfeltoverlegenhed (HSO) for genkendelse af verbalt materiale og en venstre synsfeltoverlegenhed (VSO) med hensyn til genkendelse af non-verbalt materiale (Mishkin & Forgays 1952; Petersen 1989; Rayner & Pollatsek 1989:22).[4] Fænomenet kaldes 'den visuelle lateralisationseffekt' og findes vel at mærke også hos læsere af sprog, der skrives fra venstre mod højre (Barton et al. 1965; Koriat 1985; Faust et al. 1993; Eviatar 1995; Vaid 1995:298; Eviatar 1997; Lavidor et al. 2002:263).[5] Effekten skyldes sandsynligvis, at forbindelse mellem grafem og fonem kun kan etableres i venstre hjernehalvdel (Zaidel 1977).

Ved at undersøge disse forhold hos israelske børn på forskellige læsestadier og ved læsning på forskellige sprog har man dog kunnet konstatere et skift i synsfeltoverlegenhed gennem læsetilegnelsesforløbet: Hos etsprogede begynderlæsere fandt man en klar VSO, som i 3. klasse var afløst af en klar HSO (Silverberg et al. 1980). Hos større skolebørn så man ligeledes HSO, men når eleverne i 7. klasse læste engelsk, kunne man igen konstatere VSO. Denne VSO var dog reduceret i 9. klasse, og i 11. klasse udviste eleverne også HSO ved læsning på engelsk (Silverberg et al. 1979). Disse resultater kan forklares med, at tilegnelse af nye bogstaver kræver mere af højre hjernehalvdel, og at ordbehandlingen overtages af

---

[2] Historisk set er der ikke nogen bestemt læseretning, der er mere 'naturlig' end andre. Der findes og anvendes stadig skriftsystemer, der skrives fra højre mod venstre, venstre mod højre, oppefra og ned og nedefra og op. Nogle skriftsystemer har valgfri retning og nogle skifter retning for hver linje (Sasoon 1995:19; Coulmas 1996:131). Men horisontal skrift er den hyppigst forekommende, og den vertikale skrift, der kan anvendes på kinesisk, er stærkt på vej ud (Rayner & Pollatsek 1989:56) – formentlig pga. indflydelsen fra elektronisk tekstbehandling, som i høj grad favoriserer de horisontale tekstretninger.

[3] Tilsvarende kan konstateres ved auditiv information gennem henholdsvis venstre og højre øre (Rayner & Pollatsek 1989:21).

[4] Dette er tilfældet hos højrehåndede personer. Hos en del venstrehåndede er det omvendt.

[5] Enkelte undersøgelser viser HSO ved ordgenkendelse på engelsk og VSO ved ordgenkendelse på yiddish, der skrives med det hebraiske alfabet (Mishkin & Forgays 1952; Orbach 1953). De anvendte testmetoder gjorde det dog umuligt at isolere effekten fra indlært retningsdominans, og senere studier, hvori stimuli fx organiseres vertikalt frem for horisontalt, viser som nævnt entydig HSO også for sprog, der skrives fra højre mod venstre.

venstre hjernehalvdel, når de grafiske stimuli bliver mere velkendte og i højere grad behandles som sprogligt relaterede enheder.

Det er uklart, om den registrerede HSO for verbalt materiale har betydning for læsning i forskellige læseretninger. Nogle har argumenteret, at eftersom læsning fra venstre mod højre medfører, at input optræder først i højre synsfelt og dermed i venstre hjernehalvdel, identificeres skrevne ord på sprog med denne læseretning lettere end ord på sprog, der læses fra højre mod venstre (fx Bradshaw & Nettleton 1983). Det er imidlertid en kontroversiel hypotese, som har vist sig at være overordentlig vanskelig at afprøve eksperimentelt.

Derimod har det kunnet dokumenteres, at en indlært skriftretning har meget stor indflydelse på retningssans både generelt og i relation til sproglige og ikke-sproglige grafiske stimuli: Børn etablerer tilsyneladende en retningspræference i takt med deres eksponering for tekst, og retningspræferencen ses både i deres øjenbevægelser ved scanning af visuelt materiale og ved manuel produktion eller reproduktion af stimuli. Således viste Bannatyne (1976), at øjenbevægelser i forbindelse med scanning af billedmateriale var tilfældigt fordelt på retning hos franske børn i førskolealderen, men jo ældre og mere øvede læsere, hans forsøgspersoner var, des mere var de tilbøjelige til at scanne ensrettet fra venstre mod højre.

Tværsproglige studier med børn, hvis L1 skrives med enten det latinske, arabiske eller hebraiske alfabet viser, at ikke kun læseretning, men også de individuelle bogstavers skriveretning influerer på denne retningspræference. Mens arabiske bogstaver skrives i en sammenhængende bevægelse fra højre mod venstre, består hebraiske og latinske bogstaver mest af streger, der oftest skrives fra venstre mod højre (Keenan 1972; Lieblich et al. 1975; Nachshon & Alek 1981). Lieblich, Ninio & Kugelmass (1975) sammenlignede manuel retningspræference ved tegning af en simpel horisontal streg hos hebraisk- og arabisksprogede børn. De fandt, at i den hebraisksprogede gruppe tegnede børnene fra 1.–5. klasse fra venstre mod højre i mellem 80 og 90% af tilfældene. De arabiske børn tegnede i 75% af tilfældene også fra venstre mod højre før skolealderen, men tendensen faldt drastisk i de første skoleår, og i 5. klasse viste de 100% præference for den modsatte retning. Denne højre-mod-venstre præference fortog sig dog en smule i de følgende år, hvor den hebraisksprogede gruppe samtidig forstærkede deres venstre-mod-højre præference yderligere, hvilket for begge grupper faldt sammen med, at de blev introduceret til engelsk i skolen. Således så man i 8. klasse en venstre-mod-højre retningspræference på 100% hos den hebraisksprogede gruppe, mens der er en 80% højre-mod-venstre retningspræference hos den arabisksprogede gruppe. Lignende resultater ses hos Nachshon & Alek (1981). Nachshon & Alek (1981) og Keenan (1972) har endvidere påvist en endnu stærkere venstre-mod-højre retningspræference hos engelsksprogede end hos hebraisksprogede børn.

Nachshon (1983) undersøgte retningspræference ved reproduktion af forskellige stimuli hos hebraisk-engelsk simultant tosprogede børn i 1.-6. klasse, som havde modtaget undervisning i både hebraisk og engelsk gennem hele deres skoletid. Ved reproduktion af retningsneutrale stimuli lignede resultaterne de ovenfor refererede hos hebraisksprogede børn, men ved brug af bogstaver og tal fandt man stor retningsfølsomhed. Ved reproduktion at de enkeltstående stimuli 'M', 'V', 'W', 'X' sås en generelt øget venstre-mod-højre præference med alderen. Ved multiple stimuli, hvor rækkefølgen af hebraiske, latinske og arabiske bogstaver og tal skulle rekonstrueres med klodser, sås samme tendens – undtagen for hebraiske bogstaver, som med alderen i stadig stigende grad blev organiseret fra højre mod venstre og i 6. klasse var næsten fuldstændig styret af denne retnings-

præference. I denne undersøgelse udeblev udsvinget i retningspræference efter 5. klasse, hvilket støtter hypotesen om, at det er den begyndende eksponering for engelsk, der i de tidligere undersøgelser influerer på ændringerne på dette alderstrin.

Også andre studier af læsere, der har erfaring med flere forskellige læse- og skriveretninger viser mindre ensidig retningspræference, end man ser hos læsere af det latinske alfabet. Fx har læsere af japansk – som både kan skrives horisontalt fra højre mod venstre og vertikalt i spalter fra venstre mod højre – ingen retningspræference for produktion af cirkler. Det synes således tilfældigt, om de vælger at tegne en cirkel med eller mod uret. I modsætning hertil ses en venstre-mod-højre præference ved cirkeltegning (dvs. med uret) på ca. 75% hos hebraisksprogede og meget nær 100% hos engelsksprogede forsøgspersoner (Nachshon & Hatta 2001; Fagard & Dahmen 2003).

En række studier med engelsk-, fransk-, hebraisk-, arabisk-, urdu- og hindisprogede forsøgspersoner viser desuden, at skriftretning også influerer på afstandsperception, fx bedømmelse af en linjes midtpunkt (Barrett et al. 2000) og af afstand til højre og venstre for et fikspunkt (Fagard & Dahmen 2003), ligesom det har betydning for, fra hvilken side man scanner tegninger (Hung & Tzeng 1981:85), hvor på papiret man påbegynder en tegneopgave, og for hvordan man vender tegninger af ting, der har en for- og en bagende, fx biler, cykler og dyr (Vaid 1995). Albert (1975) fandt desuden en bemærkelsesværdig forskel i verbal-spatial association hos forsøgspersoner med forskellige læseretninger på L1: I én undersøgelse registrerede han øjenbevægelser hos forsøgspersoner i forbindelse med, at han bad dem om at se enten til højre eller venstre. Han registrerede langt større fejlprocent hos personer med hebraisk eller arabisk som L1 end hos personer med et europæisk L1. I et opfølgende studium gav han forsøgspersonerne instruktioner, der involverede spejling af den spatiale association ved at lade forsøgslederen stå ansigt til ansigt med forsøgspersonen og give instruktioner som fx: 'peg på mit højre knæ med din venstre hånd'. Mens fejlfrekvensen hos forsøgspersoner med et europæisk L1 kun var 5%, var den hos personer med hebraisk som L1 28%. Albert (1975) opstillede på den baggrund en hypotese om, at læseretning influerer på integration af informationer mellem de to hjernehalvdele. Det er imidlertid indlysende, at Alberts studier meget vel kan være influeret af kulturelle forskelle, og den valgte testmetode må derfor anses for problematisk. Neurologiske studier synes dermed nødvendige for at belyse spørgsmålet yderligere.

Sammenfattende kan man sige, at det synes dokumenteret, at når man som barn har etableret én bestemt standard for læse- og skriveretning, har det afgørende indflydelse på, hvordan man generelt orienterer sig i forhold til visuelle stimuli og manuel produktion. Det står også klart, at erfaring med flere forskellige læse- og skriveretninger gør denne generelle orientering mindre ensidig. Hertil kommer nogle kontroversielle hypoteser om læseretningens indflydelse på de neurologiske vilkår for læseprocessen og for spatial behandling mellem hjernehalvdelene. Der synes imidlertid ikke at være tilstrækkelige og tilstrækkeligt metodisk gyldige data til at støtte disse sidste hypoteser, og formodentlig vil det under alle omstændigheder være overordentlig vanskeligt at finde metoder, der gør det muligt at isolere effekten af læseretning fuldstændigt.

4.2 Bogstavarkitektur

Formålet med dette afsnit er at analysere hvilke vanskeligheder de arabiske bogstavers grafiske udtryk repræsenterer for danske lørneres tilegnelse af læsefærdighed på arabisk.

Problematikken er imidlertid særdeles kompleks. For det første er det, som det fremgår af kapitel 2, almindeligvis større perceptionsenheder end de enkelte bogstaver, der behandles i læseprocessen. Og uanset dette problem tyder visse forskningsresultater på, at menneskers visuelle perception påvirkes af den kontekst, perceptionsobjektet optræder i, således at et bogstav opfattes forskelligt alt afhængigt af om det optræder alene, sammen med andre selvstændige visuelle stimuli eller som del af en meningsfuld enhed, dvs. et ord. (McClelland & Rumelhart 1981). Forskningen på dette område er desværre alt for diffus til at man kan udlede konkrete konklusioner, der har relevans for den aktuelle problematik. På baggrund af hvordan læsetilegnelsen forløber, synes det imidlertid forsvarligt i denne sammenhæng at afgrænse analysen til bogstavniveau, da bogstaverne ikke bare er reelle grafiske enheder, men netop hos begynderlæsere i meget høj grad også perceptionsmæssigt fungerer som sådan, i kraft af at læsere, der lærer et nyt, fonografisk skriftsystem, og som har forhåndskendskab til det fonematiske princip, i begyndelsen uvægerligt omkoder fonologisk og 'staver sig igennem' tekster, når de læser, jvf. afsnit 2.5.

Semiotiske modeller og bogstavgenkendelsesteorier

Der findes flere forskellige former for semiotiske modeller, der beskæftiger sig med latinske bogstaver. Disse kan kategoriseres efter to analytiske niveauer (Watt 1988; Jameson 1994): På 'dybde-niveau' beskæftiger man sig med abstrakte regler for bogstavformerne, og der er således udviklet egentlige grammatikker og formelsamlinger for deres sammensætning (fx Künnapas 1966; Watt 1975; Cox et al. 1978; Watt 1988,1994). På det deskriptive niveau beskæftiger man sig dels med den fanemiske (dvs. visuelle), dels med den kinemiske (dvs. produktive) dimension.[6] De kinemiske modeller beskriver værdier for retning og orientering i forbindelse med den manuelle frembringelse af bogstaverne, når man skriver dem (fx Watt 1980). I beskrivelserne ekspliciteres både synlige og usynlige linjesegmenter, kurver og sammenføjninger. I denne sammenhæng er udelukkende bogstavernes visuelle udtryk interessant, fordi kun denne dimension er relevant ved læsning. Hverken abstrakte eller kinemiske modeller vil derfor blive behandlet yderligere.

Hvad angår de fanemiske modeller, findes der en række analyser, der med forskellig detaljeringsgrad beskriver alfabetet ved at opsplitte bogstaverne i delkomponenter (Eden & Halle 1961; Watt 1981). Grunden til, at teoretisering over bogstavformer overhovedet giver mening er, at bogstaverne ikke blot er en række tilfældige grafiske enheder uden indbyrdes relation. Brugere af det latinske alfabet har tvært imod en ret systematisk mental repræsentation af alfabetet, hvilket kommer til udtryk ved, at vi har en meget konsistent fornemmelse af, hvilke former, der kan karakteriseres som værende 'bogstavlignende' (Watt 1975:294; Jameson 1994; Watt 1994:218). Ud over empiriske studier er bogstavernes historie og observationer af børns skrivefejl blevet brugt som indikationer på, hvilke kriterier denne fornemmelse hviler på. Fx synes de almindeligste fejl hos børn i alfabetiseringsperioden at være udtryk for, at bogstavernes grafiske udtryk overgeneraliseres: I begyndelsen skrives alle bogstaverne af og til korrekt, og af og til spejlvendt. Senere vendes de korrekt, bortset fra N, Z og S, som stadig ses i begge versioner samt bogstavet J, som i meget lang tid konsekvent fastholdes i spejlvendt version, fuldstændig ligesom grækerne spejlvendte lambda, da de importerede det fra det fønikiske alfabet. Fænomenet illustrerer, at begyndere generaliserer bogstavernes orientering, på baggrund af at hovedparten af dem

[6] Modellerne fremstår således samlet som en parallel til den generative fonologi (se fx Kenstowicz & Kisseberth 1979).

vender mod højre. J er således det eneste bogstav, der vender mod venstre, mens N, Z og S er de eneste, der er asymmetriske på den vertikale akse men ikke vender hverken den ene eller den anden vej (Watt 1975:298; 1994:219).

Mens ovenstående semiotiske modeller har til formål at beskrive den mentale systematik, der ligger til grund for vores opfattelse af bogstavformerne, har læseteoretisk funderede teorier været mere koncentreret om, hvilke processer der ligger bag vores evne til hurtigt og præcist at genkende og anvende alfabetisk materiale. Der findes to centrale teorier om genkendelse af bogstaver:

Ifølge *template-matching*-teorien gemmer vi i hjernen en repræsentation af hver enkelt bogstav, vi kan genkende. Genkendelse finder således sted ved, at bogstavets form passer til en skabelon, der findes i hukommelsen, og er altså en binær proces: genkendelse eller ikke genkendelse. Det centrale problem ved denne teori er, at den forudsætter, at vi har en skabelon for hver variation af hvert af de bogstaver, vi er i stand til at genkende. Erfaringsmæssigt er vi imidlertid i stand til at genkende bogstaver i utallige variationer af både trykte og håndskrevne skrifter. En forklaring på dette kan være, at vi 'renser' stimulus for unødvendig information, mens kun den essentielle information udnyttes. Således er det en standardiseret udgave af input, der sammenlignes med skabelonen. Der er dog stadig tale om en holistisk proces, hvor én figur skal passe til én skabelon (Rayner & Pollatsek 1989:11).

Den mere analytisk orienterede *feature-detection*-teori har derimod som udgangspunkt, at bogstaver har mange fællestræk (Rayner & Pollatsek 1989:11-15), at der kan være tale om forskellige grader af genkendelse, og at input kan opfattes flertydigt (Gibson et al. 1963; Bouma 1971; Massaro 1979). Størstedelen af forskningen på området omhandler genkendelse af det latinske alfabets majuskler. Disse består især af rette linjer samt et begrænset udvalg af kurveformer, og tilsyneladende genkender vi den grafiske enhed ved at analysere disse komponenter (Rayner & Pollatsek 1989:11-15). Ifølge *feature-detection*-teorien skelner vi altså E fra F eller L i kraft af antallet af horisontale linjer. Der foreligger flere studier, der bygger på *feature-detection*-teorien, hvori der opstilles bogstavkomponent-matricer, hvor forekomsten af de enkelte komponenter i hvert bogstav registreres, fx horisontale, vertikale og diagonale linjer og åbne og lukkede kurver. Gibson (1963) og Geyer (1973) registrerer også andre visuelle aspekter, fx symmetri, mens Lindsay & Norman (1972) registrerer rette og spidse vinkler.

Der er således en tæt sammenhæng mellem *feature-detection*-teorien og de førnævnte fanemiske, semiotiske modeller, der netop er karakteriseret ved, at de opsplitter bogstaverne – også fortrinsvis majusklerne – i delkomponenter.[7] Bogstavkomponentmatricerne berører begge områder, idet de har til formål at definere de visuelle karakteristika, der har indflydelse på, hvordan vi skelner mellem bogstaverne, men mens de som fanemiske modeller forsømmer at forklare bogstavernes udtryk – at de enkelte komponenter forekommer, beskriver kun dårligt bogstavets egentlige udtryk (Watt 1975) – synes de som perceptionsteoretiske modeller at være for semiotisk forankrede, idet de komponenter, der umiddelbart kan defineres ud fra bogstavernes visuelle udtryk, bliver undersøgelsernes

[7] Forskellen er naturligvis, at de fanemiske modeller har selve bogstavernes grafiske udtryk som udgangspunkt, mens *feature detection*-teorien fokuserer på vores perception af samme udtryk (se Watt 1975:296 for en diskussion af det visuelle vs. kognitive udgangspunkt).

A.

perceptuelle enheder

visuelle komponenter

	A	B	C	D	E	F	..
1	+				+	+	
2		+		+	+	+	
3	+						
4		+	+	+			
:							

I komponentmatricen registreres, om den enkelte komponent forekommer (Gibson 1963) eller hvilket antal gange, den forekommer (Lindsay, 1972 ; Geyer 1973) i hvert enkelt bogstav.

B.

responser

stimuli

	A	B	C	D	E	F	..
A	.743	.070	.012	.049	.012	.113	
B	.012	.765	.037	.101	.037	.049	
C	.028	.015	.651	.025	.199	.083	
D	.015	.597	.049	.239	.042	.058	
:							

I forvekslingsmatricen registreres hyppighed for obs. respons ved hver enkelt bogstav (Rumelhart 1971) hos Pew & Gardner (1965) hyppighed for forveksling af definerede bogstavpar.

Figur 4.1 Eksempel på komponentmatrix (A) og forvekslingsmatrix (B) for de latinske majuskler.

afgørende parametre. Andre modeller tager mere direkte afsæt i perceptionen af bogstaverne, og kortlægger – også i matricer – graden af forveksling imellem dem (fx Pew & Gardner 1965; Rumelhart 1971) (se figur 4.1).

Som bekendt er det meste af den tekst vi omgiver os med imidlertid fortrinsvis skrevet med minuskler, som fremstår mere kurvede end majusklerne.[8] Ifølge Tinker (1965), Bouma (1971) og Massaro et. al. (1980) genkendes minusklerne først og fremmest som en generel form, hos Bouma defineret som *"the smallest enclosing polygon without indentations"* (Bouma 1971:460), og først herefter inddrages de mindre delkomponenter i analysen. I sin undersøgelse fremprovokerede Bouma genkendelsesfejl hos en gruppe forsøgspersoner ved at præsentere bogstaverne på stor afstand og i periferien af synsfeltet, og registrerede deres responser i en forvekslingsmatrice. På den baggrund kunne han definere syv grupper af bogstaver, hvor forveksling var hyppig mellem bogstaver i samme gruppe, men meget sjælden på tværs af grupperne (se figur 4.2). Ifølge Bouma genkendes bogstavet 'e' således primært som en cirkelform, hvor den vandrette streg sekundært adskiller det fra 'c' og 'o', der er de eneste alternativer (Bouma 1971).

Genkendelse af minusklerne sker altså ifølge Boumas teori som en hierarkisk proces, som involverer både bogstavets samlede struktur og dets delkomponenter, og som ikke blot foregår som genkendelse, men også involverer udelukkelsesmetoden.

[8] Dette afspejler at man havde skiftet skrivemateriale, da de blev udviklet. Da majusklerne blev til, huggede man bogstaverne i træ eller sten, mens pen og papir var udbredt da minusklerne blev udviklet – formentlig netop for at tilpasse skriften til materialerne (Mallon 1952).

Hovedgruppe	Beskrivelse af generel form	Bogstaver i gruppen
Små	rektangulær ydre form med indhold	a, s, z, x
	cirkulær ydre form	e, o, c
	vertikal stregserie	n, m, u
	triangulær ydre form	r, v, w
Opadgående	bred form	d, h, k, b
	smal form	t, i, l, f
Nedadgående	-	g, p, j, y, q

Figur 4.2 Boumas gruppering af bogstavernes generelle former baseret på forveksling af bogstaver (Bouma 1971).

På baggrund af ovenstående kan man identificere to særskilte bogstavgenkendelsesmodeller for latinske bogstaver: En ren komponentanalysemodel, der beskriver genkendelse af majusklerne, og en hierarki-analysemodel, som beskriver genkendelse af minusklerne. Imidlertid foregår også genkendelse af majusklerne muligvis hierarkisk. I hvert fald peger forskningsresultater fra oftalmologien i den retning (Massaro & Schmuller 1975:231).

I henhold til den valgte læseprocesmodel, hvor bogstavdetektorerne som en del af det samlede ordgenkendelsessystem repræsenterer et selvstændigt konnektionistisk netværk, må man forkaste *template-matching*-teorien. Ligesom forestillingen om at ordgenkendelse sker som en konsultation af et mentalt leksikon i konnektionismen forkastes til fordel for en model, hvor mening genereres i kraft af erfaringsbaserede lagringer i det neurale netværk, må forestillingen om et mentalt sæt af statiske bogstavskabeloner betragtes som utilstrækkelig til at forklare den fleksibilitet, vi udviser i forbindelse med genkendelse af skrifttyper. *Feature-detection*-teorien, som da også er bedre dokumenteret end den førnævnte (Massaro & Schmuller 1975), kan derimod forholdsvis let inkorporeres i den valgte model og være med til at beskrive, hvad der mere nøjagtigt karakteriserer bogstav-detektionsnetværket. Man kan således forestille sig, at perceptionen foregår som en stimulation af bogstavkomponentenheder, der giver associationer om andre komponent-enheder, som de erfaringsmæssigt passer sammen med (McClelland & Rumelhart 1981; Sanocki 1987). Aktivering af en bogstavkomponent sker altså ikke isoleret men i en kontekst af de øvrige komponenter, bogstavet indeholder. Selvom denne proces ikke er ensrettet, kan *feature-detection*-teoriens forestilling om en hierarkisk genkendelsesproces godt forsvares, på samme måde som forestillingen om bogstavdetektion og ordgenkendelse som en moduleret proces i den samlede læseprocesmodel kan det. Bogstavkomponenterne bliver hermed blot procesenheder i bogstavdetektionsnetværket, nøjagtig ligesom bogstaverne er det i ordgenkendelsesnetværket. De to netværk er dermed på samme tid selvstændige og integrerede: De udfører hvert deres arbejde, men er gensidigt afhængige, forstået sådan at de to processer foregår i et løbende samspil, hvor informationen går begge veje, så bogstavgenkendelsesforløbet føder ordgenkendelsesprocessen, som omvendt påvirker og støtter bogstavgenkendelseprocessen 'oppefra'. Og på nøjagtig samme måde som erfaring med både ords holistiske grafiske udtryk og med bestemte bogstavkonstellationer udnyttes i ordgenkendelsesprocessen, udnyttes erfaring med generelle bogstavformer og komponentkonstellationer i bogstavdetektionsprocessen.

I det følgende præsenteres forskning, der kan belyse hvordan disse processer fungerer ved tilegnelse af et nyt skriftsystem.

Tilegnelse af en ny skrift

En række studier af børns opfattelse af forskellige skriftsystemer illustrerer udmærket, hvordan den systematiske repræsentation af bogstavformer i alfabetet løbende etablerer sig som kontrastive elementer, der gør det muligt at skelne mellem skriftsystemer og individuelle bogstaver eller tegn. Lavine (1972) fandt, at når amerikanske 3-årige blev bedt om at definere om forskellige former for stimuli var 'skrift' eller ej, beskrev de latinsk og hebraisk som 'skrift', mens kun ganske få ville acceptere samme betegnelse for kinesisk. Geva & Wade-Woolley (1998) påtvang engelsk-hebraisk tosprogede børn i Canada at opfatte forskellige skriftsystemer som skrift ved at præsentere det i en billedlig kontekst, hvor en fårekylling skrev på engelsk, spansk, russisk, kinesisk, vokaliseret hebraisk og cree[9] og ved at spørge: 'skriver fårekyllingen engelsk her?'. De femårige bedømte alle de nævnte former for skrift til at være engelsk, hvilket muligvis afspejler, at børn i den alder endnu ikke associerer forskellige sprog med forskellige skriftsystemer. Lidt ældre børn, der kunne det latinske alfabet og kunne skrive deres eget navn på det, opfattede kinesisk, hebraisk og cree som værende 'ikke engelsk', mens spansk og russisk stadig i deres øjne var engelsk. Kun børn der havde etableret forholdsvis gode læsefærdigheder og et enkelt førskolebarn, der havde været særligt eksponeret for spansk og derfor lagde mærke til den lille forskel i form af enkelte accenter, var i stand til at skelne spansk som værende 'ikke engelsk' (Willows & Geva 1995; Geva & Wade-Woolley 1998). Når russisk blev opfattet som engelsk af børn, der kunne hele det latinske alfabet, må det skyldes, at de ikke opfatter dette alfabet som et udtømmende udvalg af bogstavmuligheder, og at det russiske alfabet ikke er afgørende anderledes end det latinske i selve bogstavkomponenterne.

Voksne lørnere har naturligvis en mere analytisk tilgang til et nyt skriftsystem, men i de første faser af tilegnelsen af et nyt skriftsystem kan de nye former forekomme forvirrende – især når bogstavkomponenterne er fremmede:

> *"When individuals encounter orthographies that they have not yet learned to read and write, they may first view them as a disorganized bunch of squiggles. As they come to recognize and reproduce individual characters, they begin to learn the distinctive features of the particular system. This process is probably very gradual, extending over many months. Through perceptual learning they progress toward better and better differentiation of the characters of the system. What began as amorphous and confusible with something else comes to be perceived as structured and more specific..."* (Willows & Geva 1995:361)

Det er selvfølgelig en banal konstatering, at et nyt skriftsystem i begyndelsen ligner kruseduller, men efterhånden begynder at give mening. Det interessante er, at processen med at tilegne sig et nyt system til at skelne mellem de individuelle tegn er forholdsvis langsommelig. I hvert fald indicerer en undersøgelse af østasiatiske lørneres L2-læsning på engelsk, at det ikke kun er den alfabetiske ortografi, men muligvis også det latinske alfabet i sig selv, der udgør et problem for god læsetilegnelse hos disse lørnere (Jackson et al. 1994). De undersøgelser af læsekompetencer på arabisk hos L2-lørnere, der refereres i afsnit 1.2, viser ligeledes, at den arabiske skrift synes at udgøre en selvstændig barriere (Hana 1972; Khaldieh 1996).

[9] Et indiansk sprog, der skrives med et syllabisk skriftsystem af samme slags som inuktiut (ᑎᑎᕋᐅᖅ ᓄᑦᑕᐊᖅ).

Ydermere er der flere undersøgelser, der viser, at forskellige skriftsystemer animerer til forskellige visuelle scanningsstrategier, og at nogle skriftsystemer kan være objektivt vanskeligere at tilegne sig end andre:

Et lille antal studier tilbage fra 1930'erne og 1940'erne viser, at forskellige typografier påvirker øjenbevægelser under læsning forskelligt. De viser fx at i forhold til læsning af tekst skrevet med små bogstaver, kræver læsning af tekst med store bogstaver flere øjenfikseringer og længere læsepauser, mens sammenblanding af små og store bogstaver nedsætter læsehastigheden yderligere (Levy-Schoen & O'Regan 1979:17). Disse resultater indicerer, at øjenbevægelsernes hastighed nedsættes jo mindre velkendt skriften er for den enkelte læser – og ved tilegnelse af et nyt skriftsystem er skriften naturligvis mindre velkendt for læseren.

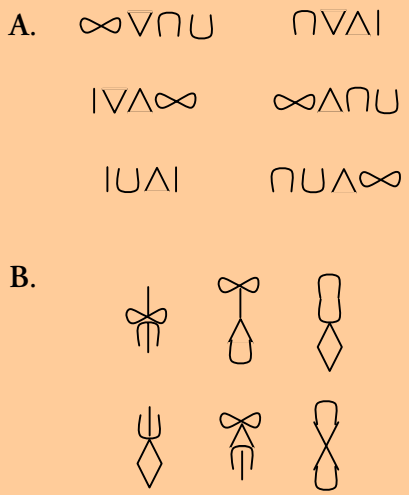
Der foreligger også undersøgelser, der viser, at voksne forsøgspersoner, hvis modersmål skrives med det latinske alfabet, scanner bogstaver anderledes end de scanner andre slags visuelle input. Hammond & Green (1982) og Green, Hammond & Supramamian (1983) præsenterede engelsksprogede forsøgspersoner for input bestående af enten enkeltstående bogstaver eller en grafisk form af bogstavlignende karakter, for derefter at lade dem søge efter tilstedeværelsen af samme objekt i en række bestående af fem lignende former. For ikke-alfabetiske former blev målobjekter hurtigst genkendt, hvis de lå i midten af rækken, mens reaktionstiden var klart langsomst, hvis de lå i en af enderne. Bogstaver derimod blev hurtigst genkendt, hvis de lå i venstre ende af rækken, og også bogstaver i højre ende og i midten blev genkendt hurtigere end bogstaver på de øvrige pladser. I en grafisk præsentation af resultaterne, der viser målobjektets position i rækken (vandret) i forhold til reaktionstid (lodret), giver genkendelse af ikke-alfabetisk materiale således en tydelig U-form, mens alfabetisk materiale giver en M-form. Undersøgelserne viste desuden, at den særegne scanning af alfabetisk materiale også findes hos børn, og altså udvikles i en meget tidlig alder (Green et al. 1983). Med udgangspunkt i disse undersøgelser lavede Randall & Meara (1988) et lignende eksperiment med arabisksprogede forsøgspersoner, hvor de sammenlignede scanning af arabiske og latinske bogstaver. Eksperimentet viste en U-form for scanning af både latinske og arabiske bogstaver, og mens forsøgspersonerne fik kortere reaktionstid jo mere erfaring de fik med at læse på engelsk, bibeholdt de vel at mærke at finde målobjekter placeret i midten hurtigere end i både højre og venstre ende af rækken. Personer med arabisk som L1 og første skriftsprog scanner altså tilsyneladende alfabetisk materiale fundamentalt anderledes end engelsksprogede personer. Ved behandling af latinske bogstaver overfører de den U-formede scanningsstrategi, som engelsksprogede personer normalt kun anvender for ikke-alfabetisk materiale. Umiddelbart synes den mest oplagte forklaring på fænomenet at være, at der er mere grafisk sammenfald mellem arabiske end mellem latinske bogstaver, hvorfor de kræver en mere detaljesøgende scanning.[10] I så fald indicerer det, at ikke blot de arabiske bogstavers fremmedartethed, men også deres objektive grafiske udtryk udgør en barriere for hurtig genkendelse hos lørnere med et latinsk-skrevet L1.[11] Under alle omstændigheder synes det dokumenteret, at de scanningsstrategier, der er etableret i den tidlige læsetilegnelse overføres ved senere

[10] Der kan dog også være andre forklaringer på undersøgelsens resultater. Fx kan det konsonantiske skriftsystem tænkes at være en medvirkende årsag i kraft af de manglende vokaler og, ikke mindst, den ikke-lineære morfologiske struktur. Randall & Meara inddrager desuden en noget traditionel, orientalistisk orienteret forklaringsmodel ved at henvise til tværkulturelle studier fra 1930-erne, der viser større opmærksomhed på detaljer blandt arabere end blandt europæere, og mener, at der kan være en sammenhæng mellem kulturelle og lingvistiske fænomener (Randall & Meara 1988:144).

[11] Netop dette spørgsmål søges belyst nærmere i det følgende afsnit om tilegnelse af arabisk skrift.

indlæring af andre skriftsystemer, og at man således ikke umiddelbart tilegner sig samme strategier som L1-læsere anvender.

Som beskrevet i afsnit 2.3 befordrer fonografiske skriftsystemer fonologisk afkodning, mens mere logografisk baserede skriftsystemer befordrer en mere leksikalsk afkodning. En række undersøgelser dokumenterer, at også skriftsystemets grafiske udtryk kan påvirke afkodningsprocessen, ligesom både skrift og ortografi tilsyneladende har betydning for læsehastigheden: Brooks (1977) testede to grupper af forsøgspersoner for læsning af skrift bestående af diskrete (dvs. ikke sammenhængende) symboler, der skulle læses horisontalt fra venstre mod højre, og læsning af de samme symboler arrangeret sammenhængende og vertikalt, så de fremstod som glyffer. Disse skulle læses oppefra og ned. For den ene gruppe var skriftsystemet alfabetisk, og symbolerne havde direkte korrespondance til engelske bogstaver og ordene til engelske ord (engelsk var modersmål for forsøgspersonerne). For den anden gruppe var skriftsystemet logografisk med korrespondance til engelske ord men ikke til lyd. Begge grupper lærte i to omgange 24 ord skrevet med den ene skrift og derefter 24 andre ord skrevet med den anden skrift, så de til sidst kunne genkende 96 ord inden for ét sekund pr. ord. Som det ses i figur 4.3, var der for gruppen, der læste den alfabetiske skrift, ikke stor forskel på, hvor hurtigt de lærte at afkodede den diskrete, horisontale skrift og den glyflignende, vertikale skrift. For gruppen, der læste logografisk skrift var forskellen derimod massiv. De glyflignende ords mere distinkte fremtræden gjorde det altså nemmere at lære dem, når de var logografer. Det glyflignende alfabet var i begyndelsen sværere at læse end det diskrete alfabet, men også her blev den distinkte form en fordel lidt efter lidt. Efter ca. 60 forsøg blev det glyflignende afkodet hurtigere end det diskrete alfabet og efter ca. 200 forsøg også hurtigere end de logografiske glyffer. Dette forløb må formodes at afspejle, at det er en fordel, at de enkelte bogstaver er klart afgrænsede i begyndelsen, hvor forsøgspersonerne forlader sig på fonologisk afkodning, men at det løbende bliver en fordel, at ordet nemt genkendes som en distinkt helhed, når forsøgspersonerne efterhånden opnår automatisk ordgenkendelse.



C.

	Alfabetisk skrift	Logografisk skrift
Diskrete Liste 1	79,6	81,5
Glyffer Liste 1	83,8	26,0
Diskrete Liste 2	50,2	42,0
Glyffer Liste 2	45,4	27,0
Diskrete Liste 1+2	37,6	28,5
Glyffer Liste 1+2	23,4	24,0

Figur 4.3 Brooks artificielle skriftsystemer. A: Horisontalt arrangerede, diskrete symboler. B: Samme symboler arrangeret vertikalt som glyffer. C: Antal forsøg for at opnå ordgenkendelse på max. ét sekund pr. ord af 24 ord pr. liste (Brooks 1977)

Undersøgelsen illustrerer, at det er muligt at tilegne sig et fremmed skriftsystem og opøve hurtig, automatisk ordgenkendelse. Men det dokumenterer også, at adgangen til et nyt skriftsystem er vanskelig: Selv med dette artificielle skriftsystem, der kun blev anvendt på et meget begrænset antal ord, hvis semantiske indhold udelukkende repræsenteredes på forsøgspersonernes modersmål, måtte forsøgspersonerne gennem et meget stort antal forsøg for at opnå automatisk ordgenkendelse. Endvidere fremgår det, at det grafiske udtryk ved det skriftsystem, der skal tilegnes, ikke er uden betydning. Her sammenlignes et horisontalt, diskret og et vertikalt, glyf-lignende arrangement af de valgte symboler, og som det fremgår læses diskrete symboler lettest i begyndelsen.

Også undersøgelser fra 'den virkelige verden' dokumenterer, at der er objektive forskelle på, hvor vanskelige forskellige skrifters tegn er at skelne. Willows & Geva (1995) rapporterer således om et studium af en japansksproget patient med erhvervet dysleksi, som havde lettere ved at genkende og reproducere *kana* end *kanji,* selvom begge systemer er i brug på japansk og indlæres sideløbende. Hvad angår de alfabetiske skriftsystemer viser flere studier, at afkodning af hebraisk og arabisk kræver mere visuo-spatial opmærksomhed end afkodning af engelsk (Share & Levin 1999; Abu-Rabia 2001; Shatil & Share 2003), og sammenlignende studier af engelsk og hebraisk skrift viser da også, at hebraiske bogstaver er vanskeligere at skelne end latinske: Shimron & Navon (1981) dokumenterede dette ved at definere parametrene 'særegenhed' (som et mål for hvor kritisk en bogstavkomponent er for genkendelse af bogstavet) og 'entydighed' (som et mål for hvor sjældent en bogstavkomponent forekommer i andre bogstaver). Også i empiriske test, hvor de målte genkendelse af bogstaver, hvis komponenter var underlagt forskellige former for manipulationer, fandt de, at manipulationerne påvirkede genkendelse af de hebraiske bogstaver mere end genkendelse af de latinske (Shimron & Navon 1980,1981). Geva & Siegel (2000) fandt også hos engelsk-hebraisk tosprogede børn, at de havde langt flere visuelle bogstavgenkendelsesfejl på hebraisk end på engelsk. Da det dermed er rimelig godt dokumenteret, at hebraiske bogstaver er objektivt vanskeligere at skelne end latinske, er det i denne sammenhæng interessant, at et komparativt studium af hebraisk og arabisk skrift hos arabisk-hebraisk tosprogede indicerer, at arabiske bogstaver er endnu vanskeligere at skelne end hebraiske (Eviatar et al. 2004).

Dermed synes det at stå klart, at selvom det naturligvis er muligt at tilegne sig gode læsefærdigheder på arabisk som L2, så er det ikke en simpel opgave. Tilegnelse af et nyt skriftsystem synes at være en kompleks opgave i alle tilfælde, og da arabiske bogstaver sammenlignet med latinske bogstaver sandsynligvis er vanskeligere at skelne, er arabisk skrift muligvis særlig vanskelig at tilegne sig.

Tilegnelse af arabisk skrift

Da semiotiske og perceptionspsykologiske undersøgelser svarende til de ovenfor beskrevne tilsyneladende ikke findes for andre skrifter end den latinske,[12] foreligger der ikke noget materiale, der umiddelbart kan anvendes til at forklare præcis hvorfor arabiske bogstaver er objektivt vanskeligere at skelne end de latinske, og i hvor høj grad arabisk skrift selvstændigt udgør en barriere for L2-læsning på arabisk.

---

[12] Bortset fra en enkelt 'generativ grammatik for kinesiske ideogrammer' refereret af Cox, Blesser & Eden (1978:430)

En lille deskriptiv, komparativ analyse af latinsk og arabisk skrift kan imidlertid illustrere, at de genkendelsesprocedurer, vi har opøvet ved læsning af latinsk skrift, ikke umiddelbart kan overføres ved læsning på arabisk. Som bogstavgenkendelsesteorierne i forening med de semiotiske modeller viser, er den veludviklede evne vi som læsere af det latinske alfabet har til hurtigt og præcist at skelne latinske bogstaver fra hinanden, baseret på genkendelse af et afgrænset udvalg af grafiske komponenter. Nedenstående opstillinger (figur 4.4 - 4.7) er inspireret af de ovennævnte studier, der med forskellig kompleksitets- og detaljeringsgrad opstiller forskellige former for komponentmatricer, men er af hensyn til overskueligheden præsenteret efter en simplere systematik. De er altså ikke direkte forankret i empiriske studier af den menneskelige perceptionspsykologi, men tjener til at give et let tilgængeligt overblik over, hvor forskelligartede bogstavkomponenter de to alfabeter repræsenterer.

Som det ses i figur 4.4 begrænser komponenterne sig for de latinske majusklernes vedkommende langt hen ad vejen til rette linjer samt halv- og helcirkler. Som det fremgår, er knap halvdelen af bogstaverne opbygget med en enkelt vertikal linje som base. Disse distancerer sig næsten alle fra hinanden i kraft af horisontale linjer, vertikale linjer eller kurver mod højre. De øvrige er mere symmetrisk opbygget, enten med to vertikale linjer som base eller konstrueret af diagonale linjer eller på baggrund af en cirkelform. For minusklernes vedkommende er over halvdelen af bogstaverne kopieret fra majusklerne (i figuren illustreret med fed skrift), mens resten er konstrueret som variationer over eksisterende former – enten roteret og/eller spejlet (fx p,q,b,d) eller med forlængelse nedad (j,y) eller opad (l,h). Kun a og e udgør selvstændige former. Minusklerne er kategoriseret i figur 4.5, hvor parametrene simpel og forlænget form kombineres med komponent-systematikken fra majusklerne, og resultatet er meget bekvemt en gruppering, der ikke falder langt fra Bourmas i figur 4.2.

Til sammenligning er de tilsvarende delkomponenter for de arabiske bogstaver systematiseret i figur 4.7 og 4.6 for hhv. medial- og finalposition. Bogstavernes udseende i initial- og selvstændig position er udeladt, da kun ه og ع/ غ i disse positioner afviger afgørende fra deres udseende i medialposition, jvf. figur 3.3. Medialformerne er kategoriseret som de latinske minuskler, tilføjet en distinktion mellem smalle og brede bogstaver. De seks bogstaver, der ikke sammenbindes mod venstre (markeret med *) udgør en tværgående gruppe, idet netop den manglende sammenbinding udgør en markant grafisk forskel i forhold til de øvrige bogstaver, når bogstaverne optræder i ord. Hovedparten af finalformerne afviger fra medialformerne i kraft af en svungen streg, der grafisk markerer ordets afslutning. Denne finale synes umiddelbart at være bogstavernes mest markante komponent, og med enkelte undtagelser er det de samme tre finaleformer, der går igen. Dette giver – på linje med den enkelte og dobbelte vertikale base, der præger hovedparten af de latinske majuskler – anledning til en selvstændig kategorisering efter finalens udseende.

Ved sammenligning af de latinske minuskler og de arabiske medialformer, som er de to skriftsystemers mest frekvente udgaver af bogstaverne, har basisformerne i de to skriftsystemer helt forskellig karakter. I det arabiske alfabet er for det første prikkerne komponenter, der indgår i halvdelen af bogstaverne. For det andet er sammenbindings-stregen en markant komponent, der præger alle bogstaverne, som derfor i sammenligning med de latinske fremstår mere som grafiske tilføjelser i form af en gennemgående bundstreg, end som selvstændige enheder. For det tredje er der ved de arabiske bogstaver i meget højere grad end ved de latinske et direkte grafiske sammenfald bogstaverne imellem:

Vertikal base →	I	I I	Ingen
Horisontaler	T F E L	H	
Diagonaler	K R	N M	A V X Z Y W
Kurver	P B D J	U	O Q C G S

Figur 4.4 De latinske majusklers grafiske komposition

	Simpel	Forlænget
Vertikaler	i r	**j** l **f** t
	u n **m**	h
Diagonaler	**v** **w** **x** **z**	**y** **k**
Kurver	**o** **c** e	**p** q b d
	s a	g

Figur 4.5 De latinske minusklers grafiske komposition

Finaleform →		Fremad på linjen	Fremad u. linjen	Bagud u. linjen	Ingen/ særlig
Vertikaler	simpel	ب ت ث	ن		
			س ش		
	for.	ك	ل		ـا
Diagonaler / åbne kurver	simpel		ى ي		د ذ
	forlænget			ح خ ج	ر ز
				ع غ	
Lukkede kurver	simpel	ف	ق		ه ة
			ص ض		
	forlænget				ط ظ
					و م

Figur 4.6 De arabiske finalers grafiske komposition

		Simpel	Forlænget
Vertikaler	Smal	نـ تـ ثـ بـ يـ	لـ ـا *
	bred	سـ شـ	
Diagonaler / åbne kurver	smal	عـ غـ	
		د * ذ *	ر * ز *
	bred	حـ خـ جـ	كـ
Lukkede kurver	smal	مـ	هـ
		فـ قـ	و *
	bred	صـ ضـ	طـ ظـ

Figur 4.7 De arabiske medialers grafiske komposition (*: sammenbindes ikke)

Med Boumas gruppering i figur 4.2 som model, er bogstaverne i figur 4.8 grupperet efter dette sammenfald. Som det fremgår, er 22 af de 28 bogstaver konstrueret over kun otte basisformer, inden for hvilke kun antallet af prikker over eller under basisformen (for dråbeformen også den vertikale streg) gør det muligt at skelne bogstaverne fra hinanden. På baggrund af bogstavgenkendelsesteorierne, der vedrører det latinske alfabet, giver det mening at forestille sig en hierarkisk genkendelsesproces, hvor basisformen er den primære genkendelsesfaktor, og hvor prikkerne er afgørende for den endelige genkendelse.

Hovedgruppe	Beskrivelse af generel form	Identisk base	Øvrige
Smalle	enkel vertikal	ـيـ ـبـ ـثـ ـتـ ـنـ	ـلـ
	triangel	ـغـ ـعـ	
	diagonal/åben kurve	ـجـ ـخـ ـحـ	ـكـ
	lukket kurve	ـقـ ـفـ	ـمـ ـهـ
Brede	vertikal stregserie	ـشـ ـسـ	
	dråbeform	ـظـ ـطـ ـضـ ـصـ	
Sammenbindes ikke	afsluttes under linjen	ـزـ ـرـ	ـوـ
	afsluttes på linjen	ـذـ ـدـ	ـاـ

Figur 4.8 Gruppering af de arabiske bogstavers generelle former baseret på identiske basisformer.

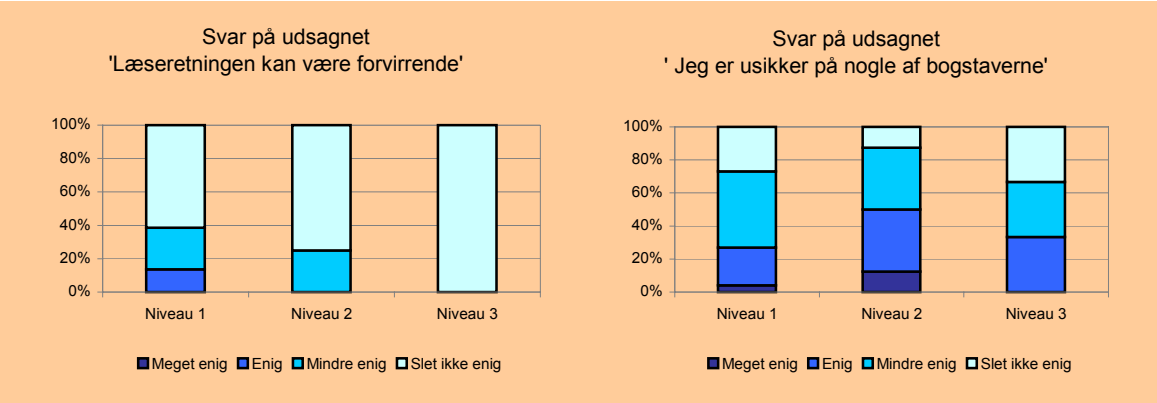
Faktisk findes der – omend beskedent – belæg for denne påstand, idet undersøgelser af børns og gymnasieelevers oplæsning på arabisk viser flere tilfælde af fejlagtig identifikation blandt yngre eller svage læsere, hvor forveksling netop fandt sted mellem bogstaver med samme grafiske base – fx ح, خ og ج (Azzam 1989; Abu-Rabia & Siegel 1995).

4.3 Sammenfatning

De præsenterede forskningsresultater illustrerer sammen med ovenstående komparative analyse af latinsk og arabisk skrift, at arabisk skrift nødvendigvis må udgøre et selvstændigt tilegnelsesaspekt ved etablering af læsefærdighed på arabisk, når man som L2-lørner møder op med en række veletablerede bogstavgenkendelsesstrategier fra et latinsk-skrevet L1.

Det ser således ud til, at etablerede scanningsstrategier fra L1 overføres til L2, uanset at dette kan være uhensigtsmæssigt. Hvad angår læseretning tyder meget på, at man med et latinsk-skrevet L1 udvikler en særlig stærk retningspræference, der er direkte modsat den, man udvikler med arabisk som L1. Præcis hvilken indflydelse dette forhold har på tilegnelsesforløbet står imidlertid ikke klart. I spørgeskemaundersøgelsen, der er udført i relation til dette forskningsprojekt og som præsenteres nærmere i kapitel 6,[13] ser man dog, at hos 37 danske lørnere af arabisk som L2 på tre forskellige færdighedsniveauer, opfattes læseretningen som betydeligt mindre problematisk end de arabiske bogstaver, hvilket fremgår af figur 4.9. Dette sammenholdt med de nævnte forskningsresultater, der viser, at børns retningspræferencer forholdsvis hurtigt ændrer sig, når de eksponeres for en ny

[13] Dette er en foregribelse af det empiriske studium, men vælges præsenteret kort allerede her, idet det har så oplagt relation til konklusionen på dette kapitel.



Figur 4.9 Uddrag af spørgeskemaundersøgelse vedrørende danske L2-lørneres opfattelse af deres læsetilegnelsesproces på arabisk (n=37).

tekstretning, indicerer, at læseretningens indflydelse må formodes at være marginal i sammenligning med de barrierer, der udgøres af bogstavarkitekturen. Som det ses af figur 4.9, opfatter de adspurgte lørnere tilsyneladende de arabiske bogstaver som en barriere selv meget sent i tilegnelsesforløbet. Kun én ud af tre på niveau 3, som repræsenterer færdiguddannede lørnere, er slet ikke enig i udsagnet 'Jeg er usikker på nogle af bogstaverne'. Set i relation til den præsenterede teori og den komparative analyse er dette måske ikke så overraskende. De bogstavdetektionskompetencer, man har opbygget gennem mange års læsning på L1, er nemlig i praksis ubrugelige i forbindelse med læsning på arabisk, og bogstavdetektionsnetværket skal udvides med et helt nyt sæt af bogstavkomponenter og 'spilleregler' for samspillet imellem dem. Denne opgave kompliceres yderligere af, at de arabiske bogstavers grafiske udtryk er præget af forholdsvis små grafiske forskelle. At skelne mellem bogstaverne synes derfor at kræve en mere detaljesøgende scanning, og i hvert fald kræver det opmærksomhed på komponenter, der tager sig helt anderledes ud, end dem vi på vores L1 per automatik scanner efter, når vi afgør at et bogstav er ét frem for et andet. Og det er muligvis et flerdimensionalt problem, for meget tyder på, at selvom vi får etableret nye funktioner i bogstavdetektionsnetværket, er vi påvirkede af en overførsel af allerede tilegnede funktioner, hvorfor vi automatisk overfører scannings- og detektionsmetoder, der har hjulpet os på L1, selvom de er uhensigtsmæssige at anvende på L2.

Teoretisk set synes arabisk skrift altså – uafhængigt af skriftsystemets sprogspecifikke dimension – at udgøre en barriere for tilegnelse af gode læsefærdigheder på arabisk som L2. Med henblik på at dokumentere dette empirisk og beskrive omfanget og karakteren af problematikken nærmere, inddrages bogstavarkitekturen i undersøgelsen af danske lørneres L2-læsning på arabisk, der præsenteres i kapitel 6. Forinden skal vi dog i det følgende kapitel se nærmere på, hvilken rolle den anden fremmede dimension ved det arabiske skriftsystem – den konsonantiske ortografi – spiller for L2-læsetilegnelsen på arabisk.

5. Ortografisk dimension: Det konsonantiske skriftsystem

Formålet med dette kapitel er at analysere hvilken rolle det konsonantiske skriftsystem spiller for den sekundære ordgenkendelse ved læsning på arabisk, med henblik på at bestemme hvilke implikationer arabisk ortografi har for læsetilegnelsen hos danske lørnere. Da der ikke foreligger nogen forskning i netop dette spørgsmål, vil der blive taget udgangspunkt i forskning vedrørende læsning på L1, ligesom forskningsresultater vedrørende hebraisk vil blive inddraget på baggrund af de retningslinjer, der i afsnit 3.4 defineres som forsvarlige ved sammenligning mellem ordgenkendelsesprocesser på arabisk og hebraisk. Kapitlet vedrører således ikke blot ordgenkendelse på arabisk, men på semitiske sprog generelt. Tilgængelige forskningsresultater vedrørende ordgenkendelse på de semitiske sprog som L1 sammenholdes til et teoretisk grundlag for, hvad der karakteriserer ordgenkendelse på disse sprog. Dette vil danne udgangspunkt for en analyse af, hvilke særlige forhold, der synes at gøre sig gældende for ordgenkendelse på semitiske sprog som L2.

5.1 Læsning på semitiske sprog som L1

Det konsonantiske skriftsystem, som anvendes ved skriftlig repræsentation af de semitiske sprog arabisk og hebraisk, indebærer at der på hvert af de to sprog findes to forskellige ortografier: vokaliseret og uvokaliseret. Mens den vokaliserede version – godt hjulpet af en forholdsvis konsekvent grafem-fonem korrespondance – skaber en meget gennemskuelig ortografi, er den uvokaliserede version på begge sprog kompleks og præget af homografi. Også det faktum, at ord på grund af rod-mønstersystematikken er præget af en relativ ensartethed, øger kompleksiteten. I dette afsnit gengives forskningsresultater, der på forskellig vis belyser, hvad disse aspekter betyder for ordgenkendelse, hvorefter de sammenfattes til en oversigt over, hvilke særlige forhold, der går sig gældende for læsning på semitiske sprog. Først præsenteres forskningsresultater, der vedrører direkte fysisk målelige aspekter af processen som øjenbevægelser og læsehastighed. Derefter præsenteres forskning, der undersøger forskelle i læsning af vokaliseret og uvokaliseret tekstmateriale, organiseret efter forskellige læseprocesstadier, dvs. ordbestemmelse (genkendelse af en bogstavstreng som et gyldigt ord), ordgenkendelse (forståelse af et ords meningsindhold) og læseforståelse (forståelse af meningsindholdet af et samlet tekststykke). Dernæst inddrages forskning i rødders og mønstres rolle for ordgenkendelse samt forskning i dysleksi, og endelig gives en analyse af, hvilken indflydelse diglossien kan have på læseprocessen. Til sidst sammenfattes resultaterne til en samlet analyse af, hvad der kunne karakterisere en specifik læseprocesteori for semitiske sprog.

Tværsproglige studier af øjenbevægelser og læsehastighed

Der findes et lille antal undersøgelser, der sammenligner øjenbevægelser og/eller læsehastighed ved læsning på arabisk eller hebraisk med læsning på et europæisk sprog. I en undersøgelse fra 1981 fandt man ved sammenligning mellem læsning af engelsk og hebraisk længere fixeringstider og kortere saccader ved læsning af hebraisk, men samme samlede læsehastighed for de to sprog (Pollatsek et al. 1981). En lignende undersøgelse af læsning på fransk og arabisk viste også samme læsehastighed på de to sprog samt flere – men ikke signifikant længerevarende – fixeringer på arabisk end på fransk. Dog holdt læserne på arabisk signifikant længere pauser, hvor øjnene stod stille. (Roman & Pavard 1987)

Forfatterne forklarede i begge undersøgelser dette med, at hebraiske/arabiske ord er mere informationsmæssigt kompakte end ord på de europæiske sprog.

En senere undersøgelse af Shimron & Sivan (1994) viser derimod stor forskel på læsehastighed. Her testedes voksne, universitetsansatte israelere i to separate undersøgelser. Forsøgspersonerne fordelte sig i to grupper med hhv. engelsk og hebraisk som modersmål. I den første undersøgelse læste gruppen med engelsk som modersmål engelsk signifikant hurtigere end gruppen med hebraisk som modersmål læste hebraisk. Ydermere viste det sig, at de hebraisksprogede forsøgspersoner læste engelsk hurtigere end de læste deres modersmål. I den anden undersøgelse fandt man imidlertid, at læsehastigheden på engelsk nok var hurtigere end læsning af uvokaliseret hebraisk men ikke hurtigere end læsning af vokaliseret hebraisk.

Hertil kommer en række studier, der viser at engelsk-hebraisk tosprogede børn i de første skoleår afkoder vokaliseret hebraisk hurtigere end den mere komplekse engelske ortografi, på trods af, at hebraisk er L2. (Geva et al. 1993; Geva & Wade-Woolley 1998; Geva & Siegel 2000).

Da de ældre undersøgelser er lavet med små populationer i forhold til de senere studier, synes disse resultater tilsammen at indicere, at læsehastigheden er langsommer jo mere kompleks en ortografi, tekstmaterialet anvender. I sammenligning med de latinske skriftsystemer (her repræsenteret af engelsk og fransk) og vokaliseret semitisk fremstår de uvokaliserede semitiske ortografier altså som den 'tungeste' variant, der kræver flest og mest langvarige øjenfixeringer og generelt læses langsomst.

Vokalernes betydning for ordbestemmelse, ordgenkendelse og læseforståelse

Navon & Shimron (Navon & Shimron 1981) var blandt de første, der gav sig til at undersøge, hvordan vokaliseringen influerer på ordbestemmelse på hebraisk. De lod forsøgspersoner lave ordbestemmelsesopgaver for vokaliserede ord og pseudoord og bad dem se bort fra vokalerne. Resultaterne viste, at hastigheden for ordbestemmelse var nedsat, når konsonanterne isoleret dannede et reelt ord, men vokaliseringen var fonetisk ukorrekt. De konkluderede, at selvom gode læsere på hebraisk ikke behøver vokalisering for at kunne foretage hurtig ordbestemmelse, kan de ikke umiddelbart ignorere vokaltegnene, når de er der (Navon & Shimron 1984). Ligeledes viste Frost (1994) ved et eksperiment, hvor hebraiske konsonantstrenge blev vist alene i 300 millisekunder, hvorefter vokaliseringen kom frem, at forsøgspersonerne var tilbøjelige til at vente på vokalerne, før de oplæste ordene, selvom de var unødvendige for afkodning. En undersøgelse af Barnea & Breznitz (1998) viser imidlertid, at det er selve bogstaverne, der er det centrale input i forbindelse med ordgenkendelse af vokaliseret skrift. I undersøgelsen brugte de par af vokaliserede hebraiske ord, som enten var identiske, havde samme bogstaver men forskellig vokalisering eller var forskellige i både bogstaver og vokalisering. De bad forsøgspersonerne om at markere, om de viste ordpar var identiske eller ej og fandt, at reaktionstiden for ord med forskellige bogstaver var signifikant hurtigere end reaktionstiden for de øvrige ord. Så selvom læsere ikke kan ignorere de diakritiske vokaltegn, når de er til stede, synes der ikke at være tvivl om, at de spiller en sekundær rolle i forhold til selve bogstaverne.

Hebraisk har været et yndet sprog at undersøge dualruteteorien på. I modsætning til sprog, der skrives med det latinske alfabet, kan man på hebraisk kontrollere graden af fonologisk

information i videnskabelige undersøgelser. Da kompetente læsere af hebraisk er trænede i at læse både uvokaliseret og vokaliseret tekst, kan de være deres egen kontrolgruppe for læsning af tekst med forskellig grad af grafisk repræsentation af fonemer, og sammenligning af læsning af de to former for tekst kan afdække, hvilken rolle fonemerne spiller for læseprocessen.[1] Denne form for undersøgelser viser ret entydigt, at ordgenkendelse ved læsning af hebraisk uvokaliseret tekst er domineret af ortografisk afkodning, og at læsning på hebraisk i denne henseende adskiller sig fra sprog, der anvender det latinske skriftsystem (Bentin et al. 1984; Koriat 1984; Frost et al. 1987). Dette skyldes ifølge Frost & Bentin (1992b:35), at præleksikalsk fonologisk behandling ikke er mulig ved læsning af uvokaliseret tekst. Den fonologiske information er ikke i tilstrækkeligt omfang til stede i den grafiske repræsentation, og læseren får dermed først adgang til fonologien, når han har konsulteret det indre leksikon. Samtidig synes afkodning af hebraisk vokaliseret tekst at være overvejende fonologisk (Frost 1994). Tværsproglige studier synes at bekræfte disse resultater. Fx viste en sammenligning af tid for ordbestemmelse ved læsning af den meget gennemskuelige serbokroatiske og den stærkt komplekse uvokaliserede hebraiske ortografi ingen signifikant forskel (Frost et al. 1987). Når læsere på hebraisk uden vokalisering kunne bestemme et ords gyldighed lige så hurtigt som læsere på serbokroatisk, hvor den grafiske repræsentation svarer fuldstændigt til den fonemiske, måtte det (jvf. dualruteteorien) skyldes, at læsere, foranlediget af den komplekse ortografi, anvender den leksikalske frem for den fonologiske vej til lexis. Endvidere viste Bentin (1989) i en undersøgelse af gentagelseseffekten af homografer og homofoner på hebraisk, at vokalisering influerer på, i hvor høj grad fonologisk og ortografisk repetition påvirker forsøgspersonernes reaktionstid ved ordbestemmelse. Han konkluderede, at vokalerne, når de er markeret, fremprovokerer en mere fonologisk medieret læseproces end læsning uden vokaler.

Nyere studier har imidlertid vist, at afkodning af uvokaliseret tekst ikke er udelukkende ortografisk. I disse studier anvendes en avanceret maskeringsteknik, hvor målord vises i ganske kort tid (14-16 ms) efterfulgt af maskerings-pseudoord, der i forskellig grad ligner målordet fonemisk og grafemisk. Her udnyttes det, at man på hebraisk har flere konsonanter, der repræsenterer samme fonem. Således kan ordet כפית [kapit] ('teske'), som består af bogstaverne [kpit] eksempelvis maskeres med pseudoordene כפיט [kpit], קפית [kpit] og קפיט [kpit] (alle fonemisk identiske men med forskellig grad af grafemisk overlap i form af antal bogstaver tilfælles med målordet) og sammenlignes med maskering med pseudoord med varierende fonemisk lighed med målordet, fx כפיז [kpiz] og כפזת [kpzt]. Disse studier viser, at jo mere maskerings-pseudoordet lydligt ligner målordet, jo bedre kan forsøgspersonerne gengive, hvilket målord, de har fået vist. Dette indicerer, at også uvokaliseret tekst i nogen grad aktiverer fonologiske processer (Gronau & Frost 1997; Frost & Yogev 2001). I en sammenligning mellem effekten af fonemisk og grafemisk maskering fremgår det dog, at grafemisk maskering havde markant større effekt end fonemisk maskering. Disse resultater afliver hypotesen om, at ordgenkendelse på hebraisk er udelukkende ortografisk, men understøtter, at ortografiske processer er dominerende ved afkodning af de uvokaliserede, semitiske ortografier.

Bentin & Ibrahim (1996) søgte at belyse problematikken med udgangspunkt i diglossi-situation på arabisk. De lod forsøgspersoner teste for læsning af ord fra MSA, skrevne ord

[1] Arabisk er lige så velegnet til dette formål som hebraisk, men som tidligere anført er læseforskning stort set ikke-eksisterende i den arabiske verden, mens der findes forholdsvis mange hebraisksprogede læseforskere både i Israel og i vestlige forskningsmiljøer.

fra den palæstinensisk talesprogsvariant og pseudoord, og bad dem bl. a. i en ordbestemmelsesopgave forkaste alle andre ord end dem fra MSA. Ordene var delvist vokaliseret, således at der ikke var nogen tilfælde af homografi. Da ordene fra talesproget blev forkastet endog langsommere end pseudoordene, konkluderede forfatterne, at ordgenkendelse er overvejende fonologisk medieret, idet forsøgspersonerne ikke ignorerede de dialektale ords fonologiske struktur, selvom det ville havde lettet ordbestemmelsesopgaven for dem. I en anden deltest skulle forsøgspersonerne kun forkaste pseudoordene og acceptere alle ord – fra MSA eller palæstinensisk talesprog – der lød som et ord, hvorved man mente at fremtvinge en fonologisk analyse af ordene. Da reaktionstiderne for at godkende ord fra MSA var længere i den test, hvor dialektale ord skulle forkastes end i testen, hvor fonologisk analyse var påtvunget, konkluderede forfatterne, at forsøgspersonerne heller ikke kunne ignorere den fonologiske struktur af ordene fra MSA. Selvom forfatternes konklusioner synes temmelig vidtrækkende – især i betragtning af, at de ikke afgrænser deres resultater til at vedrøre 'delvist vokaliseret arabisk', peger undersøgelsen i samme retning som ovennævnte konklusioner fra hebraisk: At vokalerne ikke lader sig ignorere, når de er markeret, og at vokaliseret skrift medierer fonologisk afkodning.

Flere har forsøgt at afgøre, hvorvidt vokalisering og dermed muligheden for en mere fonologisk behandling letter læseprocessen. Frost, Katz & Bentin (1987) mente på baggrund af de ovenfor nævnte tværsproglige undersøgelser – der tydeligvis gav massiv støtte til hypotesen om ortografisk kompleksitet (og dermed teorien om, at ordgenkendelse på hebraisk er overvejende leksikalsk)[2] – at vokalisering ikke letter ordgenkendelsen. Dog fandt Koriat (1985), at vokaler havde en beskeden positiv effekt på hurtig ordbestemmelse, men mest for lavfrekvens-ord. De tidligste studier var imidlertid flertydige, da man fandt, at tiden for ordbestemmelse var omtrent identisk for vokaliserede og uvokaliserede hebraiske ord, men at vokaler havde en positiv effekt på hastigheden ved højtlæsning (Navon & Shimron 1981; Shimron & Navon 1982; Bentin et al. 1984; Koriat 1984). Bentin, Bargai & Katz (1984) tog tid på højtlæsning og ordbestemmelse af uvokaliserede konsonantstrenge, hvoraf nogle kunne læses som ét og andre som to forskellige ord. Resultaterne viste, at den fonologiske flertydighed forsinkede højtlæsningen, men ikke ordbestemmelse. Disse resultater støttes af en undersøgelse af ordbestemmelse på arabisk, der viser, at vokalisering faktisk gør ordbestemmelse signifikant langsommere (Roman & Pavard 1987) samt af Bentin & Frost (1987), som testede tid for ordbestemmelse for homografer, der kunne læses som enten et høj-frekvens- eller et lav-frekvensord og de samme ord i vokaliseret form. Resultaterne viste, at ordbestemmelse var hurtigere for uvokaliserede homografer end for begge de vokaliserede (og altså ikke længere flertydige) alternativer. Vokaliseringen gjorde altså ikke ordbestemmelse hurtigere.

Endvidere viser en test af morfologisk repetition[3] med homografiske associationsord med både en højfrekvent og en lavfrekvent betydning, der blev vist umiddelbart før et målord med eller uden semantisk relation til en af associationsordets betydninger, at et semantisk relateret associationsord gav hurtigere ordbestemmelse af målordet, uanset om dette var relateret til associationsordets lav- eller højfrekvente betydning (Frost & Bentin 1992a). Da begge betydningers associationsværdi således kan måles, tyder det på, at betydningen af begge ord aktiveres inden endelig ordgenkendelse.

---

[2] Jvf. afsnit 2.3

[3] Jvf. kapitel 3, fodnote 4.

På baggrund af ovenstående resultater ræsonnerede de førende forskere på området, at ordbestemmelse synes at være baseret på en hurtig genkendelse, som ikke forudsætter detaljeret fonologisk analyse, selvom læseren savner vokaler til at skelne mellem visuelt identiske ord ved højtlæsning. (Frost et al. 1987; Frost & Bentin 1992a; Katz & Frost 1992). Nyere forskning har imidlertid sandsynliggjort, at der rent faktisk finder præleksikalsk fonologisk behandling sted ved læsning af hebraiske homografer. I test med amplitudeforvrængning[4] af hebraiske ord fandt Frost (1991), at den indflydelse, visuelle stimuli har på opfattelsen af støj som ord, når visuelle og auditive stimuli repræsenterer samme ord, kan konstateres uanset om de visuelle stimuli var vokaliserede eller uvokaliserede ord (Frost 1991). Dette blev af Frost & Kampf (1993) udnyttet til at undersøge, om homografiske ord aktiverer en eller flere akustiske former. Dette gjorde man ved med samme metode at teste, om den korrespondance mellem visuelle og auditive stimuli, som forsøgspersoner opfatter, når de to former for stimuli matcher, etableres uanset hvilket fonologisk alternativ (høj- eller lavfrekvent) der bruges til den auditive stimulus, når den visuelle stimulus er homografisk. Disse studier viser klart, at både højfrekvente og lavfrekvente ords akustiske former aktiveres ved eksponering for homografer. Dette indicerer ifølge forfatterne, at ordgenkendelse ved læsning af uvokaliserede hebraiske ord, der er homografer, aktiverer alle de fonologiske alternativer i læserens bevidsthed. Samlet set indicerer alle disse resultater, at både semantiske og fonologiske repræsentationer af både høj- og lavfrekvente betydninger aktiveres ved eksponering for hebraiske homografer.

Frost & Bentin ræsonnerede endvidere, at vokalisering ikke er essentiel for ordgenkendelse på hebraisk, da konsonantstrukturen er tilstrækkelig til at specificere et ord (Frost & Bentin 1992b:33). Flere af de ovenfor nævnte undersøgelser viser imidlertid, at vokalisering på hebraisk rent faktisk gør ordgenkendelse hurtigere, på trods af at flere tegn skal behandles. (Navon & Shimron 1981; Shimron & Navon 1982; Bentin et al. 1984; Koriat 1984; Koriat, A. 1985; Bentin & Frost 1987). Meget tyder på, at dette skyldes ophævelsen af homografien. Dette stemmer overens med en undersøgelse af Frost (1995), der viser, at reaktionstider på oplæsning af homografer bliver længere, jo mere flertydige ordene er.

Også senere studier, der tester læsning af tekst frem for isolerede ord viser, at forsøgspersoner på hebraisk læser vokaliseret tekst hurtigere og mere fejlfrit end uvokaliseret tekst (Shimron & Sivan 1994; Shimron 1999; Abu-Rabia 2001). Hvad angår arabisk er der fundet lignende resultater: Roman, Pavard & Asselah undersøgte læsning af homografer ved at lade testpersonerne læste arabiske sætninger, der var tvetydige i kraft af et initialt verbum, der, fordi det blev fremstillet i uvokaliseret form, kunne læses som enten i passiv eller aktiv modus. Ifølge forfatterne afklarede testpersoenerne tvetydigheden ved at danne hypoteser

---

[4] Amplitudeforvrængning (*amplitude-modulated noise*) er en testmetode udviklet af Frost, Repp & Katz (1988), hvor en akustisk repræsentation af et ord forvrænges til støj, hvor ordets præcise fonemiske og syllabiske struktur forsvinder, mens det stadig er muligt at tyde rytme- og trykstrukturer samt nogle konsonantklasser, hvor der fx kan skelnes mellem en klusil og en spirant, uden at man kan definere den eksakte konsonant. Denne støj kan betegnes som ordets 'akustiske form'. Når et ords akustiske form præsenteres sammen med det samme ord på skrift, opfatter forsøgspersoner støjen som en meningsfuld repræsentation af det aktuelle ord. Er de to former for stimuli derimod repræsentationer af forskellige ord, opfattes den akustiske stimulus blot som støj, også selvom stimuli ligner hinanden i det lydlige udtryk. Sådanne resultater opnås både med høj- og lavfrekvente ord men ikke med pseudoord, hvilket betragtes som dokumentation for, at et skrevet ord både genererer en automatisk fonologisk repræsentation af ordet **og** en repræsentation af ordets akustiske form. Ordets akustiske form er altså en af flere former for information, som en lytter bruger til at opnå auditiv ordgenkendelse – ligesom et ords grafiske form formodes at være det ved visuel ordgenkendelse (Frost et al. 1988; Frost 1991).

og på grundlag af kontekst afprøve disse – hvilket naturligvis var tidskrævende (Roman et al. 1985). Lignende resultater ses af en række mere udførlige undersøgelser af Abu-Rabia, hvor effekten af vokalisering og kontekst på højtlæsning testes. Disse undersøgelser viser klart, at vokalisering såvel som kontekst giver færre oplæsningsfejl, og at gode læsere har større udbytte af vokalisering og kontekst end svage læsere (Abu-Rabia & Siegel 1995; Abu-Rabia 1997b,1998,2001):

I den første undersøgelse (Abu-Rabia & Siegel 1995) testedes 20 gode og 20 svage læsere for afhængighed af kontekst ved læsning af henholdsvis vokaliserede og uvokaliserede ord. Resultaterne viser, at kontekst for såvel gode som svage læsere spiller en central rolle for læsning på arabisk. Undersøgelsens testpersoner – en gruppe ottendeklasses canadiske elever med arabisk som modersmål, som alle modtog modersmålsundervisning i arabisk - præsenteredes for 10 vokaliserede og 10 uvokaliserede sætninger. Alle sætninger indledtes af et verbum, hvilket er en almindelig sætningskonstruktion på formelt standardarabisk. Forsøgspersonen læste først sætningens første ord, mens resten af sætningen var dækket. Derefter præsenteredes han for den fulde sætning og havde mulighed for at rette sig selv ved genlæsning af sætningens første ord. Alle forsøgspersoner viste ved læsning af de uvokaliserede sætninger en naturlig forkærlighed for at læse de isolerede verber i den form, der har den største frekvens i almindelig tekst. Fx blev ordet أَكَّلَ [akkala] ('fodrede') af alle forsøgspersoner læst som أَكَلَ [akala] ('spiste'), når det blev præsenteret uden for konteksten أكل الولد كلبه [akkala al-walad kalbahu] (altså 'drengen fodrede sin hund' frem for 'drengen spiste sin hund'). På samme måde blev ordet لانَ [laana] ('blev blød') læst som لأنّ [liʔanna] ('fordi'), som naturligvis er langt hyppigere forekommende. Samtidig blev ord, der både repræsenterede den mest simple verbalform og bar den hyppigst forekommende mening i næsten alle tilfælde læst korrekt også uden for kontekst. Denne tendens slog endog igennem ved læsning af de vokaliserede sætninger, hvor især de svage læsere viste en stærk tendens til at ignorere fordoblingstegnet ّ (*shadda*), som ændrer ordets mening. Fx blev ordet سَلَّمَ [sallama] ('hilste') læst som سَلِمَ [salima] ('var rask/sund/sikker').

Ligeledes viser to senere undersøgelser (Abu-Rabia 1997b,1997a,1998), hvor gode og svage læsere præsenteres for vokaliserede og uvokaliserede enkeltstående ord og tekststykker, at vokalisering såvel som kontekst på arabisk giver færre oplæsningsfejl, og at de gode læsere har det største udbytte af begge dele.

I det seneste studium af Abu-Rabia (2001) testes voksne arabisk-hebraisk tosprogedes læsning af isolerede ord og tekststykker med og uden vokaltegn på begge sprog. Et overblik over resultaterne gives i figur 5.1. Disse resultater støtter Abu-Rabias tidligere undersøgelser, da højtlæsning af begge tekstformer på både arabisk og hebraisk er mere korrekt med end uden vokaltegn. Resultaterne for betydningen af kontekst er lidt mere spegede. På arabisk er korrektheden af højtlæsning af uvokaliserede ord i kontekst signifikant bedre end læsning af uvokaliserede isolerede ord (ligesom i de tidligere undersøgelser), men det samme er ikke tilfældet på hebraisk. Det er et overraskende resultat i betragtning af, at betydningen af kontekst ved læsning på uvokaliseret hebraisk som før nævnt fremgår klart af tidligere undersøgelser (Koriat 1984; Koriat 1985). Omvendt viser sammenligningerne af vokaliseret skrift, at kontekst ved læsning af vokaliseret tekst har en effekt på hebraisk men ikke på arabisk. Mens resultatet for arabisk kan forklares med, at effekten af den gennemskuelige ortografi udvisker kontekstens betydning for afkodning, er det særdeles overraskende, at kontekstens effekt på hebraisk slår igennem i sammenligningen mellem læsning af de vokaliserede frem for de uvokaliserede ord og

tekststykker. Man skulle jo forestille sig, at læseren trækker mere på kontekst ved læsning af uvokaliseret skrift, hvor homografien er udtalt, hvilket resultaterne fra arabisk jo også klart afspejler.[5]

Arabisk

	+ vokaler - kontekst	- vokaler - kontekst	+ vokaler + kontekst	- vokaler + kontekst
+ vokaler - kontekst				
- vokaler - kontekst	x			
+ vokaler + kontekst	is	---		
- vokaler + kontekst	---	x	x	

Hebraisk

	+ vokaler - kontekst	- vokaler - kontekst	+ vokaler + kontekst	- vokaler + kontekst
+ vokaler - kontekst				
- vokaler - kontekst	x			
+ vokaler + kontekst	x	---		
- vokaler + kontekst	---	is	x	

Figur 5.1: Korrelationsmatrix over signifikante forskelle mellem læsning af isolerede ord (- kontekst) og tekststykker (+ kontekst) på vokaliseret og uvokaliseret arabisk og hebraisk.
x = signifikant forskel; is = ikke signifikant forskel; --- = sammenhæng ikke refereret.
(Abu-Rabia 2001)

Abu-Rabias læseeksperimenter kan kritiseres metodisk, blandt andet fordi fuld vokalisering med korrekte kasusendelser, som anvendes ved højtlæsning på arabisk, langt fra er nødvendig for læseforståelsen og normalt undlades i dagligdags skrift og tale.[6] Abu-Rabia er i sine seneste undersøgelser selv opmærksom på denne problematik og forsøger at eliminere den ved at ignorere fejl i eller udeladelser af kasusendelser (Abu-Rabia 1997b:71; 2001:48). I almindelighed hører kasusendelser sig imidlertid til ved højtlæsning, og den betydelige kognitive byrde, den løbende syntaktiske analyse lægger på læseren, når der læses uvokaliseret tekst, kan tænkes at influere på det samlede oplæsningsresultat, hvad enten egentlige kasusfejl medregnes eller ej. Det fremgår ikke af metodebeskrivelserne for disse eksperimenter, om forsøgspersonerne er oplyst om, at kasusendelserne kan ignoreres, og selv om det skulle være tilfældet, er de syntaktiske overvejelser muligvis en automatiseret proces ved højtlæsning, også selvom de ikke foretages bevidst.

På trods af denne kritik synes der at være belæg for, at vokalisering er en hjælp til afklaring af homografer ved højtlæsning af både arabisk og hebraisk. Endvidere er effekten af kontekst ret overbevisende hvad angår arabisk men lidt mindre klar for læsning på hebraisk. Effekten af kontekst er ikke signifikant i den omtalte undersøgelse af Abu-Rabia (2001), men begge de tidligere nævnte studier af Koriat (1984, 1985b) viser, at kontekst i langt højere grad end vokaler letter afkodningen. Desuden fandt Roman & Pavard (1987), at vokalisering påvirker læsehastigheden langt mere ved ordbestemmelse end ved læsning af tekststykker, hvilket tyder på, at læsere i meget høj grad udnytter kontekst.

---

[5] De overraskende resultater hvad angår læsning på hebraisk hænger muligvis sammen med at hebraisk i denne undersøgelse er testpersonernes andetsprog, hvilket svækker sammenligneligheden mellem disse og de øvrige præsenterede forskningsresultater.

[6] - med en enkelt undtagelse for den ubestemte endelse i akkusativ, som anføres i skrift og desuden også anvendes ved nogle faste udtryk i daglig tale.

Flere af ovenstående undersøgelser viser endvidere, at også ordfrekvens spiller en afgørende rolle ved afkodning af homografer. I Koriats undersøgelse ser vi, at vokalisering først og fremmest støtter ordbestemmelse af lavfrekvensord (Koriat 1985), ligesom Bentin & Frost (1987) rapporterer, at ordfrekvens har en afgørende indflydelse på hastigheden for både ordbestemmelse og ordgenkendelse. I deres undersøgelse blev lavfrekvente ord, også i vokaliseret form, genkendt signifikant langsommere end både vokaliserede og uvokaliserede højfrekvente ord. De ovennævnte studier af Frost & Bentin (1992a), Frost & Kampf (1993) og Frost (1995) viste ligeledes, at selvom både højfrekvente og lavfrekvente ord aktiveres ved eksponering for homografiske ord, så er aktiveringen af den højfrekvente betydning stærkere end aktiveringen af den lavfrekvente, idet associationseffekten af højfrekvente betydninger er stærkere, ligesom højfrekvente fonologiske repræsentationer bevares længere tid i korttidshukommelsen end de lavfrekvente alternativer. Endvidere fandt Bentin & Ibrahim (1996) ved både ordbestemmelse og ordgenkendelse en langt større ordfrekvenseffekt på delvist vokaliseret arabisk, end hvad der normalt ses på europæiske sprog. Dette sammenholdt med Abu-Rabias resultater (Abu-Rabia & Siegel 1995) tyder på, at afkodning af lavfrekvente ord er særligt krævende, men lettes betydeligt, hvis der findes andre former for information – det være sig vokalisering eller kontekst. I praksis kan man da også observere, at vokalisering på både arabisk og hebraisk af og til benyttes på enkeltstående ord, der fordi de er homografer og lavfrekvente eller i den tekstuelle sammenhæng tvetydige, gør teksten vanskelig at afkode. Dette kan sammenfattes til, at såvel ordfrekvens som vokaler og kontekst kan udnyttes positivt i ordgenkendelsesprocessen ved homografi, og at kontekst synes at være den af de tre mulige ressourcer, der udnyttes bedst.

Ovenstående studier kan kritiseres for at fokusere ensidigt på højtlæsning, idet flere undersøgelser viser, at der hverken på arabisk eller hebraisk er nogen signifikant sammenhæng mellem flydende højtlæsning og læseforståelse (Abu-Rabia 2001 100; Saiegh-Haddad 2003a). Dette står i modsætning til tilsvarende forskning vedrørende læsning på engelsk, som viser så klar en sammenhæng mellem højtlæsning og læseforståelse, at højtlæsning synes at kunne anvendes som en direkte indikator for læsekompetence, herunder læseforståelse (Fuchs et al. 2001). Denne markante forskel er i lyset af ovennævnte forskningsresultater ikke overraskende: Det store behov for at udnytte kontekst på grund af den massive homografi ved læsning af uvokaliseret tekst må alt andet lige virke obstruerende på flydende højtlæsning, og på arabisk kompliceres tingene yderligere af det traditionelle krav om brug af kasusendelser ved højtlæsning. Omvendt er korrekt og flydende højtlæsning af vokaliseret tekst på både arabisk og hebraisk mulig uden egentlig læseforståelse. Dette ses hos hebraisksprogede begynderlæsere (Gholamain & Geva 1999; Geva & Siegel 2000), arabiske førskolebørn og hos muslimer, der ikke kan arabisk, men læser højt fra koranen uden nogen form for forståelse.

Men også studier af læseforståelse viser, at vokalisering støtter læseprocessen. I det før omtalte studium af Shimron & Sivan (1994) fandt de i begge eksperimenter, at læseforståelsen var bedre ved læsning af vokaliseret frem for uvokaliseret tekst. Shimron (1999) fandt også i en undersøgelse af hebraisksprogede begynderlæsere, at visse prosatekster blev forstået og deres indhold husket bedre, når de var vokaliserede. Ligeledes fandt Abu-Rabia (1999) i en undersøgelse af læseforståelse blandt arabisksprogede børn i 2. og 6. klasse, at børnene på begge klassetrin havde bedre læseforståelse ved læsning af vokaliseret frem for uvokaliseret tekst. Også et andet studium af Abu-Rabia (2001) med

gymnasieelever viste, at læseforståelsen på både hebraisk og arabisk var bedre ved læsning med vokalisering end uden.

Med hensyn til læsehastighed ved læsning efter forståelse[7] viser ét ud af de to eksperimenter i undersøgelsen af Shimron & Sivan en lille men ikke signifikant højere læsehastighed ved 'lydløs' læsning af vokaliserede i forhold til uvokaliserede hebraiske tekster (Shimron & Sivan 1994). Ligeledes indicerer de førnævnte undersøgelser af tid for ordbestemmelse og højtlæsning, at vokalisering har en positiv effekt på hastigheden ved læsning efter forståelse. I modsætning til disse resultater står en enkelt kilde – en opfølgende undersøgelse af det førnævnte studium af Roman & Pavard tilbage fra 1987. Ved en sammenligning af læsning af tekststykker på ca. 100 ord med og uden vokaler observerede de, at vokalisering reducerede læsehastigheden signifikant (Roman & Pavard 1987). Men denne undersøgelses resultater må altså siges at stå meget alene.

Rødders og mønstres rolle for ordgenkendelse

Som beskrevet i afsnit 3.5, findes der talrige psykolingvistiske studier, hvis resultater indicerer, at rødder og mønstre på semitiske sprog har separate repræsentationer hos indfødte sprogbrugere. Således kan førskolebørn manøvrere sprogligt med rødder og mønstre som selvstændige morfemer (Clark & Berman 1984; Berman 2002; Ravid 2002), hjerneskadede afasi-patienters metatesefejl berører kun rødder og aldrig mønstre (Obler 1989; Berg & Abd-El-Jawad 1996; Prunet et al. 2000), og i ordbestemmelsesopgaver viser forsøgspersoner følsomhed overfor fonotaktiske begrænsninger i rødder (Berent et al. 2001; Frisch & Zawaydeh 2001; Berent & Shimron 2002), ligesom de er mere følsomme over for morfologiske end semantiske associationsord ved morfologisk repetition (Bentin & Feldman 1990; Feldman & Bentin 1994).

Hertil kommer undersøgelser, der specifikt i relation til ordgenkendelse søger at afdække, om ord – jvf. parallelruteteorien – har selvstændige 'opslag' i det indre leksikon, eller om de har rod- og/eller mønsterbaserede repræsentationer:

Frost, Forster & Deutsch (1997) gennemførte en række test af morfologisk repetition med hebraiske nomenformer. For at teste rødders og mønstres leksikalske status, undersøgte de, hvordan repetition af henholdsvis rødder og mønstre påvirker hastighed for ordbestemmelse og genkendelse af ord og pseudoord: Først eksponerede de testpersonerne for et associationsord, der enten var identisk med målordet, havde samme rod som målordet, havde samme mønster som målordet, eller hverken havde rod- eller mønsterrelation til målordet. Umiddelbart derefter blev de eksponeret for målordet, og hastighed for ordbestemmelse (tryk på en 'ja' eller 'nej' –knap) eller ordgenkendelse (højtlæsning) blev målt. Associations- og målord havde i alle situationer samme antal bogstaver, og alle ord var højfrekvente og uvokaliserede, men ingen af dem var homografer. Alle anvendte pseudoord var konstrueret af en pseudo-rod og et validt mønster. Forsøget viste, at når associations- og målord havde samme mønster, var der ingen signifikant repetitionseffekt, men når associations- og målord havde samme rod, havde forsøgspersonerne signifikant hurtigere ordbestemmelse og ordgenkendelse end i kontrolsituationen, hvor der ingen relation var mellem associations- og målord. Reelle målord, der var identiske med associationsord, blev dog altid genkendt hurtigst, hvorimod

[7] 'Læsning efter forståelse' skal her ses som en modsætning til højtlæsning.

genkendelse af de identiske pseudoord aldrig var signifikant hurtigere end reaktionen i de øvrige situationer, hvilket synes at bekræfte, at den målte associationseffekt vedrører leksikalsk association. Forfatterne konkluderede, at hebraiske nominers rødder har en selvstændig leksikalsk status hos hebraisksprogede læsere, men at det samme ikke er tilfældet for mønstrene.

I en opfølgende undersøgelse lavede Deutsch, Frost & Forster (1998) samme slags test med hebraiske verber. Ud over at konstruere stimuli som i den foregående test, blot med verbalmønstre, lavede de yderligere en test af mønster-repetition, hvor alle målord var pseudoord, blot havde halvdelen af disse både en valid rod og et korrekt verbal-mønster, men i en sammensætning der gav et ikke-eksisterende ord. Resten af pseudoordene var ligesom i den foregående undersøgelse konstrueret på baggrund af pseudo-rødder. I ordbestemmelsesdelen af denne test skulle forsøgspersonerne kun forkaste ord med pseudo-rødder som ugyldige. Resultatet viser en markant signifikant effekt af alle de tre former for morfologisk repetition på både ordbestemmelse og ordgenkendelse, og forskellene i resultaterne for de to mønster-test er absolut marginale. Forfatterne konkluderer, at for de hebraiske verbers vedkommende har både rødder og mønstre selvstændig leksikalsk status. Og dét at morfologisk dekomposition af verbale former finder sted, selv når et ord består af en ikke-eksisterende kombination af rod og mønster, forstærker indikationerne af, at verbale mønstre behandles selvstændigt i forbindelse med visuel ordgenkendelse.

Abu-Rabia & Awwad (2004) dublerede det første af de to ovennævnte studier (Frost et al. 1997) med en undersøgelse af morfologisk repetition af nomenformer på arabisk. Abu-Rabia & Awwad rapporterer, at de ligesom Frost et al. fandt en lille, men ikke statistisk signifikant, forøgelse af hastigheden for ordbestemmelse og ordgenkendelse ved eksponering for associationsord med samme mønster som målordet. Associationsord med samme mønster havde dog signifikant indflydelse på, hvor gode forsøgspersonerne var til at forkaste pseudoord i ordbestemmelsesopgaven, idet fejlprocenten kun var halvt så stor, når associations- og målord havde samme mønster, end når der ingen relation var mellem de to ord. I modsætning til hvad der var tilfældet hos Frost et al. gav heller ikke forudgående eksponering for ord med samme rod som målordet signifikant bedre resultater, og fejlprocenten for at forkaste stimuli som pseudoord var helt modsat mønster-testen højere, når forsøgspersonerne forinden var blevet eksponeret for associationsord med samme rod. Abu-Rabia & Awwad konkluderer på den baggrund:

> *"We found that words in the nominal derivational morphology of Arabic are represented in the mental lexicon as whole words, and there is no need for morphological decomposition prior to lexical decoding."* (Abu-Rabia & Awwad 2004:16)[8]

[8] Da resultatet i skrivende stund er under udgivelse, henviser sidetalshenvisninger i relation til denne kilde til forfatterens eksemplar af artiklen.

Sammenligner man resultaterne fra de to studier nærmere, finder man imidlertid en bemærkelsesværdig diskrepans i resultaterne, som i deres fulde form ses i figur 5.2. I figur 5.3 er forskellene i de to undersøgelser vist som forskelle i ms og i procent mellem reaktionstider med og uden mønster- og rodrelation mellem associations- og målord. Mens en reduktion i reaktionstid på blot 13 ms, svarende til godt 2%, ved eksponering for rod-relaterede, gyldige associationsord er statistisk signifikant i den hebraiske test, er en reduktion på 78 ms, svarende til hhv. 13,6% og 12,9% i den arabiske test med mønster-relaterede associationsord og associationspseudoord ikke nok til at udgøre en statistisk signifikant forskel. Dette tyder på langt større standardafvigelser i den arabiske test, og som det fremgår af figur 5.2, er også gennemsnitsværdierne langt mere ensartede i den hebraiske end i den arabiske test. I den arabiske test er alle forskellene i reaktionstider mellem identiske og ikke relaterede ordpar signifikante, så associationsværdien af de identiske ord kunne altså godt måles, men mens også associationsværdien af identiske pseudoord i rod-testen var signifikant, var dette ikke tilfældet i mønster-testen. Dette bevirker, at resultaternes gyldighed hvad angår måling af leksikalsk association i den arabiske test faktisk

		Hebraisk			Arabisk		
	MØNSTRE						
		ordbestemmelse	fejl	ordgenkendelse	ordbestemmelse	fejl	ordgenkendelse
ord	identisk	541	5,2	543	605	9	416
ord	mønster-relateret	578	9,1	571	703	8	495
ord	ikke relateret	579	8,2	574	726	10,5	573
pseudoord	identisk	617	8,4	603	790	15	498
pseudoord	mønster-relateret	627	8,8	632	836	12,5	527
pseudoord	ikke relateret	624	9,6	631	850	22	605
	RØDDER						
		ordbestemmelse	fejl	ordgenkendelse	ordbestemmelse	fejl	ordgenkendelse
ord	identisk	566	6,2	583	615	12,5	443
ord	rod-relateret	578	6,7	594	686	12	644
ord	ikke relateret	591	6,4	607	708	12	631
pseudoord	identisk	640	8,9	653	732	16,8	504
pseudoord	rod-relateret	642	9,8	663	763	26	728
pseudoord	ikke relateret	639	8,3	661	813	17	682

Figur 5.2 Reaktionstider for ordbestemmelse og ordgenkendelse (ms) samt fejl (%) ved ordbestemmelse hos Frost, Forster & Deutsch (1997) (Hebraisk) og Abu-Rabia & Awwad (2004) (Arabisk).

	Hebraisk		Arabisk	
MØNSTRE	ordbestemmelse	ordgenkendelse	ordbestemmelse	ordgenkendelse
ord	1 (0,2 %)	3 (0,5 %)	23 (3,2 %)	78 (13,6 %)
pseudoord	- 3 (-0,5 %)	1 (0,2 %)	14 (1,6 %)	78 (12,9 %)
RØDDER	ordbestemmelse	ordgenkendelse	ordbestemmelse	ordgenkendelse
ord	**13 (2,2 %)**	**13 (2,1 %)**	22 (3,1 %)	- 13 (-2,1 %)
pseudoord	- 3 (-0,5 %)	- 2 (-0,3 %)	50 (6,2 %)	- 46 (-6,7 %)

Figur 5.3 Forskel i reaktionstider (ikke relateret – mønster-/ord-relateret) i ms og % hos Frost, Forster & Deutsch (1997) (Hebraisk) og Abu-Rabia & Awwad (2004) (Arabisk). Signifikante forskelle er markeret med lys baggrund.

fremstår som tvivlsom. Især i sammenligning med den tilsvarende hebraiske test, hvor gennemsnitsværdierne er langt mere regelmæssige, synes resultaterne af den arabiske undersøgelse uklare. Undersøgelsens statistiske svagheder følges op af Abu-Rabia & Awwads ovenfor citerede bastante konklusion, som der under ingen omstændigheder er belæg for at drage: At der ikke kan måles en statistisk signifikant associationsværdi, betyder jo ikke nødvendigvis, at den ikke eksisterer. Et kriterium for udvælgelse af stimuli er, at ordene er højfrekvente, og som forfatterne selv pointerer, antages det ofte, at højfrekvente ord afkodes overvejende ortografisk, hvilket indebærer at de i højere grad end lavfrekvente ord genkendes på baggrund af ordets samlede visuelle form, mens de øvrige ressourcer som fonologi og morfologi fortrinsvis aktiveres, når en sådan automatisk genkendelse ikke kan etableres (Abu-Rabia & Awwad 2004:18). Samtidig begrænser forfatternes analyse af de markante forskelle mellem resultaterne i deres egen undersøgelse og den hebraiske – på trods af, at den fungerer som model for det valgte testdesign – til følgende:

> *"The letters in Arabic are connected to one another, except for five letters that are not connected from the left; such a writing system seems to be cognitively demanding so that readers overlook morphological decomposition of words. Such an orthographic characteristic makes the Arabic language different even from Hebrew."* (Abu-Rabia & Awwad 2004:16)

Det er naturligvis muligt, at det rent faktisk er kvalitative forskelle mellem hebraisk og arabisk skrift, der er anledning til de forskellige resultater, men der kan også tænkes at være andre grunde – fx en for uensartet population eller mangelfulde afkodningsfærdigheder hos respondenterne i den arabiske undersøgelse. De uregelmæssige resultater og ikke mindst de meget store fejlprocenter i den arabiske undersøgelse kunne tyde på, at netop sådanne metodetekniske problemer ligger bag. Forfatterne til den arabiske undersøgelse gør desværre heller ikke meget ud af at analysere de forskelle, der viser sig signifikante: Hvorfor giver mønster-relaterede associationsord færre ordbestemmelsesfejl for mål-pseudoord end associationsord uden relation? Og hvorfor giver rod-relaterede associationsord flere ordbestemmelsesfejl for ord i samme kategori? En forklaring kunne være, at roden faktisk spiller en afgørende rolle for, om stimuli opfattes som et ord: Når associationsordet har samme mønster som målordet, må forskellen på de to ord nødvendigvis ligge i roden. Hvis man forestiller sig, at respondenterne opfatter mønsteret som 'skabelon' for et ord, kan det forklare, at mønster-relaterede associationsord giver mere overskud til at fokusere på roden og korrekt forkaste den som en pseudo-rod. I situationen med rod-relation (hvor associationsordet er defineret til altid at være den pseudo-rod, der bruges til at konstruere mål-pseudoordet) kan man forestille sig, at det er den forudgående aktivering af pseudo-roden sammenholdt med det valide mønster i mål-pseudoordet, der får forsøgspersonerne til i så høj grad fejlagtigt at bedømme stimuli til at være valide ord.

I hvert fald synes det på baggrund af de uensartede reaktionstider, repetitionseffekten på pseudoord og den signifikante gentagelseseffekt på fejlprocenten i ordbestemmelses-opgaverne – i denne ene undersøgelse, der vedrører arabisk – overilet helt at forkaste rødder og mønstre som værende af betydning for leksikalsk behandling af arabiske nominer. Under alle omstændigheder synes det i det mindste for læsning på hebraisk sandsynliggjort, at den semitiske morfologiske struktur spiller en ikke ubetydelig rolle for ordgenkendelse. At gentagelseseffekt for mønstre kun kan registreres på verbalformer skyldes muligvis, at der er færre verbalmønstre, og at de er mere fonologisk regelmæssige end nomenmønstrene. Om den morfologiske struktur spiller samme rolle på arabisk kan på baggrund af ovennævnte enkeltstående undersøgelse af Abu-Rabia & Awwad (2004) umiddelbart hverken be- eller afkræftes. En gentagelse af Abu-Rabia & Awwads test og en dublering af undersøgelsen af

Deutsch et al.(1998) med morfologisk repetition af verbalformer på arabisk, ville i den sammenhæng naturligvis være interessante, og sidstnævnte er ifølge Abu-Rabia & Awwad undervejs, men på nuværende tidspunkt altså endnu ikke gennemført (Abu-Rabia & Awwad 2004).

En helt anden slags undersøgelse af Ryan & Meara (1991) viser imidlertid, at konsonant-strukturer sandsynligvis er centrale for afkodning på arabisk. Forsøgspersonerne i undersøgelsen var arabiske lørnere, der modtog undervisning i engelsk. Deres læsefejl på engelsk blev sammenlignet med læsefejl hos lørnere med andre modersmål. Resultaterne viser, at de arabiske lørnere i modsætning til lørnere med europæiske L1 havde en klar tendens til at ignorere vokaler og overfokusere på konsonanternes betydning ved læsning på engelsk. Således forvekslede de ofte engelske ord med samme konsonanter, så fx *'biscuit'* blev læst som *'basket'*, *'spread'* som *'separate'*, *'pulls'* som *'plus'* og '*moments'* som '*mountains*'. Tilsyneladende konstruerede de altså fonologiske repræsentationer af ordene på baggrund af ordenes konsonantstrukturer. Dette tyder netop på en overførsel af en afkodningsstrategi baseret på rodkonsonanter, endda i et tilfælde, hvor det er langt fra hensigtsmæssigt.
På den baggrund må det – sammenholdt med de præsenterede resultater fra test med morfologisk repetition – betragtes som sandsynligt, at den konsonantiske rod gives separat opmærksomhed ved afkodning på semitiske sprog.

Studier af dysleksi

Alle eksisterende studier af voksnes dysleksi hos semitisksprogede er sammenligninger af deres læsefejl på hhv. et semitisk og et europæisk sprog, hhv. arabisk-fransk (Berg & Abd-El-Jawad 1996; Prunet et al. 2000) og hebraisk-engelsk (Obler 1989). Som eksemplificeret i afsnit 3.5 viser disse undersøgelser, at dyslektikeres læsefejl på semitiske sprog udelukkende går ud over rødderne, mens mønstrene ikke berøres.

Hvad angår læseproblemer hos børn, findes der nogle få undersøgelser, der især sigter mod at undersøge, hvad der nærmere bestemt på disse sprog karakteriserer dysleksi. Læse-vanskeligheder hos børn med europæiske L1 er et forholdsvis grundigt undersøgt problemområde, og det er efterhånden veldokumenteret, at børns læsehandicaps i mange tilfælde ikke kan forklares med hverken dårlig intelligens eller visuel dysfunktion, men gerne skyldes mangel på forskellige former for specifikt sproglige kompetencer (Dalby et al. 1992:107-113), og at mangel på fonologisk opmærksomhed er den primære grund til ordblindhed (Bentin 1992). Med det som udgangspunkt har man forsøgt at afdække, hvilke former for manglende sproglige kompetencer, der giver sig udslag i læse-vanskeligheder på semitiske sprog. Sprogenes uvokaliserede ortografier er jo – til forskel fra fx engelsk – komplekse netop i kraft af manglende fonologisk information, hvorfor det er forventeligt at fonologisk opmærksomhed spiller en mindre rolle for læsefærdighed.

Ben-Dror, Bentin & Frost (1995) sammenlignede en gruppe læsehandicappede børn med to grupper normalt læsende børn, hvor den ene matchede aldersmæssigt, mens den anden matchede ordforrådsmæssigt og således bestod af børn, der var yngre end de øvrige. Børnene, der alle var hebraisksprogede, blev testet på både perception og produktion af en række forskellige sproglige parametre, herunder evne til at kategorisere ord semantisk, identificere fonemer og bestemme morfologisk relation mellem ord på baggrund af rod og mønster og på tværs af bøjningsformer. I perceptionstestene var der hvad angår de nævnte parametre ingen signifikant forskel på de to normalt læsende grupper, mens de

læsehandicappede børn klarede sig markant dårligere i alle tre test. Den mest afgørende forskel på grupperne viste sig i den morfologiske test, hvor forskellen mellem den læsehandicappede og den jævnaldrende gruppe var ca. tre gange så stor som i de øvrige test. I produktionstestene var den læsehandicappede og den ordforrådsmæssigt matchende gruppe mere lige, mens gruppen af normalt læsende, aldersmæssigt matchende børn klarede sig markant bedre. Gruppen af læsehandicappede var stadig den svageste, undtagen i den semantiske test, hvor gruppen af yngre børn klarede sig dårligst. Også i denne test var forskellen mellem de læsehandicappede og de normalt læsende, jævnaldrende børn størst i den morfologiske test. I både perceptions- og produktionstesten var det i de morfologiske test relationen mellem rod og mønster, der skabte de store vanskeligheder for de læsehandicappede. Deres evne til at bestemme morfologisk relation på tværs af bøjningsmønstre var næsten lige så god som de andre gruppers, og i perceptionstesten var forskellen grupperne imellem slet ikke statistisk signifikant.

Abu-Rabia, Share & Mansour (2003) gennemførte en lignende undersøgelse med tre grupper arabisksprogede børn. Blot var den yngre kontrolgruppe her matchet på læsefærdighed frem for ordforråd, og der blev her testet på parametrene fonologi, syntaks, ortografi, morfologi, verbal hukommelse og visuel behandling. Mens de læsehandicappede børn ikke viste ringere kompetencer end de andre grupper hvad angår ortografi og syntaks i tale, viste de ringere resultater i testene af syntaks i skrift, verbal hukommelse og visuel behandling. Resultatet af den visuelle test viser, at manglende evne til visuel forarbejdning synes at medvirke til læsehandicap hos arabisksprogede. Dette støttes af et studium af læseudvikling hos hebraisksprogede børn, der viser, at visuel korttidshukommelse er et afgørende parameter for læsefærdighed på hebraisk (Meyler & Breznitz 1998). Disse resultater står i kontrast til, hvad der er tilfældet for læsefærdighed på europæiske sprog, hvor mangelfuld visuel behandling i et stort antal undersøgelser synes at være elimineret som gyldig indikator for almindelig ordblindhed hos børn (Dalby et al. 1992:107-113). Ligesom hos Ben-Dror et al. udviste de læsehandicappede hos Abu-Rabia et al. (2003) markant dårligere morforlogiske kompetencer, og til forskel fra hvad Ben-Dror et al. fandt, havde de også ringere resultater i de fonologiske test. Denne forskel skyldes muligvis de meget forskellige testmetoder. Hos Ben-Dror et al. bad man i perceptionstesten børnene om at svare 'ja' eller 'nej' til om et ord begyndte med et bestemt fonem, og i produktions-testen bad man dem om selv at sige ord, der begyndte med et defineret fonem. Hos Abu-Rabia et al. var opgaven langt mere kompleks. Her blev børnene bedt om at slette initiale hhv. finale fonemer fra ord eller pseudoord og producere resten af ordet mundtligt. Denne opgave kræver et helt andet abstraktionsniveau, hvorved andre kognitive kompetencer som fx hukommelse, der i samme undersøgelse viser sig at være svagere hos denne gruppe, muligvis får større indflydelse på udfaldet af testen. I en undersøgelse af arabisk-engelsk tosprogede børn i Canada fandt Abu-Rabia & Siegel (2002) også fonologiske svagheder hos de af børnene, der havde læsevanskeligheder, men i mindre grad end hos tilsvarende etsprogede børn. I denne undersøgelse var den fonologiske test dog baseret på bestemmelse af, hvorvidt lyden af et pseudoord var 'ordlignende' – en test der giver mening på engelsk men er stærkt tvivlsom på arabisk, dels pga. de snævre rammer, morfologien stiller for, hvordan ord kan tage sig ud, dels pga. de barrierer, forskelle mellem tale- og skriftsprog repræsenterer (Abu-Rabia & Siegel 2002:666).

Også Abu-Rabia & Taha (2004) sammenlignede dyslektiske, arabisksprogede børn med normale jævnaldrende og yngre børn matchet på læsekompetence. I dette studium sammenlignes læse- og stavefejl hos de tre grupper. Resultaterne viser, at de dyslektiske

børn laver samme antal og former for fejl som gruppen af yngre børn, og at deres fejl sammenlignet med fejlene hos de jævnaldrende børn afspejler større fonologiske afkodningsvanskeligheder, i form af at de i meget høj grad negligerer korte vokaler. Morfologisk betingede læsefejl, hvor børnene fastholder ordets rod men applicerer et forkert mønster er dog den mest udbredte fejltype hos alle grupperne.

Disse studier dokumenterer, at semitisksprogede dyslektiske børn, lige så vel som de europæisksprogede dyslektiske børn, udviser mangel på flere forskellige former for sproglige kompetencer. Men i forhold til hvad der gælder på europæiske sprog, synes det i langt højere grad at være mangel på morfologiske kompetencer, der udgør en hindring for etablering af læsefærdigheder på semitiske sprog. Desuden synes det dokumenteret, at læsning på arabisk og hebraisk i modsætning til læsning på de europæiske sprog stiller store krav til visuo-spatial behandling, hvilket muligvis ikke blot – som nævnt i kapitel 4 – vedrører behandling på bogstavniveau, men også kan spille en rolle for behandling af hele ord i kraft af den ikke-lineære morfologiske struktur og for begynderlæsnings vedkommende også i kraft af placeringen af tegnene for korte vokaler i et selvstændigt grafisk niveau i forhold til bogstaverne (Abu-Rabia et al. 2003:439).

Hertil kommer en undersøgelse af, hvordan forskellige kognitive kompetencer påvirker læsetilegnelsen hos hebraisksprogede begynderlæsere på et så tidligt læsetilegnelsesstadium, at eventuel dysleksi endnu ikke er mulig at diagnosticere (Shatil & Share 2003). Forfatterne fandt, at tidlig læsning, visuel-ortografisk hukommelse og fonologisk hukommelse hos førskolebørn var klare indikatorer for de samme børns ordgenkendelse i slutningen af 1. klasse. I modsætning til hvad man har fundet i samme slags undersøgelser med europæisksprogede børn (Adams 1990:65), var fonologisk opmærksomhed derimod ikke en statistisk signifikant faktor set i forhold til de øvrige kompetencer (Shatil & Share 2003:18).

Tilsammen peger disse studier på, at succesfuld læsetilegnelse på semitiske sprog stiller større krav til visuelle, ortografiske og morfologiske kompetencer og mindre krav til fonologiske kompetencer end læsetilegnelse på europæiske sprog.

Betydningen af diglossi

Et meget vigtigt aspekt af hele vokalproblematikken, som de nævnte studier ikke berører, er at vokalerne ikke blot i skrift men helt generelt synes at have en sekundær funktion i forhold til konsonanterne (Shimron & Navon 1981; Ryan & Meara 1991; Barnea & Breznitz 1998). Dette kommer tydeligt til udtryk på arabisk, hvor en stor del af variationen mellem dialekt og MSA samt dialekterne imellem netop forekommer som forskelle i de anvendte vokaler, jvf. beskrivelsen af arabisk diglossi i afsnit 3.3.

Et studium af, hvordan den hebraiske diglossi påvirker læseprocessen hos israelske børn i 1. klasse, 4. klasse og hos voksne har vist, at læsere på forskellige niveauer i forskellig grad udnytter den fonologiske information, vokalerne giver, til at afkode. Man testede alle tre grupper for højtlæsning af sætninger med målord, som vokaliseres forskelligt i tale og skrift. Respondenterne fra 1. klasse læste ujævnt. De formåede at oplæse ordene med de rigtige vokaler, men blev forvirrede, når ordenes fonologiske indhold ikke passede sammen med deres talesprogsdominerede leksikon. Respondenterne i 4. klasse læste derimod flydende, men ignorerede langt hen ad vejen vokalerne. De læste målordene i deres moderne version

og rettede ikke som læserne fra 1. klasse på sig selv. Dermed læste de i praksis mange ord forkert. De voksne respondenter læste derimod både flydende og korrekt. Selv når deres intonation afslørede, at de fandt et ords form mærkelig eller ukendt, læste de vokalerne korrekt (Ravid, D. 1996). Dette tyder på, at begynderlæserne holder sig til den fonologiske afkodning, hvilket er almindeligt for begyndere, der endnu ikke har etableret automatisk ordgenkendelse. De moderat øvede afkoder overvejende ortografisk, hvilket i denne sammenhæng ikke er en fordel – det medfører nemlig at de læser forkert. Deres afkodning kræver endnu så meget af dem, at de ikke har kognitiv kapacitet til at tage den information, der kan hentes i vokalerne, i betragtning. De øvede voksne læsere formår derimod både at have hurtig genkendelse af ordets grafiske repræsentation og gøre brug af den fonologiske information, vokalerne tilbyder.

På arabisk er diglossien imidlertid langt mere udpræget og har måske større indflydelse på læsefærdigheden. Abu-Rabias undersøgelse fra 1998, som tester 17-årige gymnasieelevers færdigheder i højtlæsning, viser, at oplæsning på arabisk uanset tekststil og vokalisering er en vanskelig opgave for gode såvel som svage læsere. Fejlprocenten ved oplæsning af uvokaliseret tekst, som jo er den form al tekst præsenteres på for den almindelige læser i dagligdagen, var meget stor for begge grupper af forsøgspersoner. De gode læsere lavede således oplæsningsfejl i 64% procent af alle oplæste ord inden for den informative tekstgenre. De svage læseres fejlprocent inden for samme tekststil var 86. Dette hænger utvivlsomt sammen med de kognitivt meget krævende processer, formel højtlæsning kræver: Læseren må på grund af den omfattende homografi i høj grad støtte sig til kontekst, hvilket i mange tilfælde kræver overblik over den fulde sætning, og hertil kommer, at applicering af korrekte kasusendelser til de enkelte ord kræver en kontinuerlig syntaktisk analyse af tekstmaterialet (Abu-Rabia 1998:116-117). Imidlertid viste samme undersøgelse, at selv ved oplæsning af korrekt vokaliseret informativ tekst, som var den tekststil, begge grupper mestrede bedst, lavede de gode læsere oplæsningsfejl i 22% af alle oplæste ord, mens de svage læseres fejlprocent var helt oppe på 79. Ifølge Abu-Rabia selv indicerer disse tal, at de mange diakritiske tegn – ikke blot vokaler, men også andre tegn som *shadda* (konsonantdublering) og *hamza* (glottal klusil) – kræver flere øjenfikseringspunkter, hvilket optager kognitiv kapacitet. De ringe læseresultater, som Abu-Rabia rapporterer, hænger meget godt sammen med en række studier fra 1980'erne, der indicerer, at 30-50% af arabiske grundskoleelever har markante læseproblemer (Bashi et al. 1981; Habib-Allah 1985). I disse undersøgelser forklares elevernes dårlige resultater med forhold som ringe pædagogik, dårligt undervisningsmateriale og uklare mål i curriculum.

Det synes dog også sandsynligt, at ikke mindst diglossi-aspektet er relevant for de beskedne resultater, arabiske elever generelt synes at levere hvad angår læsning. På grund af de markante forskelle mellem tale- og skriftsprog kan arabiske læsere ikke i samme grad som fx læsere på dansk, engelsk eller hebraisk støtte sig til identifikation af ord, der bruges i talesprog, ligesom de ved læsning har behov for at aktivere viden om syntaks, som kun anvendes på MSA. Flere forskningsprojekter af nyere dato har været med til at afdække denne problematik, fx Saiegh-Haddad (2003b), der ved at teste fonologisk opmærksomhed og afkodningsevner hos palæstinensiske førskole- og 1. klasses børn påviser, at distancen mellem børnenes talesprog og MSA udgør en hindring for tilegnelse af basale læsefærdigheder på MSA (Saiegh-Haddad 2003b:445). Saiegh-Haddad peger på de fonologiske forskelle mellem talesproget og MSA som det centrale problem:

> *"[...] though the first grade children seemed to have benefited from the increased exposure and experience with MSA structures that formal literacy instruction*

> *allows, they were still facing particular difficulty when confronted with the task of isolating standard phonological structures. This finding demonstrates one of the unique complexities present in learning to read in MSA. Whereas in other linguistic contexts, beginning readers come to the reading task with sufficient knowledge and experience with the phonological structures that they are taught to decode, Arabic native children are required to simultaneously master the oral representation of a set of diglossic structures that are not available to them from their oral language experience and to discover how these structures are mapped onto the specific orthography."* (Saiegh-Haddad 2003b:443-444)

Saiegh-Haddads konklusioner støttes af flere forskellige forskningsresultater, der viser, at eksponering for MSA i førskoleårene påvirker børnenes læseindlæring positivt.[9] Abu-Rabia (2000) undersøgte læseforståelse hos to grupper palæstinensiske børn i 1. og 2. klasse, hvor den ene gruppe i det sidste førskoleår havde fået læst historier og havde haft særlige aktiviteter på MSA i 3 x 30 minutter dagligt, mens den anden gruppe havde fået fortalt historier og lavet samme aktiviteter på dialekt. De børn, der havde været eksponeret for MSA, havde signifikant bedre læseforståelse i både 1. og 2. klasse end kontrolgruppen. Jenkins (2001) undersøgte læsefærdigheder i grundskolen hos en gruppe syriske børn, der havde haft tilknytning til en bestemt privat børnehave, hvor der kun kommunikeres på MSA med både børn og forældre. Børn, der havde tilbragt alle tre førskoleår i denne børnehave, kunne tale MSA stort set fejlfrit og havde signifikant bedre læsefærdigheder end kontrolgruppen af børn, der ikke havde fulgt eller havde fulgt et almindeligt førskoleforløb. Også børn, der har gået i koranskole, udviser bedre læsefærdigheder i de første skoleår end børn, der ikke har fulgt et førskoleforløb eller har gået i moderne børnehave uden fokus på MSA (Wagner & Spratt 1987). Effekten af koranskole synes dog langt svagere end effekten af eksponering for MSA, og andre faktorer som fx et urbant frem for et ruralt opvækstmiljø har større indflydelse på læsetilegnelsen end koranskoleundervisning i førskoleårene (Wagner & Spratt 1987:1218).

Inden for det seneste årti ses en tendens til, at diglossi-aspektet i stadig højere grad betragtes som et regulært problem i de arabiske sprog- og uddannelsesmiljøer (fx Ayari 1996; Maamouri 1998; Abu-Rabia 2002). Dette ekspliciteres klart i nedenstående citat fra en officiel rapport til Verdensbanken om menneskelig udvikling i regionen:

> *"A [...] question relates to whether Arabic diglossia really constitutes a critical variable for the spread of literacy and the quality of basic education in the Arab region or whether socioeconomic factors override most of the elements that relate for the general impact of the Arabic language on educational attainment and achievement. This author's claim is that even though it cannot by itself bear all the blame for the crisis in Arab education, Arabic diglossia is a definite aggravating factor in the low results of schooling and non-formal instruction and taking care of it, if at all possible, would greatly improve the quality of education in the region."* (Maamouri 1998:68)

---

[9] Førskoleforløb er meget almindelige i den arabiske verden, men langt hovedparten af de arabiske børn, der følger førskoleforløb, går i koranskoler, hvor de lærer at afkode korantekster, dvs. klassisk arabisk. Mange velhavende familier sender deres børn i moderne, private 'børnehaver', hvor der normalt også er forholdsvis stort fokus på indlæring, men hvor der i reglen udelukkende kommunikeres på talesprogsdialekten.

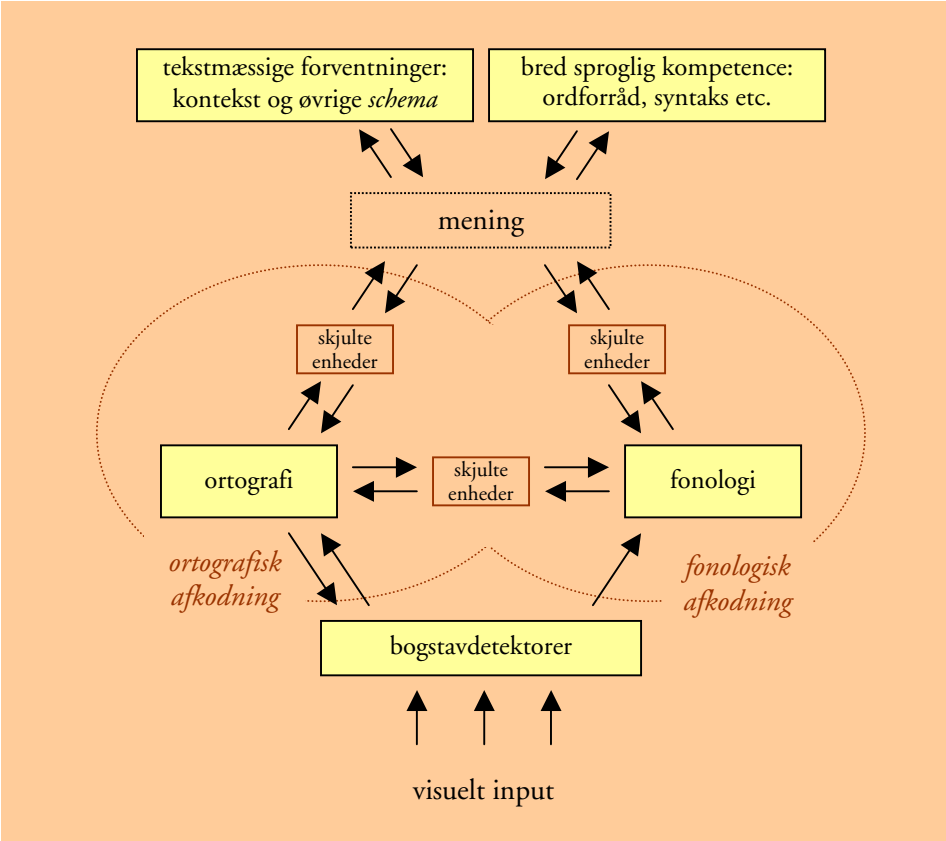
Alt i alt synes der ikke at være tvivl om, at ud over at de specifikt læsetekniske aspekter, der knytter sig til de uvokaliserede semitiske ortografier, synes at være forholdsvist kognitivt krævende, har afstanden mellem tale- og skriftsprog på arabisk en afgørende negativ indflydelse på etablering af læsefærdigheder hos indfødte sprogbrugere.

Antydninger af en læseprocesteori for semitiske sprog

Ovenstående forskningsresultater belyser ordgenkendelse på semitiske sprog fra mange forskellige vinkler og med udgangspunkt i forskellige læseteoretiske traditioner. I dette afsnit opsummeres disse resultater og relateres til den sekundære genkendelse i læseprocesmodellen fra kapitel 2, som er gengivet i figur 5.4.

Som det fremgår af ovenstående, kræver de uvokaliserede semitiske ortografier, i sammenligning med de europæiske, flere og mere langvarige øjenfixeringer og tager (måske derfor) længere tid at læse. Dette skyldes sandsynligvis, at de er mere informationsmæssigt kompakte – fx i form af artikler, præpositioner og pronominer, der internaliseres som affixer, og subjekter der impliciteres i verbalbøjninger. Det synes således sandsynliggjort, at den mentale opsplitning af disse forskellige morfemtyper kræver ekstra kognitiv kapacitet.

Endvidere er der meget der tyder på, at læseprocessen forløber forskelligt ved læsning af henholdsvis vokaliseret og uvokaliseret tekst. Det er således dokumenteret, at afkodning af uvokaliseret skrift er overvejende ortografisk, mens vokalisering genererer en mere fonologisk medieret afkodningsproces. Ved læsning af vokaliseret tekst anvendes den fulde



Figur 5.4 Model af den valgte læseteoretiske ramme (identisk med model i figur 2.5)

vokalinformation til at afkode fonologisk, og dette kan gøre afkodningen af især lavfrekvente ord hurtigere. Dette er fra et traditionelt læseteoretisk perspektiv stik modsat det forventelige. Adskillige studier af læsning på europæiske sprog viser, at fonologisk medieret afkodning er langsommere end ortografisk afkodning, hvilket fra et dualruteteoretisk udgangspunkt forklares med, at førstnævnte involverer et ekstra behandlingsniveau, jvf. figur 2.2. Men også selvom man afskriver dualruteteoriens gyldighed, er disse resultater umiddelbart overraskende, idet afkodning af vokaliseret tekst implicerer behandling af en større mængde grafisk input (de diakritiske tegn).

Også hvad angår betydningen af kontekst, kolliderer resultaterne fra semitiske sprog med den læseteoretiske tradition. Ikke blot er kontekst i langt højere grad i spil ved læsning på de semitiske sprog. Gode læsere viser sig også at være bedst til at udnytte kontekst på arabisk, og dette står i skarp kontrast til Stanovich's interaktivt-kompensatoriske teori, som er empirisk velunderbygget og meget bredt accepteret inden for læseforskningen (Stanovich 1980,1986; Just & Carpenter 1987; Rayner & Pollatsek 1989; Adams 1990; Stanovich 2000). Ifølge denne teori har kun svage læsere behov for at gøre brug af kontekst, som de støtter sig til for at kompensere for manglende afkodningsfærdigheder (Stanovich 1980).

Disse uoverensstemmelser mellem læseteori baseret på europæiske sprog og forsknings-resultater fra semitiske sprog kan alle forklares med en større redundans i de europæiske skriftsystemer. Det er fx også på dansk fuldt ud muligt at afkode uden eksakt vokalinformation: 'm•n b•h•v•r sl•t •kk• v•k•l•r f•r •t k•nn• l•s• d•nn• s•tn•ng!'. Selvom eksemplet er lidt fortegnet, idet vokalernes placering er markeret, illustrerer det, at vokalernes information sammenholdt med konsonanternes til en vis grad er redundant. Alligevel må det formodes, at selv med lidt øvelse i at læse et sådant skriftsystem, vil det være muligt at opnå større læsehastighed og mere korrekt ordgenkendelse, når vokalerne er angivet. I den optik er det ikke overraskende, at vokalerne udgør en støtte frem for en belastning ved læsning på semitiske sprog. Dette argument støttes af, at vokalernes positive effekt på afkodningshastighed svinder, når ord fremstilles i kontekst.

Også kontekstens rolle kan forklares med denne manglende redundans i de uvokaliserede semitiske ortografier: Når kun svage læsere på engelsk har behov for at gøre brug af kontekst, er det fordi den information, der er nødvendig for afkodning, for det meste allerede er til stede i den fonologiske information, hvilket ikke i nær samme grad er tilfældet på de semitiske sprog, hvor en stor andel af ordene på grund af den manglende vokalinformation er homografer, hvilket gør afkodning stærkt kontekstafhængig – fx kræver det ofte overblik over den fulde sætning at bestemme et initialt verbums form. Ved vokaliseret tekst opnår man netop samme redundans, som man har ved brug af det latinske alfabet, hvorved kontekst bliver en ekstra ressource, som det ikke er nødvendigt at udnytte, hvis man har tilstrækkelige afkodningsfærdigheder. Konsekvensen af, at semitisk skrift almindeligvis er uvokaliseret, er dermed, at kontekst kommer til at spille en helt anden rolle end ved læsning af latinsk skrift: Ved læsning på sprog, der anvender latinsk skrift er kontekst en ressource, der ved manglende afkodningsfærdigheder kan udnyttes kompensatorisk, mens gode læsere ikke behøver at gøre brug af den. Ved læsning på de semitiske sprog er gode læsere gode, netop **fordi** de støtter sig til kontekst.

Selvom afkodning af de uvokaliserede semitiske ortografier hos millioner af indfødte sprogbrugere forløber som en ukompliceret, flydende proces, synes det således at stå klart, at det har afgørende indflydelse på læseprocessen, at læseren – i kraft af det ringere omfang

af fonologisk information i forhold til hvad der gælder på latinskskrevne ortografier – har færre ressourcer at trække på. Og der synes at være grund til at formode, at resultatet er, at der trækkes på andre former for ressourcer, end dem, der anvendes ved læsning på europæiske sprog, og at afkodningsprocessen generelt kræver mere kognitiv kapacitet. At vokalisering uafhængigt af afkodningshastighed og ordgenkendelse forbedrer læseforståelse, indicerer i hvert fald, at læsning uden vokalangivelse er så kognitivt krævende, at det optager kapacitet, der – når vokalerne er til angivet – kan anvendes til i højere grad at bearbejde tekstens meningsindhold.

Hvis man på baggrund af de præsenterede forskningsresultater skal forsøge at definere nærmere, hvordan afkodningsprocessen på de semitiske sprog fungerer, må man holde sig for øje, at de refererede undersøgelser af ordbestemmelse og ordgenkendelse stort set alle afspejler, at de har dualruteteorien som udgangspunkt, hvilket har influeret på, hvordan man har fortolket dem. Når undersøgelser viste, at vokalisering ikke gjorde ordbestemmelse af homografer hurtigere, og når homografiske associationsord syntes at aktivere flere betydninger, ræsonnerede man, at ordbestemmelse finder sted før det bestemmes, hvilke vokaler der skal tilsættes, og at læseren altså ved, at der er tale om et ord, før han har bestemt sig for, hvilket ord der er tale om. På den baggrund lanceredes en udbygget læsemodel for de uvokaliserede semitiske ortografier, hvor der introduceres et 'mellemniveau' på vej til det mentale leksikon. Dette 'ekstra leksikon' består af konsonantstrenge, og eftersom ordbestemmelse ikke nødvendigvis involverer fonologisk behandling, er indgangene til det 'ekstra leksikon' muligvis udelukkende ortografiske. Under alle omstændigheder er et ords fulde fonologiske repræsentation et resultat af – frem for et middel til – ordgenkendelse (Frost et al. 1987; Frost & Bentin 1992a; Katz & Frost 1992). Sammenholdt med de dengang foreliggende forskningsresultater vedrørende betydningen af kontekst og ordfrekvens stod man således med en læseprocesmodel, hvor ordgenkendelse på semitiske sprog er baseret på en rodbaseret indgang til leksikon, og hvor man kan forestille sig, at læseren, når hun møder et flertydigt ord, trækker roden op af hatten med rødder og derpå overvejer, hvilke vokaler, det – i henhold til kontekst – giver mening at tilføje, og i tvivlstilfælde – hvis kontekst er tvetydig eller fraværende – afvejer mulighederne efter et frekvens-bestemt hierarki, hvilket jo i sagens natur i mange tilfælde fører læseren direkte til den korrekte mening.

De nyere forskningsresultater, der dokumenterer præleksikalsk fonologisk behandling, undsiger denne ordgenkendelsesmodel. Til gengæld går de fint i spænd med læseprocesmodellen i figur 5.4, som i kraft af konnektionismen jo netop hævder et kontinuerligt samspil mellem ortografiske og fonologiske processer, og hvor alskens former for relevante ressourcer, læseren har opbygget i kraft af tidligere teksteksponering, aktiveres. Her kan de tidligere undersøgelser af, hvordan de forskellige former for læserelevante ressourcer (kontekst, ordfrekvens og eventuel vokalisering) aktiveres i forhold til hinanden, problemfrit inkorporeres. Samtidig kan den konnektionistiske model også forklare, hvordan rod- og mønster-systematikken influerer på læseprocessen, ikke blot som en del af den brede sproglige kompetence (øverst i modellen), men også på ordgenkendelsesniveau, hvor den morfologiske struktur må formodes at spille en afgørende rolle som kompensation for de manglende vokaler – ikke nødvendigvis i form af en opsplitning af rødder og mønstre i selvstændige morfemer (hvilket dog i kraft af de refererede forskningsresultater synes at være tilfældet, i hvert fald på hebraisk), men som et resultat af frekvens: Det begrænsede antal mønstre giver et begrænset antal ordstrukturer, som nødvendigvis må være lagret i de skjulte enheder og dermed influere på afkodningen.

Med det som udgangspunkt kan man forsøge at definere, på hvilke områder informationsbehandlingen i ordgenkendelsesprocessen på semitiske sprog adskiller sig fra tilsvarende behandling på europæiske sprog:

Overordnet set står det klart, at fonologisk behandling ved læsning af uvokaliseret skrift spiller en mindre rolle. Det fremgår både af den refererede forskning, der har til formål at udforske hypotesen om ortografisk kompleksitet og af forskningen i dyslexi. Mens fonologiske processer er afgørende for ordgenkendelse på engelsk og fonologisk opmærksomhed er *"inescapably required"* (Adams 1990:305) for at opnå gode læsefærdigheder, er det i mindre grad tilfældet på de semitiske sprog, hvor ortografiske og morfologiske processer omvendt spiller en større rolle for etablering af gode læsefærdigheder, end det er tilfældet på de europæiske sprog. I henhold til hypotesen om ortografisk kompleksitet skyldes dette, at den fonologiske information er begrænset, og resultatet bliver, at læseren i højere grad må trække på de øvrige tilgængelige ressourcer. Dette forhold forklarer, at korrekt højtlæsning og læseforståelse på semitiske sprog, i modsætning til hvad der gælder på europæiske sprog, ikke korrelerer (Abu-Rabia 2001; Saiegh-Haddad 2003a): Mens fonologisk behandling selvsagt er en forudsætning for højtlæsning, kræver læseforståelse ikke nødvendigvis identifikation af ordenes eksakte fonologi, da ordgenkendelse i så høj grad trækker på andre ressourcer. Samtidig spiller diglossien ved læsning på arabisk muligvis også en rolle i den sammenhæng: At tale- og skriftsprogsstrukturer i så begrænset omfang overlapper hinanden, formindsker muligheden for lydlige associeringsprocesser, hvilket muligvis kan medvirke til, at læsere i så høj grad må udnytte andre ressourcer end de fonologiske.

Med udgangspunkt i den valgte læseprocesmodel, jvf. figur 5.4, kan man altså sige, at ved læsning på semitiske sprog løber informationsstrømmen i højere grad end på europæiske sprog i venstre side af modellen både på ordgenkendelsesniveau og på det øvre niveau, hvor der i langt højere grad trækkes på kontekst. Et eventuelt artikuleret output vil i modellen kunne illustreres som en ekstern forbindelse fra kassen 'fonologi' til taleorganerne, men fordi informationsstrømmen til en vis grad løber uden om denne del af modellen, kan man empirisk konstatere den svage sammenhæng mellem forståelse og højtlæsning.

Hvis man ser mere specifikt på morfologiens rolle, som jo ikke er ekspliciteret i modellen, men er internaliseret i de skjulte enheder mellem især ortografi og fonologi, kan man ud fra rent sproglige parametre konstatere, at processen på dette niveau nødvendigvis må forløbe anderledes på de semitiske end på de europæiske sprog:

Som beskrevet i kapitel 2 lagres læserens erfaringer med sprogets strukturer sig i ordgenkendelsesmodellens skjulte enheder. Ved læsning på europæiske sprog er højfrekvente og lavfrekvente bogstavsammenstillinger i denne sammenhæng essentielle, idet ordgenkendelse påvirkes af, hvilke bogstaver, der tidligere er blevet set i sammenhæng (nabofrekvenseffekt). Adams (1990) beskriver processen således:

> *"The nature of the stimulation passed along from a donating to a receiving letter depends on the frequency with which the two letters have occurred together in the reader's lifetime of reading experience. Letters that have often been seen with the donating letter will receive positive excitation; the more often they have been seen together, the stronger this positive excitation will be. Conversely letters that have rarely been seen with the donating letter will receive negative excitation, or*

> *inhibition, that is proportionate to the rareness of their co-occurrence."* (Adams 1990:109)

Implicit medfører disse processer, at vi har tendens til automatisk at opdele vores perception af lange ord i stavelser. Sandsynligheden for, at en konsonant efterfølges af en anden har nemlig tæt forbindelse til de syllabiske strukturer i vores talesprog. Er første bogstav i et ord eksempelvis et 'd', så er der langt større sandsynlighed for, at det følges af et 'r' end af et 'n', og 'dr' repræsenterer derfor en velkendt bogstavsammenstilling, mens 'dn' er sjældnere. Og da de mindre hyppige konsonantsammensætninger normalt forekommer i forbindelse med en sammensætning af to stavelser, som fx i 'midnat', vil vi – på baggrund af vores erfaring – være tilbøjelige til at opdele ordet netop der, hvis vi ikke umiddelbart kan overskue det (Adams 1990:116). Dette er naturligvis ekstra hensigtsmæssigt, når de enkelte stavelser som i eksemplet repræsenterer selvstændige betydningsdele, og strategien repræsenterer derfor endnu en ressource, som kan udnyttes ved læsning. Nabofrekvens-effekten er altså et resultat af lagring af sprogets syllabiske og morfologiske strukturer i de skjulte enheder, hvor både ortografiske og fonologiske elementer indgår og resulterer i forskellige grader af grafo- og fonotaktisk ansporing eller begrænsning.

Ud over denne lineære behandling trækker vi også på et mere holistisk billede af det enkelte ord. Dette er illustreret i nedenstående eksempel, hvoraf det fremgår, at selvom bogstaverne står i positioner, der fremkalder en åbenlys tilbagemelding om ualmindelige eller endog uacceptable bogstavkombinationer, kan vi ved hjælp af enkelte bogstaver i korrekte positioner samt på baggrund af ordlængde i mange tilfælde opnå forholdsvis hurtig ordgenkendelse.

Dtete esmepkel vseir, at når man lesær ord, man otfe har set før, bteyedr det ikke så mgeet, i hivklen rkækefløge bgotsavrene såtr i et ord. Det vgistgtie er, at det frøtse og sdiste bgosatv i odret er på de rtete padlser. Rseten kan lngie vloaypk men vi vil satdig vræe i snatd til at lsæe det.

Nrædnævere istelrelrur evirdndee, at lfakvenrvete ord er vlnakeserige at adkfoe uednr snadnåe bielnegster.[8]

Eksempel på, at lineær bogstavanalyse ikke er nødvendig for ordgenkendelse.

Selvom automatisk ordgenkendelse, som er essentiel for gode læsefærdigheder, i meget høj grad hviler på en sådan holistisk opfattelse af det enkelte ord som en samlet enhed, aktiveres den lineære bogstavanalyse alligevel (Adams 1990:111). Den fungerer som en 'hjælpemotor', der støtter processen ved at bekræfte vores opfattelse af det holistiske input, afgøre sagen i tvivlstilfælde og 'lappe huller' undervejs, hvorved læsehastigheden øges.

På de semitiske sprog er sagen imidlertid en anden, fordi den fonologiske information er begrænset, og den morfologiske struktur er anderledes: Da vokalerne er skjulte, er der ikke samme grundlag som på europæiske sprog for at opbygge en reaktion på velkendte og uvante bogstavkombinationer, og derfor udebliver konsonantsammenstillingernes

[10] I den nederste passage står: 'Nærværende illustrerer endvidere, at lavfrekvente ord er vanskeligere at afkode under sådanne betingelser.'

information om stavelser også. Det er således ikke muligt at opbygge klare positive henholdsvis negative tilbagemeldinger på grundlag af kombinationer som 'dr' og 'dn'. For det første kan 'dr' i relation til det relevante ords fonologiske udtryk i princippet være enten 'dar', 'dur', 'dir' eller 'dr'. For det andet afhænger kombinationens ortografiske genkendelighed i meget høj grad af rodens tredje konsonant. Samtidig har de bogstaver, der er en del af et ords mønster, faste pladser i ordet, og der er ingen begrænsninger på, hvilke konsonanter et sådant mønster-bogstav kan kombineres med. Hvad angår den mere holistiske ordbehandling, er der også mindre information at hente, da det begrænsede antal mønstre indebærer et grafisk mere ensartet ordforråd. Kort sagt kan semitiske ord ikke se nær så forskellige ud, som europæiske ord kan. Store grupper af ord adskiller sig udelukkende fra hinanden i kraft af de tre konsonanter, der udgør roden, og dermed er der ikke grundlag for at etablere de nævnte positive og negative tilbagemeldinger på grundlag af bogstavkombinationer.

Til gengæld giver denne stramme semitiske morfologi andre former for ressourcer, der kan udnyttes: Visse præ-, in- og suffixer kan ligesom ordlængde give tilbagemelding om, hvilke mulige mønstre, der kan være tale om. Faktisk vil et præfiks i nogle tilfælde reducere de mulige mønsterkombinationer til ganske få eller endog til et enkelt. Og når mønsteret er givet, er vokalerne det også. Genkendelse af et mønster er derfor essentiel, når en konsonantkombination som 'dr' afkodes. Det er nemlig mønsteret, der afgør, om der er tale om 'dar', 'dur', 'dir' eller 'dr'. Samtidig spiller genkendelse af roden muligvis en afgørende rolle, idet en sådan genkendelse reducerer antallet af mulige mønstre.

Forskellen på hvilken form for morfologiske strukturer, der lagres i de skjulte enheder på henholdsvis europæiske og semitiske sprog, kan også illustreres ved, at man på europæiske sprog klart kan skelne mellem pseudoord (som er ordlignende), fx 'kvir' eller 'flas' og nonord (som ikke er ordlignende), fx 'ikvr' eller 'lfas': Pseudoordene består af velkendte bogstavkombinationer og repræsenterer en acceptabel fonologisk struktur, og den lineære bogstavanalyse giver derfor positive tilbagemeldinger, selvom ordene ikke har nogen semantisk værdi. Derimod giver nonordene negative tilbagemeldinger, fordi de anvendte konsonantkombinationer støder mod de grafo- og fonotaktiske begrænsninger, vi gennem eksponering for tekst har etableret i vores ordgenkendelsessystem. På de semitiske sprog vil forskellen på et pseudo- og et nonord imidlertid ikke afhænge af bogstavsammenstillinger, men af gyldige eller ikke gyldige mønstre. Hvis mønsteret er gyldigt, vil 'ordet' blive opfattet som ordlignende, og forskellen på et ord og et pseudoord afhænger derfor udelukkende af, om ordet er konstrueret på baggrund af tre rodkonsonanter, der i kombination med det aktuelle mønster giver et rigtigt ord. Konstruktion af et nonord kræver derimod et ikke-eksisterende mønster.[11] På samme måde vil det være umuligt at konstruere en semitisksproget version af eksemplet på forrige side af, hvordan man på dansk kan læse højfrekvente ord, hvor kun første og sidste bogstav står på deres rette pladser. En tilsvarende manipulation af semitiske ord vil give et helt anderledes resultat, da man med tre rodkonsonanter og fx et infix ofte vil have materiale til et stort antal gyldige ord, fordi ombytning af rodens konsonanter tit vil resultere i en anden gyldig rod, ligesom en anden placering af et infix i nogle tilfælde vil give et andet gyldigt mønster.

---

[11] Bemærk, at der både på europæiske og semitiske sprog dog findes især låneord, der er gyldige, selvom de kolliderer med de etablerede grafo- og fonotaktiske begrænsninger, fx på arabisk 'ديمقراطية' ('demokrati') og på dansk 'ftalater'.

På den baggrund må man konkludere, at det ud fra sproglige parametre er meget usandsynligt, at rullende tripel-bogstavanalyse og nabofrekvenseffekt, som synes at være essentielle elementer i afkodningen på europæiske sprog, spiller en væsentlig rolle ved læsning på semitiske sprog. Derimod må man – både på grund af forskellene i semitisk og indoeuropæisk morfologi og på baggrund af de præsenterede forskningsresultater – formode, at mønstre og rødder på semitisk får en eller anden grad af 'separat opmærksomhed' i ordgenkendelsesprocessen, hvorved aktivering af kendskab til mulige mønstre spiller en væsentlig rolle for ordgenkendelsen. Med udgangspunkt i konnektionismen er det i hvert fald sandsynligt, at en eller anden form for 'mønsterfrekvenseffekt' lagres i ordgenkendelsessystemets skjulte enheder og udnyttes til at opnå hurtig ordgenkendelse – ikke mindst fordi den manglende redundans i skriftsystemet reducerer udvalget af ressourcer, det er muligt at trække på, hvorfor genkendelse af valide mønstre synes at være den primære kilde til vokalinformation, i hvert fald ved læsning af lavfrekvente ord.

5.2 Læsning på semitiske sprog som L2

Mens der som beskrevet i afsnit 1.2 ikke findes nogen studier, der vedrører det konsonantiske skriftsystems indflydelse på tilegnelse af læsefærdighed på arabisk som L2, findes der to undersøgelser, der beskæftiger sig med denne problematik med hebraisk som L2:

Cowan (1992) undersøgte to grupper engelske studerende, hvor den ene gruppe havde tysk og den anden hebraisk som L2. Begge grupper var indskrevet på det relevante sprogfag på andet semester. Forfatteren testede hvor godt de to grupper kunne genkalde sig ord på det relevante L2. Han fandt, at visuel eksponering for ord på L2 var afgørende for korrekt genkaldelse af et ord – ikke blot i relation til stavemåde, men også hvad angår ordets fonologiske repræsentation. Han fandt desuden langt flere genkaldelsesfejl hos lørnerne med hebraiske L2 end hos dem, der læste tysk. Fejltyperne hos de to grupper var også vidt forskellige. I den tyske gruppe havde man svært ved at huske slutningen på lange ord. I den hebraiske gruppe lavede man derimod mange vokalfejl, selvom en kontroltest af testpersonernes færdigheder i at anvende vokaliseringssystemet viste, at de mestrede det udmærket. Resultaterne viser ifølge forfatteren, at lørnerne med hebraisk som L2 ikke har tilstrækkelig fonologisk repræsentation af det enkelte ord, til at kunne lagre det i hukommelsen:

> *"...the degree to which the orthography of a second language reflects its phonology is a predictive index of the extent to which the written forms of lexical items can be used to facilitate L2 learning. [...] L2 learners are fairly dependent on the visual image of words in the target language for short and long-term retention, and hence the characteristics of target language orthographies can affect acquisition of vocabulary by literate L2 learners".* (Cowan 1992:286)

I et langt mere udførligt studium undersøgte Wade-Wooley & Geva (1998) genkendelse af ord bestående af henholdsvis ét og tre morfemer[12] samt højtlæsning af et tekststykke hos

[12] fx פרה [parah] ('ko') vs. בכיס [bakis] ('i lommen'), hvor ב repræsenterer både præpositionen ב [be] ('i') og den bestemte artikel, som normalt skrives ה [ha] men i kombination med præpositionen blot ændrer dennes vokal. Alle testord med tre morfemer var af denne type, dvs. kombinationer af en præposition og et substantiv i bestemt form.

russiske indvandrere i Israel og canadiske engelsk-hebraisk tosprogede jøder, og sammenlignede med indfødte sprogbrugere. De russiske forsøgspersoner var alle bosat i Israel og havde modtaget mellem 3 og 5 års intensiv *ulpan*-undervisning i hebraisk, hvilket vil sige, at de havde fulgt intensive israelske sprogintegrationsprogrammer. Canadierne havde mellem 5 og 13 års undervisning bag sig, men denne undervisning var først og fremmest foregået i Canada i en traditionel, religiøs sammenhæng og var påbegyndt ved skolestart. Kun et fåtal af dem havde derudover fulgt *ulpanim* i Israel.

Resultaterne af ordgenkendelsesopgaven ses i figur 5.5, hvoraf det fremgår, at forsøgspersoner med hebraisk som L1 læste både hurtigst og mest fejlfrit, og at L2-læserne lavede flere fejl end L1-læsere ved læsning af ord konstrueret af flere morfemer. Ved læsning af disse ord lavede russerne desuden signifikant flere fejl end de engelsksprogede L2-læsere. Dette kan skyldes, at selvom russerne bruger hebraisk i deres dagligdag, mens den engelsksprogede gruppe mest har erfaring med formel brug af sproget, har sidstnævnte længere tids erfaring med hebraisk tekst. Da alle de anvendte ord i opgaven var vokaliserede, kunne netop den engelsksprogede gruppes erfaringer med tekster fra deres religiøse skoling anvendes. Endvidere fandt man, at hvis man manipulerede stimuli ved med forskellig fontstørrelse at eksplicitere det præfix, der markerede præposition (og implicit også bestemthed), reduceredes fejlprocenten hos L2-grupperne, mens den øgedes hos L1-gruppen, hvorved den hos de tre grupper blev ens (3%). Ved den omtalte manipulation af stimuli blev ordgenkendelse hurtigere hos begge grupper af L2-læsere, mens det ingen effekt havde på L1-gruppen. Hvad angår hastighed havde den hebraisksprogede gruppe signifikant hurtigere ordgenkendelse end begge grupper af L2-læsere, og den engelsksprogede gruppe var signifikant hurtigere end den russisksprogede, hvilket igen kan forklares med den engelsksprogede gruppes længere tids erfaring med netop denne type tekstmateriale. Til gengæld var der ikke inden for nogen af grupperne signifikant forskel på ordgenkendelseshastighed, hvad angår antal af morfemer i de ord, der blev læst korrekt.

I tekstlæsningsopgaven, hvis resultater ses i figur 5.6, læste den hebraisksprogede gruppe markant hurtigere og mere fejlfrit end begge grupper af L2-læsere. Endvidere lavede gruppen med russisk som L1 signifikant færre læsefejl end gruppen med engelsk som L1, selvom den engelsksprogede gruppe havde modtaget undervisning i hebraisk over en meget

	L1=hebraisk		L1=engelsk		L1=russisk	
	1 morfem	3 morfemer	1 morfem	3 morfemer	1 morfem	3 morfemer
Andel læsefejl	1 %	1 %	4 %	6 %	5 %	8 %
Tid for læsning	583 ms	577 ms	711 ms	693 ms	846 ms	851 ms

Figur 5.5 Gennemsnitlige fejlprocenter og tider for højtlæsning af hebraiske ord bestående af henholdsvis 1 og 3 morfemer. Tid for læsning er kun beregnet på grundlag af korrekt oplæste ord (Wade-Woolley 1998)

	L1=hebraisk	L1=engelsk	L1=russisk
Tid for læsning	80 sek.	169 sek.	171 sek.
Antal læsefejl	1,00	18,92	13,72

Figur 5.6 Gennemsnitlig tid for læsning og antal læsefejl ved højtlæsning af tekststykke (Wade-Woolley 1998)

længere periode. Dette kan have flere forskellige forklaringer. Det kan skyldes, at den russiske gruppe har mere erfaring med den moderne, uvokaliserede tekststil, men kan i princippet også afspejle forskelle i latinsk og russisk skrift. Mens bogstavgenkendelse af latinske bogstaver som vist i kapitel 4 i meget høj grad udnytter de latinske bogstavers forskellige størrelser såsom forlængelse over og under bundlinjen, må en mere detaljeorienteret scanning gøre sig gældende på russisk, hvor alle bogstaver er lige høje. Dette forhold giver muligvis russerne en fordel ved bogstavdetektion på hebraisk, hvor bogstaverne ligeledes har et mere ensartet grafisk udtryk.

Uanset forskellene mellem de to L2-grupper viser dette studium, at læsning på hebraisk selv hos lørnere på meget højt niveau er en forholdsvis vanskelig opgave. Ordgenkendelse forsinkes af ord med præpositionspræfixer, og læsehastigheden ved læsning af sammenhængende tekst er betragteligt langsommere end hos indfødte sprogbrugere. I betragtning af, at det er påvist, at L2-læsere på nære fremmedsprog faktisk godt kan opnå læsehastighed svarende til L1-læseres (Favreau & Segalowitz 1982), er det påfaldende, at de russiske indvandrere, der er studerende ved et israelsk universitet, bruger dobbelt så lang tid som de indfødte sprogbrugere på tekstlæsningsopgaven.

5.3 Sammenfatning

Som beskrevet i afsnit 2.4, afhænger en lørners læsefærdigheder på L2 grundlæggende af nogle individuelle kognitive forudsætninger og af nogle generelle sproglige kompetencer, som kan overføres fra det ene sprog til det andet. En lørner, der er en god læser på sit L1 vil derfor have gode muligheder for også at udvikle sig til en god læser på L2, i takt med at den generelle sprogforståelse på det pågældende sprog udvikles.

Ovenstående gennemgang af forskelle i læseproces på europæiske og semitiske sprog som L1 og resultaterne af de to beskrevne studier vedrørende L2 indicerer imidlertid, at læsetilegnelse på semitiske sprog som L2 ikke 'bare' er et spørgsmål om generel sprogtilegnelse, men også involverer specifikt læserelaterede aspekter. De semitiske sprog adskiller sig således fra de nære fremmedsprog, ikke kun i kraft af et mere fremmed ordforråd og en anderledes syntaktisk struktur, men også af følgende faktorer, der har direkte indflydelse på etableringen af gode læsefærdigheder:

- Et antal uvante fonemer
- Et alfabet bestående af meget fremmedartede grafemer
- Manglende fonologisk information og deraf følgende reduceret redundans i skriftsystemet
- En fremmed morfologisk struktur
- Et stort behov for brug af kontekst til afklaring af homografer

Til gengæld er de manglende muligheder for association mellem tale- og skriftsprog på arabisk af mindre betydning, fordi lørneren normalt introduceres til MSA også som talesprog.

Følgen af de nævnte faktorer er, at det ordgenkendelsessystem, den voksne lørner har etableret og optrænet på sit L1 gennem flere år, ikke rigtigt kan bruges til noget. Alle kasserne i ordgenkendelsesmodellen skal genopbygges med repræsentationer fra det ny skriftsystem: Som beskrevet i kapitel 4 må lørneren tilpasse sine scanningsstrategier eller

udvikle andre for at kunne genkende og skelne de nye bogstaver. Hertil kommer, at der skal opbygges associationer til både kendte og ukendte lyde, som i nogle tilfælde også kan være vanskelige at skelne. Og hertil kommer, at velkendte procedurer i ordgenkendelsesprocessen obstrueres af manglende tilgængelig information: Som beskrevet i afsnit 2.5 udnytter lørnere af et nyt alfabetiske skriftsystem nærmest per automatik deres kendskab til det fonematiske princip, og når automatiseret ordgenkendelse udebliver, falder de derfor uvægerligt tilbage på denne strategi. Med udgangspunkt i læseprocesmodellen er problemet, at lørneren søger at tvinge sig vej til ordets mening gennem den højre side af modellen, men metoden har sine begrænsninger, fordi en stor del af den fonologiske information ikke er til stede. Afkodning kræver altså – ud over at det aktuelle ord findes i ordforrådet – enten direkte genkendelse af ordets grafiske udtryk, et vist mål af genkendelse af ordets morfologiske struktur, eller at der gennemføres en række hypotesetest, hvor forskellige vokalkombinationer afprøves. Sidstnævnte metode resulterer naturligvis i en meget langsommelig læseproces.

Sagen kompliceres yderligere af, at korrekt afkodning af et ord i ganske mange tilfælde også kræver, at forståelsen af kontekst er tilstrækkelig klar til at kunne udnyttes til afklaring af et ords betydning i tilfælde af homografi. I begynderfasen er dette aspekt mindre relevant, fordi ordforrådet er begrænset, og et homografisk ord derfor måske kun kendes som repræsentant for én glose. Men efterhånden som ordforrådet øges og det tekstmateriale, lørneren eksponeres for bliver mere varieret, øges behovet for at udnytte kontekst betragteligt. Efterhånden som lørneren forventes at kunne håndtere autentiske tekster, giver homografien yderligere vanskeligheder, fordi tekst, der ikke er gradueret efter allerede indlært ordforråd, gerne indeholder mange ukendte ord. Når en del af disse er homografer, som lørneren kender som ét ord, men i den relevante kontekst betyder noget andet, bliver det vanskeligt at få succes med en topstyret læsestrategi, hvor der gættes ud fra kontekst.

Det er imidlertid uvist, om L2-lørnere på længere sigt fastholder at overføre de processer, de anvender på deres L1, eller om de løbende udvikler læsekompetencer, der ligner indfødte sprogbrugeres og i så fald i hvor høj grad, de gør det. Studiet af Wade-Wooley & Geva (1998) viser, at L2-lørnere med hebraisk som målsprog selv med mange års studier i sproget bag sig og med daglig eksponering for og brug af sproget fortsat læser forholdsvis langsomt, ligesom den komplekse morfologiske struktur i form af ord sammensat af flere morfemer forhaler ordgenkendelse og er anledning til ordgenkendelsesfejl. Undersøgelsen vedrører imidlertid kun effekten af ord sammensat af præposition, bestemt artikel og substantiv.

På denne baggrund opstilles og afprøves i det følgende kapitel en række hypoteser, der vedrører de ovennævnte fem faktorer, der adskiller tilegnelse af læsefærdigheder på semitiske sprog som L2 fra tilsvarende indlæring på de nære fremmedsprog. Målet for disse test vil primært være at undersøge følgende:

- Hvor vigtig en rolle spiller de fremmede grafemer og fonemer isoleret set?
- I hvor høj grad hæmmer de manglende vokaler læsehastighed og læseforståelse? (uafhængigt af højtlæsning, som behandles af Wade-Wooley & Geva (1998)).
- I hvor høj grad formår lørnere at bruge rod-mønster-systematikken til at kompensere for den manglende fonologiske information?

Som beskrevet i afsnit 1.2 er det første punkt tidligere behandlet af Khaldieh (1996), men med et problematisk testdesign. De øvrige to spørgsmål er ikke tidligere blevet forskningsmæssigt belyst.

En undersøgelse, der kan kaste lys over disse aspekter, vil gøre det muligt at konkretisere problematikken yderligere og dermed gøre det muligt at determinere nærmere, hvordan undervisning i læsefærdighed på arabisk som L2 kan målrettes de specifikke vanskeligheder, danske lørnere står overfor.

6. Danske lørneres læsning på arabisk

Som det fremgår af kapitel 4 og 5, er der en række forhold ved det arabiske skriftsystem, som teoretisk set gør læsetilegnelse på arabisk som L2 anderledes end læsetilegnelse på de nære fremmedsprog. I dette kapitel præsenteres en undersøgelse, der har til formål at dokumentere dette empirisk og afdække de specifikke vanskeligheder nærmere.
Dette søges realiseret gennem en række undersøgelser af danske lørneres læsning på arabisk. I undersøgelsen indgår læsetest, hvor respondenterne er eller har været arabiskstuderende ved danske universiteter. I det følgende præsenteres fire hypoteser, der er formuleret på baggrund af de teoretiske afsnit, og der gives en metodebeskrivelse af, hvordan de søges afprøvet. Derefter beskrives respondentunderlaget, og der gives en kort karakteristik af de respondenter, der indgik i undersøgelsen. Derpå følger testresultaterne og analyser af disse: Først behandles de enkelte deltest kvantitativt, hvorefter resultaterne af den ene deltest udforskes nærmere gennem en kvalitativ analyse af respondenternes konkrete oplæsningsfejl. Bemærk, at analyseafsnittet ikke blot sigter mod at afprøve de opstillede hypoteser, men mod på baggrund af data at tilvejebringe så meget information som muligt om læsning på arabisk. Først i sammenfatningen, der afslutter kapitlet, sættes resultaterne direkte i relation til de konkrete hypoteser.

6.1 Hypoteser og metode

Den grafiske dimension – Test A

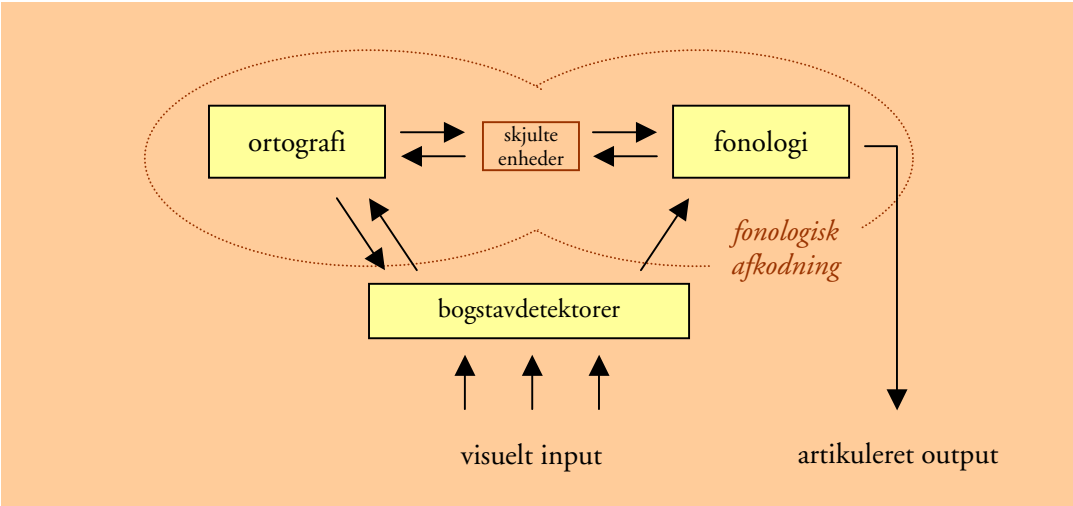
Vedrørende danske lørneres læsning af arabisk skrift kan på baggrund af kapitel 4 opstilles følgende hypotese:

Hypotese 1: Den uvante bogstavarkitektur sænker hastighed og øger fejlfrekvens for ordgenkendelse på arabisk som L2.

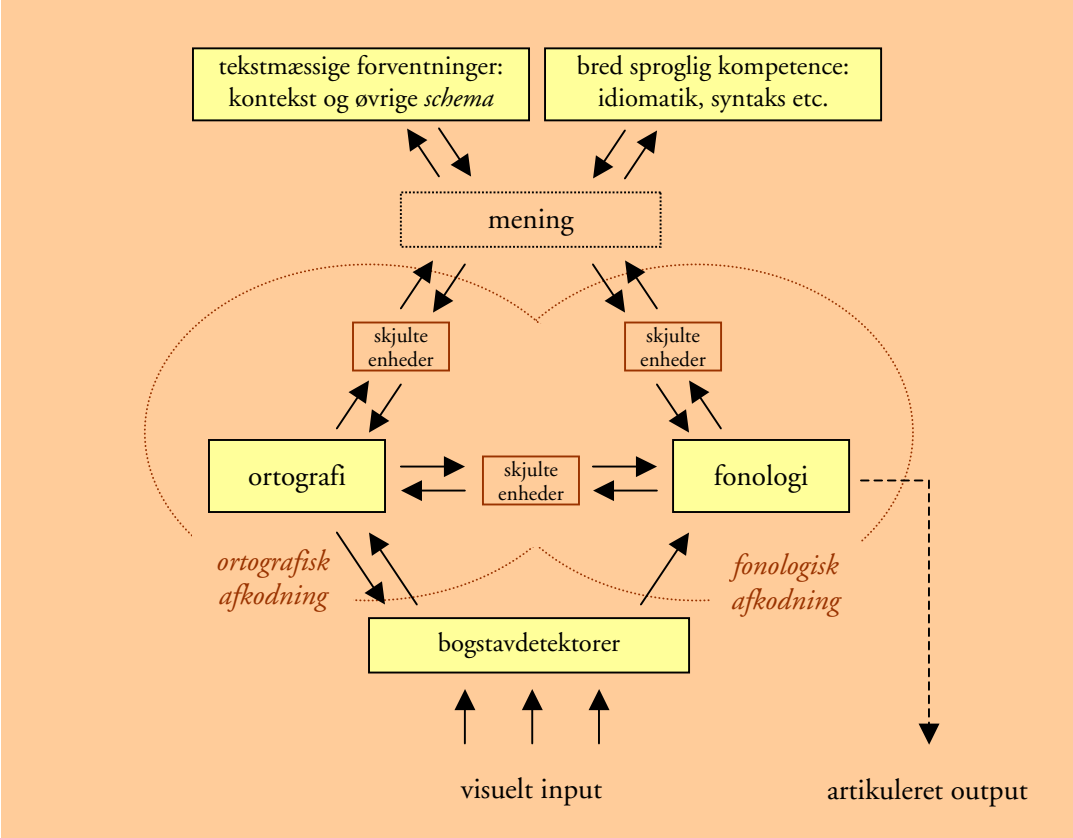
Afprøvning af hypotese 1 kræver en test, hvor læsehastighed og fejlfrekvens for ordgenkendelse registreres ved danske lørneres læsning dels af den uvante dels af den vante bogstavarkitektur, og hvor testmaterialet er renset for den semantiske funktion for at isolere effekten af den grafiske dimension. Dette kan gøres ved at sammenligne læsning af pseudoord skrevet med hhv. det arabiske og det latinske alfabet. Herved afskæres den øvre del af læseprocesmodellen, jvf. figur 6.1. Da modellen også involverer fonologi og da fremmede fonemer, som beskrevet i afsnit 2.4, formodes at influere negativt på ordgenkendelse på L2, er det nødvendigt at kontrollere for denne dimension. Derfor suppleres med følgende hypotese:

Hypotese 2: De fremmede fonemer sænker hastighed og øger fejlfrekvens for ordgenkendelse på arabisk som L2.

Afprøvning af hypotese 2 kræver en test, hvor læsehastighed og fejlfrekvens for ordgenkendelse registreres ved læsning af arabisk tekst med og uden de for danskere fremmedartede fonemer, og hvor effekten af fonologisk afkodning isoleres ved, ligesom ved afprøvningen af hypotese 1, at rense testmaterialet for den semantiske dimension. Dette kan gøres ved at sammenligne læsning af pseudoord skrevet med det arabiske alfabet, hhv. med og uden bogstaver, der repræsenterer fremmede fonemer.



Figur 6.1 Model for oplæsning af pseudoord (test A og C)



Figur 6.2 Model for læsning efter forståelse (test B). Den stiplede linje til artikuleret output markerer, at det er valgfrit for respondenterne, om de vil læse teksten højt.

De nævnte testformer kræver imidlertid visse overvejelser: For det første kan der stilles spørgsmålstegn ved, hvorvidt forskellige skriftsystemer overhovedet kan sammenlignes. Ved måling af læsning af almindelig tekst vil sproglige forskelle gøre det vanskeligt at opstille et gyldigt sammenligningsgrundlag – skal den arabiske og den danske tekst fx afstemmes efter samme antal bogstaver, fonemer, ord, morfologiske enheder eller mening? Problemet med afstemning af mening er i denne test ikke relevant, da der anvendes pseudoord, og teksterne er i dette tilfælde afstemt, så de indeholder samme antal ord, stavelser og fonemer. Pseudoord skrevet med det latinske alfabet er således konstrueret, så de i henhold til almindelige danske udtaleregler må forventes oplæst med samme trykfordeling, som den der findes i de tilsvarende arabiske pseudoord, hvor trykket afgøres af vokaliseringen.

For det andet er der spørgsmålet om, hvordan de enkelte pseudoord konstrueres. Som beskrevet i afsnit 5.1 må et arabisk pseudoord konstrueres med et gyldigt mønster. Til dette formål er anvendt en autentisk tekst, hvori alle ordrødder er erstattet enten af pseudorødder eller af rødder, der i kombination med det givne mønster ikke giver et gyldigt ord. Alle pseudoord er vokaliserede, dels fordi pseudoord uden vokalisering stiller læseren på herrens mark i forhold til udtale, når der ikke findes meningsindhold hverken i ordet eller i kontekst, dels fordi den uvokaliserede ortografi under alle omstændigheder vil umuliggøre en fornuftig afstemning med de latinskskrevne pseudoord.

For det tredje er der spørgsmålet om målemetode. Registreringen af respondenternes omkodning af det grafisk input kan kun foretages på baggrund af artikuleret output – dvs. respondenternes højtlæsning af pseudoordene. Testmetoden involverer altså højtlæsning, hvilket – som illustreret i figur 6.1 – indebærer, at den fonologiske kode overføres til artikulatoriske kommandoer, og at disse i praksis udføres. Imidlertid er der ved sammenligning af ordgenkendelse af engelske ord med forskellig frekvens, ortografisk redundans og grad af grafem-fonem korrespondance, påvist omtrent samme hastighedsforskydninger i konnektionistiske simulationers ordgenkendelse og menneskelige forsøgspersoners højtlæsning (Seidenberg & McClelland 1989:531). Dette taler for, at det er forsvarligt at bruge højtlæsning som parameter for ordgenkendelse. Rent praktisk er der ikke mulighed for at måle på læsning af enkeltord. Hertil kræves et avanceret og fintfølende apparatur, der på millisekunder kan registrere tidsafstand mellem stimulus og respons. Et sådant apparatur er ikke tilgængeligt, og der er inden for dette projekts rammer ikke midler, der muliggør anskaffelse af det. I stedet vælges en metode, hvor respondenterne læser længere pseudoordsekvenser. Denne metode har – ud over at være realiserbar – den fordel, at sekvenserne til en vis grad imiterer en reel læseproces, hvor lørneren af et nyt skriftsystem efter kort tid 'løber træt', hvorved opmærksomheden svækkes. Denne effekt vil med den valgte metode indgå i resultatet og formentlig resultere i større, og dermed mere målbare, forskelle mellem de forskellige færdighedsniveauer – og vel at mærke forskelle, der også kommer til udtryk i reelle læsesituationer.

Testmaterialet består af følgende:

- ét afsnit bestående af pseudoord af variabel længde skrevet med arabisk, vokaliseret skrift **uden** fremmede fonemer (AI). Se Appendix 1.3
- ét afsnit bestående af pseudoord af tilsvarende længde skrevet med arabisk, vokaliseret skrift **med** fremmede fonemer (AII). Se Appendix 1.4
- ét afsnit bestående af pseudoord af tilsvarende længde skrevet med latinsk skrift (AIII). Se Appendix 1.5

De arabiske afsnit er sat med fonten *Traditional Arabic*, 24 pkt., det latinskskrevne med *Times New Roman*, 20 pkt. Hvert afsnit præsenteres på et selvstændigt A4-ark i forskellig rækkefølge. Tid for læsning samt antal og typer af læsefejl hos respondenterne registreres. Herved måles:

- forskel på hastighed og fejlfrekvens for visuel ordgenkendelse ved læsning af arabisk og latinsk skrift (hypotese 1)
- forskel på hastighed og fejlfrekvens for visuel ordgenkendelse ved læsning af arabisk skrift med og uden fremmede fonemer (hypotese 2)

Desuden dokumenteres forsøgspersonernes konkrete læsefejl, hvilket muliggør en kvalitativ analyse af fordelingen af fejltyper på individer og niveauer. Til fejlregistrering benyttes en videreudviklet version af Hanas (1972) og Khaldiehs (1996), hvor Hanas bygger på Monroes system. Som beskrevet i afsnit 1.2 opererer Hana med følgende kategorier:

- vokalfejl
- konsonantfejl
- tilføjelse af lyd
- udeladelse af lyd
- gentagelse af lyd

Khaldieh beskæftiger sig kun med konsonantfejl og opererer med to kategorier:

- visuelle fejl
- fonologiske fejl

Hanas kategorisering synes at være for upræcis til dette formål: En samlet kategori af 'vokalfejl' er problematisk, fordi vokaler kan have forskellige former for status i et ord. Fx er udeladelse af arabiske endelsesvokaler muligvis almindelige, fordi de på rigtige ord ofte er kasusmarkører, som mange læsere ikke har glæde af at fokusere på, når de læser efter forståelse. Kategorien muliggør heller ikke en konsekvent skelnen mellem fejl i bogstaver og i diakritiske tegn – mens korte vokaler på arabisk skrives med diakritiske tegn, skrives de lange med bogstaver, og netop denne skelnen mellem bogstaver og diakritiske tegn er essentiel, fordi den kan afdække hvorvidt fejl hyppigst finder sted inden for den ene af disse kategorier. En samlet kategori af 'konsonantfejl' synes ligeledes uhensigtsmæssig, fordi den omfatter fejl, der kan have vidt forskellige kilder – der kan være tale om forveksling af bogstaver pga. grafisk lighed, fonologisk lighed, manglende genkendelse eller manglende evne til at udtale det rette fonem. Khaldiehs kategorisering inden for konsonantfejl er i den forbindelse umiddelbart værdifuld, men desværre ikke metodisk kvalificeret. Det er nemlig ikke muligt at vide præcis hvorfor et bogstav udtales forkert – hvis et bogstav forveksles med et andet, som ligner det grafisk, kan det være oplagt at kategorisere det som en visuel fejl, men at der rent faktisk er tale om en visuel forveksling er jo ikke givet – især fordi grafisk slægtskab mellem forskellige bogstaver i mange tilfælde hænger sammen med et fonetisk slægtskab mellem samme bogstaver, som det fremgår af figur 3.4.

På den baggrund konstrueres følgende fejlkategorier for læsning af AI og AII, dvs. passagerne med arabiske pseudoord:

Bogstavfejl:

- Individuel fejl: Foreløbigt ukategoriseret fejl i form af udeladelse eller tilføjelse af konsonant eller forveksling mellem bogstaver, som ikke umiddelbart synes at være hverken grafisk eller fonetisk beslægtede.
- Grafisk/fonetisk fejl: Forveksling mellem bogstaver, som umiddelbart synes at være grafisk og/eller fonetisk beslægtede.
- Repetition: Gentagelse af en hel stavelse. Gentagelse af konsonant betegnes i arabisk skrift ved *shadda* og kan derfor ikke udgøre en bogstavfejl. Repetition af en stavelse bestående af én konsonant og én vokal registreres som en enkelt fejl.
- Ombytning: ombytning af to bogstaver med eller uden vokal imellem.
- Lang vokal: Udeladelse, ændring eller tilføjelse af en lang vokal.
- Hamza: Udeladelse eller tilføjelse af *hamza* (stemmebåndslukke).

Diakritiske fejl:

- Kort vokal: Udeladelse, ændring eller tilføjelse af en vokal, der ikke er en endelse.
- Enkel endelse: Udeladelse, ændring eller tilføjelse af en endelsesvokal.
- Tanwiin-endelse: Udeladelse, ændring eller tilføjelse af en kasusendelse, der markerer ubestemt form (-un, -an, -in)
- Shadda: Udeladelse eller tilføjelse af *shadda* (fordoblingstegn)

Der ses bort fra enkelte udtaletekniske fejl, som er irrelevante i relation til undersøgelsens mål, da de udelukkende vedrører begynderstuderendes manglende kendskab til udtaleregler og i øvrigt under normale omstændigheder er enten morfologisk eller syntaktisk betingede, hvilket ikke giver mening ved læsning af pseudoord. Det drejer sig om følgende fejltyper:

- Manglende overførsel af de såkaldte solbogstaver til erstatning af ل ('laam') i den bestemte artikel, som hvis fx pseudoordet النُّلْكُ [an-nulku] i test AI fejlagtigt udtales [al-nulku].[1]
- Udtale af den lange vokal ا ('alif') i forbindelse med tanwiin-endelsen i akkusativ, som hvis fx pseudoordet مَنِكاً [manikan] i test AI fejlagtigt udtales [manikaan].[2]
- Udtale af den lange vokal ا ('alif') i forbindelse med den verbale pluralisendelse وا - , som hvis pseudoordet نَلَبُوا [nalabuu] i test AI fejlagtigt udtales [nalabuuaa].[3]

Kategorierne 'Indviduelle fejl' og 'Grafisk/fonetisk fejl' analyseres efterfølgende nærmere, idet de konkrete oplæsningsfejl noteres, hvilket muliggør en post hoc analyse, hvor særligt høje hyppigheder af bestemte bogstavforvekslinger kan indicere, at fejlene skyldes grafisk eller fonologisk slægtskab bogstaverne imellem frem for individuelle læsefejl.

Til brug for AIII, dvs. de latinskskrevne pseudoord, anvendes følgende fejlkategorier:

---

[1] Denne udtaleregel gælder for dentale, alveolære og postalveolære konsonanter (eksklusiv ج [ʒ], inklusiv emfatiske), der efterfølger den bestemte artikel, således at artiklens [l] erstattes af en fordobling af den efterfølgende konsonant.

[2] Den ubestemte akkusativ-endelse ledsages på denne måde altid af et 'alif', der ikke udtales.

[3] Maskulinum pluralis i 3. person perfektum og imperativ har denne bøjningsendelse, hvor 'alif' er stumt. Denne bogstavkombination i slutningen af et ord forekommer ellers ikke.

- Vokalfejl
- Konsonantfejl
- Ombytning
- Repetition

Da det latinske alfabet anvendes til at skrive en lang række sprog, hvor forskellige udtaleregler gør sig gældende, er det nødvendigt at tillade en vis variation i udtale i denne test. Følgende registrerede udtaleformer hos de marokkanske respondenter, der har kendskab til fransk hhv. spansk, accepteres på det grundlag som korrekte:

- Udeladelse af e i slutningen af et ord, fx 'koruse' læst som [korus]
- 'u' udtalt som [y]
- 'j' udtalt som [x]

Derimod registreres fx udtale af 'v' som [f] som en fejl. Ingen af de marokkanske respondenter rapporterer om særligt kendskab til et sprog, hvor denne udtaleregel eksisterer, fx tysk. Fejlen skyldes derfor formentlig, at fonemet [v] ikke findes på arabisk. Det svarer til danske lørneres fejlagtige læsning af uvante arabiske fonemer, der også registreres som fejl.

Ved sammenligning mellem læsning af det arabiske og det latinske alfabet er det på grund af forskellene mellem de to skriftsystemer og de større variationsmuligheder ved læsning af det latinske alfabet nødvendigt at konstruere en særlig fejlkategori til formålet. Da hensigten med sammenligningen er at undersøge respondenternes bogstavgenkendelse, udgår i denne kategori alle fejl relateret til de diakritiske tegn, *hamza* og længden af vokaler i test AI og AII. Kategorien kaldes 'realfejl' og omfatter altså kun de fejltyper, der registreres i alle tre test, dvs.:

- Substitution: Forveksling mellem bogstaver, dvs. summen af forvekslingsfejl under kategorierne individuelle fejl og grafisk/fonetiske fejl i test AI og AII og summen af vokal- og konsonantfejl i test AIII.
- Ombytning
- Repetition

I alle tre test gælder det, at hvis respondenten gentager oplæsning af det samme pseudoord, registreres kun den sidste oplæsning. Dette giver vedkommende mulighed for at rette sig selv uden at dette influerer på fejlregistreringen.

Den ortografiske dimension – Test B

Hvad angår det konsonantiske skriftsystems indflydelse på læseprocessen kan på baggrund af de teoretiske afsnit udledes følgende hypoteser relateret til vokalproblematikken:

Hypotese 3: Manglende vokaler har negativ indflydelse på afkodningshastighed hos læsere med arabisk som L2

Hypotese 4: Manglende vokaler har negativ indflydelse på læseforståelse hos læsere med arabisk som L2

Afprøvning af hypotese 3 kræver en test, hvor læsehastighed og forståelse testes på baggrund af et vokaliseret (BI) og et uvokaliseret (BII) tekststykke, der niveau- og omfangsmæssigt er sammenlignelige. For at imødekomme de begrænsede sproglige forudsætninger hos respondenter på niveau 1, må der udformes en særlig letlæst version til dette niveau. For at korrigere for indflydelsen fra forskelle i ordforråd mellem de to tekster på hvert niveau, må halvdelen af respondenterne læse den ene tekst (tekst A) vokaliseret og den anden tekst (tekst B) uvokaliseret, mens den anden halvdel læser tekst A uvokaliseret og tekst B vokaliseret.

Testmaterialet består dermed af følgende:

- fire korte tekster på vokaliseret arabisk (BI):
    - BI, niveau 1, tekst A. Se Appendix 1.6
    - BI, niveau 1, tekst B. Se Appendix 1.10
    - BI, niveau 2, tekst A. Se Appendix 1.14
    - BI, niveau 2, tekst B. Se Appendix 1.18
- fire korte tekster på uvokaliseret arabisk (BII):
    - BII, niveau 1, tekst A. Se Appendix 1.7
    - BII, niveau 1, tekst B. Se Appendix 1.11
    - BII, niveau 2, tekst A. Se Appendix 1.15
    - BII, niveau 2, tekst B. Se Appendix 1.19
- fire ark med *multiple choice*, der kontrollerer forståelse for hver tekst (BI og BII) på dansk til brug for de danske respondenter:
    - Svarark, niveau 1, tekst A. Se Appendix 1.8
    - Svarark, niveau 1, tekst B. Se Appendix 1.12
    - Svarark, niveau 2, tekst A. Se Appendix 1.16
    - Svarark, niveau 2, tekst B. Se Appendix 1.20
- fire ark med *multiple choice*, der kontrollerer forståelse for hver tekst (BI og BII) på arabisk til brug for kontrolgruppen af indfødte sprogbrugere:[4]
    - Svarark, niveau 1, tekst A. Se Appendix 1.9
    - Svarark, niveau 1, tekst B. Se Appendix 1.13
    - Svarark, niveau 2, tekst A. Se Appendix 1.17
    - Svarark, niveau 2, tekst B. Se Appendix 1.21

Med henblik på at kunne henføre det enkelte svarark til henholdsvis test BI og BII udformes materialet, så hvert enkelt svarark – ud over tekstversion og niveau – mærkes med hhv. 'BI' eller 'BII' afhængigt af, om testarket knytter sig til læsning af den vokaliserede eller den uvokaliserede version af teksten. Svararkene udformes altså i 16 forskellige versioner.

Teksterne tilpasset lørnere på niveau 1 (tekstniveau 1) er en historie om en families dagligdag og en anden om, hvad man kan læse i avisen. Begge tekststykker er fra lærebogen: محمد بدوي و فتحي علي يونس (١٩٨٨) الكتاب الاساسي في تعليم اللغة العربية لغير الناطقين بها، الجزء الأول، المنظمة العربية للتربية والثقافة والعلوم، تونس.صفحة ١٦٦-١٦٩(الدرس ١٠) وصفحة ٣٠٥-٣٠٩(الدرس ٢٠). Det kontrolleres, at ingen af respondenterne har anvendt denne bog i undervisningen. Teksterne er let revideret i forhold til originalversionen, så de passer i ordforråd og længde

[4] Da undersøgelsen omhandler læsning på arabisk, anvendes termen 'indfødte sprogbrugere' her og efterfølgende i betydningen indfødte **arabisksprogede** sprogbrugere – også i tilfælde, hvor der er tale om læsning af latinsk skrift.

til hinanden og til testformen. De er sat med fonten *Traditional Arabic*, 20 pkt. Tekst A er på 124 ord og 1032 tegn (ekskl. diakritiske tegn), mens tekst B er på 124 ord og 1058 tegn.

Teksterne tilpasset respondenterne på niveau 2 og 3 er bearbejdede udgaver af et par små hæfter i en serie af historier om 'Joha' – en figur fra et utal af arabiske anekdoter og fabler:
ابو المجد أحمد عبد الرحمن: جحا حزين على حماره. نوادر جحا للأطفال ٣٨. المؤسسة العربية الحديثة. القاهرة
شفيق عطا: جحا أكل السمك. نوادر جحا للأطفال ٥٩. المؤسسة العربية الحديثة. القاهرة.
Også her er teksterne sat med *Traditional Arabic*, 20 pkt. Tekst A er på 192 ord og 1358 tegn (ekskl. diakritiske tegn), mens tekst B er på 190 ord og 1354 tegn.

Hver respondent modtager historier og svarark på separate A4-ark. Som anført er svararket på dansk for danske respondenter, mens det for de indfødte sprogbrugere er oversat til arabisk. Hvert svarark består af fem *multiple choice*-opgaver med tre svarmuligheder til kontrol af forståelse. *Multiple choice* er en omdiskuteret opgavetype, men er valgt til formålet af følgende grunde:

- *Multiple choice* tester mere entydigt læseforståelse end andre testtyper som fx genfortælling, resume eller oversættelse, hvor respondentens produktive færdigheder kan influere på resultatet.
- *Multiple choice* giver en entydig måleenhed i form af antallet af korrekte besvarelser.

Testtypen giver i princippet også mulighed for at måle forskellige niveauer af forståelse, der involverer både semantiske, læsestrategiske og pragmatiske kompetencer – fx. forståelse af enkelte ord, overblik over tekstens generelle indhold og opfattelse af forfatterens tone og holdning. I denne sammenhæng er det dog ikke hensigtsmæssigt at udnytte hele dette spektrum af muligheder. Det samlede tidsforbrug til testforløbet må afpasses efter, at respondenterne og deres undervisere skal være villige til at ofre undervisningstid på formålet. Dette hensyn sætter nogle begrænsninger for teksternes omfang og antallet af opgaver. Så i stedet for at inkludere forskellige niveauer af tekstforståelse prioriteres det at dække ét tekstlæsningsniveau så godt som muligt, hvorfor alle opgaverne målrettes test af forståelse af de centrale elementer i de enkelte afsnit i hver tekst. Da alle svarark skal indeholde lige mange opgaver og alle opgaverne relatere til nogenlunde samme tekst-læsningsniveau, bliver resultatet, at antallet af opgaver begrænses til fem for hver historie. I relation til den kvantitative analyse er et så begrænset antal ikke optimalt, men på baggrund af de nævnte tidsmæssige og metodemæssige restriktioner accepteres dette testdesign.

Halvdelen af respondenterne læser den vokaliserede tekst først og den uvokaliserede tekst sidst, mens den anden halvdel læser den uvokaliserede tekst først og den vokaliserede sidst. Respondenterne instrueres i at læse hver enkelt historie færdig og derefter lægge den væk, mens de udfylder det tilhørende svarark. Det er ikke tilladt at gå tilbage og se i teksten i forbindelse med besvarelserne. Respondenterne oplyses eksplicit om, at det står dem frit for, om de vil læse teksterne højt eller 'indenad'. Tid for læsning registreres af respondenterne selv efter endt læsning af hver enkelt historie, men inden besvarelse af det tilhørende svarark. Denne tidsregistrering foregår ved hjælp af *timer*'ne i sproglaboratoriestationerne.

Teksterne på niveau 1 læses af lørnere på niveau 1. Teksterne på niveau 2 læses af lørnere på niveau 2 og 3. De indfødte sprogbrugere læser begge tekster. Formålet er en kontrollæsning af niveauforskellen. Herudover indgår deres læsning af niveau2-teksten i sammenligning med de øvrige niveauer som niveau 4.

Tid for læsning og antal korrekte besvarelser registreres. Herved måles:

- Forskel i læsehastighed ved læsning med og uden vokaler (hypotese 3)
- Forskel i forståelse ved læsning med og uden vokaler (hypotese 4)

Den ortografiske dimension – Test C

Ud over denne test af vokalernes rolle opstilles på baggrund af de teoretiske afsnit endnu en hypotese i relation til den arabiske ortografis rolle for L2-læsning:

Hypotese 5: L2-lørnere mangler kendskab til de morfologiske strukturer, som indfødte sprogbrugere støtter sig til som kompensation for de manglende vokaler.

Hypotese 5 kan afprøves med en test, hvor L2-lørnere og indfødte sprogbrugere ved højtlæsning skal vokalisere uvokaliserede pseudoord, hvis bogstavstruktur determinerer, hvilken vokalisering der er forenelig med det morfologiske system.

Testmaterialet består af følgende:

- Ni pseudoord, som har en bogstavstruktur, der gør udvalget af valide mønstre begrænset ved applicering af vokaler (C). Se appendix 1.21

De ni pseudoord er konstrueret til formålet. Disse og de mønstre, der i henhold til deres ordstrukturer er valide, ses i figur 6.3. Pseudoordene er sat med fonten *Traditional Arabic*, 24 pkt. med tre rækker á tre ord på et A4-ark.

mulige valide mønstre	ortografisk struktur	Pseudo-ord	
اِنفَعلَ	انفعل	انبفص	1
اِستفْعلَ	استفعل	استهفل	2
مفْعول	مفعول	مصبوم	3
يتفاعل	يتفاعل	يتشارل	4
اِنفِعال	انفعال	انكشال	5
فاعِل، فاعلَ	فاعل	طاهث	6
اِستِفْعال	استفعال	استفداك	7
مفاعلة	مفاعلة	مكادغة	8
اِفْتِعال	افتعال	امتفال	9

Figur 6.3 Pseudoord konstrueret til brug for test C samt deres ortografiske struktur (jvf. kapitel 3) og mulige valide mønstre

Vokalisering af pseudoordene ved højtlæsning registreres. Herved måles:

- i hvor høj grad bogstavstrukturen automatisk fremkalder de valide mønstre hos respondenterne på de forskellige færdighedsniveauer (hypotese 5)

Spørgeskemaundersøgelse

Ud over de anførte test, der sigter mod praktisk at belyse de teoretiske spørgsmål der rejses i kapitel 4 og 5, bedes respondenterne om at besvare et spørgeskema. Formålet med spørgeskemaet er at indhente følgende former for oplysninger:

- Personlige data: Respondenternes alder, køn og uddannelse til brug for en overordnet karakteristik af respondentgruppernes sammensætning.
- Kendskab til fremmedsprog: Muliggør frasortering af respondenter, der ikke matcher målgruppen – hos de danske lørnere pga. kendskab til det arabiske skriftsystem før påbegyndelse af arabiskuddannelsen og hos de marokkanske respondenter pga. mere erfaring med det latinske skriftsystem, end den der kan opnås inden for det marokkanske skolesystems rammer. Formålet hermed beskrives nærmere i afsnit 6.2.
- Indtryk og erfaringer vedrørende læsning på arabisk (danske lørnere) eller på europæiske sprog (marokkanske respondenter).
- Samtykke til at respondenternes testresultater indgår som empirisk materiale i nærværende afhandling.

Spørgeskemaet er for de danske respondenter på dansk (Appendix 1.23) og for de marokkanske respondenter på arabisk (Appendix 1.24)

Analyseforløb:

Efter testforløbet blev alle lyddata digitaliseret og CoolEdit 2000 blev anvendt som audio-software i analyseforløbet. I test A blev hver enkelt respondents tid for læsning af hver af de tre deltest opmålt ud fra det visuelle lydbølgeforløb, hvorved der kan måles på hundrededele af sekunder. Oplæsningsfejl blev registreret manuelt på udskrifter af hver deltest og derefter systematiseret efter de relevante fejlkategorier i et fejlregistreringsskema (se Appendix 2.1). I test C blev respondenternes oplæsning ligeledes registreret manuelt på en udskrift af testmaterialet. Især resultaterne i test A viste sig komplicerede og krævende at registrere pga. en meget høj fejlfrekvens hos nogle respondenter. Derfor blev alle registreringer af lyddata efterkontrolleret to gange af andre personer end forfatteren – og mindst én af disse gange af en indfødt sprogbruger. I test B forelå resultaterne umiddelbart i form af besvarelser af *multiple choice* og respondenternes egen registrering af tidsforbrug ved læsning af tekststykkerne.

For hver respondent blev personlige data, de numeriske resultater i test A og resultaterne af test B og C derefter samlet på et enkelt ark (se Appendix 2.2), hvorefter de blev indtastet i Intercooled Stata.

Endelig blev alle substitutionsfejl i test A systematiseret i et fejlskema, hvor alle de forekommende fejlkombinationer og hyppigheden af dem blev registreret for hver respondent. Fejlskemaet med alle de indtastede observationer foreligger i Appendix 4.1.

Kun substitutionsfejl blev inddraget, da kun denne fejltype kan relateres direkte til bogstavgenkendelse.

I afsnit 6.3 gennemgås de kvantitative analyser på baggrund af grafiske fremstillinger af data, som suppleres med statistisk verifikation ved hjælp af ANOVA (*analysis of variance*) med robust standardfejl gennemført i statistikprogrammet Intercooled Stata 8. Robust standardfejl anvendes for at kompensere for, at data ikke er helt normalfordelte. I test hvor der indgår mere end to grupper, testes der for den aktuelle variabels generelle effekt, ligesom der laves tovejsanalyser mellem de enkelte grupper. Signifikansniveauet er sat til de traditionelle 5%. Da populationen i niveau 3 er for lille, indgår det ikke på noget tidspunkt som selvstændigt parameter i analysen. De grafiske fremstillinger af datamaterialet er lavet i Intercooled Stata 7 og 8. Der er anvendt følgende diagramtyper:

- Punktdiagrammer med kassediagram-effekt: Angiver de eksakte observationer. Punktdiagrammer er valgt til at visualisere hovedparten af resultaterne, fordi andre diagramtyper ved så lille en population kan skjule eventuelle skævvridninger, der skyldes ganske få ekstreme observationer. Kassediagram-effekten er sat til at markere gennemsnit og standardafvigelse frem for median og kvartiler, der er standard for traditionelle kassediagrammer. Dette valg er truffet, fordi gennemsnittet relaterer til ANOVA'en, og fordi alle observationer alligevel er synlige.
- Spredningsdiagrammer: Anvendes til visualisering af kontrolanalyser mellem tværgående parametre i test A og B. Formålet hermed uddybes i afsnit 6.3.
- Boblediagrammer: Anvendes til visualisering af fejlfrekvens i test B og C. Cirklernes størrelse afspejler antallet af observationer i den relevante kategori.
- Søjlediagrammer: Angiver gennemsnitsværdier. Søjlediagrammer er anvendt til visualisering af resultaterne af spørgeskemaundersøgelsen og ved opsummering af resultaterne inden for hver deltest. Da de udelukkende gengiver gennemsnit, er det i relation til selve læsetestene ikke hensigtsmæssigt at lade dem stå alene i en undersøgelse med så lille en population, men de er velegnede til at give en let overskuelig visualisering af de samlede resultater, idet de bedre end de øvrige diagramtyper tillader et overblik på tværs af flere parametre i samme diagram.

Udskrift af datasættet, de statistiske analyser i deres fulde form og kommandoer til programmering af graferne foreligger som henholdsvis Appendix 3.1, 3.2 og 3.3. Lyddata foreligger i komprimeret format på CD-rom som Appendix 4.3.[5]

6.2 Beskrivelse af respondenterne

Som udgangspunkt var målet at inddrage alle danske studerende, der modtog undervisning i kommunikativt arabisk i forårssemestret 2002, hvor testene blev gennemført. Herved udgik studerende ved Aarhus Universitet, hvor arabiskstudiet er udelukkende filologisk anlagt. Konsekvensen af at udelukke disse lørnere er ganske vist et begrænset respondent-grundlag og dermed en svækkelse af undersøgelsens validitet, hvad angår kvantitativt baserede konklusioner. Men emnet for afhandlingen er kommunikative læsefærdigheder, og

[5] Det komplette, originale datamateriale i form af kassettebånd, skriftlige besvarelser og forfatterens analyseark er i bedømmelsesperioden tilgængeligt på Institut for Kultur, Historie og Samfundsbeskrivelse ved Syddansk Universitet.

	KU: BA i arabisk	ECTS	Timer	SDU: BA negot. i arabisk	ECTS	Timer
1. semester	Arabisk propædeutik	15	8	Sprogfærdighed 1 (propædeutik)	11	7
2. semester	Arabisk propædeutik	15	8	Sprogfærdighed 1	12	7
3. semester	Moderne litteratur	5	2	Sprogfærdighed 2	9	6
	Oversættelse dansk-arabisk	7½	2			
	Kommunikativt arabisk	7½	4			
4. semester	Moderne litteratur	5	2	Sprogfærdighed 2	10	6
	Klassisk litteratur	7½	2			
	Dialekt	7½	4			
	Kommunikativt arabisk	7½	4			
5. semester	Moderne litteratur	5	2	Sprogfærdighed 3	13	14
	Klassisk litteratur	7½	2	Historie og samfundsforhold	6	3
				Erhvervskommunikation	2	2
				Humanistisk BA-Projekt	6	
6. semester				Sprogfærdighed 3	17	16
				Religion og kultur	6	3
				Humanistisk BA-Projekt	6	
7. semester				Forhandlingsteori og -træning	6	2
				Interkulturel kommunikation	7	3

Figur 6.4 Fag med arabisk sprog som emne og/eller undervisningsmedium på bachelordelen af kandidat-uddannelsen i arabisk ved KU og Cand. negot. uddannelsen i arabisk ved SDU. Begge uddannelser er udvidet med et semester i forhold til normalforløbet pga. det propædeutiske forløb. På Cand. negot. uddannelsen foregår 5. og 6. semester i et arabisk land.

den form for læseinstruktion der praktiseres i den filologiske tilgang, er uforenelig med den afgrænsning der i afsnit 1.1 er givet af læsning som en kommunikativ kompetence.

Arabiskuddannelserne ved Københavns Universitet (KU) og Syddansk Universitet (SDU), der altså som de eneste universitære institutioner i Danmark udbyder kommunikativt arabisk, er opbygget meget forskelligt. Ved Københavns Universitet er arabiskuddannelsen en humanistisk uddannelse med stor vægt på litteraturstudier, men med indbyggede kommunikative elementer i form af fagene 'kommunikativt arabisk' og 'dialekt' på 3. og 4. semester, ligesom halvdelen af undervisningstiden i det propædeutiske forløb er afsat til samtale på MSA. Grammatiske forklaringer og litteraturanalyse foregår på dansk. Ved Syddansk Universitet er arabisk en sproglig linje under Cand. negot., som er en kombineret humanistisk og samfundsvidenskabelig uddannelse fordelt med 50% til hvert område. Al sprogundervisning er kommunikativ og så vidt det er muligt er arabisk undervisningssprog i alle sprogfag efter første semester. Endvidere er det obligatorisk at aflægge 5. og 6. semester på en uddannelsesinstitution i et arabisk land. Fordelingen af sprogfagene og den vægt, hvormed de indgår på bachelordelen af de to uddannelser, fremgår af figur 6.4. På overbygningen kan man i København vælge mellem en sproglig og en samfundsfaglig linje. På den sproglige linje undervises i arabisk sprogvidenskab, dialektologi og litteratur, hvor undervisningssproget er dansk. Kommunikativt arabisk indgår ikke. Ved Syddansk Universitet er overbygningen også linjedelt på den samfundsvidenskabelige studiedel, mens fagene 'områdestudier' og 'forhandlingsteori', hvor undervisningssproget og litteraturen er arabisk, er obligatoriske, og specialeresumeet skal også være på arabisk.

En række praktiske omstændigheder omkring de to uddannelsesforløb lægger nogle metodiske og praktiske begrænsninger på respondentgrundlaget, der reducerer dette yderligere. Disse kan sammenfattes til følgende punkter:

- Studerende har før 2. semester ikke opnået et tilstrækkeligt kendskab til skriftsystemet til at testning er realistisk.
- Efter 4. semester er undervisningen ved KU udelukkende litterært orienteret.
- På 5. og 6. semester er de studerende ved Syddansk Universitet spredt ud over Mellemøsten, hvorfor det ikke er praktisk muligt at teste dem.
- Der er på Cand. negot. uddannelsen kun optag hvert andet år, hvilket begrænser antallet af studerende og giver stor spredning i niveau.

På grund af tidsmæssige begrænsninger var det nødvendigt at gennemføre alle test i samme semester. Med henblik på at inddrage forskellige færdighedsniveauer blev det – på baggrund af ovennævnte restriktioner – besluttet at teste alle studerende på de to uddannelsers 2. semester og samle dem i et fælles færdighedsniveau (niveau 1) samt teste studerende på 4. semester ved KU (niveau 2). På grund af det spredte optag på Cand. negot. var der ikke nogen negot.-studerende på 4. semester på tidspunktet for testen. Dog blev en enkelt studerende, der egentlig læste på 6. semester, men som af personlige årsager havde været tvunget til at afbryde sit udlandsophold på 5. semester meget tidligt i dette forløb, inkluderet i denne gruppe.

Desuden blev mulighederne for at inddrage et tredje færdighedsniveau undersøgt. De studerende på overbygningen var imidlertid vanskeligere at få kontakt til og passe ind i testforløbet. Grunden hertil var først og fremmest:

- Der er kun få studerende indskrevet på overbygningsuddannelserne.
- Endnu færre følger overbygningsuddannelsernes normalforløb.
- Der er en stor spredning i niveau, dels på grund af forskelle uddannelserne imellem, dels på grund af forskellige individuelle dispositioner i forhold til fx studieophold i udlandet.

På trods af disse ulemper blev der taget personlig kontakt til alle studerende på overbygningen på Cand. negot. samt til fire studerende, der var færdiguddannet det pågældende semester, ligesom et antal færdiguddannede kandidater fra Københavns Universitet blev kontaktet. For at udvide populationen blev alle studerende på Åben Uddannelse i arabisk ved Syddansk Universitet desuden inddraget. Disse studerende var på uddannelsens 4. semester, hvilket i undervisningstid svarer til 2. semester på Cand. negot. uddannelsen, hvorfor de niveaumæssigt blev inkluderet i denne gruppe (niveau 1).

Som kontrolgruppe blev inddraget en gruppe indfødte sprogbrugere[6], der alle var studerende ved Al-Akhawayn Universitetet i Marokko. Universitetet er Marokkos eneste universitet, hvor undervisningen er betalingspligtig. Undervisningssproget er engelsk. Dette universitet blev valgt, fordi de nødvendige sproglaboratoriefaciliteter var til stede, og fordi marokkanske respondenter blev anset som egnede, fordi der i den marokkanske talesprogsvariant er en relativ god fonemisk korrespondance med MSA – fx bruges på

[6] Som beskrevet i afsnittene om diglossi i kapitel 3 og 5 afviger 'indfødte sprogbrugere' på arabisk fra det man normalt forbinder med begrebet, idet afstanden mellem talesproget og MSA gør, at sprogbrugerne først ved skolestart konfronteres med det sprog, de læser på, hvorved det er at sammenligne med et andetsprog. Da tværsproglige sammenligninger i dette studium kun vedrører læsning af pseudoord, har det dog ingen metodetekniske implikationer.

marokkansk de fonemerne 'q' og '?', som i de centrale talesprogsvarianter er næsten fraværende. De studerende var indskrevet ved en række forskellige studier og fulgte undervisning i engelsk som fremmedsprog på to forskellige hold.

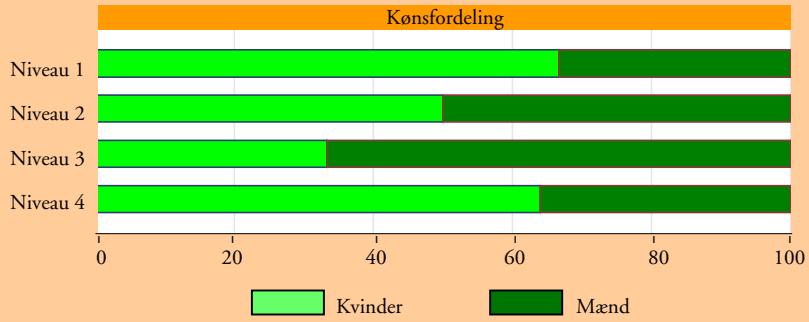
Testforløbet foregik for alle de marokkanske respondenter og for de danske respondenter på grunduddannelsen i undervisningstiden, men alle deltagere fik mulighed for at afstå fra at lade deres testresultat indgå i undersøgelsen. Ingen af de studerende benyttede sig af denne mulighed. Hvad angår de danske overbygningsstuderende og færdiguddannede, der indvilgede i at deltage, blev der aftalt test på tidspunkter, der tilgodeså deres øvrige aktiviteter. Testene blev gennemført i sproglaboratorier på henholdsvis Syddansk Universitet i Odense, Københavns Universitet Amager og Al-Akhawayn Universitetet i Ifrane, Marokko. De danske respondenter blev testet i april 2002, og de marokkanske respondenter i oktober 2003. Der blev ved alle testsessioner som udgangspunkt afsat 60 minutter til formålet, hvilket kun for enkelte studerende på det laveste færdighedsniveau var utilstrækkeligt. Disse studerende fik lov til at bruge den tid, der var nødvendig for at færdiggøre testen, og det længste testforløb varede 80 minutter inkl. instruktion. Instruktioner blev givet både mundtligt med støtte fra en overhead og skriftligt – på dansk for de danske respondenter og på arabisk for de marokkanske respondenter. Se Appendix 1.1 og 1.2. Undertegnede overværede alle test og var under hele forløbet til rådighed for spørgsmål.

I forbindelse med testene skulle alle respondenter udfylde et spørgeskema, hvori de blandt andet skulle angive deres kendskab til andre fremmedsprog. Alle danske respondenter, der før påbegyndelse af deres uddannelse havde kendskab til arabisk eller andre sprog, der anvender det arabiske alfabet, blev efter testen frasorteret. Ligeledes blev marokkanske respondenter, der havde boet i et land uden for den arabiske verden og dermed ikke nødvendigvis var sprogligt sammenlignelige med andre indfødte sprogbrugere, sorteret fra. Endelig blev data fra respondenter, der i alle deltest havde forsømt at følge instruktionerne korrekt, holdt uden for datasættet.[7] Disse nødvendige reduktioner indskrænkede desværre

Niveau 1	Cand. negot., SDU, 2. semester BA arabisk, KU, 2. semester Åben Uddannelse, SDU, 4. semester **I alt**	10 10 7 **27 respondenter på niveau 1**
Niveau 2	Cand. negot., SDU, 6. semester BA arabisk, KU, 4. semester **I alt**	1 7 **8 respondenter på niveau 2**
Niveau 3	Færdiguddannede, Cand. negot., SDU Færdiguddannede, Kandidatuddannelsen, KU **I alt**	2 1 **3 respondenter på niveau 3**
Niveau 4	Indfødte sprogbrugere **I alt**	25 **25 respondenter på niveau 4**

Figur 6.5 Oversigt over respondenter i datasættet

[7] Nogle respondenter forsømte at følge instruktionerne i én deltest men gennemførte de øvrige som anvist. I sådanne tilfælde blev data fra de deltest, der var gennemført korrekt, indeholdt i datasættet.



Aldersfordeling

Født	før 1971	1971-1975	1976-1980	1981-1985	efter 1985
Niveau 1	6	6	14	1	0
Niveau 2	0	4	4	0	0
Niveau 3	1	2	0	0	0
Niveau 4	0	0	1	18	3

Figur 6.6 Køns- og aldersfordeling hos respondenterne på hvert enkelt niveau

populationen betydeligt. Hos de danske respondenter stod studerende med arabisk baggrund for langt den største del af denne reduktion. Antallet og niveaufordelingen af de respondenter, der opfyldte de anførte krav, ses i figur 6.5. I figur 6.6 er desuden en oversigt over den køns- og aldersmæssige fordeling af respondenterne inden for hvert af niveauerne. Heraf ses, at der er en overvægt af kvinder på niveau 1 og 4. På niveau 2 er fordelingen lige, og på niveau 3 er to af de tre respondenter mænd. Aldersmæssigt er niveauerne ret forskelligt sammensat. På niveau 1 er der forholdsvis mange født før 1971. Dette skyldes at studerende på Åben Uddannelse, som er inkluderet i niveau 1, generelt er noget ældre end 'ordinære' studerende. På niveau 4 er alle respondenterne forholdsvis unge – kun én er født før 1981. Dette afspejler at de indfødte sprogbrugere alle er første- eller andetårsstuderende, og at universitetsstuderende i Marokko generelt er yngre end danske studerende.

Som det fremgår af figur 6.5, falder antallet af brugbare danske respondenter i takt med færdighedsniveauet. Dette skyldes følgende forhold:

- Det almindelige frafald i løbet af uddannelserne.
- Det nævnte fravær af Cand. negot. studerende på 4. semester.
- En stor koncentration af studerende med arabisk baggrund på det hold på Cand. negot. uddannelsen, der på tidspunktet for testen læste på overbygningen. Disse måtte jvf. ovenstående udgå af undersøgelsen.
- Kun en enkelt kandidat fra KU havde både lyst til og mulighed for at medvirke.

Som nævnt var det kun muligt at mønstre tre danske respondenter over bachelorniveau, som opfyldte de nævnte kriterier. Disse tre havde desuden meget forskellige erfaringer med undervisning i arabisk. Alligevel indgik de i undersøgelsen, hvor de blev samlet i ét niveau, idet de trods alt skønnedes at kunne komplementere undersøgelsen med en vigtig

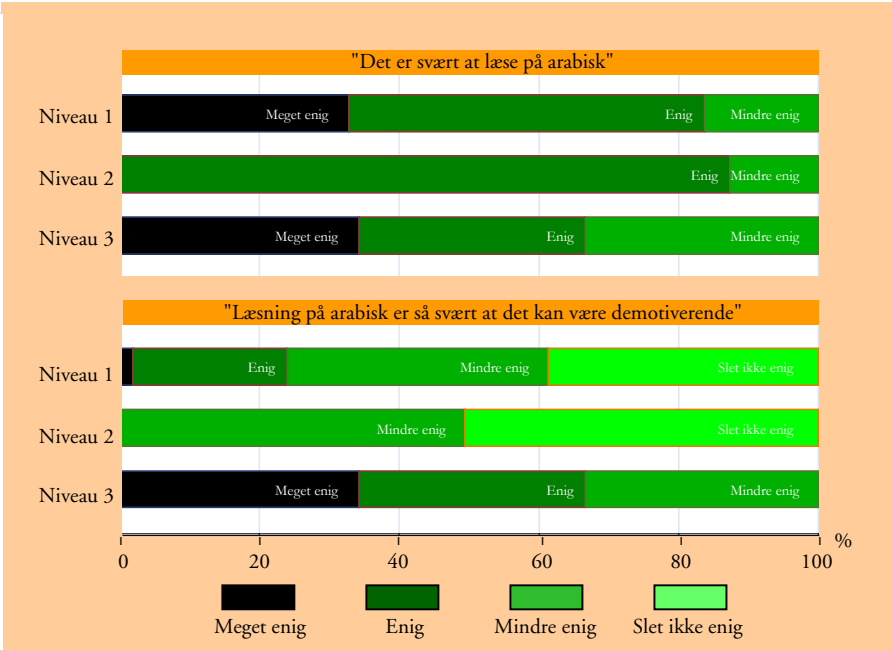
dimension i kraft af, at de repræsenterer det højeste færdighedsniveau blandt danske kandidater med arabisk som fremmedsprog. Det er dermed indlysende, at de kvantitative analyser i denne undersøgelse på grund af det lille antal respondenter mangler statistisk validitet. Som det fremgår er det beskedne respondentgrundlag imidlertid et vilkår, som det under de givne tidsmæssige og økonomiske omstændigheder ikke har været muligt at kompensere for. Den manglende statistiske validitet forhindrer da heller ikke, at undersøgelsen kan give nogle værdifulde informationer om, hvad der karakteriserer de deltagende lørneres læsning på arabisk. Det skal blot understreges, at der for analyserne må tages det forbehold, at endegyldige konklusioner vil kræve en tilsvarende undersøgelse med en større population.

6.3 Testresultater og analyser

Danske lørneres indtryk af og erfaringer med læsning på arabisk.

I figur 6.7-6.12 ses horisontale søjlediagrammer, der visualiserer resultaterne af undersøgelsen vedrørende de danske respondenters indtryk af og erfaringer med læsning på arabisk. Diagrammerne viser den procentvise fordeling af respondenter, der erklærer sig henholdsvis 'meget enig', 'enig', 'mindre enig' og 'slet ikke enig' på de udsagn vedrørende læsning på arabisk, de i spørgeskemaundersøgelsen præsenteres for. Fordelingerne af svar kan på baggrund af den lille respondentgruppe synes lidt fortegnede, især på niveau 3, hvor der jo kun er tre respondenter. Oversigterne giver ikke desto mindre et lille indblik i, hvordan danske lørnere oplever og forholder sig til at læse – og lære at læse – på arabisk.

I figur 6.7 ses reaktionerne på generelle udsagn om, at det er svært at læse på arabisk. Som det ses er der meget stor enighed om, at det faktisk er svært. Udsagnet om, at det er så svært, at det kan være demotiverende, får mindre entydig opbakning. På Niveau 1 er meningerne delte, men på niveau 2 er der slet ingen, der erklærer sig 'meget enig' eller



Figur 6.7 Respondenternes holdning til udsagn om, at det er svært at læse

'enig'. På niveau 3 er der samme svarmønster ved det første – generelle – og det andet – ret negativt ladede – udsagn. Ganske vist repræsenterer de 33%, der erklærer sig 'meget enig' kun en enkelt person, men samtidig er der ingen på niveau 3, der er helt uenige. Dette afspejler sandsynligvis, at respondenterne på niveau 3 må forholde sig til langt mere autentisk tekstmateriale end de øvrige – hvad enten det er i undervisningen eller i 'den virkelige verden'.

I figur 6.8 ses reaktionerne på udsagn der specifikt vedrører det tekstmateriale, der anvendes i undervisningen. På niveau 1 og 2 er lørnerne stort set tilfredse med sværhedsgraden af materialet, men på niveau 3 mener to af de tre respondenter, at undervisningsmaterialet er for svært. Dette afspejler nok igen den store forskel der er i materialets karakter på grund- og overbygningsuddannelserne. Mens arabiskundervisningen på grunduddannelserne først og fremmest er struktureret som fremmedsprogsundervisning, er undervisningen på overbygningerne fagligt orienterede, og materialet er autentisk og altså ikke produceret specielt til fremmedsprogslørnere.

Figur 6.8 Respondenternes holdning til udsagn vedrørende undervisningen

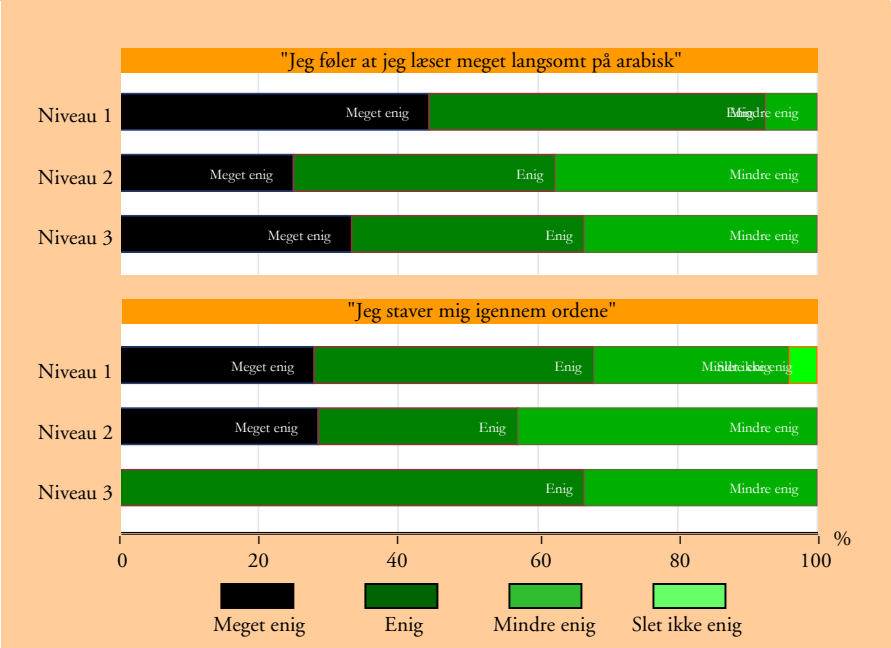
Diagrammerne i figur 6.9 uddyber, hvad der nærmere karakteriserer de vanskeligheder, lørnerne tydeligvis oplever. At læseretningen skulle være forvirrende får ikke den store tilslutning. På niveau 1 erklærer enkelte lørnere sig enige, men den holdning svinder med bedre færdighedsniveau, og på niveau 3 erklærer alle respondenterne sig uenige i udsagnet. Så er der mere opbakning til, at de fremmede sætningskonstruktioner gør læsning svær. Også her svinder enigheden med udsagnet systematisk med højere færdighedsniveau.

Endnu flere bekræfter, at de stadig er usikre på nogle af bogstaverne. Faktisk er der flere på niveau 2 end på niveau 1, der er enige i det udsagn, og kun én ud af de tre respondenter på niveau 3 er helt uenig. Man kan forestille sig, at lørnere på begynderniveau oplever en begejstring over at lære at gennemskue de fremmede grafemer – og først senere oplever de

Figur 6.9 Respondenternes holdning til udsagn vedrørende specifikke vanskeligheder

det som en frustration, at de ikke har samme hurtige og automatiske genkendelse, som de er vant til at have ved læsning af det latinske alfabet. Mest enighed er der i, at det største problem er ordforråd. Ingen erklærer sig direkte uenige, og på alle niveauer er mindst en tredjedel meget enige i dette udsagn.

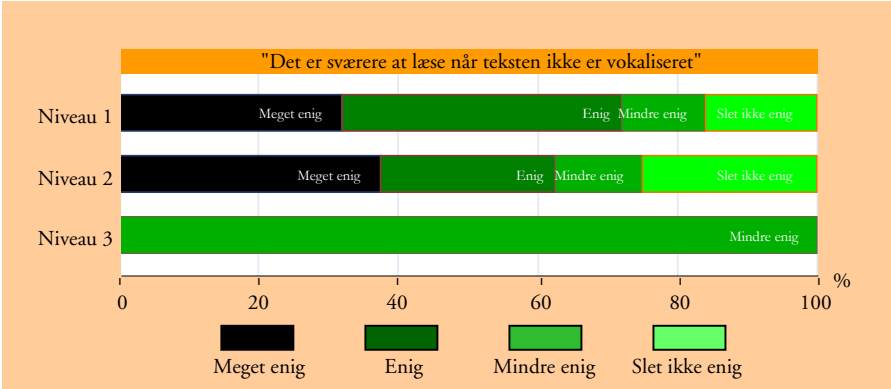
I figur 6.10 og 6.11 ses reaktioner på udsagn vedrørende aspekter, der ikke kan indgå i en sammenligning med de øvrige, men som kan forventes at bidrage til opfattelsen af, at læsning på arabisk er vanskelig. Hvad angår udsagnene om læsehastighed er der bred konsensus niveauerne imellem. Hovedparten af alle respondenterne føler, at de læser meget langsomt på arabisk. På niveau 1 erklærer 93% sig 'meget enig' eller 'enig' i udsagnet. Det mere ekspliciterende udsagn 'Jeg staver mig gennem ordene' får lidt mindre opbakning, men hovedparten af respondenterne erklærer sig dog stadig enige. Kun en enkelt



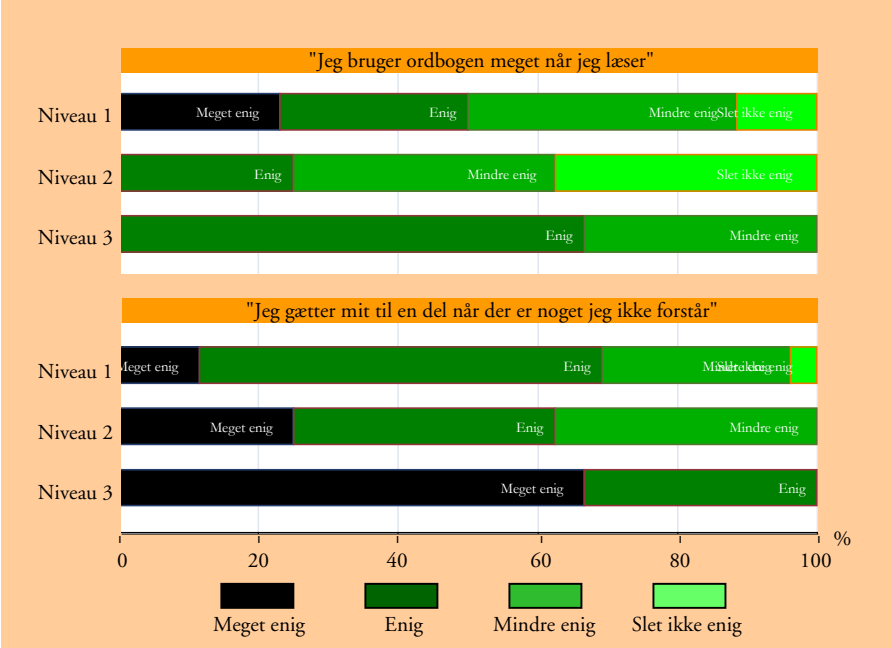
Figur 6.10 Respondenternes holdning til udsagn vedrørende læsehastighed

respondent er slet ikke enig, og mærkeligt nok befinder denne lørner sig på niveau 1. Generelt synes læsehastighed altså at være et selvstændigt 'problem' for de fleste, og for rigtig mange i så høj grad, at de ligefrem føler, at de staver sig gennem ordene.

I figur 6.11 ses på niveau 1 og 2 en nogenlunde jævn fordeling af besvarelser, men hovedparten af lørnerne erklærer sig dog 'meget enig' eller 'enig' i, at det er sværere at læse, når teksten ikke er vokaliseret. Men det synes at være meget individuelt, om og i hvor høj grad manglende vokaler anses for at gøre læsning vanskelig. På niveau 3 er alle 'mindre enig' i udsagnet – men til forskel fra lørnerne på niveau 1 og 2 er de enige med hinanden. Selvom det langt fra er entydigt, kan besvarelserne nok fortolkes derhen, at der er en tendens til, at begynderne i højere grad end lørnere på de højere niveauer opfatter de manglende vokaler som et problem – og det er vel ikke overraskende, da de manglende vokaler stadig er et nyt og fremmed fænomen for begynderne, mens de mere øvede lørnere efterhånden har lært at forholde sig til de særlige karakteristika ved det arabiske skriftsystem.



Figur 6.11 Respondenternes holdning til udsagn om vokalernes betydning



Figur 6.12 Respondenternes holdning til udsagn vedrørende læsestrategier

I figur 6.12 ses reaktionerne på udsagn vedrørende læsestrategier. De afspejler den naturlige udvikling i læsekompetence – det bliver med bedre sprogbeherskelse lettere at gætte ud fra kontekst – og vidner samtidig om, at flertallet benytter sig af de strategier, der er målet for den kommunikative læseundervisning – nemlig, at man ikke føler sig for knyttet til sin ordbog, men bruger de kompetencer, man har, og gætter sig frem til resten så godt man kan.

Respondenternes selvstændige formuleringer af, hvad de finder er det sværeste ved læsning på arabisk, viser, at ud over ordforråd, som af 15 respondenter nævnes som det største problem, er det aspekter af afkodningsprocessen, der opfattes som mest problematisk. I alt ni respondenter nævner forhold, der vedrører den fremmede grafik og fonologi. Som en af respondenterne svarer på spørgsmålet om, hvad det sværeste er ved at læse på arabisk:

> *"Simpelthen bare, at det er så fremmed fra alt hvad man kender i lyd og billede"* (Resp. nr. 1)

De fleste formulerer sig dog mere specifikt. En respondent skriver fx:

> *"Man er helt fortabt, hvis man ikke kan få bogstaverne til at give mening, dvs. man SKAL være sikker på alle bogstaver for at kunne forstå et ords betydning. Selvom jeg synes jeg er godt med i arabisk, laver mange afleveringsopgaver osv., så er det stadig svært for mig at læse – især er det svært at læse ord og tekster, man ikke kender i forvejen [...]"* (Resp. nr. 6)

Nogle udtrykker en meget eksplicit bevidsthed om, at de mangler automatisk ordgenkendelse, fx:

> *"at man i stedet for at opfatte ord som helheder staver sig igennem fra venstre og går i stå pga. konsonanternes ophobning"* (Resp. nr. 11)

> *"at jeg endnu ikke læser ordbilleder – så jeg må stave [...]"* (Resp. nr. 31)

> *"at genkende ord, dvs. ofte udtaler jeg det så forkert, at jeg ikke genkender det"* (Resp. nr. 4)

Syv respondenter nævner de manglende vokaler som det sværeste for dem. En respondent udtrykker sine vanskeligheder sådan:

> *"[... at] vokal- og sukuun [ikke-vokal] -forholdet er uklart. Hvis man kan læse et ord er det ofte fordi man genkender det og derfor kender dets lyd og vokalisering"* (resp. nr. 11)

Fire respondenter fremhæver problemer med læsehastighed, og én respondent påpeger det demotiverende ved, at læsningen går så langsomt. På spørgsmålet om, hvad det sværeste ved læsning på arabisk er, lyder svaret:

> *"[...] manglende rutine. – [jeg] mister tålmodigheden med mig selv"* (Resp. nr. 37)

Seks respondenter nævner grammatik og sætningskonstruktion som deres største problem og to mener, at det sværeste er at læse højt. Endelig nævnes det arabiske sprog generelt som værende afgørende for, at læsning på arabisk er svært for dem, fx:

> *"at sproget er så fremmed fra de europæiske sprog som jeg kender i forvejen, det er ikke muligt at gætte et ord ud fra noget som jeg kender i forvejen"* (Resp. nr. 21)

> *"[...] alle ord er næsten 100% fremmede, på europæiske sprog kan man gætte sig til meget"* (Resp. nr. 29)

Samlet set viser respondenternes reaktioner, at danske lørnere generelt finder læsning på arabisk vanskelig, og at bogstavgenkendelse, vokalisering og læsehastighed, der er nogle af de primære aspekter, læsetestene berører, også af mange af respondenterne betragtes som de primære problemområder i den forbindelse. Også kompetencer, der ligger højere i læsehierarkiet, bidrager i mange lørneres øjne til vanskelighederne – ikke mindst ordforråd, der i de lukkede svarkategorier får mest opbakning som det vanskeligste element ved læsning, og som også i den åbne svarkategori bliver nævnt af flest respondenter. Reaktionerne på udsagn vedrørende læsestrategier vidner om, at mange studerende efterhånden tager de kommunikative læsestrategier til sig og i høj grad benytter en *top-down* styret læsning, hvor de gætter sig til meget af det, de ikke umiddelbart forstår. Det er dog også tydeligt, at de vanskeligheder skriftsystemet repræsenterer faktisk skaber en del frustration hos lørnerne – mange føler tydeligvis, at ordene er svære at gennemskue, at de må stave sig gennem dem, og at det går meget langsomt. Meget tyder altså på, at emnet for dette studium er særdeles relevant for danske arabiskstuderende.

Statistiske kontrolanalyser

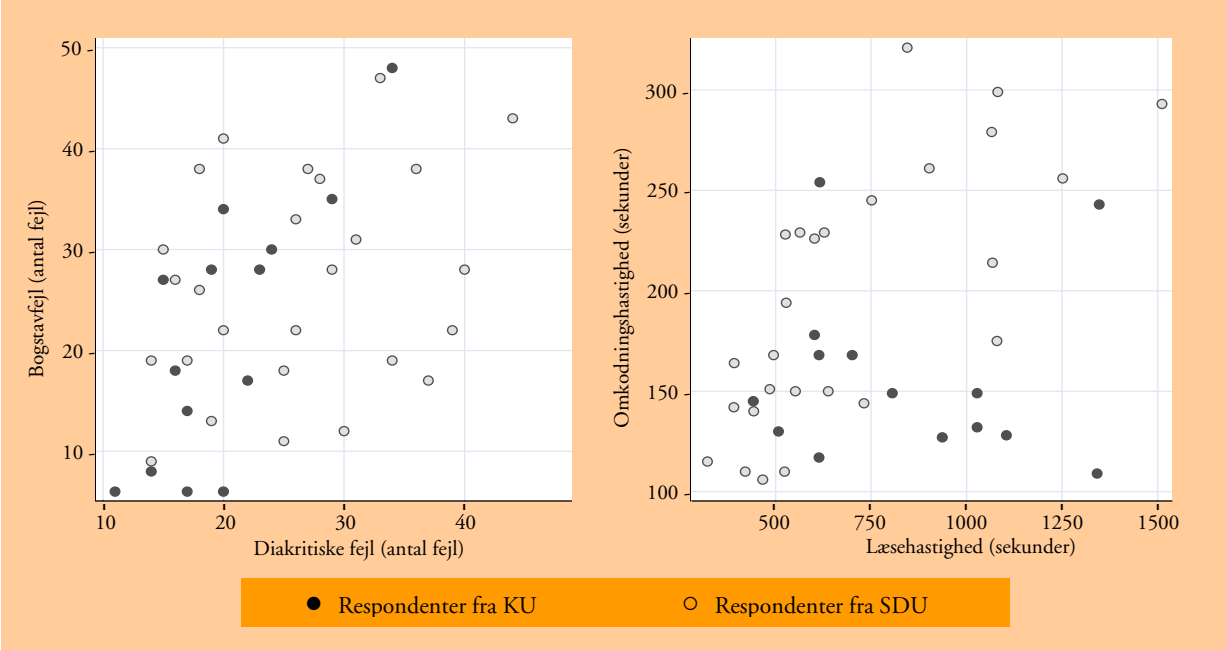
For at kontrollere for visse faktorers indflydelse på resultaterne af de egentlige læsetest, gennemføres tre statistiske kontrolanalyser. Det drejer sig om følgende faktorer:

- Respondenternes køn og alder
- Respondenternes undervisningsmiljø – dvs. henholdsvis KU og SDU
- Korrelation mellem datakategorierne omkodningshastighed og omkodningsfejl i test A og mellem datakategorierne læsehastighed og læseforståelse i test B

Hvad angår respondenternes køn og alder, køres ANOVA af henholdsvis køns og alders generelle indflydelse på resultaterne i alle deltest. Denne viser, at hverken køn eller alder har generel signifikant indflydelse. Beregningerne indgår i den statistiske analyse, der foreligger som Appendix 3.2.

Med hensyn til undervisningsmiljøets indflydelse udvælges respondenterne på niveau 1 som testgruppe, fordi denne respondentgruppe er den største, ligesom respondenterne på de øvrige niveauer er meget uens med hensyn til hvor længe de har studeret, hvilket gør dem mindre egnede til at kontrollere for, om undervisningsstedet isoleret har indflydelse på resultaterne. Der testes på fire forskellige resultatkategorier i datasættet. Disse udvælges ud fra, hvorvidt de skønnes egnede til formålet og repræsentative for testformerne generelt. Det drejer sig om de to fejlkategorier 'bogstavfejl' og 'diakritiske fejl' i test A samt de målte omkodnings- og læsehastigheder i henholdsvis test A og B. Resultaterne fra både test AI, AII, BI og BII indgår og ses arrangeret i spredningsdiagrammer i figur 6.13. Respondenter, der generelt laver få fejl/læser hurtigt, vil således placere sig i nederste venstre hjørne af diagrammerne, mens respondenter, der generelt laver mange fejl/læser langsomt, vil placere sig i øverste højre hjørne. Da resultaterne fra respondenter fra henholdsvis KU og SDU repræsenteres af forskellige markører, vil en systematisk forskel på fejlfrekvens og/eller hastighed dermed kunne ses i diagrammerne. Med hensyn til fejlfrekvens fremgår det af figuren, at flere respondenter ved SDU end ved KU laver mange diakritiske fejl, men der er ikke tale om en systematisk anderledes gruppering af resultaterne. Den lille forskel kan muligvis forklares med, at undervisningen ved SDU er mere målrettet kommunikativ, hvorfor der ikke er fokus på korte vokaler i forbindelse med læsning. Med hensyn til hastighed ses heller ikke nogen markant forskel på respondenterne fra KU og SDU. Det er respondenter fra SDU, der generelt klarer sig både bedst og dårligst, hvilket er statistisk forventeligt, idet der er flest af dem.

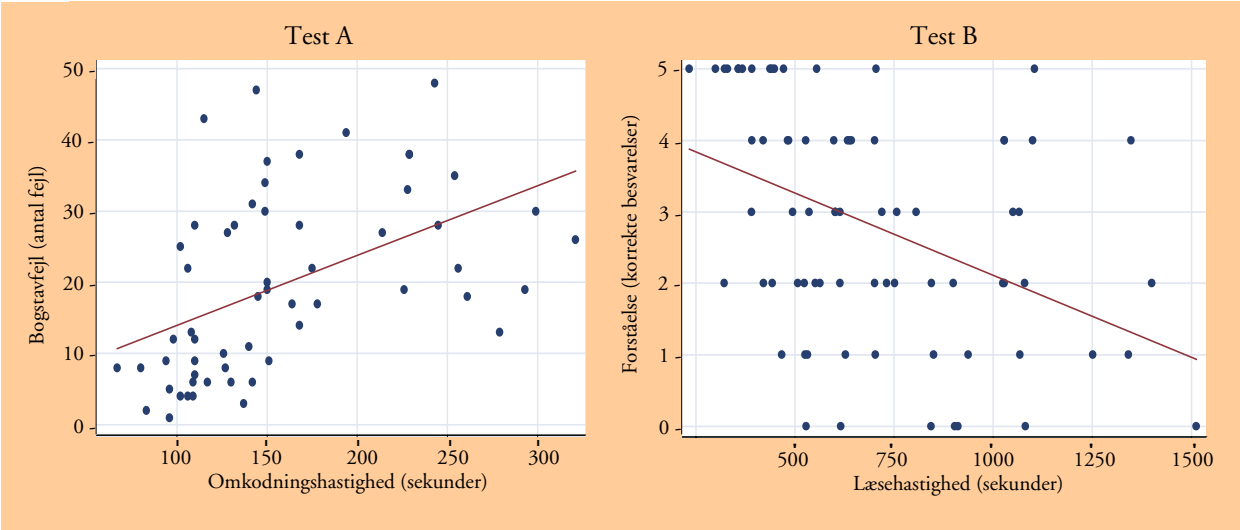
Alt i alt synes der ikke at være tegn på, at studerende fra det ene universitet på dette niveau generelt klarer testene bedre eller dårligere end studerende fra det andet universitet.



Figur 6.13 Udvalgte testresultater for respondenterne på niveau 1 fordelt efter studiested

Med hensyn til eventuel korrelation mellem hastighed og korrekthed i test A og B har testdesignet den indbyggede svaghed, at der kan være et indbyrdes *pay-off* mellem hastighed og korrekthed. I test A kan man således tænke sig, at respondenter, der vælger at læse langsomt, kan opnå at læse mere fejlfrit, mens respondenter, der vælger at koncentrere sig mindre om at læse korrekt, har mulighed for at opnå større hastighed. Det samme gør sig gældende i test B i forholdet mellem hastighed og forståelse. Denne sammenhæng vil i sagens natur altid eksistere, men respondenterne er udtrykkeligt blevet bedt om at læse så hurtigt – og i test A samtidig så korrekt – som muligt. Hvis de imidlertid alligevel har vægtet det ene hensyn meget tungere end det andet, vil deres individuelle dispositioner forvanske resultaterne, da udfaldet i den ene kategori i så fald vil influere på udfaldet i den anden kategori. Til kontrol af om der eksisterer sådanne sammenhænge i data, måles omkodningshastighederne mod fejlkategorierne 'bogstavfejl' og 'diakritiske fejl' fra både AI og AII, og læsehastighederne mod læseforståelsen fra både BI og BII. Alle danske respondenter medtages i analysen. Hvis den nævnte sammenhæng eksisterer, vil det kunne registreres som en negativ korrelation mellem de to kategorier i test A (hvor fejl er en negativ faktor) og som en positiv korrelation mellem de to kategorier i test BI (hvor forståelse er en positiv faktor). Analysen viser, at i test A er korrelationskoefficienterne 0,481 for hastighed mod bogstavfejl og 0,144 for hastighed mod diakritiske fejl. I test B er korrelationskoefficienten –0,421. I figur 6.14 er resultaterne for hastighed-bogstavfejl i test A og hastighed-forståelse i test B visualiseret i korrelationsdiagrammer suppleret med lineære regressionslinjer, der illustrerer de generelle tendenser. I det første diagram ses den positive korrelation mellem omkodningshastighed og bogstavfejl. Som det fremgår, er der altså **ikke** tale om, at de hurtige laver mange fejl og de langsomme få, men tværtimod er der den generelle sammenhæng, at de respondenter der læser hurtigt også er dem, der læser mest korrekt, mens de der læser langsomt også laver mange fejl. I det andet diagram ses en klar negativ korrelation mellem læsehastighed og læseforståelse. Analogt med resultatet fra test A ses det, at de respondenter, der læser hurtigt generelt set også er dem, der har den bedste forståelse.

Der er altså ikke tegn på, at den uønskede sammenhæng eksisterer, og det ser dermed ud til at være helt uproblematisk, at hastighed og korrekthed i de to test måles på samme gennemlæsning.



Figur 6.14 Korrelationsdiagrammer over hastighed og korrekthed i test A og B. Omfatter alle danske respondenters resultater i de anførte kategorier i test AI, AII, BI og BII.

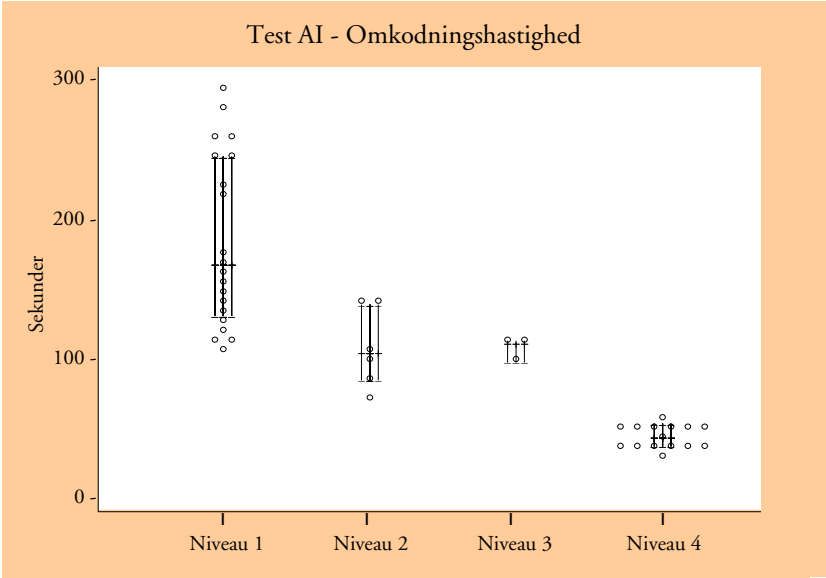
Test A

I figur 6.15-6.17 ses diagrammer, der viser de konkrete observationer af tid for læsning af passagerne med arabiske pseudoord uden fremmede fonemer (AI), med fremmede fonemer (AII) og pseudoord med latinsk skrift (AIII). Som det fremgår af figur 6.15 og 6.16 læser alle lørnere med arabisk som fremmedsprog de arabiske pseudoord langsommere end de indfødte sprogbrugere. Det fremgår også tydeligt, at lørnere på niveau 2 omkoder betragteligt hurtigere end lørnere på niveau 1. Derimod omkoder lørnere på niveau 3 tilsyneladende ikke hurtigere end lørnere på niveau 2. Tiderne hos de tre lørnere på niveau 3 samler sig derimod ret entydigt omkring gennemsnitshastigheden for lørnere på niveau 2. Den statistiske analyse bekræfter, at niveau generelt har en indflydelse på tid for læsning af de arabiske pseudoord ($p<0,001$ i både AI og AII). Også tovejsanalyserne viser signifikante forskelle i omkodningshastighed mellem niveauerne 1, 2/3 og 4 i både AI og AII ($p<0,001$ i alle test). Observationerne indicerer, at mens lørnere fra det første til det andet studieår opnår markant bedre omkodningshastighed, stagnerer fremskridtet herefter, og selv færdiguddannede kandidater omkoder signifikant langsommere end indfødte sprogbrugere.

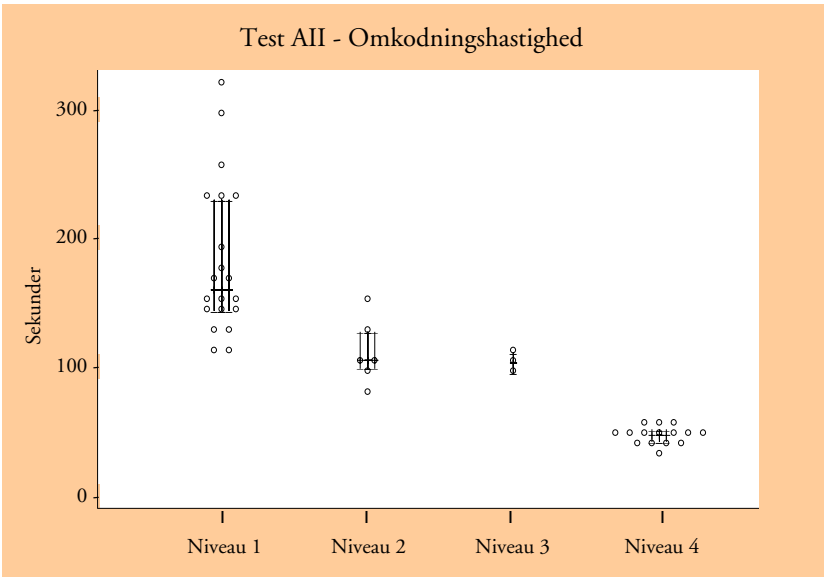
Ved læsning af pseudoord skrevet med latinsk skrift er der som forventet ingen signifikant forskel mellem grupperne af respondenter med dansk som modersmål ($p=0,476$). Derfor måles disse respondenter som en samlet gruppe imod gruppen af indfødte sprogbrugere. Herved ses, at danske respondenter ikke læser de latinskskrevne pseudoord signifikant hurtigere end de marokkanske respondenter ($p=0,076$).

Hvad angår tiderne for læsning på tværs af testtyperne, ses observationerne for de enkelte niveauer i figur 6.18-6.21. Af disse fremgår det, at der for alle danske respondenter er markant forskel på læsning af arabiske og danske pseudoord. Den statistiske analyse bekræfter, at testtype generelt har signifikant indflydelse på omkodningshastigheden ($p<0,001$ for både niveau 1 og 2/3), ligesom tovejsanalyserne viser signifikante forskelle mellem læsning af arabiske og danske pseudoord, uanset om man tester AI og AII separat eller samlet mod AIII ($p<0,001$ i alle tilfælde). De indfødte sprogbrugere læser ikke overraskende de latinske pseudoord langsommere end de arabiske, men forskellen i omkodningshastighed mellem testtyperne er meget mindre end hos de danske respondenter. Den statistiske analyse viser imidlertid, at også for denne gruppe har testtype generel signifikant indflydelse på omkodningshastighed ($p=0,001$), og tovejsanalyserne viser klart signifikante forskelle mellem læsning af arabiske og latinske pseudoord ($p=0,001$ mellem AI og AIII, $p=0,026$ mellem AII og AIII, $p=0,003$ mellem AI+AII og AIII). Med hensyn til læsning af arabiske pseudoord med og uden fremmede fonemer synes der ikke for nogen af niveauerne umiddelbart at være nogen markant forskel mellem resultaterne i de to testtyper. Den statistiske analyse bekræfter dette indtryk for de danske respondenter ($p=0,850$ for niveau 1 og $p=0,569$ for niveau 2/3). Meget overraskende er forskellen hos de indfødte sprogbrugere statistisk signifikant ($p=0,004$).

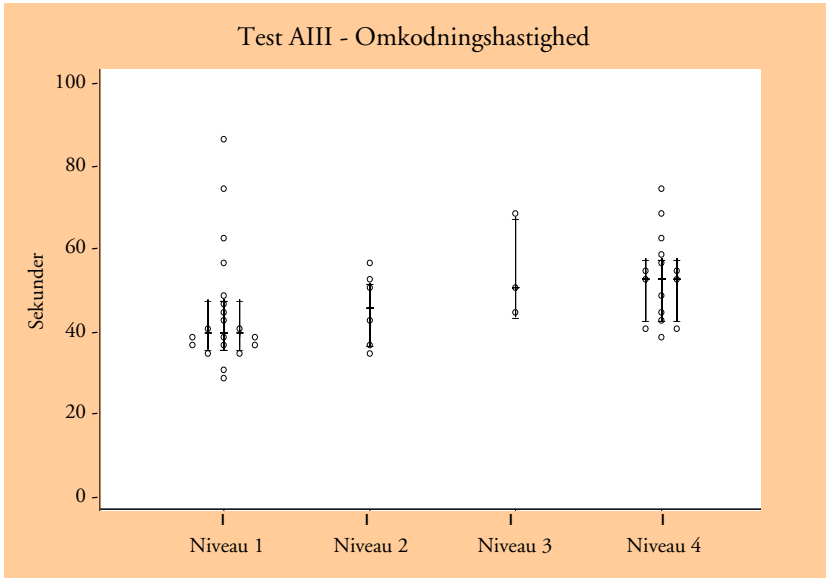
Mens det altså ikke kan påvises, at oplæsningen af pseudoord hos de danske respondenter forsinkes af fonemer, der er særegne for arabisk, influerer disse fonemer tilsyneladende på oplæsningshastigheden hos de indfødte sprogbrugere.



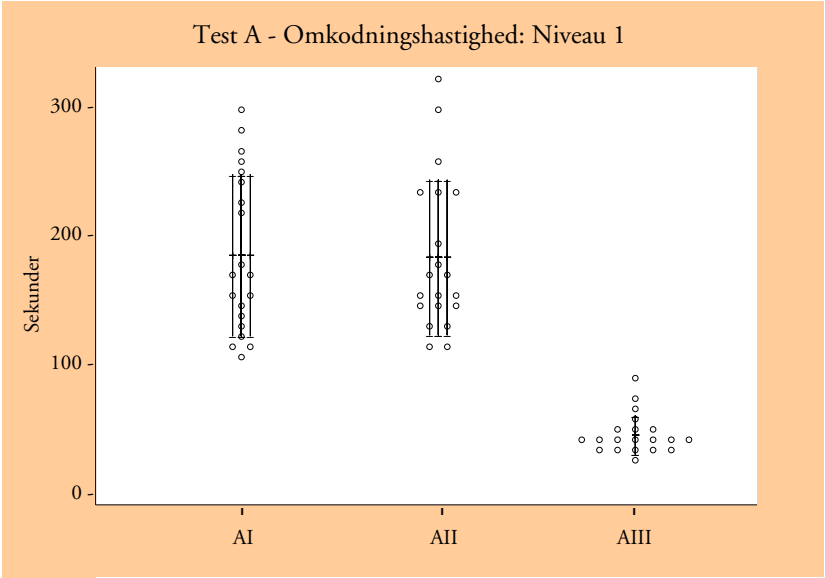
Figur 6.15 Tid for læsning af arabiske pseudoord uden fremmede fonemer.



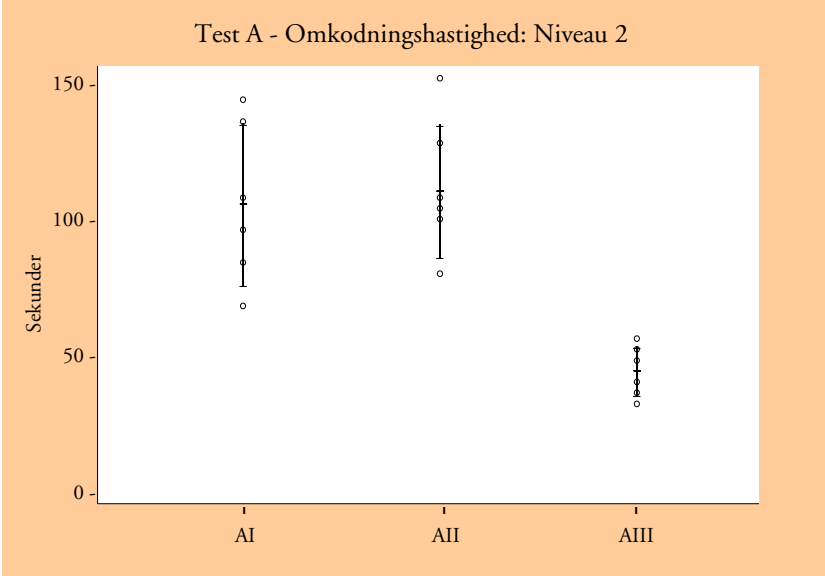
Figur 6.16 Tid for læsning af arabiske pseudoord med fremmede fonemer.



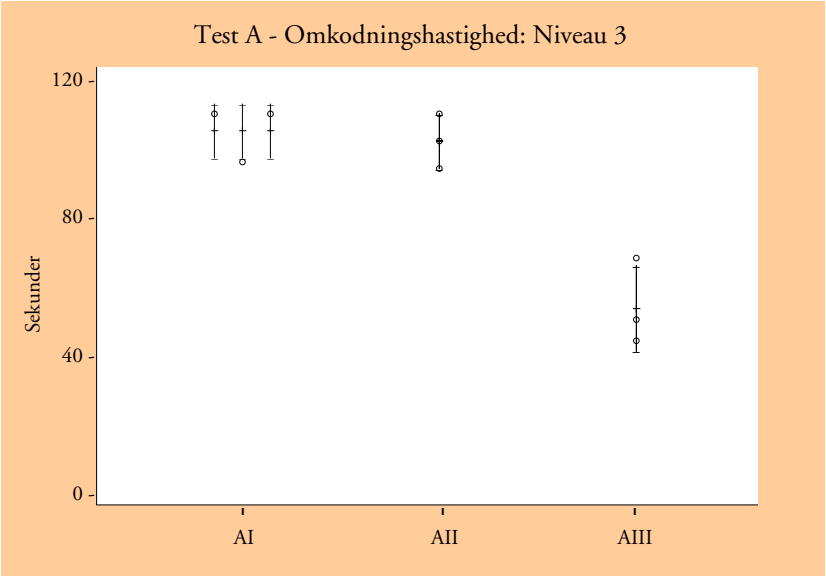
Figur 6.17 Tid for læsning af pseudoord med latinsk skrift



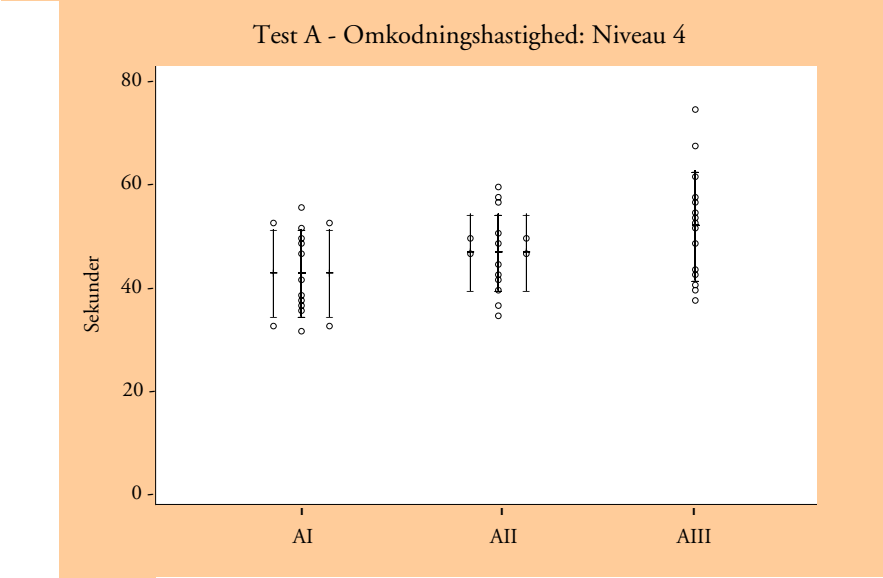
Figur 6.18 Tid for læsning af pseudoord på tværs af testtype, niveau 1



Figur 6.19 Tid for læsning af pseudoord på tværs af testtype, niveau 2



Figur 6.20 Tid for læsning af pseudoord på tværs af testtype, niveau 3



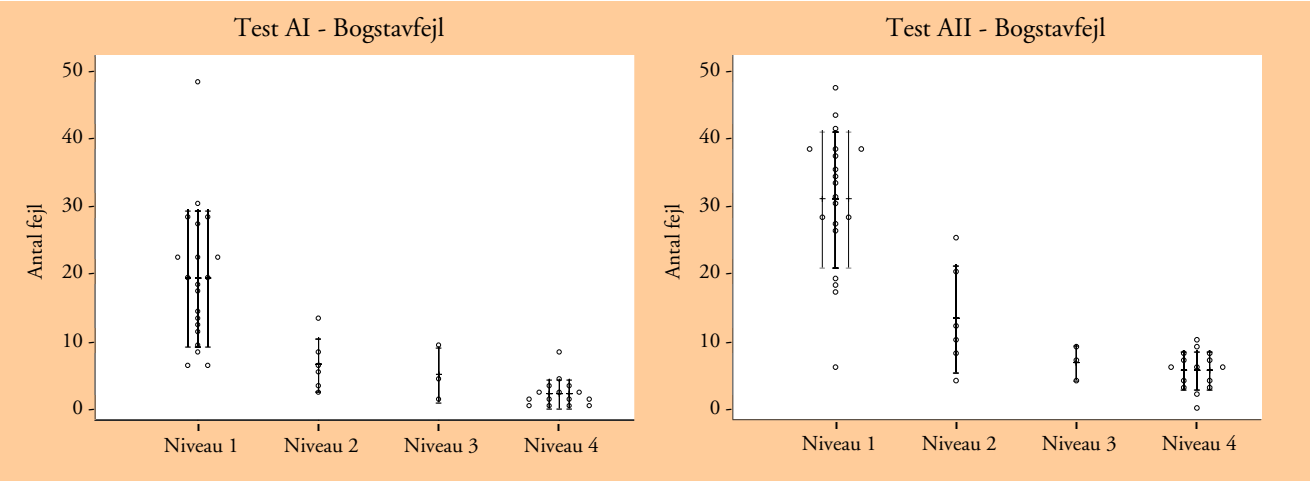
Figur 6.21 Tid for læsning af pseudoord på tværs af testtype, niveau 4

Tilsammen synes disse resultater at bekræfte den teoretiske analyses indikation af, at det arabiske skriftsystem udgør en afgørende barriere for voksne danske lørneres læsehastighed på arabisk. Fonemer der er særegne for arabisk, synes ikke at have nogen indflydelse på omkodningshastigheden. Det har de fremmede grafemer derimod, idet de danske lørneres læsning af arabiske pseudoord sammenlignet med både de indfødte sprogbrugeres og med deres egen læsning af latinske pseudoord er signifikant langsommere.

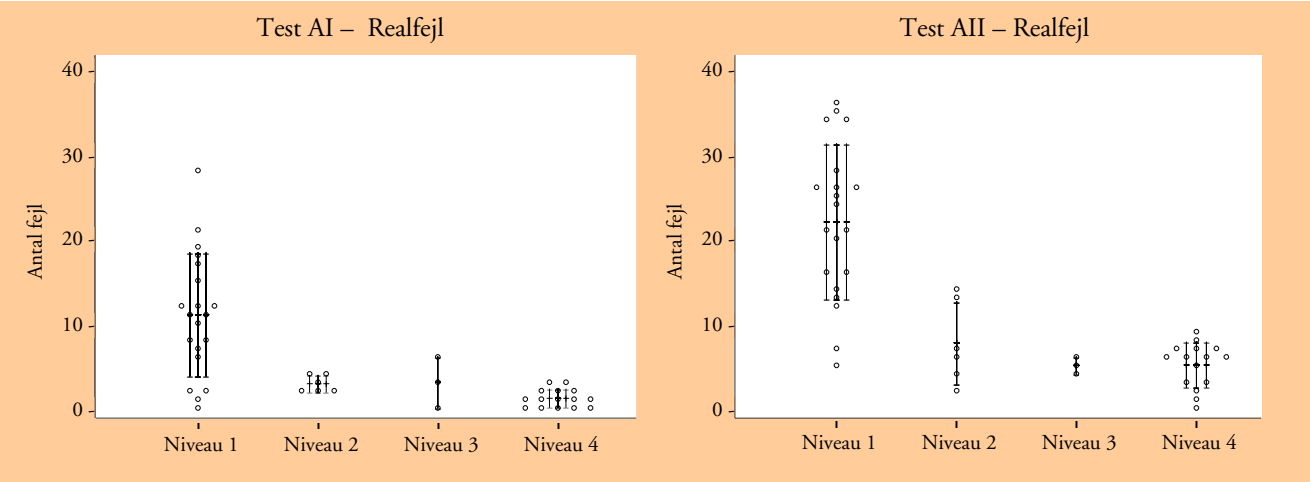
Omvendt læser de indfødte sprogbrugere de latinske pseudoord signifikant langsommere end de læser arabiske pseudoord, men de læser dem ikke signifikant langsommere end danskerne gør. Dette afspejler højst sandsynligt, at de indfødte sprogbrugere har indlært det latinske alfabet i en tidlig alder, hvorved de har haft mange år til at opbygge en hurtig og automatiseret bogstavgenkendelse, der næsten er sammenlignelig med den der ses hos sprogbrugere, hvis modersmål skrives med det latinske alfabet. At de indfødte sprogbrugeres omkodningshastighed i modsætning til de danske respondenters lader sig påvirke af fonemer, der kun for de danske respondenter burde være fremmede, er bemærkelsesværdigt. En forklaring kan være, at de bogstaver, der repræsenterer disse fonemer er grafisk og/eller fonetisk særligt vanskelige at skelne, men da de danske respondenter har langsom bogstavgenkendelse og/eller omkodning af alle bogstaverne, slår effekten kun igennem hos de indfødte sprogbrugere. Denne mulige fortolkning uddybes nærmere i analysen af de konkrete oplæsningsfejl nedenfor.

I figur 6.22-6.24 ses diagrammer, der viser antallet af respondenternes oplæsningsfejl ved læsning af de arabiske pseudoord fordelt efter de i afsnit 6.1 beskrevne fejlkategorier 'bogstavfejl', 'realfejl' og 'diakritiske fejl'.

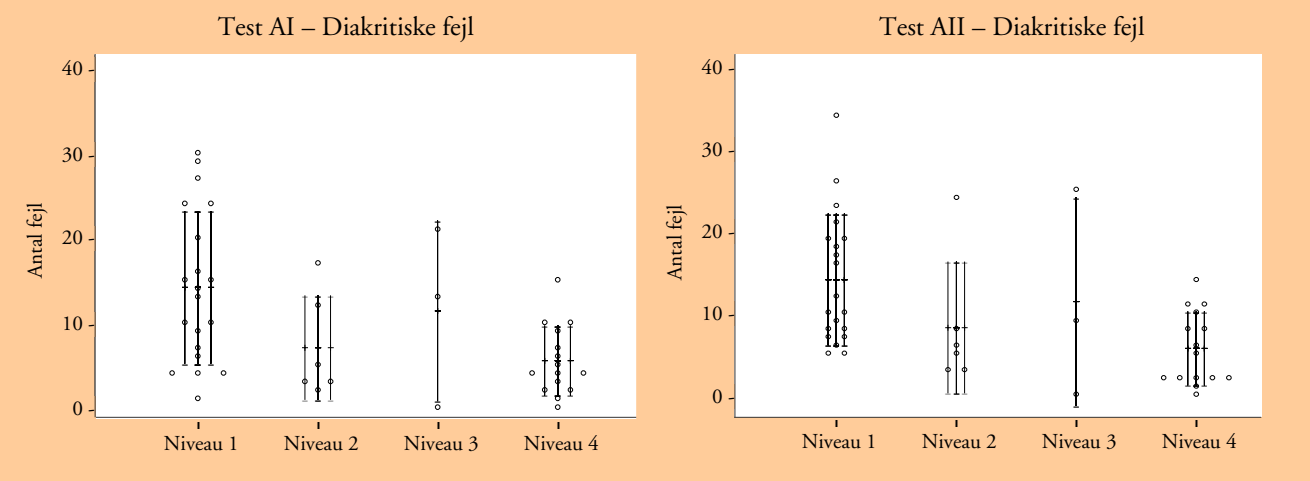
Hvad angår bogstavfejl ses i figur 6.22 i test AI, hvor der ikke er bogstaver, der repræsenterer fremmede fonemer, en udvikling fra niveau 1 til niveau 2, der er sammenlignelig med udviklingen i omkodningshastighed, idet antallet af fejl reduceres betragteligt fra niveau 1 til niveau 2. Tages de tre observationer på niveau 3 i betragtning, ses der efterfølgende



Figur 6.22 Bogstavfejl ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII)



Figur 6.23 Realfejl ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII)



Figur 6.24 Fejl i diakritiske tegn ved læsning af arabiske pseudoord uden fremmede fonemer (AI) og med fremmede fonemer (AII)

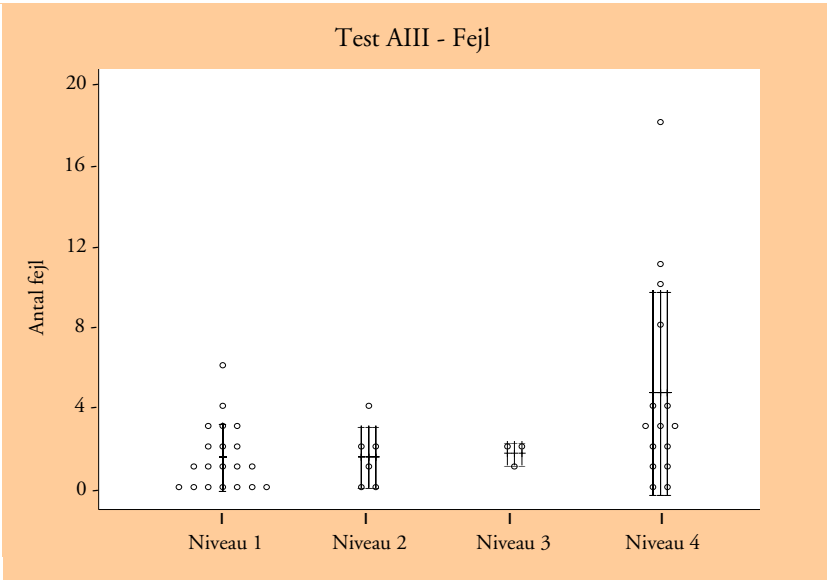
endnu en lille fremgang (fra gennemsnitligt 6,1 til 4,7 fejl). I test AII, hvor der indgår fremmede fonemer, er disse forskelle mellem niveauerne næsten tilsvarende, men udviklingen hos de danske lørnere synes dog en anelse mere gradueret. Forskellen mellem niveau 2 og 3 er større end i test AI, mens der næsten ikke ses nogen forskel mellem niveau 3 og 4. Den statistiske analyse bekræfter, at niveau generelt har indflydelse på antallet af bogstavfejl (p<0,001 i både AI og AII), ligesom tovejsanalyserne viser, at der er signifikante forskelle både mellem niveau 1 og 2 (p<0,001 i både AI og AII) og mellem niveau 2/3 og 4 (p=0,007 i AI og p=0,030 i AII). Desværre kan der som tidligere anført ikke måles på niveau 3 separat. Da realfejl er en reduceret udgave af bogstavfejl, jvf. definitionen i afsnit 6.1, ses ikke overraskende omtrent samme billede som ved disse, hvilket fremgår af figur 6.23.

Med hensyn til antallet af oplæsningsfejl på tværs af testtyperne i figur 6.22 og 6.23, ses en klar forskel mellem resultaterne i AI og AII – især på de lavere niveauer. Men også hos lørnerne på niveau 3 og 4 ses et større antal fejl i AII end i AI. Tilstedeværelsen af fremmede fonemer synes altså at påvirke fejlfrekvensen for bogstavfejl – og i endnu højere grad realfejl – på alle niveauer. Den statistiske analyse bekræfter, at forskellene på antallet af bogstav- og realfejl ved læsning af arabiske pseudoord henholdsvis med fremmede fonemer (AI) og uden fremmede fonemer (AII) er signifikante på alle niveauer (p<0,001 for både bogstavfejl og realfejl på niveau 1, p=0,025 for bogstavfejl og p=0,016 for realfejl på niveau 2/3, og p=0,001 for bogstavfejl og p<0,001 for realfejl på niveau 4). Dermed står det klart, at de fonemer, der for de danske lørnere er fremmede, påvirker antallet af oplæsningsfejl hos alle respondentgrupper – også hos de indfødte sprogbrugere, hvis omkodningshastighed som vist oven for også lader sig influere af disse fonemer.

Hvad angår fejl i de diakritiske tegn ses i figur 6.24, at forskellen niveauerne imellem er mindre udtalt end ved de andre fejltyper. Gennemsnittet for fejl på niveau 3 er i både AI og AII højere end for niveau 2, men dette skyldes først og fremmest et stort antal fejl hos en enkelt studerende på niveau 3, som på grund af den meget lille population på dette niveau påvirker gennemsnittet markant. Endvidere er observationerne på niveau 2, 3 og 4 mere spredte end i de øvrige fejlkategorier. Det tyder på, at fejl i de diakritiske tegn i forhold til bogstavfejlene er mere individuelt end niveaumæssigt betingede. Den statistiske analyse viser dog, at niveau generelt set har indflydelse på antallet af fejl i de diakritiske tegn (p<0,003 i alle test).[8] Tovejsanalyserne indicerer imidlertid, at resultaterne på niveau 1 er udslagsgivende for dette resultat, da der ingen signifikante forskelle er mellem de øvrige niveauer, uanset om man måler niveau 2 og 3 sammen (p=0,268 i AI og p=0,26 i AII) eller udelukker niveau 3 (p=0,557 i AI og p=0,441 i AII). Tovejsanalyserne viser ydermere, at der kun er signifikant forskel mellem niveau 1 og 2 i AI (p=0,028), mens forskellen i AII ikke er signifikant (p=0,111). Ser man på resultaterne på tværs af de to testtyper i figur 6.15, synes de fremmede fonemer ikke at influere systematisk på antallet af oplæsningsfejl af de diakritiske tegn, hvilket er forventeligt og bekræftes af den statistiske analyse (p=0,95 for niveau 1, p=0,615 for niveau 2/3 og p=0,892 for niveau 4).

Omkodningsfejl i test AIII, der indeholder pseudoord skrevet med det latinske alfabet, fremgår af figur 6.25. Her er gennemsnittene for niveau 1, 2 og 3 stort set ens, hvilket er

[8] Da de få observationer på niveau 3 i høj grad bidrager til det visuelle indtryk af, at fejl i de diakritiske tegn er mere individuelt end niveaumæssigt betingede, gennemføres de statistiske test både med niveau 2 og 3 samlet mod de øvrige niveauer og med test, hvor niveau 3 helt udelades. 'Alle test' relaterer således til begge disse analyser for både AI og AII.

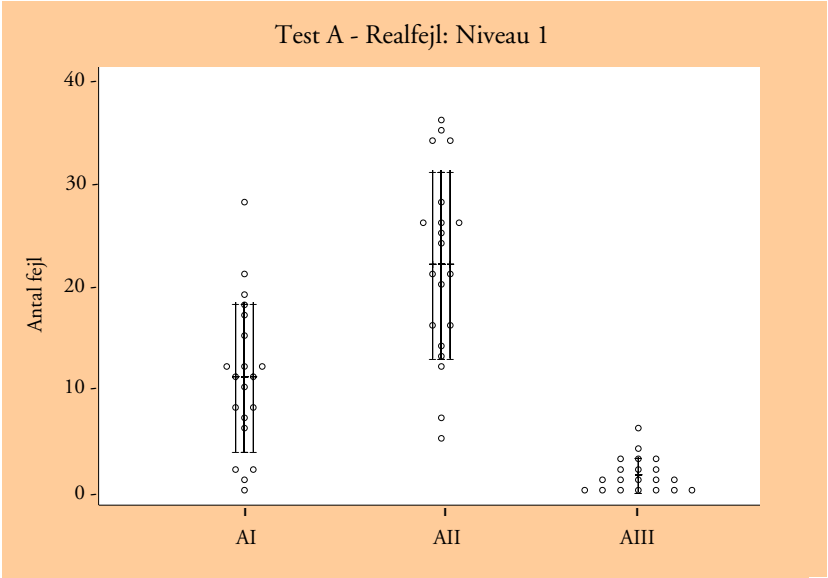


Figur 6.25 Omkodningsfejl ved læsning af pseudoord med latinsk skrift

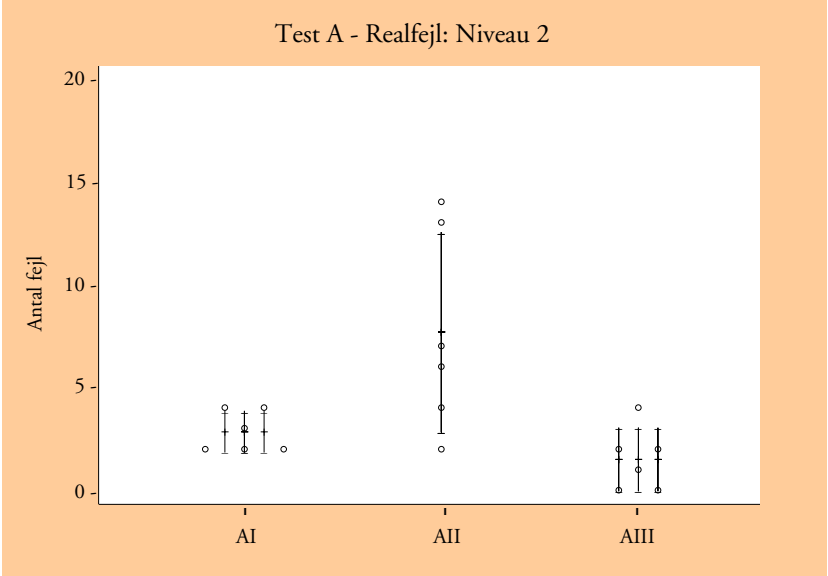
forventeligt, da alle respondenterne i disse grupper har dansk som modersmål. Den statistiske analyse bekræfter, at der ingen forskel er mellem disse (p=1), hvorfor de i den statistiske sammenligning med niveau 4 behandles som en enkelt gruppe. På niveau 4 er observationerne ret spredte, så selvom størstedelen af dem ligger i samme område som for de danske respondenter, læser kun ganske få pseudoordene helt fejlfrit, og enkelte respondenter laver mange fejl, hvilket trækker gennemsnittet op. Den statistiske analyse viser da også, at forskellen mellem de danske og de arabiske respondenters fejlfrekvens er signifikant (p=0,021).

Oplæsningsfejl på tværs af skriftsystemerne – dvs. realfejl i test AI og AII og fejl i test AIII ses i figur 6.26-6.29. Måles der på skriftsystemerne generelt – dvs. hvis resultaterne i AI og AII måles samlet imod resultaterne i test AIII – viser den statistiske analyse, at forskellen i fejlfrekvens er signifikant for både niveau 1 (p<0,001) og niveau 2/3 (p<0,001), men ikke for niveau 4 (p=0,244). Alle danske lørnere laver altså flere omkodningsfejl ved læsning af arabiske pseudoord end ved læsning af pseudoord skrevet med det latinske alfabet, mens de arabiske respondenter ikke kan påvises at lave flere omkodningsfejl ved læsning af latinske pseudoord end ved læsning af arabiske pseudoord.

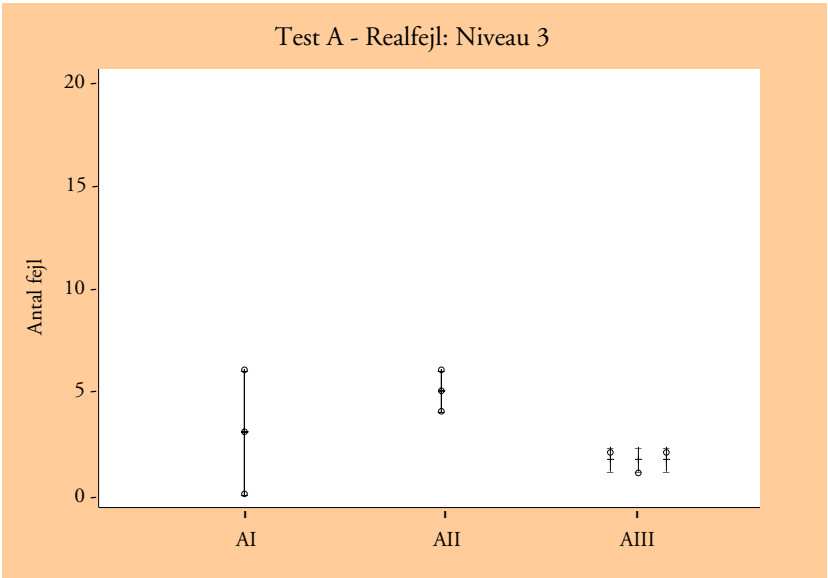
I tråd med de ovenstående resultater vedrørende bogstavfejl, og som det fremgår af den grafiske fremstilling i figur 6.26-6.29, er der hos de danske lørnere imidlertid stor forskel, hvis man separat sammenligner latinsk skrift (AIII) med henholdsvis arabisk skrift uden fremmede fonemer (AI) og arabisk skrift med fremmede fonemer (AII): På niveau 1 ses en klar forskel mellem læsning af latinsk skrift og læsning af begge de arabiske testtyper. Denne forskel reduceres på niveau 2 og 3 – mellem AI og AIII sker reduktionen fra niveau 1 til 2, mens den mellem AII og AIII sker mere gradueret – både fra niveau 1 til 2 og fra niveau 2 til 3. Hvis fejlfrekvensen ved læsning af arabisk skrift uden fremmede fonemer statistisk måles alene mod fejlfrekvensen ved læsning af latinsk skrift, er forskellen signifikant på niveau 1 (p<0,001) og niveau 4 (p=0,016) men ikke signifikant på niveau 2/3 (p=0,34). I modsætning til, hvad der er tilfældet for resultaterne af omkodningshastighed, opnår de danske lørnere altså, hvad angår korrekt fonologisk omkodning, tilsyneladende færdigheder, der ikke påviseligt afviger fra dem de har ved læsning af det latinske alfabet, når blot de ikke



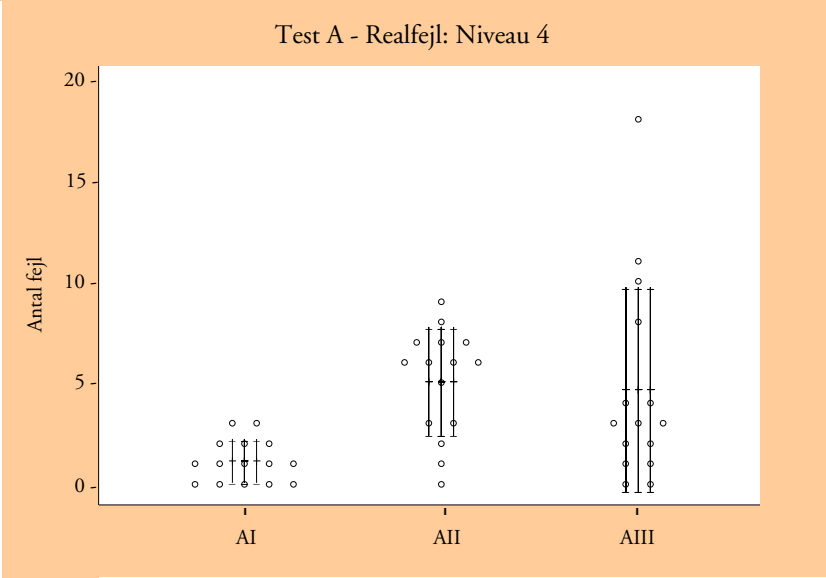
Figur 6.26 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 1



Figur 6.27 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 2



Figur 6.28 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 3



Figur 6.29 Realfejl på tværs af testtype ved læsning af pseudoord, niveau 4

konfronteres med de meget fremmedartede fonemer. Dermed synes det sandsynliggjort, at den højere fejlfrekvens hos de danske lørnere ved læsning af arabisk skrift i forhold til latinsk skrift, hos begyndere skyldes både grafiske og fonologiske barrierer, mens den senere i indlæringsforløbet i højere grad skyldes de fremmede fonemer end den fremmede grafik.

Som det ses i figur 6.29, er der også på niveau 4 en klar forskel på observationerne mellem både AI – AIII og AII – AIII, blot selvfølgelig med flest fejl i AIII. Gennemsnitsværdierne for AII og AIII er dog nogenlunde ens. Ved læsning med bogstaver, der for de danske lørnere repræsenterer fremmede fonemer, laver de indfødte sprogbrugere altså gennemsnitligt lige så mange fejl som de gør ved læsning af det latinske alfabet. Der er dog store individuelle forskelle i, hvor mange omkodningsfejl de indfødte sprogbrugere laver i AIII. Men det er bemærkelsesværdigt, at hovedparten af de marokkanske respondenter læser de latinske pseudoord mindst lige så fejlfrit som de læser arabiske pseudoord med de særegne arabiske fonemer, som jo i forhold til de latinske pseudoord burde være dem mindre fremmede. Den statistiske analyse bekræfter, at testtype generelt har indflydelse på resultatet ($p<0{,}001$), og at forskellen er signifikant mellem AI og AIII ($p=0{,}016$), men ikke mellem AII og AIII ($p=0{,}782$). Som anført ovenfor er forskellen mellem AI og AII – de to arabiske testtyper – desuden signifikant ($p<0{,}001$).

Sammenfattende ser det ud til, at danske lørneres læsning på arabisk er stærkt påvirket af det fremmede skriftsystem uafhængigt af sproglige aspekter som ordforråd og grammatik. Hos begyndere er skriftsystemet en afgørende barriere, der resulterer i, at bogstavgenkendelsen både er langsommelig og giver anledning til mange omkodningsfejl. Da forudsætningerne for automatisk ordgenkendelse på dette niveau ikke er til stede, må dette nødvendigvis influere på lørnernes læsefærdighed generelt – herunder læseforståelse. Lørnernes færdigheder udvikles dog meget i de første studieår. Gennemsnitshastigheden for læsning af pseudoordene i denne test næsten fordobles fra 1. til 2. studieår, og omfanget af omkodningsfejl reduceres markant. Antallet af bogstavfejl mere end halveres og også antallet af fejl i de diakritiske tegn falder, men udviklingen er for disse fejl dog mere individuelt betingede. Hvad angår omkodningshastighed stagnerer fremskridtet efter 2.

studieår, og selv færdiguddannede kandidater omkoder markant langsommere end indfødte sprogbrugere, som i gennemsnit læser dobbelt så hurtigt. Hvad angår omkodningsfejl fortsætter den gunstige udvikling gennem hele tilegnelsesforløbet. Denne fortsatte udvikling sker tilsyneladende først og fremmest i kraft af en stadig større fortrolighed med de fonemer, lørnerne ikke kender fra dansk eller de nære fremmedsprog. Tilegnelsesforløbet for de kendte fonemer går hurtigere, og noget tyder på, at kun tilstedeværelsen af de fremmede fonemer forhindrer de danske lørnere i at opnå en lige så korrekt omkodning som de indfødte sprogbrugere. Den grafiske dimension er altså – naturligvis – en større barriere i begyndelsen end senere i forløbet, hvor den først og fremmest influerer på omkodningshastighed, men tilsyneladende ikke på omkodningsfejl. Omkodningsfejl skyldes efter 1. studieår derimod i højere grad de fremmede fonemer end den fremmede grafik.

At omkodningshastighed på langt sigt i højere grad end korrekt omkodning begrænses af det fremmede skriftsystem er ikke overraskende. Analogt med hvad der er tilfældet ved læseindlæring på L1, kommer korrekt bogstavgenkendelse normalt først i tilegnelsesforløbet. Korrekt bogstavgenkendelse er en forholdsvis enkel og afgrænset opgave, som mange børn mestrer allerede før skolestart. Hurtig omkodning af strenge på flere bogstaver kræver imidlertid en meget hurtig og automatiseret genkendelse, og er dermed en forholdsvis kompleks kognitiv kompetence, som det kræver meget øvelse at etablere. Det forklarer imidlertid ikke umiddelbart, hvorfor hastighedsforøgelsen hos lørnerne i denne undersøgelse stagnerer efter 2. studieår. Hvis man sammenligner med undersøgelsens indfødte sprogbrugere, ses det, at de generelt set omkoder arabisk og latinsk skrift omtrent lige godt – både hvad angår fejl og hastighed. Dette kan forklares med, at de i en tidlig alder har lært at læse på sprog, der anvender det latinske alfabet. Normalt stifter marokkanske elever bekendtskab med det latinske alfabet i 3. klasse, hvor de påbegynder undervisning i fransk som 1. fremmedsprog. Nogle respondenter, der har fulgt undervisning i privatskole, har haft fransk allerede fra skolestart. Og muligvis er det ikke kun omfanget af eksponering men også lørnerens alder ved tilegnelse, der influerer på resultatet. At den hurtige udvikling i omkodningshastighed hos de danske lørnere i denne undersøgelse stagnerer efter 2. studieår **kunne** indicere, at der for bogstavgenkendelse som for fonemgenkendelse findes noget der ligner en øvre tærskel for, hvor automatiseret en genkendelse det er muligt at opnå, når fremmedsproget indlæres i voksenalderen. Det er velkendt, at det er særdeles sjældent, at lørnere opnår fuldstændig fonemgenkendelse og -produktion på et fremmedsprog, der indlæres efter puberteten, og i betragtning af de kognitive paralleller, der ses mellem auditiv og visuel sprogbehandling (fx Massaro 1975,1979), er det ikke utænkeligt, at der med hensyn til bogstavgenkendelse findes lignende kognitive restriktioner for indlæring. Dokumentation for denne hypotese om, at alder ved tilegnelse af et skriftsystem er et afgørende parameter for, hvor høj en læsehastighed man opnår, vil dog kræve en undersøgelse med en større population af lørnere, der først som voksne har stiftet bekendtskab med det arabiske skriftsystem, men som har været massivt eksponeret for arabisk skrift over en længere årrække – fx i forbindelse med bosættelse i et arabisk land. En anden forklaring kunne være, at lørnerne efter 2. studieår ikke læser ret meget på arabisk. I betragtning af, at studerende på Cand. negot.-uddannelsen efter 2. studieår læser et år i et arabisk land og har flere fag, hvor al litteratur er på arabisk, ligesom studerende ved KU efter 2. studieår læser arabisk litteratur, synes det umiddelbart at være en usandsynlig forklaring. Ud fra pædagogiske kriterier kan der imidlertid godt argumenteres for, at netop det faglige/litterære tekstfokus efter

bacheloruddannelsen svækker læsetræningsaspektet. Denne mulighed vil blive diskuteret nærmere i kapitel 7.

De indfødte sprogbrugere udviser meget overraskende usikkerhed over for pseudoord, der indeholder fonemer, der for de danske lørnere er fremmede. Da effekten kan måles på både hastighed og fejlfrekvens er det sandsynligt, at resultatet ikke skyldes en svaghed i testdesign, men må forklares med særlige grafiske eller fonetiske karakteristika ved de bogstaver, der repræsenterer disse fonemer. Den kvalitative analyse af respondenternes konkrete omkodningsfejl, som følger efter gennemgangen af resultaterne for test B og C, vil belyse dette nærmere.

En opsummerende oversigt over de danske respondenters resultater i den kvantitative analyse af test A ser således ud:

- De læser på alle niveauer arabisk skrift meget langsommere end indfødte sprogbrugere.
- De læser på alle niveauer arabisk skrift meget langsommere end de læser latinsk skrift.
- De forbedrer deres omkodningshastighed markant i løbet af de første to studieår, men derefter går udviklingen i stå.
- De læser tilsyneladende arabiske pseudoord med og uden fremmede fonemer lige hurtigt.
- De laver på alle niveauer flere fejl end indfødte sprogbrugere.
- De læser på alle niveauer latinsk skrift mere fejlfrit end de læser arabisk skrift.
- De formindsker antallet af bogstavfejl markant i løbet af de to første studieår, og fremgangen fortsætter.
- Udviklingen går langsommere for læsning med end for læsning uden fremmede fonemer.
- De laver flere bogstavfejl ved læsning af pseudoord med end ved læsning uden fremmede fonemer.
- Oplæsningsfejl skyldes i begyndelsen både den fremmede grafik og de fremmede fonemer. Senere synes de først og fremmest at skyldes de fremmede fonemer og i mindre grad den fremmede grafik.
- Fejl i de diakritiske tegn er efter begynderstadiet mere individuelt end niveaumæssigt betingede.

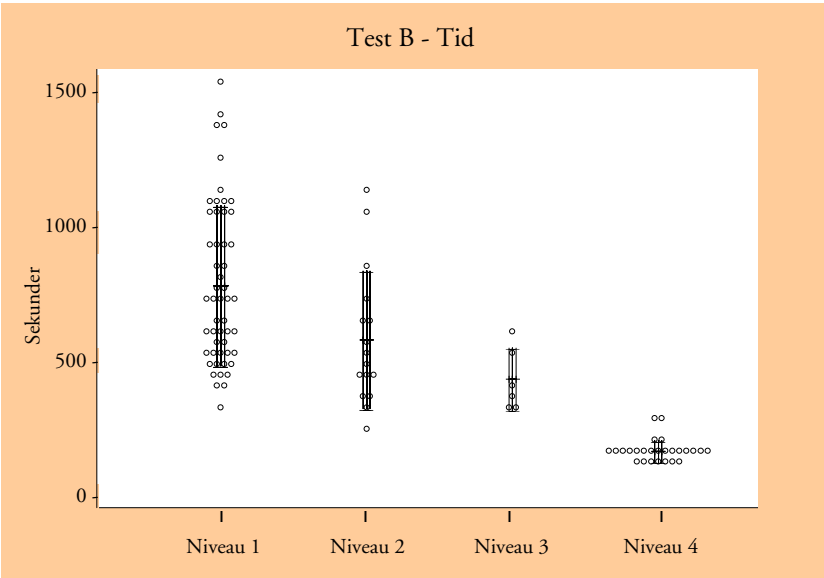
For de indfødte sprogbrugere kan resultaterne opsummeres som følger:

- De læser ikke latinsk skrift lige så fejlfrit som de danske lørnere, men tilsyneladende lige så hurtigt.
- De læser arabisk skrift endnu hurtigere end de læser latinsk skrift – men ikke med færre fejl.
- Til forskel fra hvad der er tilfældet for de danske respondenter, påvirkes både deres hastighed og fejlfrekvens af de fonemer, der i princippet kun burde være fremmede for de danske lørnere.

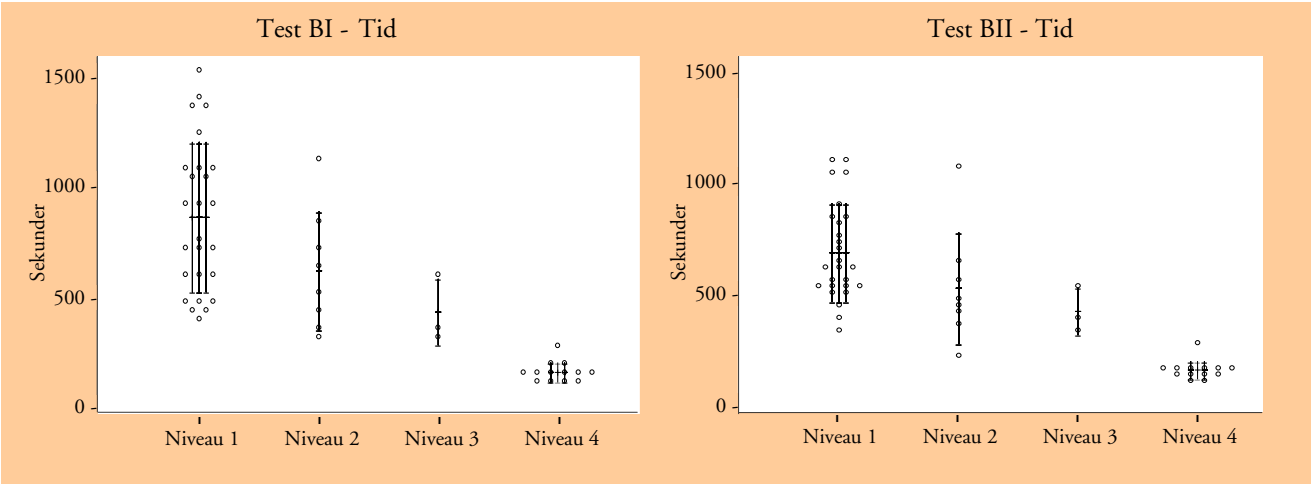
Test B

I figur 6.30 ses alle observationer af hastighed for læsning af de arabiske historier både med og uden vokalisering. Som beskrevet i afsnit 6.1 læser lørnere på niveau 1 andre historier end lørnere på niveau 2 og 3, da det er nødvendigt at tilgodese de mangelfulde sprogfærdigheder hos begynderne. De indfødte sprogbrugere læser alle fire historier, men kun deres resultater for læsning af de tekster, lørnerne på niveau 2 og 3 måles på, fremgår af denne grafiske sammenligning. Af denne samlede fremstilling ses en jævn reduktion i tidsforbruget over de fire niveauer, hvilket tyder på, at teksterne der er produceret til niveau 1, på tilfredsstillende vis tager højde for disse lørneres sproglige kompetencer. Selvom de danske lørnere – i takt med at de får mere undervisning – bliver hurtigere læsere på arabisk, synes de færdiguddannede kandidater stadig at læse betydeligt langsommere end de indfødte sprogbrugere, men den jævne udvikling af læsehastigheden set i forhold til udviklingen i hastigheden for omkodning af pseudoord kunne tyde på, at lørnere på højt niveau fortsætter med at udvikle læsekompetencer, der ligger i den øvre del af læsehierarkiet, selvom de bremses af den fremmede grafik. Om forskellen på niveau 2 og 3 primært skyldes den fremmede grafik eller manglende generelle sproglige færdigheder, kan denne undersøgelse ikke afklare.

Den statistiske analyse bekræfter, at niveau generelt har indflydelse på læsehastigheden ($p<0{,}001$), og tovejsanalyserne viser signifikante forskelle både mellem niveau 1 og niveau 2/3 ($p=0{,}005$) og mellem niveau 2/3 og niveau 4 ($p<0{,}001$) Også hvis man måler mellem niveau 1 og 2 alene, er forskellen signifikant ($p=0{,}046$). Det skal dog bemærkes, at den forskel der her ses og måles mellem niveau 1 og 2, ikke forholder sig eksakt til forskellen i læsehastighed mellem respondenterne på de to niveauer, idet de jo læser forskellige tekster. Men de indfødte sprogbrugeres kontrollæsning af alle de forskellige testmaterialer i test B viser – som det fremgår nedenfor i figur 6.36 – at de indfødte sprogbrugere læser teksterne, der er tilpasset niveau 1, hurtigere end de læser teksterne tilpasset niveau 2. Dermed synes det sikkert at antage, at læsehastigheden hos respondenterne på niveau 2 faktisk er større end hos begynderne, hvilket naturligvis også er forventeligt.



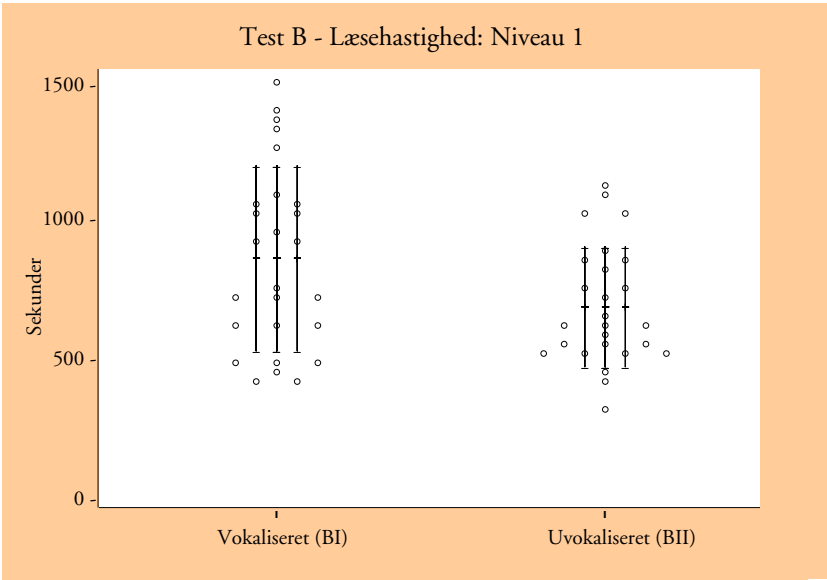
Figur 6.30 Tid for læsning af arabiske tekststykker både med og uden vokalisering



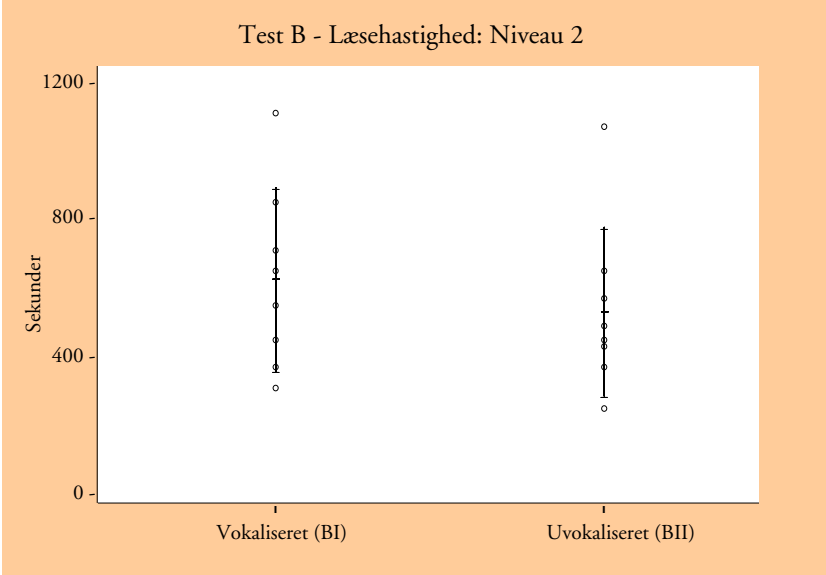
Figur 6.31 Tid for læsning af arabiske tekststykker med vokalisering (BI) og uden vokalisering (BII)

I figur 6.31 ses de samme resultater som i figur 6.30 fordelt på de to testtyper. Som det ses er progressionen niveauerne imellem nogenlunde den samme ved læsning med og uden vokalisering, og opsplitningen af observationerne ændrer da heller ikke afgørende på resultaterne i den statistiske analyse, som også i dette tilfælde bekræfter, at niveau generelt influerer på læsehastigheden ($p<0{,}001$ i både BI og BII), ligesom tovejsanalyserne viser signifikante forskelle både mellem niveau 1 og niveau 2/3 ($p=0{,}005$ i BI og $p=0{,}019$ i BII) og mellem niveau 2/3 og 4 ($p<0{,}001$ i både BI og BII).

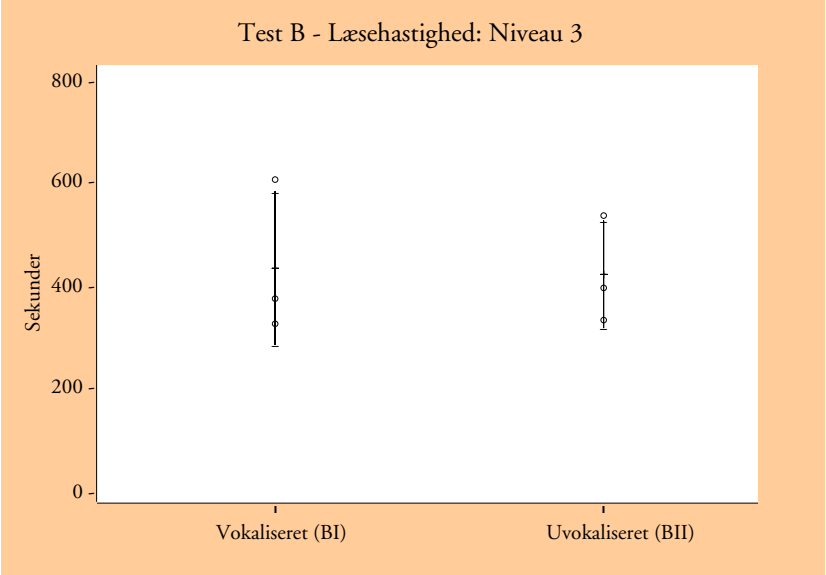
Hvad angår tid for læsning på tværs af de to testtyper fremgår det af figur 6.31 – og endnu tydeligere af figur 6.32-6.35, som bedre illustrerer forskellene mellem testtyperne på hvert enkelt niveau – at forskellene på tidsforbrug ved læsning af de to testtyper er størst på niveau 1, mindre på niveau 2, meget lille på niveau 3. På niveau 4 er der stort set ingen forskel. Gennemsnitshastigheden for de to testtyper adskiller sig på niveau 4 således kun



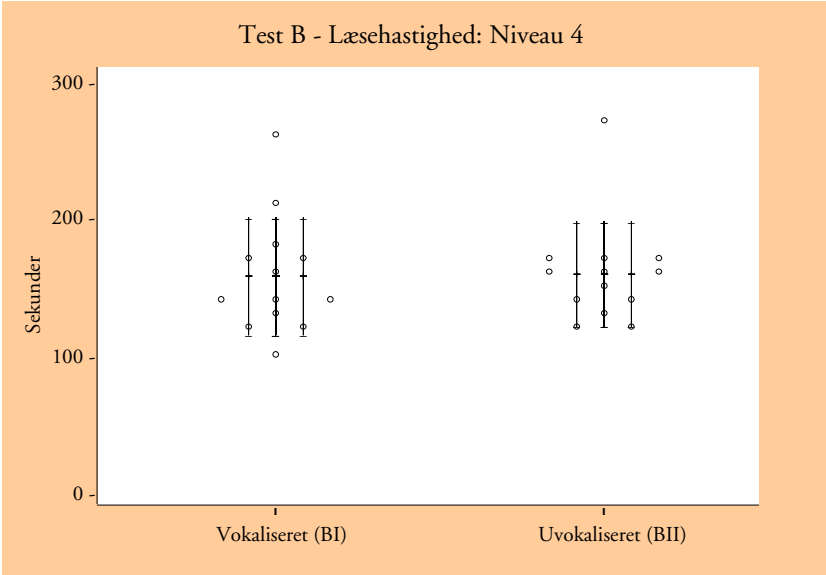
Figur 6.32 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 1



Figur 6.33 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 2



Figur 6.34 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 3

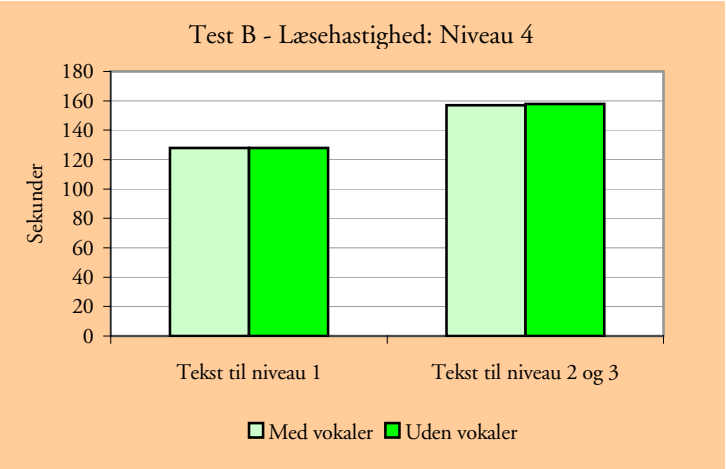


Figur 6.35 Tid for læsning af arabisk tekststykke med og uden vokaler, niveau 4

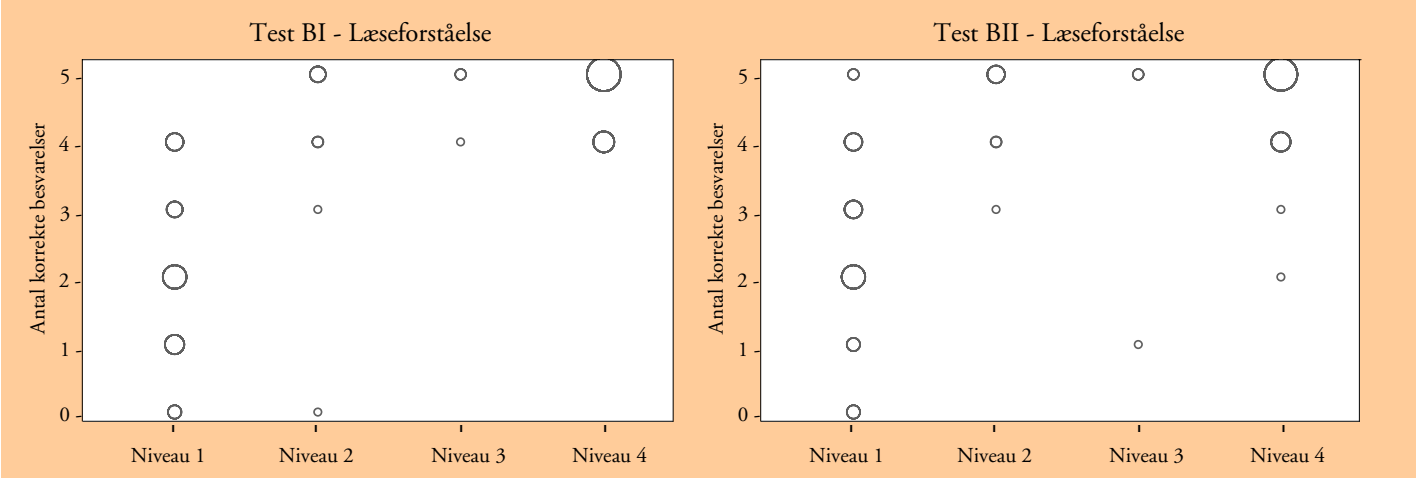
med 2 sekunder. De grafiske fremstillinger viser altså et mønster, hvor forskellen i læsehastighed med og uden vokaler reduceres, jo bedre sprogfærdighed, respondenterne har. Men forskellen er modsat den forventede, idet den uvokaliserede tekst læses hurtigere end den vokaliserede tekst. Den statistiske analyse bekræfter, at forskellen mellem testtyperne er signifikant på både niveau 1 (p=0,004) og niveau 2 (p=0,038). Også hvis niveau 2 og 3 måles sammen, er forskellen signifikant (p=0,039). På niveau 4 er forskellen ikke signifikant (p=0,781).

En plausibel forklaring på disse resultater er, at selvom vokalerne bidrager med fonologisk information og dermed gør skriftsystemet mere gennemskueligt, tillader at skelne mellem homografer og gør læsningen mindre kontekstafhængig, kan denne hjælp ikke opveje den ekstra kognitive byrde, som samtidig tilføjes i kraft af den yderligere grafiske information, vokalerne **også** repræsenterer. Denne forklaring bestyrkes af, at den forsinkelse, som den ekstra grafik afstedkommer hos begynderne, reduceres i takt med respondenternes større sprogfærdighed.

I figur 6.36 ses de indfødte sprogbrugeres gennemsnitlige læsehastighed ved læsning af alle de fire tekster, der er produceret til denne undersøgelse, hvoraf de to er særligt tilpasset respondenterne på niveau 1, mens de andre er bestemt for respondenterne på niveau 2 og 3. Som det fremgår læser de indfødte sprogbrugere den letteste tekst hurtigst og gennemsnitsværdierne for den uvokaliserede og den vokaliserede version er fuldstændig ens. Den lidt sværere tekst læser de lidt langsommere, og den uvokaliserede udgave læses i gennemsnit 1 sekund langsommere end den vokaliserede. Som beskrevet ovenfor er denne forskel ikke signifikant. Som det fremgår af kapitel 5, viser andre undersøgelser af indfødte sprogbrugeres læsning med og uden vokaler, at de normalt læser vokaliseret tekst hurtigst. At det samme ikke kan dokumenteres i denne undersøgelse, kan højst sandsynligt begrundes med, at teksterne er produceret til fremmedsprogslørnere, hvorfor de ordforrådsmæssigt er forsimplede i forhold til autentiske tekster. For de indfødte sprogbrugere er teksterne derfor højst sandsynligt så lette at læse, at den positive effekt, vokalerne normalt repræsenterer, ikke slår igennem. Så selvom hastighedsforskellen mellem testtyperne ved læsning af de lidt sværere tekster i dette tilfælde er meget lille og ikke statistisk signifikant, kan man udmærket forestille sig, at jo mere kognitivt krævende tekster, indfødte sprogbrugere udsættes for, des mere vil vokalisering udgøre en hjælp i afkodningen og dermed øge læsehastigheden.



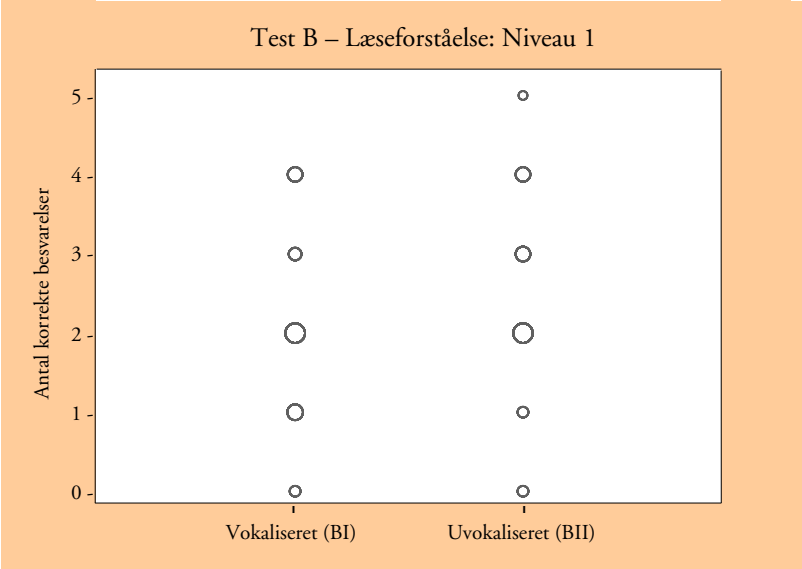
Figur 6.36 Kontrolgruppens gennemsnitlige læsehastighed ved læsning af de to typer testmateriale i test B



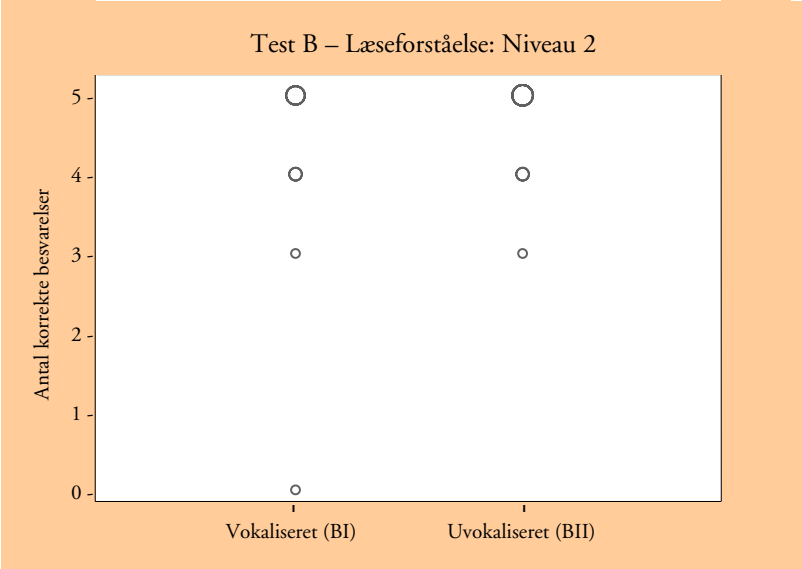
Figur 6.37 Læseforståelse ved læsning af arabiske tekststykker med vokaler (BI) og uden vokaler (BII). Cirklernes størrelse afspejler antallet af observationer i den relevante kategori.

I figur 6.37 ses boblediagrammer, der illustrerer respondenternes læseforståelse ved læsning af de arabiske historier med og uden vokalisering. Som det må forventes, stiger andelen af korrekte besvarelser med højere færdighedsniveau, hvilket fremgår tydeligst i test BI, hvor ingen respondenter på niveau 3 og 4 har mindre end fire ud af fem korrekte besvarelser. I test BII er udviklingen mindre entydig, men det generelle billede synes dog at være det samme. Den statistiske analyse bekræfter, at niveau generelt har effekt på læseforståelse ($p<0{,}001$ i både BI og BII). Tovejsanalyserne viser signifikante forskelle mellem niveau 1 og niveau 2 ($p=0{,}006$), men ikke mellem de øvrige niveauer, hvad enten man måler niveau 2 og 3 samlet mod niveau 4 ($p=0{,}172$) eller man måler niveau 2 separat mod de indfødte sprogbrugere ($p=0{,}16$). Dette betyder selvfølgelig ikke nødvendigvis, at der ikke er forskel på læseforståelse mellem respondenterne på niveau 2, 3 og 4, men blot at der ikke i data er fundet tilstrækkeligt grundlag til at kunne dokumentere det, når der måles mellem disse niveauer separat.

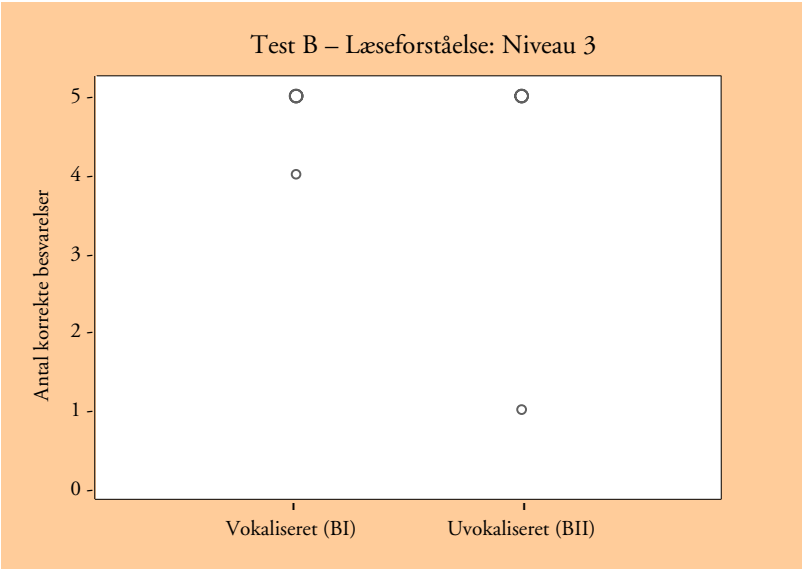
Hvad angår læseforståelse på tværs af de to testtyper, ses resultaterne klarere illustreret i figur 6.38-6.41. Af figurerne fremgår det, at læseforståelsen på niveau 1 og 2 er en smule bedre ved læsning af uvokaliseret end ved læsning af vokaliseret skrift. På niveau 3 og 4 er tendensen omvendt, men på niveau 3 afgøres udfaldet af en enkelt respondents mange fejlagtige besvarelser, og resultatet er derfor ikke særlig informativt, og på niveau 4 er den gennemsnitlige forståelse næsten ens for de to testtyper. På baggrund af den grafiske fremstilling kan man måske alligevel driste sig til at sige, at resultaterne indicerer et mønster, der ligner det ovenstående ved resultaterne for læsehastighed i samme test. Resultaterne for niveau 1 og 2 er modsat det forventede: Vokalerne giver ikke bedre læseforståelse – tværtimod. Men for de højere niveauer synes tendensen – omend den er meget beskeden – at være den modsatte, således at vokalerne bidrager til en bedre forståelse. Forklaringen kunne være den samme som givet ovenfor med hensyn til læsehastighed: På begynderstadiet udgør den øgede grafiske information, som vokalerne repræsenterer, et øget træk på de kognitive ressourcer. Derfor hæmmer de læseforståelsen frem for at støtte den. Hos lørnere på højt niveau og hos indfødte sprogbrugere kan vokalerne derimod godt udnyttes til at skabe bedre forståelse. Denne fortolkning af resultaterne underbygges dog ikke særlig klart af data, og den statistiske analyse viser da også, at forskellene i forståelse ved læsning af henholdsvis vokaliseret og uvokaliseret skrift ikke er signifikant på nogen af



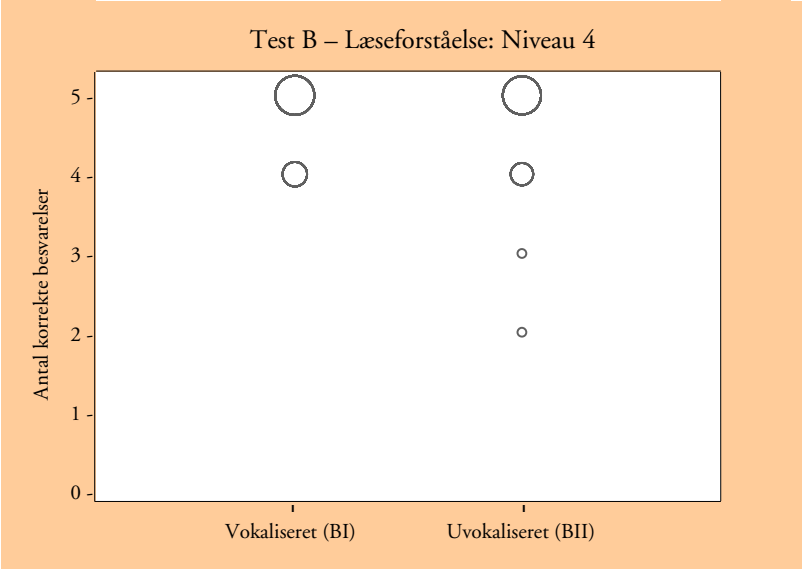
Figur 6.38 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 1



Figur 6.39 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 2



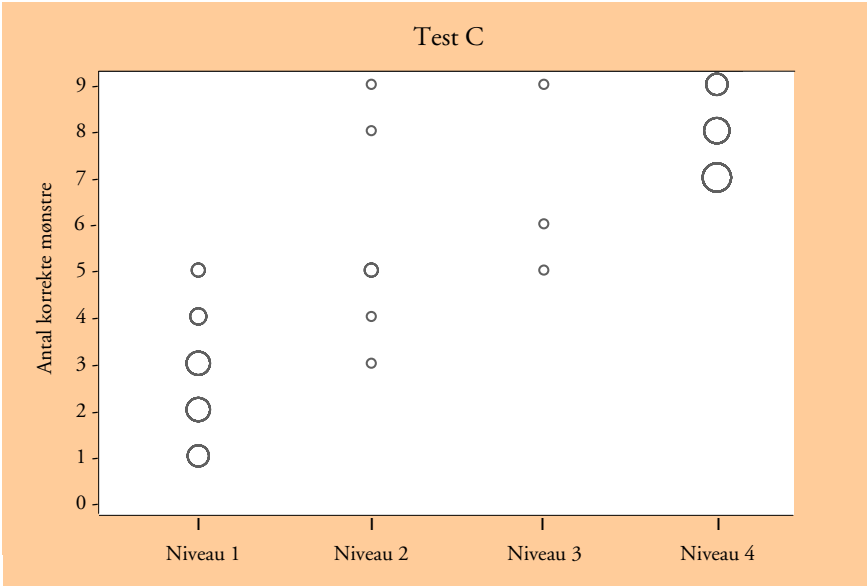
Figur 6.40 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 3



Figur 6.41 Læseforståelse ved læsning af arabiske tekststykker med og uden vokaler, niveau 4

niveauerne (p=0,316 for niveau 1, p=0,294 for niveau 2, p=0,731 for niveau 2/3 og p=0,41 for niveau 4). I denne test spiller vokalerne altså ikke nogen statistisk signifikant rolle for læseforståelsen. Det betyder selvfølgelig ikke nødvendigvis, at vokaler ingen betydning har for forståelse generelt. En svaghed ved denne undersøgelse er den begrænsede population, der gør det mindre sandsynligt, at der opstår et mønster i resultaterne. Dette problem forstærkes i denne deltest af et testdesign, hvor læseforståelsen testes på kun fem opgaver. Under de givne betingelser, hvor mange aspekter af lørnernes læsning søges belyst, og hvor respondenterne må afse undervisningstid til formålet, ville det dog vanskeligt kunne lade sig gøre at tilrettelægge testen meget anderledes. Som beskrevet i metodeafsnittet, har det – i betragtning af de tidsmæssige begrænsninger – været nødvendigt at lade testmaterialet bestå af forholdsvis korte tekster, som hverken teoretisk eller praktisk muliggør en omfattende og dybtgående testning, der kunne give en mere detaljeret afdækning af respondenternes forståelse af teksterne.

På trods af de nævnte begrænsninger synes resultaterne af test B tilsammen at vise, at vokalisering ikke som forventet giver højere læsehastighed og bedre læseforståelse hos danske lørnere. I hvert fald ikke i de første studieår. Tvært imod synes der at være en modsat sammenhæng. Med hensyn til læseforståelse er resultaterne for usikre at konkludere ud fra, men hvad angår læsehastighed synes det sandsynliggjort, at vokalerne – på trods af den fonologiske information, de repræsenterer – bliver en byrde frem for en hjælp. På baggrund af resultaterne i test A kan dette forklares med den manglende fortrolighed med det arabiske skriftsystem. Grafikken kræver simpelt hen så meget kognitiv kapacitet at afkode, at afkodningsmekanismen overbelastes af, at der tilføjes yderligere grafik. Resultatet er, at vokalinformationen ikke kan udnyttes, og overbelastningen resulterer i en ringere – eller i hvert fald mere langsommelig – afkodning.



Figur 6.42 Antal korrekte mønstre appliceret på uvokaliserede pseudoord.

Test C

I figur 6.42 ses et boblediagram, der viser resultaterne af test C, hvor respondenterne oplæser uvokaliserede pseudoord og dermed selv skal applicere vokaler i henhold til pseudoordenes struktur. Som det ses, er der over niveauerne en jævn stigning i, hvor mange af pseudoordene, der vokaliseres i henhold til valide mønstre. Den statistiske analyse bekræfter, at niveau generelt har signifikant indflydelse på antallet af korrekte mønstre, og tovejsanalyserne viser signifikante forskelle, både mellem niveau 1 og niveau 2/3 ($p<0{,}001$) og mellem niveau 2/3 og niveau 4 ($p=0{,}016$). Også hvis man måler niveau 1 mod niveau 2 separat er forskellen mellem dem signifikant ($p=0{,}002$). Denne test viser altså, at lørnere på begynderstadiet har overordentlig vanskeligt ved at vokalisere korrekt på baggrund af ordstruktur. Kompetencen forbedres med større sprogfærdighed, men indfødte sprogbrugere vokaliserer i denne undersøgelse stadig mere korrekt end lørnere på højt niveau.

Som det ses, laver også mange indfødte sprogbrugere enkelte fejl. En gennemgang af de konkrete fejl viser, at de indfødte sprogbrugeres vokaliseringsfejl først og fremmest er influeret af dialektale vokaliseringsmønstre – jvf. figur 6.3 drejer det sig især om pseudoord nr. 1 og 7, som hver især bliver vokaliseret forkert af syv af de 21 indfødte sprogbrugere. Pseudoord nr. 4 og 8 læses begge forkert af 4, pseudoord nr. 5 og 6 læses forkert én gang hver, mens nr. 2, 3 og 9 læses korrekt af alle indfødte sprogbrugere. Fejlene hos de danske lørnere synes umiddelbart mere tilfældige. Dog er nogle ord tilsyneladende lidt vanskeligere end andre: De mindst problematiske pseudoord er nr. 3, 6 og 9, som alligevel læses forkert af knapt halvdelen af de danske lørnere. Nr. 1, 2, 7 og 8 læses oftest forkert, og pseudoord nr. 8 læses fejlagtigt af langt de fleste – kun seks af de 31 danske lørnere vokaliserer det korrekt. Desuden viser gennemgangen af de konkrete fejl i test C, at de danske lørnere – især på begynderniveau – generelt udviser en forkærlighed for at benytte *fatha* – dvs. applicere vokalen [a].

Kvalitativ analyse af omkodningsfejlene i test A

Som et supplement til de kvantitative analyser kan det være interessant at se nærmere på, præcis hvilke slags omkodningsfejl, respondenterne laver ved læsning af pseudoord. En overordnet gennemgang af omkodningsfejlene viser, at hos de danske lørnere er der klart flest tilfælde af substitution af bogstaver – dvs. fejl der skyldes, at ét bogstav læses som et andet – både i test AI og AII. Både tilføjelser og udeladelser af bogstaver forekommer også hos alle danske lørnere – tilføjelser lidt hyppigere end udeladelser. Repetitioner og ombytninger er sjældne og forekommer hos de fleste lørnere slet ikke. Ved de lange vokaler forekommer både udeladelser, ændringer og tilføjelser, og tilføjelser er inden for denne specifikke fejlgruppe den mest hyppige – og på niveau 1 er den særdeles udbredt. Dette skyldes utvivlsomt usikkerhed omkring vokalernes længde – ofte indsættes en lang i stedet for en kort vokal. Mange udeladelser af lange vokaler skyldes formentlig samme usikkerhed i forhold til vokallængde, idet lange vokaler erstattes af korte. Det er dog heller ikke helt ualmindeligt at én lang vokal substitueres med en anden. Fejl i *hamza* forekommer kun sjældent, og når det sker er det oftest en tilføjelse. Også hvad angår fejlene i de diakritiske tegn er substitution den hyppigste fejltype. Udeladelser forekommer især ved endelser – både simple endelser og *tanwiin*-endelser udelades næsten konsekvent hos nogle lørnere. Det er der for så vidt ikke noget mærkeligt i. Mange lørnere sorterer simpelt hen endelsesvokaler fra, fordi de i talesprog i mange tilfælde ikke udtales. Tilføjelser og udeladelser af fordoblingstegnet *shadda* forekommer, men i forhold til hyppigheden af fejl i de øvrige diakritiske tegn er denne type fejl ikke særlig udbredt. De marokkanske respondenter laver en del substitutionsfejl i test AII, men derudover er de fleste af deres oplæsningsfejl fejl i de diakritiske tegn – mest i form af tilføjelser. I test AIII er ombytninger den hyppigste fejltype hos de danske lørnere. Udeladelser og tilføjelser forekommer kun ganske få gange. Ved substitution er vokal- og konsonantfejl nogenlunde lige repræsenteret. Hos de marokkanske respondenter ses nogenlunde lige mange substitutions- og ombytningsfejl. En høj koncentration af vokalfejl findes hos enkelte af respondenterne, mens andre slet ikke laver vokalfejl. Repetition af bogstaver i AIII findes hverken hos de danske eller de marokkanske respondenter.[9]

Til brug for en mere dybtgående analyse af de konkrete oplæsningsfejl udvælges udelukkende bogstavsubstitutionsfejl. Denne fejltype er ikke blot den mest hyppigt forekommende hos de danske lørnere – det er også den eneste fejltype, der direkte kan relateres til det teoretiske udgangspunkt. Bogstavsubstitutioner kan nemlig umiddelbart sammenkædes med bogstavgenkendelsesteorierne, der er præsenteret i kapitel 4, mens udeladelser, tilføjelser og ombytninger er vanskeligere at forklare. Desuden udgår de marokkanske respondenters vokalfejl i test AIII af analysen, fordi de må antages primært at skyldes fonetiske fejl, da der i deres modersmål kun opereres med tre vokaler.

Undersøgelsen kan opdeles i følgende to delstudier:

- En analyse af de indfødte sprogbrugeres fejl, der kan give bogstavgenkendelsesteorierne, som blev præsenteret i kapitel 4, et empirisk grundlag, selvom populationsgrundlaget naturligvis er meget begrænset.

[9] Det totale sæt af enkeltobservationer foreligger kun som originalt analysemateriale. Det er derfor ikke inkluderet i Appendix 2, men foreligger til gennemsyn i bedømmelsesperioden, jvf. afsnit 6.1.

- En analyse af de danske lørneres fejl, der gør det muligt at finde eksempler på, hvordan de scanningsstrategierne hos læsere, der først lærer det arabiske skriftsystem i voksenalderen, influerer på bogstavgenkendelsen.

I figur 6.43 ses en oversigt over de indfødte sprogbrugeres bogstavforvekslinger. De arabiske bogstavfejl er kategoriseret efter, hvorvidt det grafiske input og det fonetiske output repræsenterer bogstaver inden for samme grafiske base, jvf. figur 4.8. De latinske bogstavfejl er opstillet i vilkårlig rækkefølge. Da testmaterialet ikke er kontrolleret i relation til antallet af forekomster af hvert enkelt bogstav, giver antallet af forekomster af hver fejltype ikke grundlag for indbyrdes sammenligninger, men antallet af respondenter, der laver samme type fejl, kan give en ide om, at den vedkommende fejltype er stærkt repræsenteret. Bemærk i øvrigt, at der ikke umiddelbart kan sammenlignes mellem forekomsten af fejl på tværs af de to skriftsystemer, da opgørelsen bygger på dobbelt så meget arabisk som latinsk tekstmateriale, og da de anvendte fejlkategorier ikke er reguleret med henblik på at skabe et brugbart sammenligningsgrundlag for de to skriftsystemer (som det er tilfældet ved kategorien realfejl).

	Niveau 4															Antal resp. der laver denne fejl
Resp. ID	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**52**	**53**	**54**	**57**	**58**	**59**	**60**	**61**	
Forveksling på baggrund af identisk base																
ب ← ن							1									1
ب ← ت														1		1
ت ← ث				1	3											2
ع ← غ				1									2			2
ق ← ف									1							1
خ ← ح						1			1	1	1	1		1		6
د ← ذ					1								1			2
ذ ← د		1	2	2	1	2	1					3				7
ض ← ظ									2	2	4			2	1	5
ظ ← ض			3	1		1		1		1	1	1		1	1	9
ظ ← ط									1			1				2
Forveksling på tværs af baser																
ز ← ذ						2	1	3	2						1	5
خ ← غ			1													1
ك ← ق									1							1
ر ← غ								1								1
ة ← ا					1			1								2
Forveksling af latinske bogstaver																
f → v						1										1
v → f			1													1
r → f										1						1
t → f			1													1
t → k										1						1
t → b															1	1
k → h			1				2									2
m → n				2												1

Figur 6.43 Oversigt over substitutionsfejl hos de indfødte sprogbrugere

Det umiddelbare indtryk af oversigten over de arabiske bogstavfejl er, at ud af alle mulige kombinationer af forvekslinger mellem bogstaver er kun forholdsvis få repræsenteret, men til gengæld forekommer mange af dem flere gange. Dette underbygger, at bogstav-genkendelsesfejl ikke er tilfældige.

Som det ses, er hovedparten af fejlene forvekslinger af bogstaver med samme base, hvor kun antallet af prikker (og for dråbeformen evt. en vertikal streg) udgør forskellen mellem bogstavernes grafiske udtryk. Dette bestyrker umiddelbart analysen i kapitel 4, der bygger på Boumas gruppering af latinske bogstaver. Dog er der grund til at tage visse forbehold. For nogle af de mest hyppige fejl kan der nemlig være tale om dialektal indflydelse: I den marokkanske talesprogsvariant udtales ذ ofte som د og ظ altid som ض, og derfor kan disse fejl meget vel være resultater af overførsel af den dialektale fonologi. Samtidig kan fejl, der går den anden vej, altså fra د til ذ og fra ض til ظ skyldes overfokusering på netop disse forskelle mellem bogstavernes fonetiske værdier i henholdsvis talesprogsvarianten og MSA (hyperkorrektion). En tredje fonetisk værdi, der normalt ændres i marokkansk dialekt i forhold til MSA er ث , der gerne udtales ت. Som det ses forekommer der kun fejl, der går den anden vej. Det kan ikke udelukkes, at også disse fejl skyldes hyperkorrektion frem for forveksling af bogstavernes grafiske udtryk. Bemærk, at hos en stor del af respondenterne forekommer forveksling af خ og ح. – Nøjagtig samme fejl, som er rapporteret af Azzam (1989) og Abu-Rabia (1995).

Hvad angår fejl på tværs af de grafiske baser, er det mest bemærkelsesværdige en stor forekomst af tilfælde, hvor ز læses som ذ. Havde det været omvendt (at ذ blev læst som ز) og havde respondenterne været mellemøstlige, kunne også dette forklares med dialektal indflydelse. På marokkansk udtales ذ imidlertid aldrig som ز. Forvekslingen synes altså at være grafisk betinget. Og dette giver faktisk yderligere støtte til den teoretisk baserede klassifikation i figur 4.8. Her udgør bogstaver, der ikke sammenbindes mod venstre nemlig en selvstændig hovedgruppe, idet den manglende sammenbindingsstreg må forventes at have stor indflydelse på genkendelighed og dermed klassifikation. Når ز læses som ذ, kan det skyldes, at de begge tilhører gruppen af bogstaver, der ikke sammenbindes, og at de begge har én prik. Baserne ligner også hinanden, idet den eneste forskel imellem dem er, at ز forlænges under den bundlinje, som sammenbindingsstregen repræsenterer.

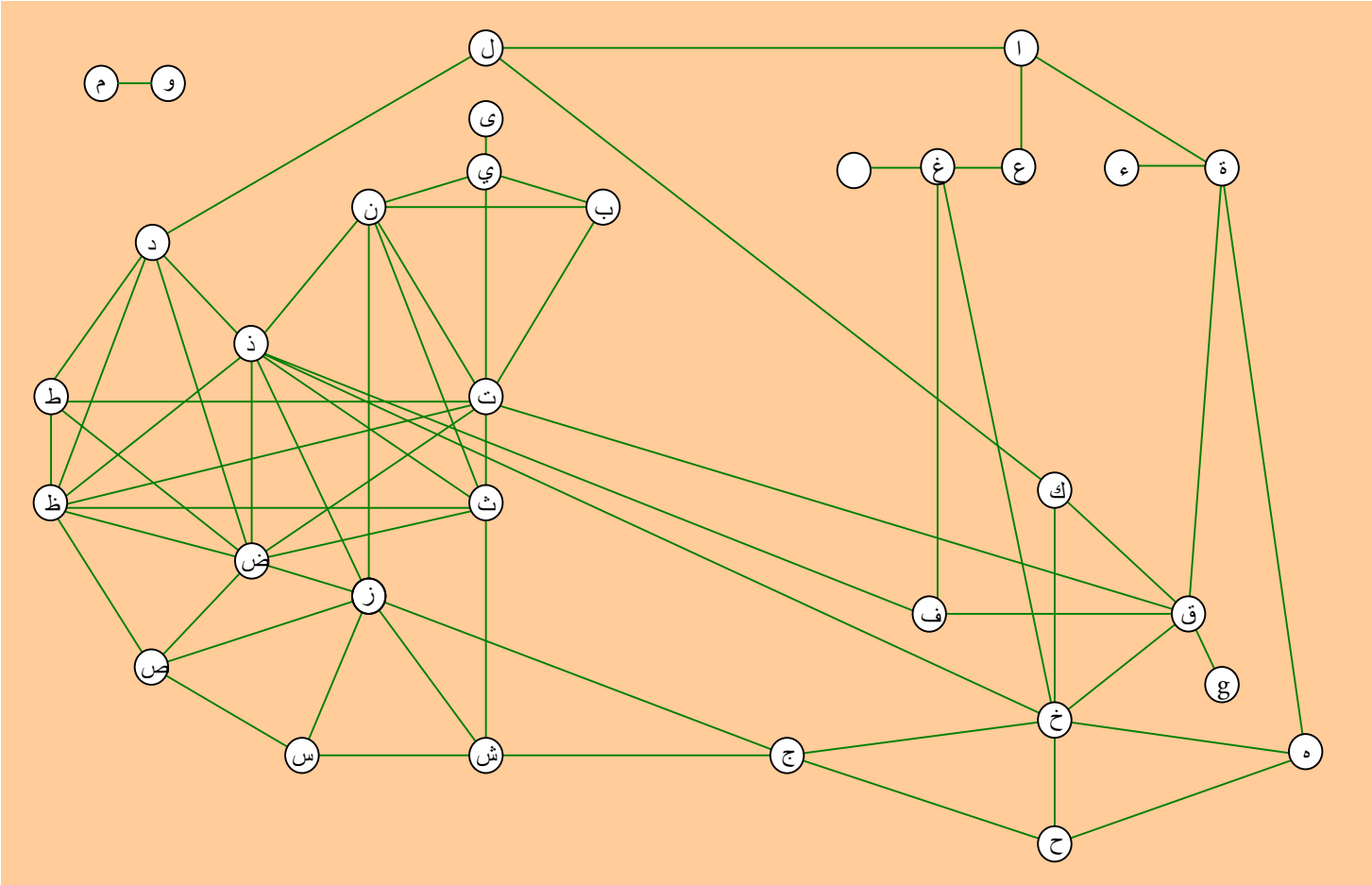
I relation til den usikkerhed, de marokkanske respondenter udviser over for de fonemer, der er fremmede for de danske lørnere, kan resultaterne give en ide om, hvad der resulterer i dette udfald. Det store antal forvekslinger mellem bogstaverne ض og ظ – og det faktum, at forvekslingen i så høj grad går begge veje – skyldes sandsynligvis **både** den dialektale indflydelse **og** de grafiske ligheder mellem disse bogstaver: Da bogstaverne i den marokkanske talesprogsvariant har samme fonetiske værdi, og da de samtidig grafisk ligner hinanden så meget, blandes de simpelt hen sammen. Hertil kommer den ret hyppige forveksling af خ og ح – som også begge repræsenterer fonemer, der er fremmede for de danske lørnere. At netop denne fejltype er så hyppig – hvilket ikke blot ses i denne undersøgelse men som nævnt også er refereret andetsteds – kan ikke umiddelbart forklares. Mange bogstavers værdi determineres udelukkende af en prik, så hvorfor forekomsten af netop denne forveksling tilsyneladende er ret hyppig, er ikke lige til at gennemskue.

Hos de danske respondenter er der langt flere observationer at analysere, og da der ikke er grund til at formode, at der findes samme systematik bag de danske lørneres fejl som bag de indfødte sprogbrugeres, udformes et såkaldt display over alle observerede bogstav-forvekslinger. Dette ses i figur 6.44. Umiddelbart kunne man måske forvente også her at se en systematik, der relaterede til bogstavkategorierne i figur 4.8. Men ved første øjekast er en sådan systematik ikke synlig. Der er dog ikke på den baggrund anledning til at formode, at de indlysende grafisk sammenfald, der er grundlaget for kategoriseringen i figur 4.8, er irrelevante. Men displayet i figur 6.44 vidner om, at der i det mindste må ligge andre former for lighed til grund for nogle af forvekslingerne. Derfor fjernes de forbindelser, der i displayet går mellem bogstaver med identisk grafisk base. Det reviderede display ses i figur 6.45. Også her er det svært at få øje på en klar systematik, men ser man nærmere på forbindelserne, kan man se, at to former for grafiske sammenfald er repræsenteret i mange af de tilbageværende forvekslingspar: For det første er der tilsyneladende mange tilfælde af forveksling af bogstaver med forskellig grafisk base men med samme antal og placering af prikker. Noget tyder altså på, at prikkerne ikke altid overses, men at der tværtimod af og til overfokuseres på dem. Hertil kommer forbindelser mellem bogstaverne ا, ل, د og ك/ـك, som også i høj grad deler et basalt grafisk udtryk i kraft af den vertikale streg og **fraværet** af prikker. Fjerner man nu også alle disse forbindelser, ses resultatet i figur 6.46. Her ses fem klart adskilte grupperinger af bogstaver, der indbyrdes er genstand for forvekslinger. Forklaringen på de tilbageværende grupperinger synes af være fonologisk gruppering. Det drejer sig om følgende grupper:

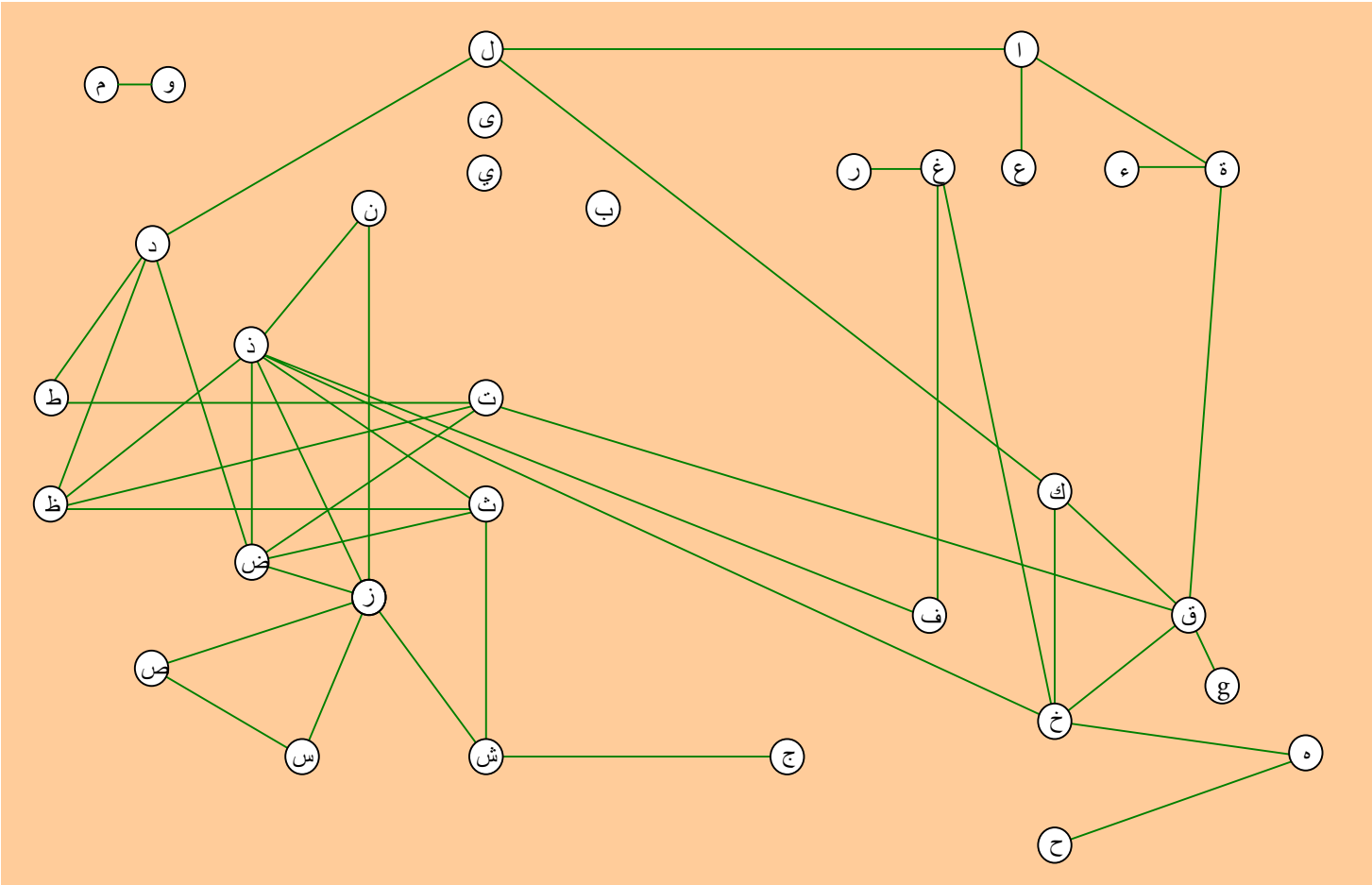
- Bogstaver, der repræsenterer dentale klusiler (man kunne kalde dem d/t-lyde).
- Bogstaver, der repræsenterer alveolære og postalveolære spiranter (s/ʒ-lyde).
- Bogstaver, der repræsenterer velære, uvulære og faralyngale spiranter (k/h-lyde)
- Bogstaver, der går på tværs af fonetiske grupperinger, men som af danskere normalt vil opfattes som r-lyde
- Bogstaver, der går på tværs af fonetiske grupperinger, men som af danskere normalt vil opfattes som a-lyde

Dermed kan bogstavsubstitutionerne kategoriseres efter en systematik som den, der ses i figur 6.47. Kategorierne bygger på de observerede bogstavsubstitutioner, og som det ses hører nogle bogstaver til i flere kategorier, idet grafisk lighed i mange tilfælde afspejler fonetisk lighed som beskrevet i kapitel 3.

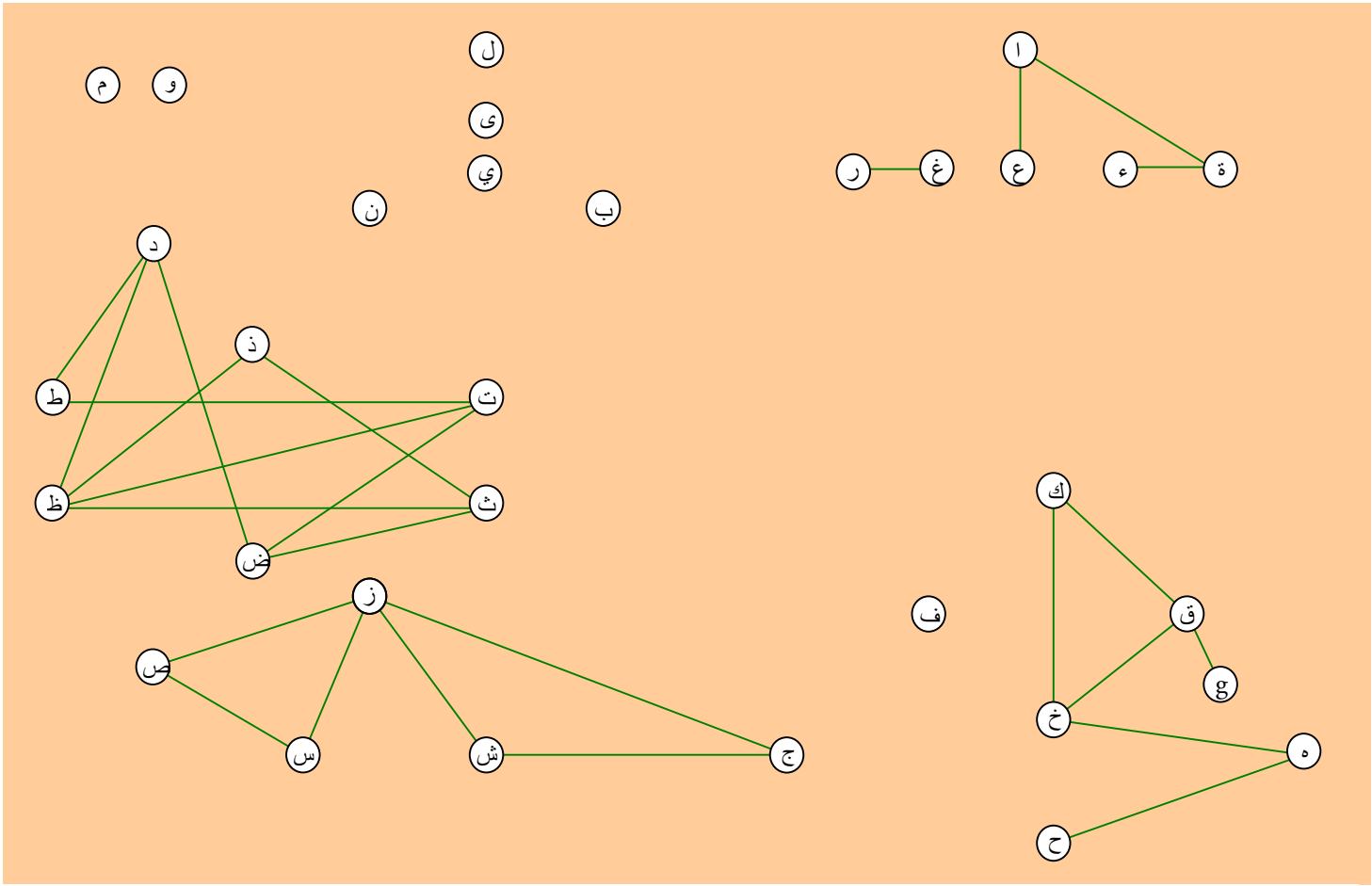
En fuld oversigt over alle substitutionsfejl hos de danske respondenter foreligger i Appendix 4.1. I figur 6.48 og 6.49 ses en beskåret oversigt, der omfatter alle de bogstavsubstitutioner, der er observeret mere end én gang hos de danske lørnere. Observationerne er inddelt efter hvorvidt de repræsenterer bogstavsubstitutioner, der jvf. oversigten ovenfor er oplagt grafisk eller fonetisk betingede – eller som kan være begge dele. Øverst i figuren ses forvekslinger af bogstaver, der jvf. figur 4.8 har identisk base. Som det ses, forekommer der en del af den slags fejl. Mærkeligt nok er forveksling mellem خ [x]og ح, som er meget hyppig hos de indfødte sprogbrugere, helt fraværende hos de danske lørnere – faktisk er der ikke mange fejl, hvor forvekslingen kommer til udtryk som flere eller færre prikker. I de fleste tilfælde er antallet af prikker korrekt, men placeringen af dem forkert.



Figur 6.44 Display over samtlige forvekslingsfejl hos de danske lørnere



Figur 6.45 Display over forvekslinger af bogstaver, der ikke har samme grafiske base



Figur 6.46 Display over forvekslinger af bogstaver, der hverken har åbenlyst grafisk sammenfald i base eller samme antal prikker

Grafemiske kategorier:

ـب	ـبـ	تـ	ثـ	يـ	نـ		
ـحـ	ـحـ	خـ	جـ				
ـصـ	ـصـ	ضـ	ظـ				
ـعـ	ـعـ	غـ					
ـفـ	ـفـ	قـ					
ه	ه	ة					
ى	ى	ي					
.	ذ	ز	ن	ض	غ	ف	خ
..	ق	ت	ة				
.:	ث	ش					
ا	ا	ل	ـك	د			
م	م	و					

Fonetiske kategorier:

t/d	ث	ذ	ظ	ت	ط	د	ض
s/ʒ	س	ص	ز	ش	ج		
k/h	ك	خ	ق	g	ح	ه	
r	ر	غ					
a	ا	ع	ء	ة	ـ		

Figur 6.47 Grafiske og fonetiske kategorier, inden for hvilke bogstavsubstitution forekommer hos de danske lørnere

	Niveau 1																				Niveau 2					Niveau 3				Antal resp.		
Resp. ID	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25	36	37	38	på niveau		
Udelukkende grafisk lighed																														1	2	3
ب ← ن	3	1					2	1			1	1															1			6		1
ن ← ب												1								1						1				2	1	
ت ← ن							1											1												2		
ي ← ن	2																													1		
ي ← ب								1	2		1																			3		
ب ← ي									1		1		1					1												4		
ي ← ت							1																1							1	1	
خ ← ج				1		2										1	4	1	1				2	1		1			2	6	3	1
ج ← خ									1			1												1						2	1	
ج ← ح					1																					1				1	1	
ص ← ض									2		2																			2		
غ ← ع													1																	1		
ف ← ق																	1	1												2		
ة ← ه														3						3										2		
ه ← ة	2	1																												2		
ي ← ى				1																2								1		2		1
ز ← ذ										1		1																		2		
ذ ← ز																																
ف ← غ	2																													1		
ق ← ة												1				1												1		2		1
ش ← ث		1	1	2			1			1							1							1						6	1	
ث ← ش	1	2										1	1																	4		
ل ← ا	2		2	2	2	1			4				1					1		3		1								9	1	
ا ← ل								1														1								1	1	
ل ← ك				1		1																								2		
ل ← د					1																					1				1	1	
ل ← د																					1										1	
	Niveau 1																				Niveau 2					Niveau 3				Niveau		
Resp. ID	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25	36	37	38	1	2	3
Udelukkende fonetisk lighed																																
ق ← خ				1	1						1																			3		
ق ← ك	7	2			5			5					1			2				1								1		7		1
ق ← g			4																											1		
خ ← ه											1	1	2	2		2	1	1												7		
ح ← ه		2		1			2	1	1											2				3						6	1	
ك ← خ					1											1														2		
ك ← ق													2				1				2									2	1	
ه ← ح						1												1												2		
ج ← ش		2																												1		
ز ← س							1	1	2							1		1									1			5		1
ز ← ث			1					1																						2		
ش ← ج					1	1						1		1																4		
س ← ز											1	2					1													3		
ض ← د		1					1				1		2	1		1		1				1		1		1				7	3	
ض ← ث								1	1																					2		
ط ← ت		1			1		1	1	2																					5		
ط ← د												1			1												1			2		1
ظ ← ذ									1			1					1													3		
ظ ← د						1											1													2		
ظ ← ت					2		1		1			1																		4		
ظ ← ث																1								2						1	1	
ر ← غ					1																									1		
غ ← ر								1			1			3				3												4		
ث ← ذ									2																					1		
ع ← ا			1					1					2													1				3	1	
ا ← ع	1	1							1																					3		
ء ← ة	1	1		1								1																		3		
- ← ة		1	3			4																								3		
ا ← ة									2	4			2							1										4		

Figur 6.48 Forveksling af bogstaver der udelukkende har grafisk eller fonetisk lighed, og som forekommer mere end én gang.

	Niveau 1																				Niveau 2					Niveau 3				Niveau		
Resp. ID	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25	36	37	38	1	2	3
Både grafisk og fonetisk lighed																																
ت ← ث																		1							1					1	1	
ث ← ت				1									1				1													3		
ح ← خ											1						1													2		
خ ← ح						1	1						1		2			1												5		
د←ذ													1							1								4		2		1
ذ←د		2					3				4			1				1							1				1	5	1	1
س ← ش	2	3	2		4					1						1		1		7					1		1	1		8	1	2
ص←س													2			1		1												3		
ض←ظ												1		1			1		1	1							2		1	5		3
ظ←ض																												1	1			2
ض←ط								1		1										1										3		
ط ← ظ												1								1										2		
ظ←ط		4		1	1	1	1	4	1	1	4		1	3												1				11	1	
غ ← خ													1			1														2		
خ ← غ								1																						1		

Figur 6.49 Forveksling af bogstaver der både har grafisk og fonetisk lighed, og som forekommer mere end én gang

At lørnerne ikke mangler fokus på antallet af prikker ses også i den næste kategori: Her er antallet af prikker og placeringen af dem den samme ved de bogstaver, der forveksles, mens bogstavernes grafiske baser er forskellige. Det er bemærkelsesværdigt, at der slet ingen tilfælde er, hvor ز læses som ذ, hvad der jvf. ovenstående er en hyppig fejl hos de indfødte sprogbrugere. Endelig er der et antal tilsyneladende grafisk betingede forvekslinger mellem bogstaverne ا, ل, د og ك (ـك), som til gengæld kun forekommer hos de danske lørnere og ikke hos de indfødte sprogbrugere. Tilsammen tyder disse forskellige fejltyper hos danske og marokkanske respondenter på, at den nævnte selvstændige kategorisering af bogstaver, der ikke forbindes til venstre, ikke er etableret hos de danske lørnere. Derfor undgår de forveksling af ز og ذ, men laver til gengæld typer af fejl, som de indfødte sprogbrugere slet ikke laver – som fx at forveksle ا og ل. At ا så ofte læses som ل som det er tilfældet, kan muligvis skyldes, at ا jo faktisk er grafisk identisk med et latinsk l.

Under de nævnte fejl, der er forvekslinger af bogstaver med grafisk lighed, ses forvekslinger mellem bogstaver, der forekommer inden for de i figur 6.47 anførte fonetiske kategorier. Her skyldes bogstavsubstitution altså sandsynligvis i højere grad fejlagtig udtale end fejlagtig bogstavgenkendelse. Ikke overraskende involverer mange af disse fejl de såkaldt fremmede fonemer. I nogle tilfælde ses overfokusering på de fremmede lyde, således at bogstaver, hvis fonetiske værdi findes på dansk, udtales som et af de mere fremmedartede – som når fx ك [k] udtales som خ [x] eller ق [q], eller når غ [ʁ] udtales som ر [r]. De forholdsvis gode resultater med udtale af bogstavet ع er overraskende, da det erfaringsmæssigt er en vanskelig lyd at producere for mange begyndere. En vis overfokusering på denne lyd ses da også – fire lørnere udtaler mindst én gang et helt 'almindeligt' ا som ع. Det står således klart, at der ikke blot er tale om, at fremmede fonemer udtales forkert, men om en generel tendens til at sammenblande fonemer inden for de anførte fonetiske kategorier, hvad enten de enkelte fonemer er fremmede eller ej – måske fordi kategorien omfatter vanskelige lyde, som lørnerne så at sige er 'på vagt overfor'.

I figur 6.49 ses bogstavforvekslinger, som på grund af både grafiske og fonetiske ligheder ikke er kan placeres som enten grafisk eller fonetisk betingede fejl. Et kvalificeret gæt kunne

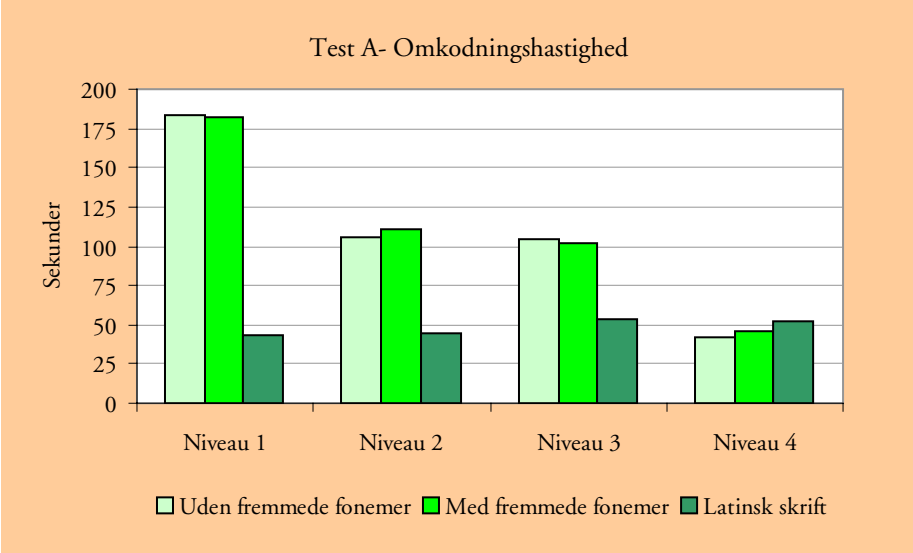
være, at den hyppige forveksling af س [s]og ش [ʃ]skyldes de to bogstavers grafiske lighed – lyden [ʃ] er jo velkendt fra engelsk. Den meget store forekomst af ظ læst som ط skyldes muligvis både grafik og fonetik. Der er for det første et stort grafisk sammenfald – og det at fonemet er fremmed, gør ondt værre. Hovedparten af lørnerne på niveau 1 laver netop denne fejl mindst én gang. Og bogstaverne med dråbeform ser generelt ud til at give flere former for vanskeligheder. Da de alle repræsenterer de såkaldt emfatiske fonemer, er det ikke overraskende, at de i mange tilfælde udtales som deres mere velkendte modstykker, men tilsyneladende sker der derudover en vis sammenblanding af bogstaverne, der har dråbeformen som base – fx. udtales ظ både som ذ [ð], د [d], ت [t] og ث [θ] – og det kan skyldes, at ظ fejllæses både grafisk og fonetisk, så ظ fx grafisk forveksles med ط, som så fonetisk forveksles med ت. Man kan også forestille sig, at sådanne forvekslinger ikke skal begrundes med så faste én-til-én forvekslinger, men simpelt hen er udtryk for en mere eller mindre diffus fornemmelse af bogstavernes grafiske og fonetiske 'identiteter' – og at bogstaver, der både deler grafisk udtryk og fonetisk kategori, ikke blot er vanskelige at skelne, men muligvis ganske enkelt mangler en unik repræsentation hos mange lørnere. Resultatet bliver, at disse lørnere har en svag og ret upræcis opfattelse af, hvilke fonetiske værdier de enkelte bogstaver repræsenterer, hvilket igen forhindrer, at de får opbygget en klar grafisk distinktion mellem dem, fordi de mangler et utvetydigt fonetisk prædikat i lørnerens bevidsthed. Og dette kan selvfølgelig også på længere sigt forhindre etableringen af hurtig og præcis bogstavgenkendelse.

6.4 Sammenfatning

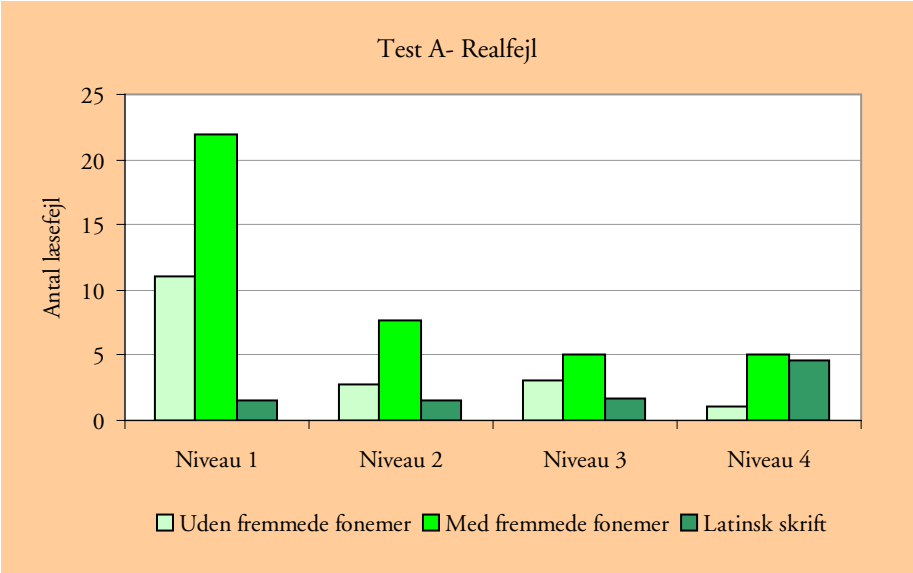
I figur 6.50-6.52 ses diagrammer, der giver et overblik over resultaterne af test A. Af figur 6.50 fremgår det tydeligt, at de danske respondenter på alle niveauer (niveau 1-3) omkoder arabisk skrift meget langsommere, end de omkoder latinsk skrift, ligesom de omkoder meget langsommere end kontrolgruppen af indfødte sprogbrugere (niveau 4). Det fremgår også, at de omkoder arabisk skrift med og uden fremmede fonemer omtrent lige hurtigt. Af figur 6.51 fremgår det, at de på alle niveauer laver flere fejl ved omkodning af arabisk skrift end ved omkodning af latinsk skrift, ligesom de laver flere fejl end de indfødte sprogbrugere. På alle niveauer laver de flere fejl i pesudoord med end i pseudoord uden fremmede fonemer, og i figur 6.52 ses det, hvordan udviklingen i omfanget af især bogstavfejl forløber mere gradueret ved læsning med end uden de særegne arabiske lyde.

Under forudsætning af, at omkodning er en gyldig indikator for ordgenkendelse – hvilket der argumenteres for i de teoretiske kapitler – kan disse resultater relateres til de to hypoteser, der vedrører test A:

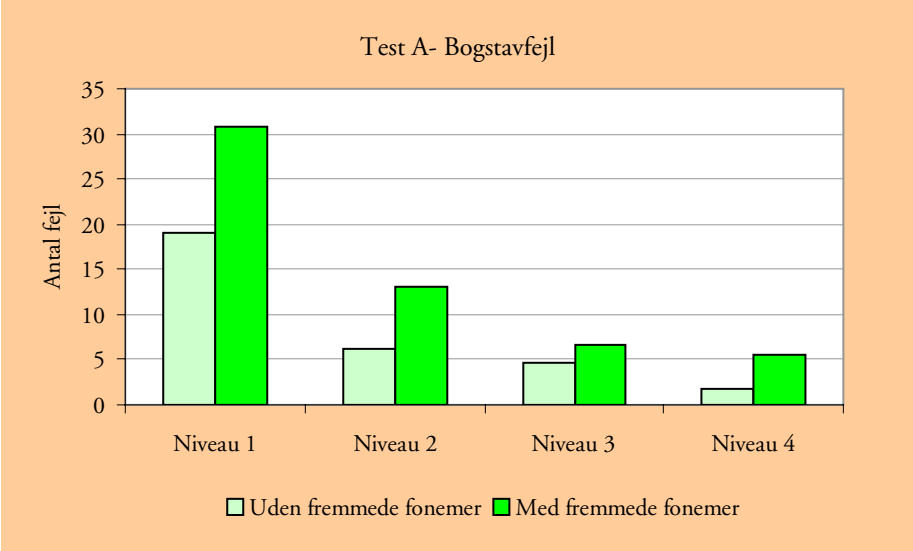
- Hypotese 1: Den uvante bogstavarkitektur sænker hastighed og øger fejlfrekvens for ordgenkendelse på arabisk som L2
- Hypotese 2: De fremmed fonemer sænker hastighed og øger fejlfrekvens for ordgenkendelse på arabisk som L2



Figur 6.50 Gennemsnitlige omkodningshastigheder på alle niveauer og ved alle testtyper i test A



Figur 6.51 Gennemsnitligt antal realfejl på alle niveauer og ved alle testtyper i test A



Figur 6.52 Gennemsnitligt antal bogstavfejl ved læsning af arabiske pseudoord i test A

Ovenstående resultater støtter hypotese 1. Den del af hypotese 2, der vedrører fonemernes indflydelse på hastighed, må forkastes, mens resultaterne understøtter hypotesen om, at fremmede fonemer øger fejlfrekvensen.

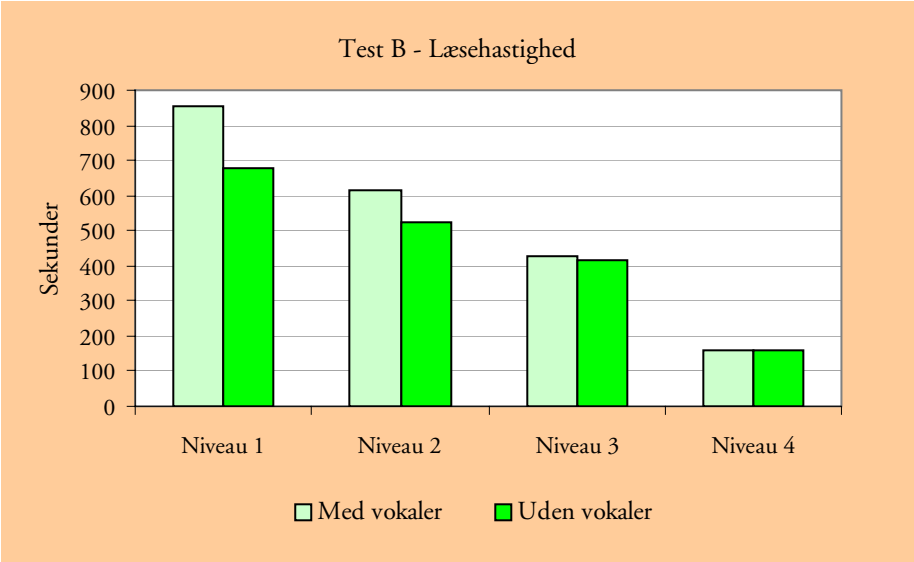
Hvis man ser på udviklingen i lørnernes omkodningsfærdigheder gennem tilegnelsesforløbet kan disse konklusioner nuanceres med følgende betragtninger: I figur 6.50 ser man, at lørnernes omkodningshastighed forbedres markant fra niveau 1 til niveau 2, men derefter går udviklingen i stå, og i figur 6.51 og 6.52 kan man se, at også med hensyn til deres omkodningsfejl formindskes antallet markant i løbet af de to første studieår, men især antallet af bogstavfejl ved læsning af pseudoord med fremmede fonemer mindskes gennem hele studieforløbet. Udviklingen går således langsommere for læsning med end for læsning uden fremmede fonemer. De fremmede fonemer repræsenterer altså betydelige vanskeligheder – og det ikke kun for begynderne. Faktisk indicerer de samlede resultater i test A, at oplæsningsfejl i begyndelsen skyldes både den fremmede grafik og de fremmede fonemer, mens de senere i tilegnelsesforløbet først og fremmest ser ud til at skyldes de fremmede fonemer og i mindre grad den fremmede grafik. Også den kvalitative analyse viser, at omkodningsfejl skyldes både grafemisk og fonetisk forveksling. Den viser også, inden for hvilke grafemiske og fonetiske kategorier disse forvekslinger finder sted. Med hensyn til fejl i de diakritiske tegn viser analyserne, at de tilsyneladende er mere individuelt end niveaumæssigt betingede.

I figur 6.53 og 6.54 ses diagrammer, der på samme måde giver overblik over resultaterne i test B. Af figur 6.53 fremgår det, at de danske lørnere generelt læser vokaliseret arabisk langsommere end uvokaliseret arabisk. I figur 6.54 ser man, at læsning af vokaliseret arabisk også giver færre korrekte besvarelser i forståelsestesten på niveau 1 og 2, mens det modsatte er tilfældet på niveau 3. Forskellene i forståelsestesten er imidlertid ikke statistisk signifikante, og vokalisering har altså ingen påviselig effekt på læseforståelsen i denne undersøgelse. De hypoteser, der relaterer til test B lyder som følger:

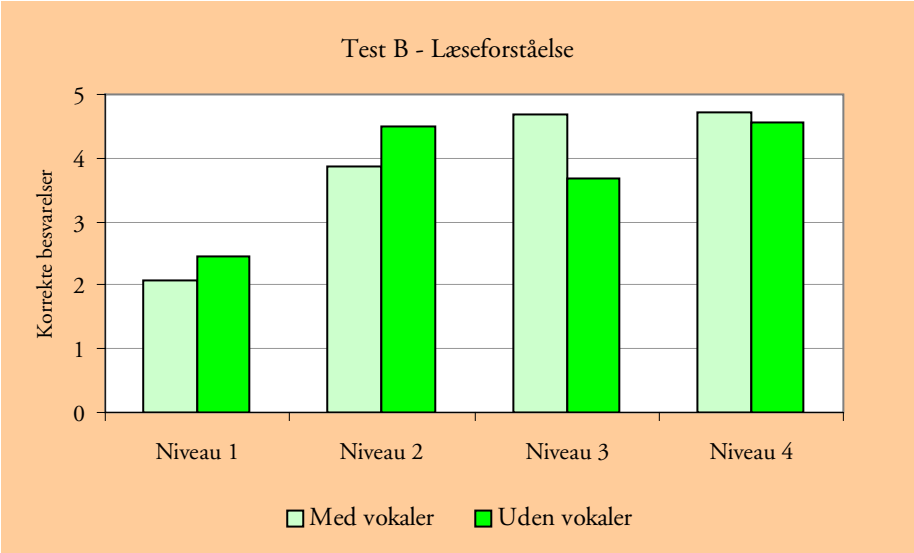
- Hypotese 3: Manglende vokaler har negativ indflydelse på afkodningshastighed hos læsere med arabisk som L2
- Hypotese 4: Manglende vokaler har negativ indflydelse på læseforståelse hos læsere med arabisk som L2

Både hypotese 3 og 4 må altså forkastes. Sammenhæng mellem vokalisering og hastighed er nemlig den stik modsatte. Og med hensyn til udviklingen mellem færdighedsniveauerne ser man, at på begynderniveau læser de danske lørnere den uvokaliserede tekst betragteligt hurtigere end den vokaliserede. Denne hastighedsforskel ved læsning med og uden vokaler svinder dog med bedre sprogfærdighed, og andre studier viser, at indfødte sprogbrugere rent faktisk kan udnytte vokaler til at opnå større læsehastighed og bedre læseforståelse, jvf. afsnit 5.1. Resultaterne indicerer, at det fremmede skriftsystem er så stor en kognitiv belastning, at den yderligere grafiske information, vokalerne repræsenterer, bliver en forhindring frem for en hjælp, men efterhånden som lørnernes læsefærdigheder styrkes, bliver belastningen mindre udtalt. At udnytte vokalisering kræver altså et kognitivt overskud, som L2-lørnere – især de der er på begynderniveau – endnu ikke har.

Dette resultat er overraskende, da den modsatte sammenhæng normalt antages, nøjagtig som det kommer til udtryk i hypoteserne. Som det ses i figur 6.11, rapporterer mange respondenter også i spørgeskemaundersøgelsen, at det er sværest at læse uvokaliserede



Figur 6.53 Gennemsnitlige læsehastigheder på alle niveauer og ved begge testtyper i test B



Figur 6.54 Gennemsnitligt antal korrekte besvarelser i forståelsestestene på alle niveauer og ved begge testtyper i test B

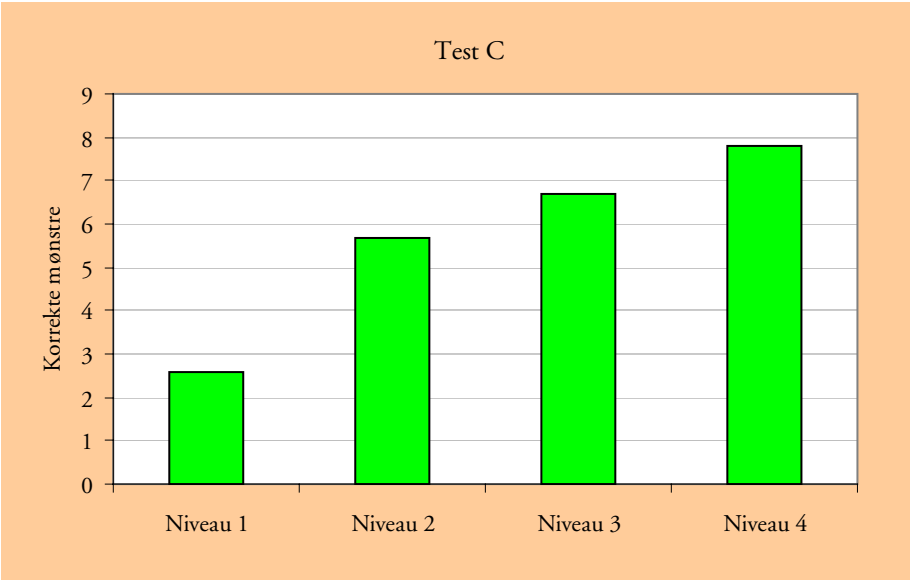
tekster. Men en enkelt af respondenterne giver faktisk i den åbne svarkategori udtryk for, at hun har gennemskuet sagens sammenhæng, idet hun blandt andet skriver:

> *"[Jeg synes] faktisk at vokaliserede tekster er sværere (de virker mere forvirrende) end uvokaliserede tekster"* (Resp. nr. 6)

Dette udsagn afspejler umiddelbart en holdning til sværhedsgrad, som må formodes ikke blot at vedrøre læsehastighed men også læseforståelse. Og i betragtning af, at det samlede billede i figur 6.54 'forstyrres' af at et stort antal fejl hos en enkelt studerende på niveau 3 ved læsning uden vokaler, kan man godt forestille sig, at der rent faktisk er en sammenhæng mellem vokalisering og forståelse, svarende til den, der er beskrevet vedrørende forholdet

mellem vokalisering og læsehastighed. Men det kan altså ikke dokumenteres i denne undersøgelse.

I figur 6.55 ses det gennemsnitlige antal korrekte besvarelser i C fordelt på niveau. Som det ses, er de danske lørnere generelt mindre gode end de indfødte sprogbrugere til at applicere korrekte mønstre på ukendte ord. Dette må skyldes, at de ikke umiddelbart genkender ordets morfologiske struktur og dermed det mønster, der giver den korrekte vokal-kombination.



Figur 6.55 Gennemsnitligt antal korrekte mønstre appliceret på uvokaliserede pseudoord i test C

Hypotesen, der vedrører test C lyder:

- Hypotese 5: L2-lørnere mangler kendskab til de morfologiske strukturer, som indfødte sprogbrugere støtter sig til som kompensation for de manglende vokaler.

Denne hypotese understøttes af undersøgelsens resultater. Ser man på progressionen over de færdighedsniveauer der testes, ses fra niveau 1 til niveau 3 en jævn fremgang, hvor lørnerne bliver stadig bedre til at genkende de relevante ordstrukturer og dermed applicere de korrekte mønstre. De færdiguddannede opnår imidlertid ikke helt lige så gode færdigheder som de indfødte sprogbrugere, selvom disse også laver fejl, der dog tilsyneladende især skyldes overførsel af dialektale træk.

Tilsammen giver resultaterne et bredt billede af de vanskeligheder, det arabiske skriftsystem udgør for L2-læseren: Test C viser, at lørnerne på de lavere færdighedsniveauer ikke umiddelbart genkender ordenes strukturer, hvorfor de ikke har en chance for at vokalisere korrekt, når de ikke har automatisk ordgenkendelse. Men test B viser, at på trods heraf er vokalisering ikke en hjælp for begynderne. Og test A kan forklare hvorfor: Fordi skriftsystemet lægger en så stor belastning på lørnenes kognitive kapacitet, at der ikke er overskud til at udnytte vokalerne. Hvad L2-læserne mangler er simpelt hen automatisering.

Nøjagtig ligesom det at lære at cykle kræver, at man først lærer at holde balancen og koordinere sine ben – og først når de ting er automatiserede, kan man overskue at holde øje med trafikken. På samme måde må L2-læseren have en hurtig og automatiseret genkendelse af grafemerne og deres fonetiske værdier, før den information, vokalerne repræsenterer, kan udnyttes positivt. Men det betyder ikke nødvendigvis, at vokalisering er irrelevant for begynderlæsere. Med cykel-analogien som eksempel kan man jo heller ikke sige, at færdselsreglerne er irrelevante for den, der begynder at cykle. Manglende kendskab til dem kan tvært imod være fatal. På samme måde vil man – hvis man helt udelader vokalisering for begynderlæsere – risikere at hindre adgangen til nye ord.

Denne problematik og mulighederne for at tilgodese behovet for automatisering bedre end det er tilfældet med de pædagogiske metoder, der anvendes i dag, diskuteres i kapitel 7. Her sættes resultaterne ind i en bredere fremmedsprogspædagogisk ramme, der kan danne grundlag for konkrete anbefalinger for nye tiltag inden for undervisning i læsning på arabisk som fremmedsprog.

7. Konklusioner og pædagogiske perspektiver

Formålet med denne afhandling har været at undersøge, hvilken indflydelse de fremmede elementer i det arabiske skriftsystem har på visuel ordgenkendelse hos danske lørnere. Med udgangspunkt i en tilpasset konnektionistisk ordgenkendelsesmodel koncentrerer afhandlingen sig om primær og sekundær genkendelse som adskilte problemstillinger, hvor 'primær genkendelse' relaterer til bogstavdetektion og altså vedrører den fremmede grafiske dimension i form af det arabiske alfabet samt de fremmede fonemer, nogle af bogstaverne repræsenterer – mens 'sekundær genkendelse' relaterer til den fremmede ortografiske dimension i form af den rodbaserede, konsonantiske morfologi og fraværet af korte vokaler i almindeligt tekstmateriale.

Hvad angår den primære genkendelse, viser forskningsoversigten, at mens læseretningen må regnes for at have marginal betydning, spiller bogstavarkitekturen tilsyneladende en afgørende rolle for læsetilegnelsen, idet tilegnelse af et nyt skriftsystem i voksenalderen efter alt at dømme er en kompleks opgave, der kræver meget træning. De scanningsstrategier, man etablerer i sin tidlige læsetilegnelse, overføres gerne til det nye skriftsystem, selvom det ikke er hensigtsmæssigt, og selv efter lang tids læsetræning tillægger man sig ikke umiddelbart de samme strategier som L1-læsere anvender. For lørnere, hvis L1 skrives med det latinske alfabet, synes det arabiske skriftsystem at være særlig vanskeligt at lære, fordi de arabiske bogstaver er konstrueret med helt anderledes bogstavkomponenter end latinske bogstaver, ligesom de tilsyneladende er objektivt vanskeligere at skelne. Det empiriske studium støtter hypotesen om, at den arabiske skrift – uafhængigt af den sproglige dimension – udgør en barriere for hurtig og korrekt ordgenkendelse på arabisk. Ganske vist forbedres lørnernes færdigheder markant fra niveau 1 (som repræsenterer studerende på 2.semester) til niveau 2 (som repræsenterer studerende på 4. semester), men derefter stagnerer fremgangen i omkodningshastighed, og selv på niveau 3 (som repræsenterer færdiguddannede kandidater) læser de i gennemsnit kun halvt så hurtigt som kontrol-gruppen af indfødte sprogbrugere. De fremmede fonemer har derimod ingen indflydelse på omkodningshastigheden. Hvad angår omkodningsfejl, falder antallet markant fra niveau 1 til 2, og for omkodning uden fremmede fonemer flader udviklingen også ud efter niveau 2, uden dog at gå helt i stå. Lørnere på niveau 3 laver over dobbelt så mange fejl som indfødte sprogbrugere ved omkodning af tekstmateriale, hvori der ikke forekommer tegn, der repræsenterer de fremmede fonemer. Grafemerne i sig selv synes altså at være en afgørende barriere også for korrekt omkodning. De fremmede fonemer giver også anledning til mange omkodningsfejl især hos begynderne, men tilegnelsen af disse fortsætter over alle niveauerne, og lørnere på niveau 3 omkoder bogstaver, der repræsenterer disse fonemer lige så korrekt som indfødte sprogbrugere, som også laver mange af denne slags omkodningsfejl – tilsyneladende især pga. overførsel af dialektal fonologi. Den kvalitative analyse viser, at omkodningsfejl især finder sted mellem bogstaver, som kun adskiller sig ved antallet og placeringen af prikker – hvilket er forventeligt, da det er en bogstavkomponent, lørnerne ikke er vant til at scanne efter. Der sker også mange forvekslinger af bogstaver, der på anden måde deler grafisk udtryk, mens de resterende fejl finder sted inden for fem afgrænsede fonetiske kategorier.

Hvad angår den sekundære genkendelse, viser forskningsresultater vedrørende læsning på arabisk og hebraisk som L1, at vokalisering med diakritiske tegn animerer til fonologiske afkodningsprocesser, mens uvokaliseret skrift pga. den manglende fonologiske information

giver anledning til en mere ortografisk afkodning. Den manglende fonologiske information og den udtalte homografi ved uvokaliseret skrift gør læsning stærkt kontekstafhængig og dermed forholdsvis kognitivt krævende, og vokalisering giver derfor større læsehastighed og bedre læseforståelse. På grund af den morfologiske struktur og den manglende vokalisering i almindelige tekster har man ved læsning på arabisk i det hele taget færre ressourcer at trække på end ved læsning af latinsk skrift. Afkodningsprocessen forløber derfor tilsyneladende meget anderledes. På indoeuropæiske sprog aktiveres processer som tripelbogstavanalyse og nabofrekvenseffekt, der på grund af den konsonantiske morfologi ikke giver mening på arabisk. På arabisk må man formodes til gengæld at udnytte kendskabet til ords rødder og mønstre samt den stramme ordstruktur, der begrænser antallet af mulige vokalkombinationer ved et givent ord. At læsning er stærkt kontekstafhængig og trækker på sprogets morfologiske struktur, må forventes at udgøre en barriere for læsning på arabisk som L2. Dette kan imidlertid ikke entydigt bekræftes af den empiriske undersøgelse. Ganske vist viser resultaterne, at kendskabet til hvilke mønstre, der kan appliceres på hvilke ordstrukturer, øges gennem tilegnelsesforløbet, men vokalisering, som tilfører fuld fonologisk information og udrydder det store antal homografer, har ikke en positiv effekt på hverken læsehastighed eller læseforståelse. Tværtimod er læsehastigheden lavere ved læsning af vokaliseret end ved læsning af uvokaliseret tekst, især hos begynderne, men forskellen svinder med højere færdighedsniveau. Den mest sandsynlige forklaring er, at den ekstra grafiske information repræsenterer så voldsom en kognitiv byrde for det i forvejen hårdt belastede afkodningssystem, at vokalinformationen ikke kan udnyttes, før der er etableret så gode afkodningsfærdigheder, at der er overskud til at levne den opmærksomhed. Hvad angår læseforståelse kan samme tendens spores, men resultaterne er ikke entydige, og forskellene i læseforståelse ved læsning af de to teksttyper er ikke statistisk signifikante. Konklusionen må dog være, at den barriere, den grafiske dimension repræsenterer, overskygger vokalproblematikken hos lørnere i de første år af tilegnelsesforløbet.

For resultaterne af den empiriske undersøgelse må der tages det forbehold, at antallet af respondenter er begrænset, og at endelige konklusioner derfor vil kræve en tilsvarende undersøgelse med en større population. Men konklusionerne på de teoretiske analyser og de empiriske resultater indicerer som nævnt, at det arabiske skriftsystem er en afgørende barriere for tilegnelsen af læsefærdighed på arabisk, og at de undervisningsmetoder, der anvendes i arabiskundervisningen derfor bør målrettes mod de sprogspecifikke aspekter, der gør læsning på arabisk anderledes at lære end læsning på de nære fremmedsprog, der er udgangspunkt for de pædagogiske metoder, som anvendes i dag. I betragtning af, at læsning er en særdeles kompliceret proces, og at børn bruger mange år på at opbygge gode læsekompetencer på deres modersmål, er det nærliggende at antage, at etablering af gode læsefærdigheder på arabisk som fremmedsprog kræver en særlig og målrettet indsats. I de følgende afsnit sættes denne konklusion ind i en praktisk pædagogisk sammenhæng, hvor der tages udgangspunkt i de metoder, der er fremherskende i dag.

7.1 Kommunikativ undervisning i læsefærdighed på arabisk

Ser man på de lærebøger, der er udviklet til undervisning i arabisk, vil det fremgå, at træning i ordgenkendelse i dag ikke er særlig højt prioriteret. I figur 7.1 ses en overordnet oversigt over indholdet af repræsentative kapitler af to meget anvendte lærebogssystemer til undervisning i arabisk som L2. Mens الكتاب الاساسي fra 1992 er ret traditionelt opbygget med læsestykker, skriftlige opgaver og grammatik, er الكتاب في تعلم العربية fra 1995 i højere grad

En typisk lektion i الكتاب الاساسي ، الجزء الثاني – fra 1992 (15 sider):

Læsestykke med farverige illustrationer og mellem 50 og 100 nye ord: 5 sider
Opgaver til at øve det nye ordforråd: 7 sider
Grammatik med øvelser: 3 sider
Læseøvelser – kun få nye ord: **ingen**

En typisk lektion i الكتاب في تعلم العربية ، الجزء الأول – fra 1995 (25 sider):

Gloselister med ca. 40 nye ord: 2 sider
Video/lytteøvelser med spørgsmål til det generelle indhold: 1 side
Opgaver til at øve det nye ordforråd: 2 sider
Diktat: 1 side
Samtaleøvelser: 2 sider
Grammatik med øvelser: 12 sider
Skriveøvelse: 1 side
Læseøvelser – autentiske tekster, ca. 90% nye ord med spørgsmål: 3 sider
Læseøvelse – kun få nye ord: 1 side (**6-10 linjer**)

Figur 7.1 Eksempler på prioritering af læsetræning i to lærebogssystemer til undervisning i arabisk som L2 (بدوي ١٩٩٢; Brustad, 1995 #436)

konstrueret i henhold til moderne kommunikative metoder. Ganske vist fylder grammatiske forklaringer og opgaver forholdsvis meget, men mundtlighed og ikke mindst lyttefærdighed er langt mere i fokus end i bogen fra 1992. Og læseøvelserne er meget udtalt tilrettelagt med forholdsvis meget autentisk tekst. Der er dog også et tekststykke konstrueret til det relevante færdighedsniveau, som giver mulighed for træning i ordgenkendelse, men det fylder uendelig lidt i et stort kapitel, hvor de øvrige færdigheder trænes intensivt.

Kommunikativ sprogundervisning i arabisk er ikke en selvfølge. Den filologiske undervisning er stadig den mest fremherskende i universitære undervisningsmiljøer, men visse steder har man efterhånden fulgt den generelle tendens i fremmedsprogsundervisningen mod at satse på at give lørnerne kommunikative frem for sprogvidenskabelige og litterære kompetencer. Denne udvikling har naturligt nok baseret sig på de teoretiske tendenser, der har præget fremmedsprogsundervisningen generelt. En forenklet historisk oversigt over fremherskende metoder ses i figur 7.2. Den tidligste fremmedsprogsundervisning var – ligesom hovedparten af arabiskundervisningen er det i dag – baseret på oversættelse og grammatik, men omkring 1930 udvikledes både i USA og Europa nye metoder, der især baserede sig på indlæring af det mest anvendelige ordforråd og satte samtalefærdigheder i centrum. I kølvandet på den generative og funktionalistiske lingvistik blev en naturlig tilgang til sproget idealiseret og ekspliciteret i budskabet om, at sprog bør tilegnes gennem naturlig og eksperimenterende interaktion med sproget frem for at indlæres gennem analytiske og bevidste manipulationer af sprogregler i formelle sammenhænge. I praksis har dette skift manifesteret sig som en række forskellige pædagogiske modeller, der hører til under paraplybetegnelsen 'kommunikativ fremmedsprogspædagogik'. Fælles for de kommunikative metoder er, at de ser sprogtilegnelsen som en kreativ og medskabende proces, hvor der er fokus på aktiv og kreativ sprogbrug, og undervisningens indhold defineres i princippet af elevernes individuelle sproglige handlingsbehov (Zimmerman 1997:12-15; Holm & Laursen 2000:85-100; Pedersen 2001). I 1990'erne kom det ensidige fokus på 'naturlig' sprogtilegnelse under pres.

		Metode	Centrale aspekter	Grammatik	Kritikpunkter
historisk lingvistik	19. årh.	Grammatik-oversættelses-metoden	Grammatik Oversættelse Litteraturstudier	Eksplicit	Manglende kommunikative færdigheder
	1874 -	Den direkte metode (Berlitz-metoden)	Dialog Udtale	Induktiv	Manglende grammatik og læsefærdighed
strukturalisme	1930 -	Læsemetoden	Strukturer Ordforråd ud fra frekvens	Systematisk graduering	Overfokusering på ordforråd og graduering
generativisme	1945 -	Den audio-lingvale metode	Strukturer Udtale Automatisering	Eksemplificeret og automatiseret gennem øvelser	Overføring af L1-indlæring på L2-indlæringsstrategi
	1970 -	Funktionel kommunikativ metode	Hverdagsdialog med udgangspunkt i sprogfunktioner	Eksemplificeret og automatiseret gennem øvelser	Manglende kreativ sprogbrug
funktionalisme	1980 -	Praktisk kommunikativ metode	Lørnerautonomi Autentisk/naturlig kommunikation	Induktiv Hypotesedannelse og -afprøvning	Manglende fokus på korrekt sprog
??	1990 -	Nye kommunikative metoder	Tasks og IT-værktøjer Kulturelle aspekter Motivation	?	?

Figur 7.2 Pædagogiske traditioner i fremmedsprogsundervisningen. (Baseret på Zimmerman 1997; Holm 2000:85-100; Pedersen 2001:12-15).

Empiriske studier viste, at studerende der blev eksponeret for meningsfuldt sprog over en længere periode, vedvarende havde problemer med at tilegne sig visse grammatiske strukturer. På den baggrund hævdede en ny generation af forskere, at undervisning, der udelukkende fokuserer på naturlig kommunikation, ikke formår at udvikle tilstrækkeligt kendskab til målsproget, og at der derfor er behov for at integrere sprogstrukturer i den kommunikative undervisning (Robinson 1996; Doughty & Williams 1998). Dette behov ses tydeligt imødekommet i fx الكتاب في تعلّم العربية (1995), der i dag synes at være det mest udbredte lærebogssystem til undervisning i kommunikativt arabisk.

Specifikt i relation til undervisning i læsefærdighed har udviklingen i den fremmedsprogs-pædagogiske disciplin medført store forandringer. Idealiseringen af 'naturlig sprogbrug' har medført, at færdigheder i at håndtere autentisk tekst er i fokus. Og mens der inden for L1-læseforskningen er etableret en generel erkendelse af, at evne til at afkode er essentiel for at kunne udvikle effektiv læseforståelse, har ordgenkendelsesprocesser ikke fået nogen særlig opmærksomhed i fremmedsprogspædagogikken. Derimod har L2-læsning først og fremmest drejet sig om at give lørnerne bevidste strategier til at opnå maksimal forståelse af indhold på baggrund af tekst, de ikke har tilstrækkelige sproglige forudsætninger for at kunne forstå det fulde indhold af. Dermed er topstyrede læsestrategier blevet det centrale omdrejningspunkt for L2-læseundervisningen, og med hensyn til materialer prioriteres autencitet meget højt. Som Day & Bamford udtrykker det:

"To say that simplified texts have a bad name in language teaching is an understatement. If authentic texts are seen as natural, interesting, relevant, and

> *pedagogical sound, texts-made-simple are generally considered to be just the opposite: stilted, unnatural, unreal, bland, and a pedagogical dead end."* (Day & Bamford 1998:56)

Universitetsmiljøer, hvor der undervises i kommunikativt arabisk, er ingen undtagelse. Inden for disse miljøer betragtes opgøret med den filologiske tilgang som en meget værdifuld nyskabelse, der stadig er i sin vorden; og fordi de kommunikative principper langt fra er en selvfølge, er der så meget desto mere fokus på dem – og dermed på brugen af autentisk materiale – hos fortalerne for brug af moderne undervisningsmetoder (se fx راموني ١٩٩٤; Al-Batal 1995; Alosh 1995; Hammoud 1996; Khaldieh 1999). I relation til læseundervisning skriver Khaldieh fx:

> *"Authentic texts have the characteristic of communicating meaning rather than teaching language features. They invoke in learners types of cognitive strategies that are geared toward meaning-getting processing. This issue is in contradiction with what is believed in FL [foreign language] in general and AFL [Arabic as a foreign language] in particular"* (Khaldieh 1999:146)

> *"Development of instructional materials for reading may deprive learners of the opportunity to apply cognitive strategies in full; a challenge necessary for these learners to develop strategic actions where they could hypothesize, plan, evaluate, and confirm their meaning."* (Khaldieh 1999:146)

Også ved Syddansk Universitet, som har været frontløber for kommunikativ undervisning i arabisk i Danmark, har autencitet været et afgørende indsatsområde. Dette fremgår eksempelvis af Helle Lykke Nielsens beskrivelse af den metode, der ligger til grund for Cand. negot.-studiet i arabisk:

> *"(...) as much teaching material as possible should be based on authentic sources such as newspapers, magazines, television etc. By authentic material we mean material which is made by Arabs for Arabs"* (Nielsen 1993:114)

At autentiske tekster kan være vanskelige at forstå for lørnere, der har begrænsede sprogfærdigheder, imødekommes af den kommunikative tilgang med topstyrede læsestrategier, hvor lørnerne trænes i at angribe teksten *top-down* og få det ud af den, de nu kan: Hvad handler teksten overhovedet om? Hvad ved jeg om emnet i forvejen? Er der illustrationer eller andet, der kan give nogle hints? (Janzen & Stoller 1998; Farell 2001). Denne fremgangsmåde blev en naturligt integreret del af den kommunikative tilgang i arabiskundervisningen. Og i bestræbelserne på lægge distance til den traditionelle tekstnære undervisning, blev *bottom-up* læsning nærmest kørt ud på et sidespor. I samtlige forsknings-publikationer vedrørende læsning på arabisk som L2, hvor der gives pædagogiske anbefalinger, argumenteres der ensidigt for udbredelse af *top-down* læsestrategier:

> *"(...) one useful and practical instructional strategy would be to enable readers to develop some understanding of the text by drawing on their life-long experiences, beliefs and knowledge. (...) Another important consideration in preparing students for reading is to familiarize them with different text genres."* (Aweiss 1993:13)

> *"We can utilize the differences in students' perception of the different tasks in teaching. For example, top-down strategies can be taught to establish the context..."* (Fragman 1999:80)

Khaldieh (1996) påpeger dog i sin artikel om ordgenkendelse hos amerikanske lørnere, at der skal arbejdes med både *top-down* og *bottom-up* strategier (Khaldieh 1996:146). Men i sin senere artikel fra 1999 skriver han:

> *"(...) learners should be familiarized with text genres and should be encouraged to consciously read at the discourse level rather than on the word and sentence level.. Most important would be enriching texts with background knowledge, knowledge about strategies and their utility, guidance to specific strategies and information to be used in solving a comprehension problem."* (Khaldieh 1999:148)

I sin seneste artikel fra 2001, hvori Khaldieh konkluderer, at ordforråd er afgørende for læseforståelse, berører han kort problematikken med dårlig ordgenkendelse:

> *"The mastery and control of vocabulary should stress visual processes, word form and morphology, and noun classes that are important to word identification at the intermediate level of reading proficiency"* (Khaldieh 2001:426)

Men på det konkrete niveau berører han i den forbindelse kun instruktion i rod-mønster systemet som et pædagogisk indsatsområde. *Top-down* strategierne bliver derimod ikke forbigået:

> *"The issue of learners' background knowledge of Arabic culture was not addressed in this study. However, prediscussion activities about cultural differences are central to helping students understand the text fully. In addition, students need to understand features of style and rhetoric so that they will be able to make connections among the several parts of the text on the paragraph level rather than just locally on the sentence level"* (Khaldieh 2001:427)

Denne prioritering af *top-down*-læsning afspejles blandt andet i fx الكتاب في تعلّم العربية (1995), jvf. figur 7.1. I lyset af konklusionerne på denne afhandling synes der imidlertid at være grund til at sætte større fokus på skriftsystemet som et selvstændigt pædagogisk problemområde i arabiskundervisningen. I det følgende afsnit diskuteres nogle muligheder for, hvordan dette kan gøres i praksis.

7.2 Pædagogiske overvejelser og anbefalinger

De primære begrundelser for at anvende autentisk materiale og træne *top-down* læsestrategier er følgende:

- Hvis lørnere forventes at skulle kunne håndtere målsproget i en autentisk sammenhæng, må den bedste måde at forberede dem på være at eksponere dem for autentisk sprog under deres uddannelse (Williams 1984:25) – dvs. målet er at efterkomme et kommunikativt behov.
- At autentiske materialer forventes at være interessante, engagerende og kulturelt oplysende for lørnerne (Nuttall 1996) – dvs. man går ud fra, at de fremmer motivationen for at læse.

Fra en kommunikativ synsvinkel er det selvfølgelig fornuftigt, at lørnerne tilegner sig strategier til at tackle tekster, de ikke umiddelbart forstår. Og med udgangspunkt i den teoretiske ramme, der defineres i kapitel 2, er en naturlig, autentisk proces også en oplagt indgang til tilegnelse fra et konnektionistisk perspektiv. Men udvikling af det

konnektionistiske netværk fordrer en løbende oplæring, og eksponering for autentisk materiale fører ikke nødvendigvis til en progression, der svarer til den naturlige, kognitive sprogtilegnelsesproces. Tekster, der er for svære til, at lørneren i praksis kan læse og forstå dem, muliggør ikke den nødvendige opøvelse af automatisk ordgenkendelse, som er så essentiel for etablering af gode læsefærdigheder. Og autentiske tekster er ofte alt for svære for L2-lørnere på begynder- og mellemniveau.

Samtidig er det oplagt, at brug af autentiske tekster og et ensidigt fokus på opøvelse af *top-down* læsestrategier i sig selv kan virke demotiverende for nogle lørnere, fordi tekst-materialet næsten altid vil opleves som svært, og fordi de i næsten alle læsesituationer bliver konfronteret med deres manglende sproglige kompetencer. Og motivation for at læse på arabisk må anses for at være en afgørende faktor, fordi det er så væsentligt, at lørnerne læser meget, hvis de skal blive gode læsere. Dermed er der tilsyneladende i L2-læse-undervisningen i arabisk behov for at sætte større fokus på at styrke lørnernes motivation for at læse og at træne deres ordgenkendelse.

I relation til L2-læsning er motivationsaspektet ikke særlig velbelyst. Der findes to publicerede forsøg på at konkretisere, hvad der påvirker motivation til at læse på et L2. De bygger begge på den generelle *expectancy-value* motivationsteori (Wigfield 1994; Dörnyei 2001), men med meget forskelligt resultat. Day & Bamford (1998) kommer frem til en liste over motivationselementer, der i høj grad vedrører eksterne faktorer, mens Mori (2002) i stedet peger på personlige faktorer hos lørneren. En oversigt over de faktorer, der opereres med i de to modeller ses i figur 7.3 og 7.4. Som det fremgår kan en del af de nævnte faktorer ikke umiddelbart påvirkes i undervisningssituationen – fx lørnernes behov for målopfyldelse, deres tidligere erfaringer med læsning og deres sociokulturelle miljø. Andre faktorer er mulige at påvirke, men er meget generelle og vedrører al L2-undervisning – fx undervisningsmiljø, holdninger til L2 og den fremmede kultur, samt lørnernes iver efter at opnå færdigheder og gode karakterer. Endelig er der de faktorer, der knytter sig direkte til læsning. Disse kan påvirkes ved at:

- Lade lørnerne læse meget
- Give lørnerne gode læseerfaringer
- Anvende tekster, der er lette at forstå for lørnere på det pågældende niveau
- Anvende materialer, der er interessante, spændende og attraktive i layout
- Give lørnerne let adgang til læsemateriale

Som det ses er de fem anbefalinger tæt forbundne: Anvender man lette og interessante tekster, vil det ofte give lørnerne gode læseerfaringer, og er læsemateriale samtidig let tilgængeligt, vil det tilsammen gøre lørnerne mere tilbøjelige til at læse meget. Og hvis de læser meget, bliver de bedre læsere, hvorved de får lettere ved at få gode læseerfaringer. Det synes altså helt afgørende for lørnernes motivation for at læse, at materialerne er interessante og lette at læse – men det kan i mange tilfælde være vanskeligt at tilgodese begge behov på samme tid.

Figur 7.3 Elementer, der er afgørende for motivation til at læse på L2, efter Day & Bamford (1998)

Figur 7.4 Elementer, der er afgørende for motivation til at læse på L2, efter Mori (2002)

Som beskrevet ovenfor, fokuserer man i den kommunikative L2-læseteori overvejende på, at materialerne skal være interessante. På baggrund af det teoretiske og empiriske materiale i denne afhandling kan man dog argumentere for, at behovet for letlæst materiale bør have langt større prioritet i læseundervisningen på arabisk som L2, hvis hensynet til etablering af ordgenkendelse – og til motivation – skal tilgodeses. Lørnere, der skal lære at læse et nyt skriftsystem, er nemlig ekstra sårbare over for vanskelige tekster, fordi de, som beskrevet i kapitel 4, er tvunget til at læse *bottom-up* – de har simpelt hen ikke kognitivt overskud til andet – de kan ikke overskue teksten, skimme udvalgte passager og hurtigt danne sig et overblik over indholdet. Og hvis de skal opnå gode læsefærdigheder, skal automatisk ordgenkendelse opøves – gennem mange timers læsetræning, hvor læseprocessen ikke obstrueres af, at tekstmaterialet er for svært.

Derfor synes det problematisk, at der i L2-læseforskningen er en veletableret forestilling om, at *top-down* læsestrategier kan kompensere for ringe afkodningsfærdigheder, jvf. afsnit 2.4. Dette gælder ikke kun for undervisning i arabisk, men er et generelt fænomen inden for L2-undervisning. Et eksempel er den pædagogiske vejledning for andetsprogs-undervisningen i dansk for udlændinge – lørnere der også ofte møder op med et andet skriftsystem i bagagen:

> *"For kursister, som ikke har automatiseret den systemiske viden [dvs. læsehierarkiets lavere niveauer], vil det sandsynligvis være en fordel at læse tekster, hvor de kan bruge deres viden om verden, den skematiske viden [dvs. læsehierarkiets højere niveauer]. Det vil være lettere for dem at læse tekster, hvor de kan danne sig en mening ved at trække på deres baggrundsviden, hvor de kan læse top-town eller topstyret."* (Andersen & Hermansen 2000:6)

Endnu mere problematisk bliver det, når lørnernes naturlige fokus på det tekstuelle niveau problematiseres som værende 'forkert' eller 'uhensigtsmæssigt'. Khaldieh, der som vist i kapitel 1 er en af de ganske få, der beskæftiger sig forskningsmæssigt med læsning på arabisk som L2, har eksempelvis følgende betragtninger om L2-lørnernes læsning:

> *"[The less proficient] readers seem to have been focusing on reading locally at the sentence level. They were very occupied trying to construct the meaning of the sentence or clause at hand; their deficiency in language ability seems to have played a major role in emphasis on the local point rather than trying to connect it with the general framework of ideas of the text."* (Khaldieh 1999:127)

> *"The less proficient non-native readers […] seemed to have been involved with a bottom-up model of reading with little consideration for evaluation of meaning generated. Therefore, evaluation of meaning as it develops during reading should be emphasized. Learners should be made aware of such an important phase in reading, and be encouraged to practice. They should be trained to focus on the identification of the source of the problem to foster their awareness towards problem recognition."* (Khaldieh 1999:145)

Khaldieh påpeger altså, at de svage læseres vanskeligheder bunder i mangelfulde læse-strategier. De skal blive bedre til at læse *top-down*. Og det er muligt, at intensiv træning i at gætte og aktivere baggrundsviden kan give læserne bedre forståelse af svære tekster. I lyset af det teoretiske materiale, der er præsenteret i denne afhandling, virker det dog umiddelbart problematisk, at man udelukkende søger at instruere lørnere i at kompensere for deres manglende afkodningsfærdigheder og undlader at gøre en målrettet indsats for at optræne

samme færdigheder. – Som nævnt ovenfor fokuserer alle pædagogiske anbefalinger i den forskningsbaserede litteratur på *top-down* strategier: Ikke i en eneste af publikationerne gives der nogen konkrete ideer til, hvordan lørnere kan trænes i at styrke de visuelle processer i forbindelse med læsning. Og som vist i figur 7.1, er optræning af ordgenkendelse på det nye skriftsystem heller ikke prioriteret i den pædagogiske fremgangsmåde, der praktiseres i de mest anvendte lærebøger. Materialerne synes næsten at afspejle, at skriftsystemet ikke opfattes som et selvstændigt problemområde. Det meget intensive fokus på opøvelse af *top-down* læsestrategier i الكتاب في تعلّم العربية tyder på, at forfatterne går ud fra, at det indledende forløb, hvor skriftsystemet gennemgås – i dette lærebogssystem på 10 lektioner (Brustad et al. 1995) – udstyrer lørnerne med tilstrækkelige kvalifikationer til at håndtere arabisk skrift. Denne afhandlings teoretiske kapitler samt den empiriske undersøgelse indicerer, at dette langt fra er tilfældet. Tværtimod synes det at stå klart, at hurtig afkodning af skrift kræver etablering af komplekse, automatiserede processer, og at det nye skriftsystem i sig selv er en afgørende barriere for at opnå hurtig ordgenkendelse og dermed gode læsefærdigheder. Lørnere vil derfor have fordel af at læse lettere tekster, der kan hjælpe dem til at opøve den automatisering, de mangler. Meget svære tekster tvinger nemlig netop lørneren til at fokusere på koden, konsultere ordbogen eller give op, fordi læseprocessen bryder sammen. Faktisk er det sandsynligt, at tekster, der sprogligt er for svære i forhold til lørnerens niveau, ikke træner nogen former for læsekompetence overhovedet (Nuttall 1996:177). Værdifuld undervisningstid, der kunne være anvendt til at træne lørnerens afkodningsfærdigheder, går til spilde, og måske endnu værre: Lørneren kan få en oplevelse af nederlag og af, at fremmedsproget bare er for svært, hvilket som ovenfor anført kan skade motivationen for at læse i det hele taget.

Derfor synes der at være grund til igen at lægge mere vægt på *bottom-up*-læsning i arabiskundervisningen. – Ikke for at vende tilbage til arkaiske undervisningsformer, men for – i en tidssvarende udformning – at imødekomme behovet for træning i ordgenkendelse. Imidlertid er der ingen grund til at gøre spørgsmålet om den optimale undervisning i læsefærdighed til et valg mellem enten *top-down* eller *bottom-up* læsning. Kritikken af den kommunikative L2-undervisnings ensidige fokus på *top-down* strategier gør jo ikke de hensyn, metoden søger at imødekomme, irrelevante. Samtidig efterlader en diskussion mellem de to nærmest modsatrettede tilgange til tekst et stort hul 'i midten' af læsehierarkiet - og her er ordforrådet helt centralt. Som det fremhæves af både det teoretiske og det empiriske materiale i denne afhandling, er ordforråd et helt afgørende element i læseprocessen – uanset om man anskuer den *bottom-up* eller *top-down*. I ordgenkendelses-processen er ordforråd essentielt: Hvis et ord ikke kendes, kan det selvsagt heller ikke genkendes. Og ved applicering af læsestrategier er ordforråd også helt centralt for succes, for hvis ikke man forstår en vis portion af en tekst, er der ingen kontekst at gætte ud fra:

> *" Whatever the effect of reading strategies is, it is short-circuited if the vocabulary is below the threshold, i.e., below the minimum of 3,000 word families, or 5,000 lexical items."* (Laufer 1997:31)
>
> *"The major problem with using contextual cues for word identification is that they are unreliable. And this is all the more true since it is the less frequent words that contribute the most meaning to the passage."* (Adams 1990:217)

Ligesom lørnerne i denne undersøgelse udpeger manglende ordforråd som deres primære svaghed i forbindelse med læsning på arabisk, viser flere undersøgelser, der vedrører andre målsprog, at ordforråd spiller den mest centrale rolle for læseprocessen:

> *"It has been consistently demonstrated that reading comprehension is strongly related to vocabulary knowledge, more strongly than to the other components of reading"* (Laufer 1997:20)

Der er da heller ingen tvivl om, at udvidelse af ordforråd på L2 er et bevidst og konkret mål for velsagtens al L2-undervisning. Men i selve undervisningsforløbet behandles ordforråd gerne som en implicit faktor (Coady 1997) – man går simpelt hen ud fra, at det kommer af sig selv, når man lytter, taler, læser og skriver på målsproget, på samme måde som da man lærte sit modersmål. Især læsning opfattes normalt som en vigtig kilde til udvidelse af ordforrådet på L2 (Coady 1997:274; Waring & Takaki 2003), og flere studier viser da også, at man samler en del nye ord op i forbindelse med læsning (Day et al. 1991; Dupuy & Krashen 1993; Grabe & Stoller 1997). Megen læsning fører altså til større ordforråd, mens et større ordforråd fører til bedre (og dermed normalt også mere) læsning. Og som beskrevet i kapitel 5 er afkodning på arabisk på grund af de manglende vokaler stærkt kontekstafhængig, men lørnere på de lavere færdighedsniveauer er endnu for svage læsere til at støtte sig til kontekst, når de læser. Samtidig viser den empiriske undersøgelse, at vokalisering er en belastning frem for en støtte – igen især for læsere på de lavere færdighedsniveauer. Men samtidig er vokalerne essentielle for etableringen af nyt ordforråd. Hvis ikke ordet har en konkret og fuldstændig fonologisk repræsentation, er det svært at fastholde i hukommelsen og kan ikke udnyttes mundtligt.

Vi står altså over for en række mere eller mindre paradoksale forhold, der er tæt indbyrdes forbundne. Spørgsmålet er, om man kan iværksætte strategier, der kan påvirke disse sammenhænge positivt og give resultater på alle de forskellige niveauer, der er involveret i forbindelse med læsning – og altså sikre, at sammenhængene udvikler sig til en god frem for en ond cirkel. 'Paradokserne' kan kort opremses som følger:

- Man bliver bedre til at læse arabisk ved at læse meget. Men inden man er god til at læse, er man ikke så tilbøjelig til at gøre det, fordi det er svært.
- Man lærer nye arabiske ord ved at læse på arabisk. Men inden man kan mange ord er man ikke så tilbøjelig til at gøre det, fordi det er svært.
- Man har behov for at støtte sig til kontekst, når man læser på arabisk. Men inden man er blevet en god læser, er det vanskeligt at overskue kontekst.
- Man har brug for vokaler for at kunne identificere kendte ord og lære nye. Men vokalisering udgør en kognitiv belastning, der svækker læseprocessen.
- Man har brug for at læse autentiske tekster for at synes, at det er interessant at læse på arabisk, men autentiske tekster på arabisk er ofte så svære, at det alligevel ikke er interessant.

Umiddelbart kan det på baggrund af ovenstående opstilling synes umuligt at angive retningslinjer for, hvordan man undgår at spolere lørnernes læseglæde og læsetilegnelse. Men ud fra betragtningen om, at lørnernes motivation for at blive gode læsere er drivkraften til at sætte gang i en god cirkel, der udvikler lørneren mod bedre færdigheder, kan man påpege nogle indsatsområder, der sandsynligvis kan vise sig værdifulde:

For det første er der ingen tvivl om, at lørnerne skal læse meget – først og fremmest fordi det er forudsætningen for etablering af automatisk ordgenkendelse, som er forudsætningen for god læsning. Den primære hindring for at læse meget er, at det er svært. Svaret herpå er at læse en masse, der er meget let. På de lavere færdighedsniveauer er lette tekster simpelt

hen forudsætningen for, at det der sker mellem lørneren og teksten, faktisk kan kaldes læsning. Day & Bamford (1998:16) foreslår en reformulering af Stephen Krashens formel for input i L2-indlæringen *i+1*[1] til *i–1*. Når det gælder L2-læsning må hensynet til indlæring af nye sproglige elementer vige til fordel for hensynet til at opøve automatisk ordgenkendelse, og input bør derfor være en lille smule lettere, end hvad det er muligt for lørneren at forstå på det aktuelle færdighedsniveau. Dette princip er så meget desto mere relevant, når L2 er arabisk eller et andet sprog, der skrives med et anderledes skriftsystem end L1. At lette tekster kan være vanskelige at konstruere, så de samtidig er interessante, må opvejes af andre faktorer i motivationsmodellerne – fx det konkrete udbytte ved at optræne sine færdigheder og de gode læseoplevelser, der ligger i at forstå en tekst på et sprog, som skrives med tegn, som de fleste i ens omgangskreds betragter som kruseduller.

For det andet skal der satses meget på udvikling af ordforråd. Men ovenstående punkter indicerer, at etableringen af ordforråd ikke skal kobles sammen med læsetræningen – teksterne bør som nævnt være lette og dermed netop ikke indeholde nyt ordforråd. Flere teoretikere har da også peget på, at eksplicit undervisning i ordforråd og indlæring af konkrete strategier til tilegnelse af ordforråd bør prioriteres i L2-undervisningen (Coady 1997; Zimmerman 1997). Dette er givetvis kontroversielt i miljøer, hvor man satser på kommunikativ undervisning, men flere empiriske studier støtter faktisk denne ide. Fx har det vist sig, at de fleste af de ord, man lærer, når man læser, glemmer man hurtigt igen (Waring & Takaki 2003), og komparative studier af indlæring af ordforråd udelukkende som en implicit faktor ved læsning sammenlignet med udbyttet af at supplere læsningen med eksplicit undervisning i ordforråd viser, at det direkte fokus på ordene, deres betydning og deres brug giver langt bedre og mere varige resultater (Paribakht & Wesche 1997).

For det tredje skal man konfronteres med ord både med og uden vokaler. Og mange lærebogssystemer anvender da også en form for *ad hoc* vokalisering, hvor nye ord vokaliseres et begrænset antal gange, hvorefter det forventes, at lørneren kender dem. Dette praktiseres eksempelvis i de lærebogssystemer, der er præsenteret i figur 7.1. I begge systemer virker vokaliseringsstrategien velovervejet: I الكتاب الاساسي (1992) præsenteres alle tekststykker vokaliseret, men her er tekststykkerne også indgangen til de nye ord, som jo er vanskelige at lære uden vokalisering. De tilhørende opgaver er til gengæld uvokaliserede eller delvist vokaliserede. I الكتاب في تعلّم العربية (1995) er ordlisterne, der indleder hvert kapitel vokaliserede, hvorved også ordenes udtale præsenteres, men al sammenhængende tekst er uvokaliseret – sandsynligvis fordi man vil undgå at gøre lørneren afhængig af vokalisering og for at give materialet et så autentisk præg som muligt. I lyset af resultaterne af denne afhandlings empiriske undersøgelser synes sidstnævnte strategi at være fornuftig: Vokaliseringen knytter sig til indlæring af ord og ikke til læsning, hvor den ser ud til at udgøre en belastning frem for en hjælp. Så længe ordet ikke genkendes automatisk, vil læseren ganske vist være tvunget til at afkode fonologisk, og automatisk ordgenkendelse kræver, at man har set ordet mange gange. Men fonologisk afkodning forudsætter ikke nødvendigvis fuld vokalisering – konsonanterne indeholder jo fonologisk information, der kan være tilstrækkelig til at lede læseren på rette spor, hvis ordet er kendt.

---

[1] Krashens hypotese er, at for at ny tilegnelse kan finde sted, må input være 'en lille smule mere' end *i*, dvs. *i+1*, hvor *i* er det, der mestres på det aktuelle færdighedsniveau (Krashen 1985).

Tilsammen tyder dette på, at det er en fornuftig strategi at være meget bevidst om, præcis hvilke færdigheder, man ønsker at fremme med et givent materiale og at træne ordgenkendelse og ordforråd separat så vidt muligt. Decideret læsetræning kan dermed praktiseres med tekster, hvor stort set alt ordforråd er kendt – dermed vil lørnerne sandsynligvis i mindre grad opleve de manglende vokaler som et problem, teksterne vil være lette at forstå, og lørnerne vil være mere tilbøjelige til at læse mere – hvorved de træner deres ordgenkendelse og bliver bedre læsere.

Endvidere kan der gøres meget for at forene principperne om relevans/autencitet og simplicitet. At finde autentiske tekster, der samtidig er lette, kan være vanskeligt, men autencitet og simplicitet behøver ikke nødvendigvis at udelukke hinanden:

> *"For purposes of the foreign language classroom, an authentic text (...) is one whose primary interest is to communicate meaning. In other words, such a text can be (...) intended for a language learner group. The relevant consideration here is not for whom it is written but that there has been an authentic communicative objective in mind."* (Swaffar 1985)

Ser man på resultaterne af den empiriske undersøgelse, synes denne pointe ikke mindst at være relevant i forhold til arabiskstuderende på de højere færdighedsniveauer. Undersøgelsen viser således, at lørnernes fremskridt i omkodningshastighed stagnerer efter niveau 2, på trods af, at de mellem niveau 2 og 3 eksponeres for et meget stort omfang af materiale, de skal læse på egen hånd, idet de på overbygningen følger enten samfundsvidenskabelige fag, hvor arabisk er undervisningsmedium (SDU) eller studier i arabisk litteratur (KU). I begge tilfælde er tekstmaterialerne ofte autentiske, og fokus er på det faglige indhold frem for sprogtilegnelse. Selvom populationen er for lille til at drage endelige konklusioner, kunne dette resultat af undersøgelsen tyde på, at lørnere på overbygningsuddannelserne fortsat har behov for læsetræning, der fokuserer på afkodning og ordgenkendelse, og at dette behov ikke tilstrækkeligt imødekommes. Ud fra den betragtning kunne man på overbygningsuddannelserne søge at introducere undervisnings-materialer, der formidler det samme stof som de relevante autentiske materialer, men som er sprogligt tilpassede lørnernes niveau. Inden for mange andre fremmedsprog har man udviklet letlæste versioner af den vigtigste litteratur, der findes på målsproget, og også samfundsvidenskabelige emner kunne udmærket formidles med materialer udviklet til det niveau, de pågældende studerende befinder sig på. Sådanne materialer ville sandsynligvis øge læselysten hos de studerende, hvorved de ville læse mere. Suppleret med en selvstændig satsning på udvikling af et målrettet ordforråd ville det støtte lørnernes videre sproglige udvikling – ikke mindst i relation til læsning – og undervisningstiden ville sandsynligvis være bedre udnyttet.

Et andet muligt indsatsområde, som også er i konflikt med den kommunikative undervisningsform, er eksplicit undervisning i kendskab til de morfologiske strukturer i arabisk. Nogle steder er denne form for undervisning ugleset, måske ikke mindst som en reaktion på, at den er så fremherskende i den filologiske tilgang. Men i det ovenfor beskrevne lærebogsmateriale الكتاب في تعلّم العربية (1995) – som muligvis er det mest kommunikativt anlagte af de systemer, der er i udbredt anvendelse – lægges der faktisk megen vægt på dette aspekt. Og måske med rette. I hvert fald viser empiriske studier fra andre fjerne fremmedsprog, at målrettet undervisning i målsprogets ortografiske strukturer giver særdeles gode resultater i læsefærdighed hos voksne lørnere (Koda 1997). Af den empiriske undersøgelse i denne afhandling fremgår det, at lørnerne gennem deres studieforløb bliver stadig bedre til at anvende disse strukturer i praksis. Dette kan afspejle,

at studerende ved SDU anvender ovennævnte materiale, hvor emnet behandles, ligesom studerende ved KU i de filologiske fag modtager målrettet undervisning i sprogets morfologiske strukturer. Undersøgelsen viser dog også, at lørnere på niveau 3 ikke mestrer mønstrene helt så godt som de indfødte sprogbrugere. Der findes ikke noget empirisk belæg – heller ikke i denne afhandling – for at denne form for undervisning har effekt på lørnernes læsefærdighed på arabisk, men på baggrund af det teoretiske materiale, der er præsenteret i kapitel 5, er der grund til at formode, at det faktisk forholder sig sådan.

Andre mulige tiltag, som på baggrund af afhandlingens konklusioner kan formodes at være virksomme, er øvelser, der målrettet træner lørnerne i grafem-fonem-forbindelser. Dette praktiseres i dag udelukkende i indledende forløb. På baggrund af resultaterne af den kvalitative undersøgelse i test A er det oplagt at foreslå at disse øvelser i højere grad rettes specifikt mod de fem fonetiske kategorier, inden for hvilke lørnerne begår meget hyppige fejl. På begynder- og mellemniveau kunne man endvidere forsøge sig med samme former for læsetræningsøvelser, som på L1 især bruges til børn med læsevanskeligheder og -handicaps. Som et eksempel kan nævnes *Repeated Readings* (RR), hvor lørnerne læser en tekst igen og igen, indtil de kan læse den med lethed (uassisteret RR), eller hvor lørneren assisteres ved samtidig at lytte til en oplæsning af teksten på lydbånd (assisteret RR). Metoderne har vist moderat positive resultater i læsehastighed på engelsk som L2 hos japanske lørnere på de lavere færdighedsniveauer (Taguchi 1997; Taguchi & Gorsuch 2002), men disse undersøgelser har alvorlige begrænsninger både i relation til metodevalg og til populationernes størrelse. I betragtning af, at der er overvældende dokumentation for metodernes effekt på L1 (Dowhower 1994), er der dog grund til at tage dem i betragtning.

Endelig skal det ikke fornægtes, at *top-down* læsestrategier er gode at mestre. Det er selvfølgelig fornuftigt at instruere studerende i, hvordan de strategisk kan angribe tekster, de ikke umiddelbart forstår, og fravriste teksterne så meget mening som muligt – uden at give op, fordi de ikke er i stand til at læse dem fra ende til anden. Den form for læsetræning er altså stadig relevant. Men som vist, er der mange indikationer på, at den bør have en mindre central plads, end den har i den kommunikative læseundervisning i arabisk i dag. Læsetræningens primære mål må være at lære lørnerne at læse på arabisk. – At lære at kompensere for manglende færdigheder bør være et element – men et sekundært element – i forhold til hensynet til at lørnerne etablerer god fonem-grafem-korrespondance, opøver automatisk ordgenkendelse og får oplevelsen af at kunne forstå arabiske tekster, så de får lyst til at læse mere.

En konkret forhindring for at gennemføre den anbefalede revision af læseundervisningen i arabisk er desværre, at der næsten intet udbud er af letlæste tekster. Mens man på dansk og formentlig også alle andre vesteuropæiske sprog har et forholdsvis stort udvalg af lette tekstmaterialer til både børn og voksne om emner, der formodes at interessere forskellige aldersklasser, er udvalget på arabisk særdeles begrænset. Lette tekster til voksne findes stort set kun i de lærebogsmaterialer, der er udviklet til undervisning i arabisk som L2. Dem er der kun få af, og som det fremgår ovenfor, er mængden af brugbar tekst i de materialer der findes, meget begrænset. Børnebøger er normalt ikke egnede til voksne læsere, for indholdet interesserer dem sjældent, og ordforrådet er ikke det samme som L2-lørnere har og/eller er ikke tilstrækkeligt forsimplet. Desuden er udvalget af arabiske børnebøger i sig selv ikke særlig stort. Så hvis arabiskstuderende skal have let adgang til et stort udvalg af letlæste tekster om relevante emner på forskellige færdighedsniveauer, skal der udvikles nye materialer. I arabiskundervisningen på SDU – hvor man som supplement til lærebogs-

materialerne allerede har igangsat undervisningsaktiviteter, der i højere grad tilgodeser de behov for fornyelse, der i denne afhandling peges på – imødekommes materialebehovet ved at lade lørnerne skrive til hinanden og lade deres tekstkreationer indgå i undervisningen. Hermed bibeholdes autenciteten, idet teksterne har et formidlende formål, emnerne har indlysende relevans for de studerende, og det sproglige niveau er givet af deres egne kompetencer. Sådanne tiltag løser mange problemer, men er ikke tilstrækkelige, fordi de er for lokalt forankrede. Hvis lørnernes behov skal tilgodeses, må letlæste noveller, romaner, landebeskrivelser, samfundsvidenskabelige analyser og andre former for relevant litteratur være tilgængelige. Første skridt må imidlertid være, at der i undervisningsmiljøet opstår en bred erkendelse af, at behovet eksisterer – erfaringen er jo, at når et behov for undervisningsmateriale er erkendt, sørger engagerede undervisere og tekstforfattere gerne for at efterkomme det.

Litteraturliste

Abu-Rabia, S. (1996) The role of vowels and context in reading of highly skilled native Arabic readers. *Journal of Psycholinguistic Research.* 25

Abu-Rabia, S. (1997a) The need for cross-cultural considerrations in reading theory: The effects of Arabic sentence context in skilled and poor readers. *Journal of Research in Reading.* 20:2.

Abu-Rabia, S. (1997b) Reading in Arabic orthography: The effect of vowels and context in reading accuracy of poor and skilled native Arabic readers. *Reading and Writing: An Interdisciplinary Journal.* 9

Abu-Rabia, S. (1998) Reading Arabic texts: Effects of text type, reader type and vowelization. *Reading and Writing: An Interdisciplinary Journal.* 10

Abu-Rabia, S. (1999) The effect of Arabic vowels on the reading comprehension of second- and sixth-grade native Arab children. *Journal of Psycholinguistic Research.* 28:1.

Abu-Rabia, S. (2000) Effects of exposure to literary Arabic on reading comprehension in a doglossic situation. *Reading and Writing: An Interdisciplinary Journal.* 13

Abu-Rabia, S. (2001) The role of vowels in reading semitic scripts: Data from Arabic and Hebrew. *Reading and Writing: An Interdisciplinary Journal.* 14

Abu-Rabia, S. (2002) Reading in a root-based morphology language: The case of Arabic. *Journal of Research in Reading.* 25:3.

Abu-Rabia, S. & J. Awwad (2004) Morphological structures in visual word recognition: The case of Arabic. *in press.*

Abu-Rabia, S., D. Share & M. S. Mansour (2003) Word recognition and basic cognitive processes among reading-disabled and normal readers of Arabic. *Reading and Writing: An Interdisciplinary Journal.* 16

Abu-Rabia, S. & L. S. Siegel (1995) Different orthographies different context effects: The effect of Arabic sentence context in skilled and poor readers. *Reading Psychology: An International Quarterly.* 16

Abu-Rabia, S. & L. S. Siegel (2002) Reading, syntactic, orthographic, and working memory skills in bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research.* 31:6.

Abu-Rabia, S. & L. S. Siegel (2003) Reading skills in three different orthographies: The case of trilingual Arabic-Hebrew-English-speaking children. *Reading and Writing: An Interdisciplinary Journal.* 16

Abu-Rabia, S. & H. Taha (2004) Reading and spelling error analysis of native Arabic dyslexic readers. *in press.*

Adams, M. J. (1990) *Beginning to read. Thinking and learning about print*. Cambridge. MIT Press.

Ager, S. (2001) Writing systems. *http://www.omniglot.com/writing/index.htm*, Omniglot. 16.08.2001.

Akamatsu, N. (1998) L1 and L2 reading: The orthographic effects of Japanese on reading in English. *Language, Culture and Curriculum.* 11:1.

Akamatsu, N. (1999) The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing: An Interdisciplinary Journal.* 11

Akamatsu, N. (2003) The effects of first language orthographic features on second language reading in text. *Language Learning.* 53:2.

Al-Batal, M. (1995) Issues in the teaching of the productive skills in Arabic, i: M. Al-Batal (red.): *The Teaching of Arabic as a Foreign Language. Issues and Directions*. American Association of Teachers of Arabic.

Albert, M. (1975) Cerebral dominance and reading habits. *Nature.* 256

Alderson, J. C. (1984) Reading in a foreign language. A reading problem or a language problem?, i: J. C. Alderson & A. H. Urquhart (red.): *Reading in a Foreign Language*. New York. Longman.

Alosh, M. M. (1995) Computer-assited language learning for Arabic: Rationale and reserach potential, i: M. Al-Batal (red.): *The Teaching of Arabic as a Foreign Language. Issues and Directions*. American Association of Teachers of Arabic.

Amayreh, M. M. (2003) Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research.* 46:3.

Andersen, K. & H. Hermansen (2000) *Læsning og læsekurser i andetsprogsundervisningen*. Amtscentret for Undervisning, Århus Amt.

Assink, E. M. H. (1997) Visual processes in word recognition by poor readers: Relative use of visual, phonological, and orthographic cues, i: C. K. Leong & R. M. Joshi (red.): *Cross-Language Studies of Learning to Read and Spell. Phonologic and Orthographic Processing*. Kluwer Academic Publishers.

Atkinson, D. (2002) Toward a sociocognitive approach to second language acquisition. *The Modern Language Journal.* 86

Aweiss, S. (1993) Cognitive processes in foreign language reading: Reasoning operations, comprehension monitoring, strategy use, and knowledge sources. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 26

Ayari, S. (1996) Diglossia and illiteracy in the Arab World. *Language, Culture and Curriculum.* 9

Azzam, R. (1989) Orthography and reading of the Arabic language, i: P. G. Aaron & R. M. Joshi (red.): *Reading and Writing Disorders in Different Orthographic Systems*. Dordrecht. Kluwer Academic Publishers.

Badry, F. (1983) *Acquisition of derivational rules in Morrocan Arabic: Implications for the development of standard Arabic as a second language through literacy*. PhD Dissertation, University of California, Berkeley.

Bamford, J. & R. R. Day (1998) Teaching reading. *Annual Review of Applied Linguistics.* 19

Bannatyne, A. (1976) *Language, Reading and Learning Disabilities*. Springfield, Illinois. Charles C. Thomas.

Bar-Lev, Z. (1983) Hebrew hieroglyphics. *Visible Language.* 17:4.

Bar-Lev, Z. (1989) Glyphs. *Bulletin of Higher Hebrew Education.* 4:1.

Bar-Lev, Z. (1990) Glyphs. A shortcut to reading Hebrew. *Pedagogic Reporter.* 40:4.

Bar-Lev, Z. (1991) Two innovations for teaching Arabic. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 24

Barnea, A. & Z. Breznitz (1998) Phonological and orthographic processing of Hebrew words. *The Journal of Genetic Psychology.* 159:4.

Baron, J. (1977) Mechanisms for pronouncing printed words: Use and acquisition, i: D. Laberge & S. J. Samuels (red.): *Basic Processes in Reading: Perception and Comprehension*. Erlbaum.

Baron, J. & J. Hodge (1978) Using spelling-sound correspondences without trying to learn them. *Visible Language (special issue edited by D. W. Massaro).* 12

Barrett, A. M., M. Kim, G. P. Crucian & K. M. Heilman (2000) Spatial bias: effects of early reading direction on Korean subjects. *Neuropsychologia.* 40

Barton, M. I., H. Goodglass & A. Shai (1965) Differential recognition of tachistoscopically presented English and Hebrew words in right and left visual fields. *Perceptual and Motor Skills.* 21

Bashi, J., S. Kahan & D. Davis (1981) Learning achievements in Arab elementary schools in Israel. Jerusalem, Hebrew University

Bat-El, O. (2002) Semitic verb structure within a universal perspective, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Béland, R. & Z. Mimouni (2001) Deep dyslexia in the two languages of an Arabic/French bilingual patient. *Cognition.* 82

Ben-Dror, I., S. Bentin & R. Frost (1995) Semantic, phologic, and morphologic skills in reading disabled and normal children: Evidence from perception and production of spoken Hebrew. *Reading Research Quarterly.* 30:4.

Benmamoun, E. (2002) The role of the imperfective template in Arabic morphology, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Bentin, S. (1989) Orthography and phonology in lexical decision: Evidence from repetition effects at different lags. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 15:1.

Bentin, S. (1992) Phonological awareness, reading, and reading acquisition: A survey and appraisal of current knowledge, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Bentin, S., N. Bargai & L. Katz (1984) Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 10:3.

Bentin, S. & L. B. Feldman (1990) The contribution of morphplogical and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology.* 42:4.

Bentin, S. & R. Frost (1987) Processing Lexical ambiguity and visual word recognition in a deep orthography. *Memory and Cognition.* 15:1.

Bentin, S. & R. Frost (1995) Morphological factors in visual word recognition in Hebrew, i: L. B. Feldman (red.): *Morphological Aspects of Language Processing*. Erlbaum Associates Publishers.

Bentin, S. & R. Ibrahim (1996) New evidence for phonological processing during visual word recognition: The case of Arabic. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 22:2.

Berent, I. & J. Shimron (1997) The representation of Hebrew words: Evidence from the Obligatory Contour Principle. *Cognition.* 64:1.

Berent, I. & J. Shimron (2002) What is a root?, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Berent, I., J. Shimron & V. Vaknin (2001) Phonological constraints on reading: Evidence form the obligatory contour principle. *Journal of Memory and Language.* 44

Berg, T. & H. Abd-El-Jawad (1996) The unfolding of the suprasegmental representations: A crosslinguistic perspective. *Journal of Linguistics.* 32

Berman, R. A. (2002) Children's lexical innovations, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Berninger, V. W. (1994) Introduction to The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues*. Kluwer Academic Publishers.

Berninger, V. W. (1995) Introduction to The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and Writing, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and writing*. Kluwer Academic Publishers.

Berninger, V. W. & R. D. Abbott (1994) Multiple orthographic and phonological codes in literacy acquisition: An evolving research program, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge, vol. 1: Theoretical and Developmental Issues*. Kluwer Academic Press.

Besner, D., E. Davelaar, D. Alcott & P. Parry (1984) Wholistic reading of alphabetic print: Evidence from the FDM and the FBI, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.

Besner, D. & N. Hildebrandt (1987) Orthographic and phonological codes in the oral reading of Japanese Kana. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 13:2.

Boeree, G. C. (2000) The language families of the world. *http://www.ship.edu/~cgboeree/languagefamilies.html*, Shippensburg University. 16.08.2001.

Bolozky, S. (2002) Tre 'roots' of denominative Hebrew verbs, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Borer, H. (2002) Computing argument structure, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Bossers, B. (1991) On thresholds, ceiling and short-circuits: The relation between L1 reading, L2 reading and L2 knowledge. *AILA Review.* 8

Bouma, H. (1971) Visual recognition of isolated lower-case letters. *Vision Research.* 11

Bourdin, B. & M. Fayol (2000) Is graphic activity cognitively costly? A developmental approach. *Reading and Writing: An Interdisciplinary Journal.* 13

Bradshaw, J. L. & N. C. Nettleton (1983) *Human Cerebral Assymetry*. Prentice-Hall.

Brooks, L. (1977) Visual patterns in fluent word identification, i: A. S. Reber & D. Scarborough (red.): *Toward a Psychology of Reading*. Erlbaum.

Brooks, L. & A. Miller (1979) A comparison of explicit and implicit knowledge of an alphabet, i: P. A. Kolers et al. (red.): *Processing of Visible Language*. Plenum Press.

Brown, T. & M. Haynes (1985) Literacy background and reading development in a second language, i: H. Carr (red.): *The Development of Reading Skills*. San Francisco. Jossey-Bass.

Brustad, K., M. Al-Batal & A. Al-Tonsi (1995a) *Alif-Baa. Introduction to Arabic Letters and Sounds*. Washington D. C. Georgetown University Press.

Brustad, K., M. Al-Batal & A. Al-Tonsi (1995b) *Al-Kitaab fii Taallum al-Arabiyya. A Textbook for Beginning Arabic. Part One*. Washington D.C. Georgetown University Press.

Cain, K., J. Oakhill & P. Bryant (2000) Phonological skills and comprehension failure: A test of the phonological processing deficit hypothesis. *Reading and Writing: An Interdisciplinary Journal.* 13

Carello, C., M. T. Turvey & G. Lukatela (1992) Can theories of word recognition remain stubbornly nonphonological?, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Carr, T. H. & B. A. Levy (1990) *Reading and its Development*. New York. Academic Press.

Chomsky, N. (1957) *Syntactic Structures*. Mouton.

Clark, E. V. & R. A. Berman (1984) Language structure and language use in the acquisition of word-formation. *Language.* 6

Coady, J. (1995) Research on ESL/EFL vocabulary acquisition: Putting it in context, i: T. Huckin et al. (red.): *Second Language Learning and Vocabulary Learning*. Ablex Publishing.

Coady, J. (1997a) L2 vocabulary acquisition through extensive reading, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge University Press.

Coady, J. (1997b) L2 vocabulary acquisition: A synthesis of the research, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge.

Cohen, A. D. (1998) *Strategies in Learning and Using a Second Language*. Longman.

Coltheart, M. (1984) Writing systems and reading disorders, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.

Coltheart, M. (1986) Graphemics and visual word recognition, i: G. Augst (red.): *New Trends in Graphemics and Orthography*. de Gruyter.

Coltheart, M. & V. Coltheart (1997) Reading comprehension is not exclusively reliant upon phonological representation, i: A. Caramazza (red.): *Access of Phonological and Orthographic Lexical Forms. Evidence from Dissociations in Reading and Spelling*. Psychology Press.

Coulmas, F. (1996) *The Blackwell Encyclopedia of Writing Systems*. Oxford. Blackwell Publishers.

Coulmas, F. (2003) *Writing Systems. An Introduction to their Linguistic Analysis*. Cambridge University Press.

Cowan, J. R. (1992) A model of lexical storage. Evidence from second language learners' orthographic errors, i: P. Downing et al. (red.): *The Linguistics of Literacy*. Amsterdam. John Benjamins.

Cox, C. H., B. A. Blesser & M. Eden (1978) The graphical context of printet characters. *Visible Language.* 12:4.

Cummins, J. (1984) *Bilingualism and special education*. Texas. Pro-Ed.

Cummins, J. (1991) Conversational and academic language proficiency in bilingual contexts., i: J. H. Hulstijn & J. F. Matter (red.): *Reading in Two Languages. AILA Review 8*. Amsterdam. Free University Press.

Cunningham, A. E., K. E. Perry & K. E. Stanovich (2001) Converging evidence for the concept of orthographic processing. *Reading and Writing: An Interdisciplinary Journal.* 14

Cziko, G. A. (1980) Language competence and reading strategies: A comparison of first and second-language readers. *Language Learning.* 30

Dalby, M. A., C. Elbro, M. Jansen & T. Krogh (1989) *Bogen om læsning II, - set i et udviklingsperspektiv*. Munksgaard. Danmarks Pædagogiske Institut.

Dalby, M. A., C. Elbro, M. Jansen & T. Krogh (1992) *Bogen om læsning III, - om læsehandicappede og læsehandicap*. Munksgaard. Danmarks Pædagogiske Institut.

Dalby, M. A., C. Elbro, M. Jansen, T. Krogh & W. P. Christensen (1983) *Bogen om læsning I, - forudsætninger og status*. Munksgaard. Danmarks Pædagogiske Institut.

Day, R. & J. Bamford (2002) Top ten principles for teaching extensive reading. *Reading in a Foreign Language.* 14:2.
Day, R., C. Omura & M. Hiramatsu (1991) Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language.* 7:2.
Day, R. R. & J. Bamford (1998) *Extensive Reading in the Second Language Classroom*. Cambridge University Press.
Derwing, B. L. (1992) Orthographic aspects of linguistic competence, i: P. Downing et al. (red.): *The Linguistics of Literacy*. Amsterdam. John Benjamins.
Deutsch, A., R. Frost & K. I. Forster (1998) Verbs and nouns are organizid and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 24:5.
Deutsch, A. & K. Rayner (1999) Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes.* 14:4.
Doughty, C. & J. Williams, Eds. (1998) *Focus on Form in Classroom Second Language Acquisition*. Cambridge, Cambridge University Press.
Dowhower, S. L. (1994) Repeated reading revisited: Reasearch into practice. *Reading & Writing Quarterly: Overcoming Learning Difficulties.* 10
Dupuy, B. & S. Krashen (1993) Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning.* 4
Durgunoglu, A. Y. & B. Öney (1987) A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing: An Interdisciplinary Journal.* 11
Dörnyei, Z. (2001) *Teaching and Researching Motivation*. Longman.
Eden, M. & M. Halle (1961) On the formalization of handwriting, i: C. Cherry (red.): *Information Theory: Fourth London Symposium*. Washington D. C. Butterworths.
Elbro, C. (2001) *Læsning og læseundervisning*. Gyldendals Uddannelse.
El-Hassan, S. A. (1978) Educated spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts. *Archivum Linguisticum.* 8
Ennaji, M. (2001) De la diglossie à la quadriglossie. *Langues et Linguistique.* 8
Ephratt, M. (1997) The psychological status of the root in Modern Hebrew. *Folia Linguistica.* XXXI/1-2
Ethnologue (2001) Internet version, 14th edition. *http://www.sil.org/ethnologue/*, SIL International. 22.08.2001.
Eviatar, Z. (1995) Reading direction and attention: Effects on lateralized ignoring. *Brain and Cognition.* 29
Eviatar, Z. (1997) Language experience and right hemisphere tasks: The effect of scanning habits and multilingualism. *Brain and Language.* 58
Eviatar, Z., R. Ibrahim & D. Ganayim (2004) Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology.* 18:1.
Fagard, J. & R. Dahmen (2003) The effects of reading-writing direction on the assymetry of space perception and directional tendencies: A comparison between French and Tunisian children. *Laterality.* 8:1.
Farah, M. J. (1989) Knowledge from text and pictures: A neuropsychological perspective, i: H. Mandl & J. R. Levin (red.): *Knowledge Acquisition from Text and Pictures*. Amsterdam. Elsevier.
Farell, T. S. C. (2001) Teaching reading strategies: 'It takes time!' *Reading in a Foreign Language.* 13:2.
Farid, M. & J. Grainger (1996) How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language.* 53

Faust, M., S. Kravetz & H. Babkoff (1993) Hemispheric specialization or reading habits: Evidence from lexical decision research with Hebrew words and senences. *Brain and Language.* 44

Favreau, M., M. Komoda & N. S. Segalowitz (1980) Second language reading: Implications for the word superiority effect in skilled bilinguals. *Canadian Journal of Psychology.* 34

Favreau, M. & N. S. Segalowitz (1982) Second language reading in fluent bilinguals. *Applied Psycholinguistics.* 3

Feldman, J. M. (1978) Wh* n**ds v*w*ls?, i: L. C. Ehri et al. (red.): *The Recognition of Words*. Newark, Delaware. International Reading Association.

Feldman, L. B. & D. Andjelkovic (1992) Morphological analysis in word recognition, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Feldman, L. B. & S. Bentin (1994) Morphological analysis of disrupted morphemes: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology.* 47:2.

Feldman, L. B. & M. T. Turvey (1983) Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance.* 9

Fender, M. (2003) English word recognition skills of native Arabic- and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics.* 24

Ferguson, C. A. (1959) Diglossia. *Word.* 15

Fischler, M. A. & O. Firschein (1987) *Intelligence. The Eye, the Brain, and the Computer*. Addison Wesley Publishing.

Foorman, B. R. (1994) Phonological and orthographic processing: Separate but equal?, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues*. Kluwer Academic Press.

Foorman, B. R. (1995) Practiced connections of orthographic and phonological processing, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and Writing*. Kluwer Academic Publishers.

Forster, K. I. & C. Davis (1984) Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 10

Fragman, A. (1999) Summarizing and translation in teaching Arabic reading comprehension in Israel. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 32

Frisch, S. A. & B. A. Zawaydeh (2001) The psychological reality of OCP-place in Arabic. *Language.* 77:1.

Frith, U. (1986) Psychologishe Aspekte des ortographischen Wissens: Entwicklung und Entwicklungsstörung, i: G. Augst (red.): *New Trends in Graphemics and Orthography*. de Gruyter.

Frost, J. (1998) *Læsepraksis - på teoretisk grundlag*. København. Dansk Psykologisk Forlag.

Frost, R. (1991) Phonetic recoding of print and its effect on the detection of concurrent speech in amplitude modulated noise. *Cognition.* 39

Frost, R. (1992) Orthography and phonology. The psychological reality of orthographic depth, i: P. Downing et al. (red.): *The Linguistics of Literacy*. Amsterdam. John Benjamins.

Frost, R. (1994) Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 20:1.

Frost, R. (1995) Phonological computation and missing vowels: Mapping lexical involvment in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 21:2.

Frost, R. & S. Bentin (1992a) Processing phonological and semantic ambiguity: Evidence from semantic priming at different SOA's. *Journal of Experimental Psychology: Learning, Memory and cognition.* 18:1.

Frost, R. & S. Bentin (1992b) Reading consonants and guessing vowels: Visual word recognition in Hebrew orthography, i: L. Katz & R. Frost (red.): *Orthography, Phonology, Morphology and Meaning*. Amsterdam. Elsevier.

Frost, R., A. Deutsch & K. I. Forster (2000) Decomposing morphologically complex words in nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 26:3.

Frost, R., K. I. Forster & A. Deutsch (1997) What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 23:4.

Frost, R. & M. Kampf (1993) Phonetic recoding of phonologically ambiguous printed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 19:1.

Frost, R. & L. Katz (1989) Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory and Cognition.* 17:3.

Frost, R., L. Katz & S. Bentin (1987) Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance.* 13

Frost, R., B. Repp & L. Katz (1988) Can speech perception be influenced by a simultaneous presentation of print? *Memory and Language.* 27

Frost, R. & O. Yogev (2001) Orthographic and phonological computation in visual word recognition: Evidence from backward masking in Hebrew. *Psychonomic Bulletin & Review.* 8:3.

Fuchs, L. S., D. Fuchs, M. K. Hosp & J. R. Jenkins (2001) Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading.* 5

Garson, J. (2002) Connectionism, i: E. N. Zalta (red.): *The Stanford Encyclopedia of Philosophy (winter 2002 edition)*. Electronic version: http://plato.stanford.edu.

Gelb, I. J. (1952) *A Study of Writing*. University of Chicago Press.

Geva, E. (1995) Orthographic and cognitive processing in learning to read English and Hebrew, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Kluwer Academic Publishers.

Geva, E. (1999) Introduction: Linguistic processes in reading across orthogaphies. *Reading and Writing: An Interdisciplinary Journal.* 11

Geva, E. & L. S. Siegel (2000) Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing: An Interdisciplinary Journal.* 12

Geva, E. & L. Wade-Woolley (1998) Component processes in becoming English-Hebrew biliterate, i: A. Y. Durgunoglu & L. Verhoeven (red.): *Acquisition of Literacy in a multilingual context: A cross cultural perspective*. Hillsdale. Erlbaum.

Geva, E., L. Wade-Woolley & M. Shany (1993) The concurrent development of spelling and decoding in two different orthographies. *Journal of Reading Behavior.* 25:4.

Geva, E. & D. Willows (1994) Orthographic knowledge is orthographic knowledge is orthographic knowledge, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues*. Kluwer Academic Publishers.

Geyer, L. H. & C. G. Dewald (1973) Feature lists and confusion matrices: Fact or artifact. *Perception & Psychophysics.* 14

Gholamain, M. & E. Geva (1999) Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning.* 49:2.

Gibson, E. J., H. Osser, W. Schiff & J. Smith (1963) An analysis of critical features of letters. Tested by a confusion matrix*Cooperative Research Project No. 639: A Basic Research Program on Reading*. Washington D.C. U.S. Office of Education.

Glinert, L. (1991) *Modern Hebrew. An Essential Grammar*. London. Routledge.

Goodman, K. S. (1968) The psycholinguistic nature of the reading process, i: K. S. Goodman (red.): *The Psycholinguistic Nature of the Reading Process*. Detroit. Wayne State University Press.

Goodman, K. S. (1976) Reading: A psycholinguistic guessing game, i: H. Singer & R. B. Ruddell (red.): *Theoretical Models and Processes of Reading*. Newark. International Reading Association.

Goodman, K. S. (1982) *Language and Literacy*. 1. Routledge.

Gough, P. B. (1972) One second of reading, i: J. F. Kavanagh & I. G. Mattingly (red.): *Language by ear and eye*. Cambridge. MIT Press.

Grabe, W. (1991) Current developments in second language reading research. *TESOL Quarterly.* 25:3.

Grabe, W. & F. Stoller (1997) Reading and vocabulary development in a second language: A case study, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge University Press.

Grabe, W. & F. Stoller (2002) *Teaching and Researching Reading*. Longman.

Grainger, J. (1992) Orthographic neighborhoods and visual word recognition, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Green, D., E. Hammond & S. Supramamian (1983) Letters and shapes: developmental changes in search strategies. *British Journal of Psychology.* 74

Green, D. W. & P. Meara (1987) The effects of script on visual search. *Second Language Research.* 4

Gregory, R. L. (1966) *Eye and Brain*. World Uni´versity Library.

Grivelet, S. (2001) Introduction to 'Diagraphia: Writing Systems and Society'. *International Journal of the Sociology of Language.* 150

Gronau, N. & R. Frost (1997) Prelexical phonologic computation in a deep orthography: Evidence from backward masking in Hebrew. *Psychonomic Bulletin & Review.* 4:1.

Habib-Allah, M. (1985) Reading and reading comprehension. *Reading Circles.* 21

Hammond, E. & D. Green (1982) The detection of targets in letter and non-letter arrays. *Canadian Journal of Psychology.* 36

Hammoud, S.-D. (1996) A survey of current classroom practices among teachers of Arabic. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 29

Hana, S. A. (1972) Quantitative measurement of errors and remedial instruction in reading Arabic. *Al-lisan al-arabi.* 9:1.

Hanley, J. R. & H. S. Huang (1997) Phonological awareness and learning to read Chinese, i: C. K. Leong & R. M. Joshi (red.): *Cross-Language Studies of Learning to Read and Spell. Phonologic and Orthographic Processing*. Kluwer Academic Publishers.

Hanna, S. A. (1964) Quantitative measurement of errors in reading Arabic. *The Modern Language Journal.* 48

Hanna, S. A. (1966) Remedial instruction in reading Arabic. *The Modern Language Journal.* 50

Harris, R. (1990) Quelques réflexions sur la tyrannie de l'alphabet, i: C. Sirat et al. (red.): *L'ecriture: Le Cerveau, L'oeil et la Main*. Brepols - Turnhout.

Hauge, T. (1992) *Fremmedsproglig læsetræning*. Odense Working Papers in Language and Communication, no. 3. Syddansk Universitet.

Hauge, T. (1993) Some theoretical considerations on L2 reading and their potential implications for instruction, i: T. Hauge (red.): *Læseforståelse - Reading Comprehension. Psykolingvistiske og litterære aspekter*. Odense Working Papers in Language and Communication, no. 5. Syddansk Universitet.

Haywood, J. A. & H. M. Nahmad (2000) *A New Arabic Grammar of the Written Language*. London. Lund Humphries.

Heath, J. (2002) Arabic derivational ablaut, processing strategies, and consonantal "roots", i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Henderson, L. (1982) *Orthography and Word Recognition in Reading*. Academic Press.

Henderson, L. (1984) Writing systems and reading processes, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.

Henderson, L. (1986) From morph to morpheme: the psychologist gaily trips where the linguist has trodden, i: G. Augst (red.): *New Trends in Graphemics and Orthography*. de Gruyter.

Hirsh, D. & P. Nation (1992) What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language.* 8:2.

Holm, L. & H. P. Laursen (2000) *Andetsprogsdidaktik*. Dansklærerforeningen.

Hoosain, R. (1995) Getting at the sound and meaning of logographic and alphabetic scripts, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Kluwer Academc Publishers.

Hung, D. L. & O. J. L. Tzeng (1981) Orthographic variations and visual information processing. *Psychological Bulletin.* 90:3.

Hung, D. L., O. J. L. Tzeng, W. L. Lee & J. M. Chang (1994) Orthography, reading disability, and cerebral organization, i: W. C. Watt (red.): *Writing Systems and Cognition*. Kluwer Academic Publishers.

Hung, D. L., O. J. L. Tzeng & A. K. Y. Tzeng (1992) Automatic activation of linguistic information in Chinese character recognition, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Haastrup, K. (1989) Fremmedsproglig reception: Lytte- og læseforståelse, i: G. Kasper & J. Wagner (red.): *Grundbog i fremmedsprogspædagogik*. Gyldendal.

Haastrup, K. (1993) Introduktion til læsning på et fremmedsprog, i: T. Hauge (red.): *Læseforståelse - Reading Comprehension. Psykoligvistiske og litterære aspekter*. Odense Working Papers in Language and Communication, no. 5. Syddansk Universitet.

Ibrahim, M. (1986) Standard and prestige language: A problem in Arabic sociolinguistics. *Anthropological Linguistics.* 28

Jackson, N. E., H. Chen, L. Goldsberry, A. Kim & C. Vanderwerf (1999) Effects of variations in orthographic information on Asian and American readers' English text reading. *Reading and Writing: An Interdisciplinary Journal.* 11

Jackson, N. E., W.-H. Lu & J. Daushen (1994) Reading chinese and reading English: Similarities, differences, and second language reading, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues*. Kluwer Academic Publishers.

Jameson, K. (1994) Empirical methods for evaluating generative semiotic models: An application to the Roman majuscules, i: W. C. Watt (red.): *Writing Systems and Cognition*. Kluwer Academic Publishers.

Janzen, J. & F. L. Stoller (1998) Integrating strategic reading in L2 instruction. *Reading in a Foreign Language.* 12:1.

Jassem, Z. A. (1994) *Lectures in English and Arabic Sociolinguistics*. Kuala Lumpur. Pustaka Antara.

Jenkins, J. (2001) *The Effect of a Preschool MSA Immersion Program on Arab Children's primary School Reading and Composition Scores*. Unpublished thesis, Brigham Young University, Department of Linguistics.

Jensen, H. (1970) *Sign, Symbol, and Script*. London. George Allen and Unwin Ltd.

Johnson, N. F. (1992) On the role of cohorts or neighbors in visual word recognition, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Just, M. A. & P. Carpenter (1987) *The Psychology of Reading and Language Comprehension*. Allyn and Bacon Inc.

Kasper, G. & J. Wagner (1989) Teorier om fremmedsprogsindlæring, i: G. Kasper & J. Wagner (red.): *Grundbog i fremmedsprogspædagogik*. Gyldendal.

Katz, L. & L. B. Feldman (1981) Linguistic coding in word recognition: Comparisons between a deep and a shallow orthography, i: A. M. Lesgold & C. A. Perfetti (red.): *Interactive Processes in Reading*. Erlbaum.

Katz, L. & L. B. Feldman (1983) Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 9

Katz, L. & R. Frost (1992a) Orthography, phonology, morphology, and meaning: An overview, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Katz, L. & R. Frost (1992b) The reading process is different for different orthographies: The orthographic depth hypothesis, i: L. Katz & R. Frost (red.): *Orthography, Phonology, Morphology and Meaning*. Amsterdam. Elsevier.

Keenan, V. (1972) Effect of Hebrew and English letters on children's perceptual set. *Journal of experimental child psychology.* 13:1.

Kenstowicz, M. J. & C. Kisseberth (1979) *Generative Phonology: Description and Theory*. New York. Academic Press.

Khaldieh, S. (1996) Word recognition of Arabic as a foreign language by American learners: The role of phonology and script. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 29

Khaldieh, S. (1999) An investigation into reader - text interaction: Native, proficient, and less proficient non-native readers of Arabic. *Al-Arabiyya. Journal of the American Association of Teachers of Arabic.* 32

Khaldieh, S. (2001) The relationship between I*raab, lexical knowledge, and reading comprehensionn of nonnative readers of Arabic. *The Modern Language Journal.* 85:3.

Kjertmann, K. (2002) *Læsetilegnelse - ikke kun en sag for skolen*. København. Alinea.

Koda, K. (1988) Cognitive processes in second language reading: Transfer of L1 reading skills and strategies. *Second Language Research.* 4

Koda, K. (1990) The use of L1 reading strategies in L2 reading: Effects of L1 orthographic structures on L2 phonological recoding strategies. *Studies in Second Language Acquisition.* 12

Koda, K. (1992) The effects of lower-level processing skills on FL reading performance: Implications for instruction. *The Modern Language Journal.* 76

Koda, K. (1995) Cognitive consequenses of L1 and L2 orthographies, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Dordrecht. Kluwer Academic Publishers.

Koda, K. (1996) L2 word recognition research: A critical review. *The Modern Language Journal.* 80:4.

Koda, K. (1997) Orthografic knowledge in L2 lexical processing: A cross-linguistic perspective, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge.

Koda, K. (1998) The role of phoenemic awareness in second language reading. *Second Language Research.* 14:2.

Koga, K. & R. Groner (1989) Intercultural experiments as a research toll in the study of cognitive skill acquisition: Japanese caracter recognition and eye movements in non-Japanese subjects, i: H. Mandl & J. R. Levin (red.): *Knowledge Acquisition from Text and Pictures*. Amsterdam. Elsevier.

Koriat, A. (1984) Reading without vowels: Lexical acces in Hebrew, i: H. Bouma & D. G. Bouwhuis (red.): *Attention and Performance X. Control of Language Processes*. Lawrence Erlbaum Associates.

Koriat, A. (1985a) Lateralization effects in reading reading pointed and unpointed Hebrew words. *British Journal of Psychology.* 76

Koriat, A. (1985b) Lexical acces for low and high frequency words in Hebrew. *Memory and Cognition.* 13:1.

Krantz, G. (1990) *Learning Vocabulary in a Foreign Language*. Göteborg. Acta Universitatis Gothoburgensis.

Krashen, S. (1982) *Principles and Practice in Second Language Acquisition*. Oxford. Pergamon Press.

Krashen, S. (1985) *The input hypothesis: Issues and implications*. London. Longman.

Kulbrandtstad, L. I. (1996) *Lesning på andrespråk*. Oslo. Scandinavian University Press.

Künnapas, T. (1966) Visual perception of capital letters. *Scandinavian Journal of Psychology.* 7

Laufer, B. (1996) The lexical threshold of second language reading comprehension: What it is and how it relates to L1 reading ability, i: K. Sajavaara & C. Fairweather (red.): *Approaches to Second Language Acquisition*. University of Jyväskylä.

Laufer, B. (1997) The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge.

Laufer, B. & Y. Yano (2001) Understanding unfamiliar words in a text: Do L2 learners understand how much they don't understand? *Reading in a Foreign Language.* 13:2.

Lavidor, M., A. W. Ellis & A. Pansky (2002) Case alternation and length effects in lateralized word recognition: Studies of English and Hebrew. *Brain and Cognition.* 50

Lavine, L. O. (1972) *The development of perception of writing in pre-reading children: A cross-cultural study*, Department of Human Development. Cornell University.

Leong, C. K. (1995) Orthographic and psycholinguistic considerations in developing literacy in Chinese, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Kluwer Academic Publishers.

Leong, C. K. & R. M. Joshi (1997) Relating phonologic and orthographic processing to learning to read and spell, i: C. K. Leong & R. M. Joshi (red.): *Cross-Language Studies of Learning to Read and Spell*. Kluwer Academic Publishers.

Levy-Schoen, A. & K. O'regan (1979) The control of eye movements in reading, i: P. A. Kolers et al. (red.): *Processing of Visible Language. Vol. 1*. Plenum Press.

Li, L. (2002) The role of phonology in reading Chinese single characters and two-character words with high, medium and low phonological regularities by Chinese grade 2 and grade 5 students. *Reading Reseach Quarterly.* 37:4.

Lieblich, A., A. Ninio & S. Kugelmass (1975) Developmental trends in directionality of drawing in Jewish and Arab Israeli children. *Journal of Cross-Cultural Psychology.* 6:4.

Lindhardsen, V. (2002) Læseforståelse - den oversete færdighed i sprogprøverne. *Sprogforum.* D

Lindsay, P. H. & D. A. Norman (1972) *Human Information Processing: An Introduction to Psychology*. New York. Academic Press.

Linguasphere (1999) Linguasphere Website. *http://www.linguasphere.org/*, Linguaspere Observatory. 2001.

Liu, I.-M. (1995) Script factors that affect literacy: Alphabetic vs. logographic languages, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Kluwer Academic Publishers.

Lubow, R. E., Y. Tsal, A. Mirkin & G. Mazliah (1994) English and Hebrew letter report by English- and Hebrew-reading subjects: Evidence for stimulus control, not hemispheric assymetry. *Brain and Cognition.* 25

MacNamara, J., M. Feltin, M. Hew & M. Klein (1968) An analytic comparison of reading in two languages. *Irish Journal of Education.* 2

Magrath, D. R. (1988) Teaching non-Latin alphabets through communication: The case of Arabic. *IRAL.* 26:3.

Mahfoudi, A. (2002) Processing Semitic languages: Are some letters/phonemes more equal than others? Proceedings in the International Symposium on: The Processing of Arabic. April 2002, 18th-20th., University of Manouba, Tunisia.

Mallon, J. (1952) *Paléographie Romaine*. Madrid. Instituto Antonio de Nebrija de Filologia.

Mann, V. A. (2000) Introduction to special issue on morphology and the acquisition of alphabetic writing systems. *Reading and Writing: An Interdisciplinary Journal.* 12

Mansouri, F. (2000) *Grammatical Markedness and information Processing in the Acquisition of Arabic as a Second Language*. Lincom Europa.

Massaro, D. W. (1975) Language and information processing, i: D. W. Massaro (red.): *Understanding Language. An Information-Processing Analysis of Speech Perception, Reading, and Psycholinguistics*. New York. Academic Press.

Massaro, D. W. (1979) Reading and listening, i: P. A. Kolers et al. (red.): *Processing of Visible Language*. Plenum Press.

Massaro, D. W. & J. Schmuller (1975) Visual features, perceptual storage, and processing time in reading, i: D. W. Massaro (red.): *Understanding Language. An Information-Processing Analysis of Speech Perception, Reading, and Psycholinguistics*. New York. Academic Press.

Massaro, D. W., G. A. Taylor, R. L. Venezky, J. E. Jastrzembski & P. A. Lucas (1980) *Letter and Word Perception*. Amsterdam. North-Holland.

Massaro, D. W., R. L. Venezky & G. A. Taylor (1979) Orthographic regularity, positional frequency and visual processing of letter strings. *Journal of Experimental Psychology: General.* 108

Mattingly, I. G. (1992) Linguistic awareness and orthographic form, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. North-Holland.

McCarthy, J. (1986) OCP effects: Gemination and antigemination. *Linguistic Inquiry.* 17

McClelland, J. L. (1999) Connectionist cognitive modelling*MIT Encyclopedia of Cognitive Science*. Electronic version: http://cognet.mit.edu.

McClelland, J. L. & D. E. Rumelhart (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review.* 88:5.

Meiseles, G. (1980) Educated spoken Arabic and the Arabic Lanugage continuum. *Archivum Linguisticum.* 11

Metsala, J. L., K. E. Stanovich & G. D. A. Brown (1998) Regularity Effects and the phonological deficit model of reading disabilities: A meta-analytic review. *Journal of Educational Psychology.* 90

Meyler, A. & Z. Breznitz (1998) Developmental associations between verbal and visual short-term memory and the acquisition of decoding skill. *Reading and Writing: An Interdisciplinary Journal.* 10

Miceli, G., B. Benvegnù, R. Capasso & A. Caramazza (1997) The independence of phonological and orthographic lexical forms: Evidence from aphasia, i: A. Caramazza (red.): *Access of Phonological and Orthographic Lexical Forms. Evidence from Dissociations in Reading and Spelling*. Psychology Press.

Miozzo, M. & A. Caramazza (1998) Varieties of pure alexia: The case of failure to access graphemic representations. *Cognitive Neuropsychology.* 15:1-2.

Mishkin, M. & G. D. Forgays (1952) Word recognition as a function of retrieval locus. *Journal of Experimental Psychology.* 43

Mitchell, D. C., F. Cuetos & D. Zagar (1990) Reading in different languages: Is there a universal mechanism for parsing sentences?, i: D. A. Balota et al. (red.): *Comprehension Processes in Reading*. Lawrence Erlbaum Associates.

Mitchell, T. F. (1980) Dimensions of style in a grammar of educated spoken Arabic. *Archivum Linguisticum.* 11

Mitchell, T. F. (1986) What is educated spoken Arabic? *International Journal of the Sociology of Language.* 61

Monsell, S., M. C. Doyle & P. N. Haggard (1989) Effects of frequency on visual word recognition. *Psychologcal Review.* 76

Mori, S. (2002) Redefining motivation to read in a foreign language. *Reading in a Foreign Language.* 14:2.

Morton, J. & S. Sasanuma (1984) Lexical Acces in Japanese, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.

Moscati, S., A. Spitaler, E. Ullendorf & W. Von Soden (1980) *An Introduction to the Comparative Grammar of the Semitic Languages*. Wiesbaden. Otto Harrasowitz.

Muljani, D., K. Koda & D. R. Moates (1998) The development of word recognition in a second language. *Applied Psycholinguistics.* 19

Maamouri, M. (1998) Language Education and Human Development. Arabic Diglossia and its impact on the quality of Education in the Arab Region. Mediterranean Development Forum, Marrakech, Morocco, International Literacy Institute, University of Pennsylvania, USA.

Nachshon, I. (1983) Directional preferences of bilingual children. *Perceptual and Motor Skills.* 56

Nachshon, I. (1985) Directional preferences in oersception of visual stimuli. *International Journal of Behavioral Neuroscience.* 25

Nachshon, I. & M. Alek (1981) The development of directional preferences: cross-cultural differences. *Psychologia.* 24

Nachshon, I. & T. Hatta (2001) Directional tendencies of Hebrew, Japanese, and English readers. *Perceptual and Motor Skills.* 93

Nassaji, H. & E. Geva (1999) The contribution of phonological and orthographic processing skills to adult ESL reading: Evidence from native speakers of Farsi. *Applied Psycholinguistics.* 20

Navon, D. & J. Shimron (1981) Does word naming involve grapheme to phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning & Verbal Behavior.* 20

Navon, D. & J. Shimron (1984) Reading Hebrew: How necessary is the graphemic representation of vowels?, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology and Linguistics*. Lawrence Erlbaum Associates.

Nazir, T. A., N. Ben-Boutayab, N. Decoppet, A. Deutsch & R. Frost (2004) Reading habits, perceptual learning, and recognition of printed words. *Brain and Language.* 88

Neill, W. T., D. V. Hillard & E. A. Cooper (1988) The detection of lexical ambiguity: Evidence for context-sensitive parallel access. *Journal of Memory and Language.* 27

Neisser, U. (1967) *Cognitive Psychology*. New York. Appleton Century Crofts.

Nerius, D. (1986) Zur Bestimmung und Differenzierung der Prinzipen der Orhographie, i: G. Augst (red.): *New Trends in Graphemics and Orthography*. de Gruyter.

Ney, J. W. & B. A. Pearson (1990) Connectionism as a model of language learning: Parallels in foreign language teaching. *The Modern Language Journal.* 74:4.

Nielsen, H. L. (1992) *Kommunikativ kompetens og fjerne fremmedsprog*. Ph.D. -afhandling, Odense Universitet.

Nielsen, H. L. (1993) The Odense programme of teaching Arabic communicatively: A preliminary evaluation of aims, achievements, and problems, i: L. E. Andersen (red.): *Middle East Studies in Denmark*. Odense. Odense University Press.

Norris, D. (1990) Connectionism: A case for modularity, i: D. A. Balota et al. (red.): *Comprehension Processes in Reading*. Hillsdale. Lawrence Erlbaum Associates.

Nuttall, C. (1996) *Teaching Reading Skills in a Foreign Language*. Heinemann.

Obler, L. K. (1989) Hebrew orthography and dyslexia - A note, i: P. G. Aaron & R. M. Joshi (red.): *Reading and Writing Disorders in Different Orthographic Systems*. Dordrecht. Kluwer Academic Publishers.

Oller, J. W. & J. Tullis (1973) Reading skills of non-native speakers of English. *International Review of Applied Linguistics in Language Teaching.* 11

Orbach, J. (1953) Retinal locus as a factor in the recognition of visually perceived words. *American Journal of Psychology.* 65

Pagh, J. (1998) Sprogbogen - om verdens sprog og deres slægtskab. *http://hjem.get2net.dk/sprogbog/index.htm*. 20.08.2001.

Paribakht, T. S. & M. Wesche (1997) Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Oxford University Press.

Parkinson, D. B. (1994) Testing native speakers. Implications for teaching Arabic to non-native speakers, i: R. M. Rammuny & D. B. Parkinson (red.): *Investigating Arabic*. Greyden Press.

Patel, P. G. (1995) Brahmi scripts, Orthographic units and reading acquisition, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy. Reading and Learning to Read Alphabets, Syllabaries and Characters*. Kluwer Academic Publishers.

Pedersen, M. S. (2001) Task Force - Et bud på kommunikativ sprogundervisning. *Sprogforum.* 20

Perfetti, C. A. & S. Zhang (1991) Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 17:4.

Perfetti, C. A., S. Zhang & I. Berent (1992) Reading in English and Chinese: Evidence for a "universal" phonological principle, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Petersen, J. (1987) *Visuel Perception og Læsning I*. København. Danmarks Lærerhøjskole.

Petersen, J. (1989) Bør et skriftsprogs læseretning have betydning for valg af første skriftsprog ved undervisning af tosprogede elever?, i: A. Holmen & J. N. Andersen (red.): *Skriftsproget og de Sproglige Mindretal*. København. Danmarks Lærerhøjskole.

Pew, R. W. & G. T. Gardner (1965) Upublicerede data præsenteret i Fisher, Monty & Glucksberg, 1969: Visual confusion matrices: Fact or artifact. *The Journal of Psychology.* 71

Pollatsek, A., S. Bolozky, A. D. Well & K. Rayner (1981) Assymetries in the perceptual span for Israeli readers. *Brain and Language.* 14

Pollatsek, A. & K. Rayner (1990) Eye movements and lexical access in reading, i: D. A. Balota et al. (red.): *Comprehension Processes in Reading*. Hillsdale. Lawrence Erlbaum Associates.

Powell, M. A. (1981) Three problems in the history of Cuneiform writing: Origins, direction of script, literacy. *Visible Language.* 15

Prowse, P. (2002) Top ten principles for teaching extensive reading: A response. *Reading in a Foreign Language.* 14:2.

Prunet, J.-F., R. Béland & A. Idrissi (2000) The mental representation of Semitic words. *Linguistic Inquiry.* 31:4.

Ramsey, W. (1999) Connectionism. Philosophical issues*MIT Encyclopedia of Cognitive Science*.

Ramsey, W., S. Stich & J. Garon (1990) Connectionism, eliminativism and the future of folk psychology. *Philosophical Perspectives.* 4

Randall, M. & P. Meara (1988) How Arabs read Roman letters. *Reading in a Foreign Language.* 4:2.

Rapp, B., L. Benzig & A. Carmazza (1997) The autonomy of lexical orthography, i: A. Caramazza (red.): *Access of Phonological and Orthographic Lexical Forms. Evidence from Dissociations in Reading and Spelling*. Psychology Press.

Ravid, D. (1996) Accessing the mental lexicon: Evidence from the incompability between representation of spoken and written morphology. *Linguistics.* 34

Ravid, D. (2002) A developmental perspective on root perception in Hebrew and Palestinian Arabic, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.

Ravn, I. (1992) Kognition - computerprogrammeret, netværkserfaret eller handlingsfrembragt? *Litteratur & Samfund.* 47-48

Rayner, K. (1997) Understanding eye movements in reading. *Scientific Studies of Reading.* 1:4.

Rayner, K. & A. Pollatsek (1989) *The Psychology of Reading*. Lawrence Erlabaum Associates.

Richman, H. B. & H. A. Simon (1989) Context effects in letter perception: Comparison of two theories. *Psychological Review.* 96:3.

Ridgway, T. (1994) Reading theory and foreign language reading comprehension. *Reading in a Foreign Language.* 10:2.

Ridgway, T. (1997) Threshold of the background effect in foreign language reading. *Reading in a Foreign Language.* 11:1.

Robinson, P. (1996) Learning simple and complex language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition.* 18

Roman, G. (1982) *Approche expérimentale de la lecture en Arabe*. Thése de Doctorat de Troisième Cycle, Université d'Aix-Marseille I.

Roman, G. & B. Pavard (1987) A comparative study: How we read Arabic and French, i: J. K. O'regan & A. Levy-Schoen (red.): *Eye Movements: From Physiology to Cognition*. Amsterdam. Elsevier.

Roman, G., B. Pavard & B. Asselah (1985) Traitement perceptif des phrases ambigues en arabe. *Cahiers de Psychologie Cognitive.* 5:1.

Rumelhart, D. E. (1971) A Multicomponent theory of confusion among briefly exposed alphabetical characters. CHIP 22 from the Center for Human Information Processing, University of California, 1971. Præsenteret i Massaro & Schmuller: Visual features, perceptual storage, and processing time in reading, i: D. W. Massaro (red.): *Understanding Language*. Academic Press.

Ryan, A. & P. Meara (1991) The case of invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language.* 7:2.

Ryding, K. C. (1991) Proficiency despite diglossia: A new approach for Arabic. *The Modern Language Journal.* 75:2.

Saiegh-Haddad, E. (2003a) Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1) - English (L2) readers. *Reading and Writing: An Interdisciplinary Journal.* 16

Saiegh-Haddad, E. (2003b) Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics.* 24:3.

Sampson, G. (1985) *Writing Systems*. Stanford University Press.

Sanocki, T. (1987) Visual knowledge underlying letter perception: Front-specific, schematic tuning. *Journal of Experimental Psychology: Human Perception and Performance.* 13:2.

Sasanuma, S. (1984) Can surface dyslexia occur in Japanese?, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.

Sasoon, R. (1995) *The Acquisition of a Second Writing System*. Oxford. Intellect.

Scheerer, E. (1986) Orthography and lexical acces, i: G. Augst (red.): *New Trends in Graphemics and Orthography*. de Gruyter.

Segalowitz, N. S., C. Poulsen & M. Komoda (1991) Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. *AILA Review.* 8

Seidenberg, M. S. (1990) Lexical Acces: Another theoretical soupstone?, i: D. A. Balota et al. (red.): *Comprehension Processes in Reading*. Hillsdale. Lawrence Erlbaum Associates.

Seidenberg, M. S. (1992) Beyond orthographic depth in reading: Equitable division of labor, i: R. Frost & L. Katz (red.): *Orthography, Phonology, Morphology, and Meaning*. Amsterdam. Elsevier.

Seidenberg, M. S. & J. L. Mcclelland (1989) A distributed, developmental model of word recognition and naming. *Psychological Review.* 96:4.
Seidenberg, M. S., M. K. Tanenhaus, J. M. Leiman & M. Bienkowski (1982) Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology.* 14
Shakir, A. (1996) Reading and translating Arabic as a foreign language: A discourse-based approach. *Turjuman.* 5:1.
Share, D. & I. Levin (1999) Learning to read and write in Hebrew, i: M. Harris & G. Hatano (red.): *Learning to read and write. A cross-linguistic perspective*. Cambridge. Cambridge University Press.
Sharkey, N. E. (1990) A connectionist model of text comprehension, i: D. A. Balota et al. (red.): *Comprehension Processes in Reading*. Hillsdale. Lawrence Erlbaum Associates.
Shatil, E. & D. L. Share (2003) Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology.* 86:1.
Shimron, J. (1980) The distribution of visual information in the vertical dimension of Roman and Hebrew letters. *Visible Language.* 14:1.
Shimron, J. (1993) The role of vowels in reading: A review of studies of English and Hebrew. *Psychological Bulletin.* 114:1.
Shimron, J. (1999) The role of vowel signs in Hebrew: Beyond word recognition. *Reading and Writing: An Interdisciplinary Journal.* 11
Shimron, J. (2002) Semitic languages: Sre they really root based?, i: J. Shimron (red.): *Language Processing and Acquisition in Languages os Semitic, Root-based, Morphology*. Amsterdam/Philadelphia. John Benjamins.
Shimron, J. & D. Navon (1980) The distribution of visual information in the vertical dimension of Roman and Hebrew letters. *Visible Language.* 14:1.
Shimron, J. & D. Navon (1981) The distribution of information within letters. *Perception & Psychophysics.* 30:5.
Shimron, J. & D. Navon (1982) The dependence of graphemes and on their translation to phonemes in reading: A developmental perspective. *Reading Research Quarterly.* 17
Shimron, J. & S. Sivan (1994) Reading proficiency and orthography: Evidence from Hebrew and English. *Language Learning.* 44:1.
Silverberg, R., S. Bentin, T. Gaziel, L. K. Obler & M. L. Albert (1979) Shift of visual field preference for English words in native Hebrew speakers. *Brain and Language.* 8
Silverberg, R., H. W. Gordon, S. Pollack & S. Bentin (1980) Shift of visual field preference for Hebrew words in native speakers learning to read. *Brain and Language.* 11
Simpson, G. B. (1984) Lexical anbiguity and its role in models of word recognition. *Psychological Bulletin.* 96:2.
Sirat, C. (1987) La morphologie humaine et la direction des écritures. *Académie des Inscriptions & Belles-Lettres.* Comptes rendus des séances de l'année 1987, janvier-mars
Sirat, C. (1990) Introduction, i: C. Sirat et al. (red.): *L'écriture: le cerveau, l'oeil et la main*. Turnhout. Brepols.
Sirat, C. (1994) Handwriting and the writing hand, i: W. C. Watt (red.): *Writing Systems and Cognition*. Kluwer Academic Publishers.
Smolensky, P. (1999) Connectionist Approaches to Language*MIT Encyclopedia of Cognitive Science*. Electronic version: http://cognet.mit.edu.
Snow, M. A. (1998) Trends and issues in content-based instruction. *Annual Review of Applied Linguistics.* 18

Spolsky, B. & E. Shohamy (1999) *The Languages of Israel. Policy, Ideology and Practice*. Multilingual Matters.
Stanovich, K. E. (1980) Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly.* 16
Stanovich, K. E. (1986) Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Annual Review of Applied Linguistics.* 12
Stanovich, K. E. (2000) *Progress in Understanding Reading*. New York. Guilford.
Suaréz, A. & P. Meara (1989) The effect of irregular orthography in the processing of words in a foreign language. *Reading in a Foreign Language.* 6:1.
Swaffar, J. K. (1985) Reading authentic texts in a foreign language: A cognitive model. *The Modern Language Journal.* 69:1.
Tabossi, P. (1991) Understanding words in context, i: G. B. Simpson (red.): *Understanding Word and Sentence*. Elsevier Science Publishers.
Taguchi, E. (1997) The effects of repeated readings on the development of lower identification skills of FL readers. *Reading in a Foreign Language.* 11:1.
Taguchi, E. & G. Gorsuch (2002) Transfer effects of repeated EFL reading on reading new passages: A preliminary investigation. *Reading in a Foreign Language.* 14:1.
Taylor, I. (1981) Writing systems and reading, i: G. E. Mackinnon & T. G. Waller (red.): *Reading Research. Advances in Theory and Practice*. Academic Press.
Tinker, M. A. (1965) *Bases of Effective Reading*. Minneapolis. University of Minnesota Press.
Turvey, M. T., L. B. Feldman & G. Lukatela (1984) The Serbo-Croatian orthography constrains the reader to a phonologically analytic strategy, i: L. Henderson (red.): *Orthographies and Reading. Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Lawrence Erlbaum Associates.
Tzeng, O. J. L. (1981) Overview: Relevancy of experimental psychology to reading instruction, i: O. J. L. Tzeng & H. Singer (red.): *Perception of Print: Reading Research in Experimental Psychology*. Lawrence Erlbaum Associates.
Urquhart, S. & C. Weir (1998) *Reading in a Second Language*. Longman.
Vaid, J. (1995) Script directionality affects nonlinguistic performance: Evidence from Hindi and Urdu, i: I. Taylor & D. R. Olson (red.): *Scripts and Literacy*. Kluwer Academic Publishers.
Vellutino, F. R. & D. M. Scanlon (1986) Linguistic coding a metalinguistic awareness: Their relationship to verbal memory and code acquisition in poor and normal readers, i: D. B. Yaden & S. Templeton (red.): *Metalinguistic Awareness and Beginnin Literacy*. Portsmouth. Heinemann.
Vellutino, F. R., D. M. Scanlon & H. Chen (1995) The increasingly inextricable relationship between orthographic and phonological coding in learning to read: Some reservations about current methods of operationalizing orthographic coding, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and writing*. Kluwer Academic Publishers.
Vellutino, F. R., D. M. Scanlon & M. S. Tanzman (1994) Components of reading ability: Issues and problems in operationalizing word identification, phonological coding and orthographic coding, i: G. R. Lyon (red.): *Frames of reference for the assessment of learning disabilities: New views of on measurement issues*. Baltimore. Brookes.
Venezky, R. L. (1970) *The Structure of English Orthography*. Mouton.
Verhoeven, L. (1991) Acquisition of biliteracy, i: J. H. Hulstijn & J. F. Matter (red.): *Reading in Two Languages. AILA Review 8*. Amsterdam. Free University Press.

Wade-Woolley, L. (1999) First language influences on second language word reading: All roads lead to Rome. *Language Learning.* 49:3.
Wade-Woolley, L. & E. Geva (1998) Processing inflected morphology in second language word recognition: Russian-speakers and English-speakers read Hebrew. *Reading and Writing: An Interdisciplinary Journal.* 11
Wagner, D. A. & J. E. Spratt (1987) Cognitive consequences of contrasting pedagogies: The effect of Quranic preschooling in Morocco. *Child Development.* 58:5.
Waring, R. & M. Takaki (2003) At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language.* 15:2.
Watt, W. C. (1975) What is the proper characterization of the alphabet? I: Desiderata. *Visible Language.* 9
Watt, W. C. (1980) What is the proper characterization of the alphabet? II: Composition. *Ars Semeiotica.* 3:3.
Watt, W. C. (1981) What is the proper characterization of the alphabet? III: Appearance. *Ars Semeiotica.* 4
Watt, W. C. (1988) What is the proper characterization of the alphabet? IV: Union. *Semiotica.* 70
Watt, W. C. (1994) Curves and angles, i: W. C. Watt (red.): *Writing Systems and Cognition*. Kluwer Academic Publishers.
Watt, W. C. (2002) What is the proper characterization of the alphabet? V: Transcendence. *Semiotica.* 138:131-178.
Whitney, P. & D. A. Warning (1991) The role of knowledge in comprehension, i: G. B. Simpson (red.): *Understanding Word and Sentence*. Elsevier Science Publisers.
Wigfield, A. (1994) Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review.* 6
Wijk, A. (1966) *Rules of Pronunciation for the English Language*. Oxford University Press.
Williams, E. (1983) Communicative reading, i: K. Johnson & D. Porter (red.): *Perspectives in Communicative Language Teaching*. London. Academic Press.
Williams, E. (1984) *Reading in the language classroom*. London. Macmillan.
Willows, D. & E. Geva (1995) What is *visual* in orthoraphic processing?, i: V. W. Berninger (red.): *The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and Writing*. Kluwer Academic Publishers.
Wright, W. (1896) *A Grammar of the Arabic Language, 3. edition*. Beirut. Librairie du Liban.
Yang, L. (1997) Tracking the acquisition of L2 vocabulary, i: J. Coady & T. Huckin (red.): *Second Language Vocabuolary Acquisition*. Oxford University Press.
Young, A. W. & A. W. Ellis (1985) Different mehods of lexical acces for words presented in left and right visual hemifields. *Brain and Language.* 24:2.
Zaidel, E. (1977) Lexical organization in the right hemisphere, i: P. Buser & A. Rongeul-Bisdel (red.): *Cerebral correlates of concious experience*. Amsterdam. Elsevier.
Zimmerman, C. B. (1997) Historical trends in second language vocabulary instruction, i: J. Coady & T. Huckin (red.): *Second Language Vocabulary Acquisition*. Cambridge.

Arabiske og hebraiske kilder:

ابو المجد أ: جحا حزين على حماره. نوادر جحا للأطفال ٣٨. المؤسسة العربية الحديثة. القاهرة شفيق عطا: جحا أكل السمك. نوادر جحا للأطفال ٥٩. المؤسسة العربية الحديثة. القاهرة.

بدوي، س. (١٩٧٣) المستويات العربيش المعاصرة في مصر. القاهرة.

بدوي، م.، م.ح. عبد اللطيف و م.البطل (١٩٩٢) الكتاب الاساسي في تعليم اللغة العربية لغير الناطقين بها، الجزء الثاني ، المنظمة العربية للتربية والثقافة والعلوم، تونس

بروستاد، ك.، م. البطل و ع. التونسي (١٩٩٥) ألف باء. مدخل الى حروف العربية وأصواتها
بروستاد، ك.، م. البطل و ع. التونسي (١٩٩٥) الكتاب في تعلم العربية. الجزء الأول.

باكلا، محمد بن حسن (٢٠٠٠) تجربة المملكة العربية السعودية في تعليم اللغة العربية للناطقين باللغات الأخرى. في: الحربي و عازر(إعداد): اللغة والثقافة العربية في عالم بلا حدود. جامعة كويت.

راموني، ر. (١٩٩٤) تنمية القدرات المعرفية والحضارية والاتصالية من خلال نصوص عربية أصيلة ومواد سمعية بصرية. i: Rammuny & Parkinson (eds.): Investigating Arabic. Greyden Press

דהאן, ח. (1997) מילון עברי-אנגלי אנגלי- עברי. אקדמון. האוניברסיטה העברית. ירושלים.

אלון, ק. ן ק. אלון (1991) כך לומדים עברית. תרגילי הפועל. שלב א'. הדפסה שניה. ישראל.

Resumé

Det overordnede formål med denne afhandling er at tilføre fremmedsprogspædagogikken for arabiskundervisningen et moderne læseteoretisk perspektiv ved teoretisk og empirisk at undersøge, hvilken indflydelse de fremmede elementer ved det arabiske skriftsystem har på visuel ordgenkendelse hos danske lørnere.

Denne problemstilling er relevant, fordi læseprocesteorier der relaterer til læsning på såvel modersmål som fremmedsprog, er udviklet i en snæver europæisk sammenhæng. Læseprocesser er traditionelt er blevet anset for at være universelle, og de pædagogiske metoder til undervisning i læsefærdighed på fremmedsprog er derfor blevet anvendt i arabiskundervisningen uden særlige hensyn til sprogspecifikke karakteristika. Inden for de seneste år har ny empirisk forskning imidlertid sat spørgsmålstegn ved teoriernes universalitet, og samtidig har arabisk i kraft af globaliseringen fået ny status: Arabisk er i dag et efterspurgt fremmedsprog, der tiltrækker mange studerende, hvorfor der er behov for målrettede undervisningsmetoder, som tager højde for, at lørnere skal tilegne sig et anderledes skriftsystem med fremmedartede tegn, der repræsenterer nye lyde og anvender et reduceret fonografisk kodningsprincip i forhold til det latinske alfabet.

Dette har ikke tidligere været eksplicit anerkendt. Derfor er målet med afhandlingen at dokumentere behovet for målrettede undervisningsmetoder, vise hvordan det manifesterer sig i lørnernes afkodningsfærdigheder, samt at pege på mulige pædagogiske retningslinjer, der kan imødegå det. Dette gøres på baggrund af en teoretisk og empirisk undersøgelse af, hvilken indflydelse de fremmede elementer i det arabiske skriftsystem har på visuel ordgenkendelse hos danske lørnere.

I afhandlingens første del defineres afhandlingens læseteoretiske udgangspunkt. I den forbindelse opstilles en læseprocesmodel, der i ordgenkendelsesfasen tager højde for, at bogstavgenkendelse er en selvstændig del af processen. Til dette formål anvendes en tilpasset konnektionistisk læseprocesmodel, der adskiller sig fra den originale ved at være opdelt i en primær fase, hvor bogstavdetektorer muliggør afkodning af selve skriften, dvs. skriftsystemets grafiske dimension, og en sekundær fase, hvor den leksikalske og fonologiske information i henhold til ortografien, dvs. skriftsystemets sprogspecifikke dimension, behandles og skaber mening. Denne model muliggør en tvedelt analyse af, hvad der karakteriserer ordgenkendelse på arabisk som fremmedsprog, hvor den grafiske og den ortografiske dimension betragtes som selvstændige problemfelter.

Derpå præsenteres det arabiske skriftsystem og den tilhørende fonologi med henblik på at beskrive hvilke forskelle, der er tale om, og der gives en kortfattet kontrastiv analyse af arabisk og hebraisk, der viser, at forskning i læsning på hebraisk – et andet semitisk sprog, der har været genstand for langt mere læseforskning – på visse områder kan generaliseres til arabisk. Mens det arabiske og det hebraiske alfabet er forskellige, er den morfologiske struktur nemlig den samme på de to sprog: Næsten alle ord er bygget op af en rod bestående af tre konsonanter, som bærer en verbal grundbetydning samt et mønster, der kan tilføre information om fx kausalt eller refleksivt aspekt og derved danne andre verber, eller inden for nomenformerne danne ord med relation til rodens semantiske værdi. Mønstrene består af forskellige kombinationer af præ- in- og suffixer, hvilket resulterer i en ikke-lineær morfologisk struktur. Kun lange vokaler markeres i skriften, mens korte vokaler

kan markeres som diakritiske tegn, men normalt udelades, hvilket på begge sprog resulterer i mange homografer.

Derefter gives en oversigt over forskning fra en række forskellige forskningsdiscipliner, som kan kaste lys over, hvad der nærmere bestemt karakteriserer de to nævnte faser i læsning på arabisk som fremmedsprog: Hvad angår den primære genkendelse, der retter sig mod skriftsystemets grafiske dimension, må læseretning anses for at have marginal indflydelse. Bogstavgenkendelse formodes derimod at spille en afgørende rolle, idet tilegnelse af en ny skrift generelt synes at være en kompleks opgave, der kræver et langt tilegnelsesforløb. Hvad angår den sekundære genkendelse, der retter sig mod skriftsystemets ortografiske dimension, gives først en oversigt over forskning i læsning på semitiske sprog som modersmål. Heraf fremgår det blandt andet, at vokalisering med diakritiske tegn medfører fonologiske afkodningsprocesser, mens uvokaliseret skrift giver anledning til en mere ortografisk afkodning pga. den manglende fonologiske information. Denne manglende fonologiske information og den udtalte homografi ved uvokaliseret skrift gør læsning stærkt kontekstafhængig og dermed forholdsvist kognitivt krævende, og vokalisering giver derfor større læsehastighed og bedre læseforståelse. På grund af den morfologiske struktur og den manglende vokalisering i almindelige tekster, har man ved læsning på arabisk i det hele taget færre ressourcer at trække på end ved læsning af latinsk skrift, og afkodningsprocessen forløber derfor tilsyneladende meget anderledes: På indoeuropæiske sprog aktiveres processer som tripel-bogstavanalyse og nabofrekvenseffekt, der i kraft af den konsonantiske morfologi ikke giver mening på arabisk. På arabisk trækker man i stedet på kendskabet til ords rødder og mønstre, og den stramme ordstruktur, der begrænser antallet af mulige vokalkombinationer ved et givent ord.

For lørnere af arabisk som fremmedsprog kompliceres læsetilegnelsesprocessen altså tilsyneladende både af et skriftsystem bestående af et sæt af meget fremmedartede grafemer - hvoraf nogle repræsenterer meget fremmede fonemer – og en anderledes morfologisk struktur, der kombineret med den manglende vokalinformation resulterer i et stort antal homografer, hvilket gør læsning stærkt kontekstafhængig. På den baggrund søges det gennem et empirisk studium af danske lørneres læsning på arabisk at afdække følgende:

- Hvilken rolle spiller de fremmede grafemer og fonemer?
- I hvor høj grad hæmmer de manglende vokaler læsehastighed og læseforståelse?
- I hvor høj grad formår lørnerne på samme måde som indfødte sprogbrugere at anvende rod-mønster-systematikken i den semitiske morfologi til at kompensere for manglende vokalinformation?

I undersøgelsen indgår lørnere på tre færdighedsniveauer. Alle danske universitets-studerende, der modtager kommunikativ undervisning i arabisk, er som udgangspunkt inddraget, men den endelige population begrænses af, at arabiskstuderende ved Århus Universitet ikke modtager kommunikativ undervisning og derfor falder uden for målgruppen, mens et stort antal af de studerende på både Syddansk Universitet og Københavns Universitet har arabisk som modersmål, og deres testresultater må udgå af analysen. Den samlede population udgøres af 27 respondenter på niveau 1 (2. semester), 8 respondenter på niveau 2 (4. semester), 3 respondenter på niveau 3 (færdiguddannede kandidater) og en kontrolgruppe af 25 indfødte sprogbrugere, der i sammenligning med de øvrige indgår som niveau 4. Med en så begrænset respondentgruppe, især på niveau 3, mangler undersøgelsen statistisk validitet. Alligevel kan den give værdifulde informationer

om, hvad der karakteriserer respondenternes læsning på arabisk. Endegyldige konklusioner vil imidlertid kræve en tilsvarende undersøgelse med en større population.

Undersøgelsen er konstrueret som tre deltest. I test A, hvor de fremmede grafemer og fonemers rolle afdækkes, omkoder respondenterne arabiske pseudoord med og uden fonemer, der er fremmede for danske lørnere, samt pseudoord skrevet med det latinske alfabet. De tre former for testmateriale er tilpasset i omfang af samme antal ord, stavelser og fonemer. I test B undersøges vokalers rolle for læsehastighed og –forståelse. Den ene historie læses vokaliseret af den ene halvdel af respondenterne og uvokaliseret af den anden halvdel og omvendt. I test C undersøges lørnernes kendskab til de morfologiske strukturer, som indfødte arabiske sprogbrugere støtter sig til som kompensation for de manglende vokaler, ved at respondenterne mundtligt applicerer vokaler på uvokaliserede pseudoord, hvis struktur definerer ét eller to valide vokaliseringsmønstre.

Ud over at teste de hypoteser, de tre deltest retter sig imod, undersøges de omkodningsfejl der opstår i test A kvalitativt. Analyserne suppleres desuden af en spørgeskemaundersøgelse, der primært har til formål at afdække hvilke indtryk og erfaringer, respondenterne har gjort sig med hensyn til læsning på arabisk.

Resultaterne af test A viser, at arabisk skrift i sig selv udgør en barriere for hurtig og korrekt ordgenkendelse. Ganske vist forbedres lørnernes færdigheder markant fra niveau 1 til niveau 2, men derefter stagnerer fremgangen i omkodningshastighed, og respondenter på niveau 3 læser i gennemsnit kun halvt så hurtigt som kontrolgruppen af indfødte sprogbrugere. De fremmede fonemer har derimod ingen indflydelse på omkodnings-hastigheden. Også antallet af omkodningsfejl falder markant fra niveau 1 til niveau 2, og også her flader udviklingen derefter ud – uden dog at gå helt i stå, som det er tilfældet for hastighed. På niveau 3 laver respondenterne over dobbelt så mange fejl som indfødte sprogbrugere ved omkodning af de fonetisk uproblematiske pseudoord. Grafemerne i sig selv synes altså at være en afgørende barriere også for korrekt omkodning. De fremmede fonemer giver også anledning til mange omkodningsfejl især på niveau 1, men tilegnelsen af disse fortsætter gennem hele studieforløbet, og på niveau 3 omkoder respondenterne bogstaver, der repræsenterer disse fonemer lige så korrekt som indfødte sprogbrugere, som også laver mange omkodningsfejl af denne type – tilsyneladende især pga. overførsel af dialektal fonologi. Den kvalitative analyse af de konkrete oplæsningsfejl i test A viser, at omkodningsfejl især finder sted mellem bogstaver, hvor den eneste forskel imellem dem er antallet og placeringen af prikker, ligesom der laves mange fejl mellem bogstaver, der på anden måde deler grafisk udtryk. De resterende fejl falder i fem afgrænsede fonetiske kategorier.

Resultatet af test B viser, at vokalisering ikke giver hurtigere læsehastighed og bedre læseforståelse. Tvært imod er læsehastigheden langsommere ved læsning af vokaliseret end ved uvokaliseret tekst. Dette er især tilfældet på niveau 1, men forskellen svinder med højere niveau, og kontrolgruppen af indfødte sprogbrugere læser de to testtyper lige hurtigt. Da andre studier af læsning på arabisk og hebraisk som modersmål viser, at vokalisering normalt giver både hurtigere læsning og bedre læseforståelse, er den mest sandsynlige forklaring på dette uventede resultat, at den ekstra grafiske information, vokalerne repræsenterer, lægger så tung en kognitiv byrde på lørnernes i forvejen belastede afkodningssystem, at vokalinformationen ikke kan udnyttes. Først når der er etableret så gode afkodningsfærdigheder, at der er overskud til at levne vokalerne opmærksomhed, kan

de anvendes positivt. Hvad angår læseforståelse ses samme tendens, men resultaterne er ikke entydige, og forskellene i læseforståelse ved læsning af de to teksttyper er ikke statistisk signifikante. På trods af at korte vokaler tilfører fuld fonologisk information og udrydder det store antal homografer, har de altså ikke positiv effekt på hverken læsehastighed eller læseforståelse, fordi den barriere, den grafiske dimension repræsenterer, overskygger vokalproblematikken hos lørnere i de første år af tilegnelsesforløbet.

I test C testes respondenternes kendskab til, hvilke valide vokaliseringsmønstre der passer med de enkelte ordstrukturer. Resultatet viser en jævn forbedret kompetence gennem tilegnelsesforløbet. Respondenterne på niveau 3 har dog ikke lige så gode færdigheder som kontrolgruppen af indfødte sprogbrugere, selvom også disse laver fejl – tilsyneladende igen i kraft af overførsel af dialektale træk.

Sammenfattende kan det konkluderes, at især grafiske, men også ortografiske elementer i det arabiske skriftsystem samt til en vis grad den fremmede fonologi, hæmmer ordgenkendelse hos danske lørnere, som derfor behøver målrettet undervisning, der kan træne deres afkodningsfærdigheder på arabisk. Indledende forløb, der introducerer lørnere til det arabiske alfabet og dets lyde, giver ikke automatisk adgang til gode læsefærdigheder på arabisk. Læsekompetence etableres primært gennem mange timers øvelse over flere år, hvor ord efterhånden genkendes automatisk og ikke kræver en langsommelig omkodning – det gælder for lørnere, der lærer et nyt skriftsystem lige så vel som det gælder for børn, der lærer at læse på deres modersmål.

De mest anvendte lærebogssystemer til undervisning i kommunikativt arabisk tilgodeser ikke i tilstrækkelig grad dette behov for øvelse i afkodning. De hyppigst forekommende former for læseøvelser fokuserer på topstyrede læsestrategier som skimning og gæt på baggrund af kontekst – strategier, der er værdifulde til at håndtere tekst, man ikke umiddelbart forstår, men ikke træner lørnerens ordgenkendelse. På den baggrund anbefales det blandt andet, at lørnere eksponeres for en stor mængde tekstmateriale uden nyt ordforråd og fremmede grammatiske strukturer, der er så let at forstå, at den kognitive kapacitet kan koncentreres om ordgenkendelse frem for om analytiske processer.

English summary

The overall aim of this thesis is to provide the teaching of Arabic as a foreign language with a modern, theoretical approach related to the acquisition of reading skills. This is done by a theoretical and empirical investigation of how aspects of the Arabic writing system which are foreign to learners, are influencing visual word recognition in Arabic. All learners included in the empirical study – except from respondents in the control group – are Danes, but the conclusions are applicable to all learners of Arabic, whose native language is written with the Roman alphabet.

The reason why this is relevant is that established reading theories related to reading in both native and foreign languages are developed within a narrow European context. Reading processes are traditionally considered universal, and the pedagogical methods used in teaching reading in foreign languages have been incorporated in the teaching of Arabic without special concern for the language specific features, the Arabic language and writing system represent. But during the last decades, new research has questioned the universal character of reading theory, and following the wake of globalisation, the Arabic language has gained a new status: Today Arabic is a popular foreign language to learn, and Arabic programmes within universities are attracting a large number of students. Therefore, there is a growing need for new teaching methods targeting the specific needs of learners, who are facing a different writing system with foreign signs representing new sounds, and which – compared to the Roman alphabet - applies a coding principle with reduced phonological information.

This need has not been explicitly recognised so far. Thus, the aim of the thesis is both to document this need, to give a characteristic of how it manifests itself in learners' decoding skills, and to give pedagogical directions to meet the requirements. This is done through a theoretical and empirical investigation of which role the foreign elements in the Arabic writing system plays in Danish learners' visual word recognition.

In the first part of the thesis, the theoretical framework for how reading processes are conceived in the thesis is defined. An adjusted connectionist model of word recognition considering letter recognition as a differentiated part of the process is developed. It is distinguished from the original model by a modulation into a primary recognition phase – where letter detectors are applied in order to decode the script, that is, the graphic dimension of the writing system – and a secondary recognition phase – where orthographic and phonological information, that is, the language specific dimension of the writing system – is processed and creates meaning in the readers' mind. This model makes it possible to conduct a separate analysis concerning the graphic and the orthographic dimensions of the Arabic writing system during word recognition in Arabic as a foreign language.

Then, the Arabic writing system and the matching phonology is presented in order to describe the differences that are at stake. A brief analysis of differences between Arabic and Hebrew is given to show how far research from Hebrew – another Semitic language which has been subject to much more reading research – can be generalised to cover Arabic as well. While the Arabic and Hebrew alphabets are different, the morphological structure is the same in the two languages: Almost all words are constructed on the base of three

consonants carrying a core meaning of action, and a pattern adding e.g. causal or reflexive aspect and this way constructs different verbs or nominal words related to the semantic value of the root. The patterns consist of different combinations of pre- in- and suffixes, which result in a non-linear morphological structure. Only long vowels are represented by letters, while short vowels can be marked with diacritics, but are normally left out. In both Arabic and Hebrew this results in a large number of homographs.

Then, a selection of research from a number of different research disciplines is presented in order to shed light on the nature of the two different word recognition phases in relation to reading in Arabic as a foreign language: In the primary phase, which relates to the writing system's graphic dimension, the reading direction seems to be of marginal importance. But letter architecture plays a significant role, as acquisition of a new script is apparently a complex process that requires much training. Regarding secondary recognition, which relates to the orthographic dimension of the writing system, an overview of research about native speakers' reading in Semitic languages is given. This research shows that vowelisation with diacritics prompts phonological decoding processes, while unvoweled script leads to a more orthographic decoding because of the lack of full phonological information. This lack of phonological information and the high number of homographs makes reading heavily context dependent, thus cognitively demanding, and vowelisation therefore results in increased reading speed and better comprehension. Because of the morphological structure and the lack of vowels in ordinary texts, readers of Arabic have generally less resources to draw on than readers of Roman script, thus the decoding process apparently takes quite a different course. In Indo-European languages processes like triple-letter analysis and neighbour-frequency effect are activated, but these processes are not relevant in Arabic because of the morphological structure. Instead, word recognition in Arabic draws on knowledge of words' roots and patterns, while the restricted word structure narrows down the number of possible vowel combinations applicable to a given word.

Thus, foreign language learners' acquisition of reading skills in Arabic is challenged by a writing system consisting of a set of very foreign graphemes – of which some represent very foreign phonemes – and a different morphological structure which, together with the lack of vowel information, results in a large number of homographs, which makes reading heavily context dependent. Based on this, the empirical study of Danish learners' reading in Arabic seeks to reveal:

- What roles do the foreign graphemes and phonemes play?
- To what degree do the missing vowels inhibit reading speed and comprehension?
- To what degree are learners able to apply the system of root and pattern in Semitic morphology to compensate for the lack of vowel information?

In the study, three different proficiency levels are included. At the starting point, all Danish university students receiving communicative teaching of Arabic were considered. But the final population was reduced, because students of Arabic at Aarhus University are not within the target group, because they are exclusively taught according to the philological tradition. Also, the results of a large number of the students at both University of Southern Denmark and Copenhagen University had to be excluded from the analysis because of Arabic being their mother tongue. The final population consists of 27 respondents at level 1 (2. semester), 8 respondents at level 2 (4. semester), 3 respondents at level 3 (students having finished their Masters degree) and a control group of 25 native speakers, which in

comparison with the others are regarded as level 4. With such a limited group of respondents, especially at level 3, the study lacks statistical validity. Even though, it can give valuable information about what characterises the respondents' reading in Arabic. But final conclusions would demand an equivalent study with a larger population.

The study is organised as three different tests. In test A, where the role of the foreign graphemes and phonemes is investigated, the respondents are decoding passages of Arabic pseudo words with and without phonemes which are foreign to Danish learners, and pseudo words written in the Roman alphabet. The proportions of the three types of material are adjusted through the same number of words, syllables and phonemes. In test B the vowels' role in reading speed and comprehension is tested. The respondents read two stories of same length and difficulty, of which only one is voweled. To counterbalance, one story is read voweled by half of the respondents and unvoweled by the other half and vice versa. In test C, learners' knowledge of the morphological structures, that native Arabic speakers use in order to compensate for the lack of vowels, is examined. The respondents are to orally apply vowels to unvoweled pseudo words having a structure that allows only one or two different vowel patterns.

Apart from testing the hypotheses targeted by the three tests, the recoding errors occurring in test A are investigated qualitatively. The analyses are also supplemented by a questionnaire in order to get information about the respondents' experiences and impressions from reading in Arabic.

The results of test A show, that Arabic script in itself is an obstacle to rapid and correct word recognition. Learners' skills are strongly improved from level 1 to level 2, but then the development in recoding speed stops, and respondents at level 3 recode with approximately half the speed of the control group of native speakers. The foreign phonemes, on the other hand, have no influence on recoding speed. Also, the number of recoding mistakes drops significantly from level 1 to level 2 and then the development flats out – but without stopping completely as is the case with recoding speed. At level 3, respondents make twice as many mistakes as native speakers when recoding the phonetically neutral pseudo words. Thus, graphemes seem to be an important obstacle to correct as well as to quick recoding. The foreign phonemes also result in many mistakes, especially at level 1, but the acquisition of these phonemes continue throughout the course of study, and at level 3, respondents decode letters which represent these phonemes approximately as correctly as the native speakers, who also make a number of these mistakes – apparently mostly because of transfer of dialectal phonology. The qualitative analysis of the actual recoding mistakes in test A show, that recoding mistakes are particularly confusions between letters, where the only difference between them is the number and placement of dots, just as many confusions happen between other letters of graphic similarity. The rest of the confusions fall in five demarcated phonetic categories.

The result of test B shows, that vowelisation does not improve neither reading speed nor comprehension. On the contrary, reading speed is slowed down when text is voweled compared to unvoweled. This is particularly the case at level 1, but the difference is diminishing with higher level, and the control group of native speakers reads the two types of text at the same speed. As other studies of reading in Arabic and Hebrew as a native language show that vowels normally improve both reading speed and comprehension, the most probable explanation to this unexpected result is, that the extra graphical information

that vowels represent, adds a heavy cognitive burden on the decoding system, which is already heavily charged, and because of this cognitive overload, vowel information cannot be utilized. Only when good decoding skills have been established to a point, where a cognitive surplus is available and allows the reader to pay attention to vowels, these can be used advantageously. Regarding reading comprehension, the same tendency is apparent, but the results are not unequivocal and difference in comprehension when reading the two types of text is not statistically significant. So, despite the fact that short vowels add full phonological information and eradicate the large number of homographs, they do not positively affect neither reading speed nor comprehension, because the obstacle that the graphic dimension represents overshadows the vowel issue through the learners' first years of acquisition.

In test C, the respondents' knowledge of what valid vowel patterns fit with what word structures is tested. The result shows a steady improvement in competence over the levels. The respondents at level 3 do not achieve as good results as the control group of native speakers, though native speakers also make a number of mistakes – apparently again because of transfer of dialectal features.

The conclusion is, that especially graphical, but also orthographical, elements of the Arabic writing system and to a certain degree also the foreign phonology, inhibit Danish learners' word recognition, and teaching should target these issues by offering explicit training of decoding skills in Arabic. Introductory courses, which introduce learners to the Arabic alphabet and its sounds, do not automatically give them access to good reading skills in Arabic. Reading competence is primarily established through years of training, where words little by little are recognised automatically so they do not demand a slowly progressing recoding process – this goes for learners learning a new writing system as well as for children learning to read in their mother tongue.

The most widely used materials for teaching Arabic communicatively do not sufficiently consider this need for training of decoding skills. The most common kind of reading practice is focused on the establishment of top-down reading strategies like skimming and guessing from context – strategies which are valuable to master in order to handle text, which is not immediately and fully understood, but which do not train learners' word recognition. In this light it is – among other things – recommended, that learners are exposed to a multitude of text material without new vocabulary and foreign grammatical structures, which are so easily understood, that learners' cognitive capacity can be directed to word recognition instead of analytical processes.

APPENDIX 1

TESTMATERIALE

Læsetest

Denne test indeholder følgende:

- Tre tekster med arabiske og latinske pseudo-ord. Skal læses højt.
- Ni pseudo-ord med arabiske bogstaver. Skal læses højt.
- To små historier på arabisk med spørgsmål på dansk om tekstens indhold. Du bestemmer selv, om du vil læse højt.

Fremgangsmåde:

Start:
1. Tænd for båndoptageren (*i Odense: tryk "attendance"*).
2. Isæt båndet – side A.
3. Påbegynd optagelse (*i Odense: tryk "drill"*).

Test A:
1. Vend det første testark og læs med det samme.
2. Læs teksten højt.
3. Gentag med de næste to testark – hold evt. en lille pause mellem hvert ark.

Test C:
1. Vend testarket.
2. Læs de ni ord højt i den rækkefølge, tallene angiver.
3. Sluk båndet og vend det til side B.
4. Spol båndet tilbage.

Test B:
1. Påbegynd optagelse (*i Odense: tryk "drill"*).
2. Vend testarket.
3. Nulstil timeren på pulten foran dig.
4. Læs teksten så hurtigt som muligt (men langsomt nok til at du forstår den).
5. Når du er færdig, så sluk båndet og skriv din tid fra timeren på **svararket**.
6. Aflever historien til Gunna eller læg den uden for rækkevidde.
7. Udfyld svararket. Har du ingen anelse om det rigtige svar, så lad være med at gætte – bare spring det over.
8. Gentag med den næste historie (husk at nulstille timeren)

Afslutning:
1. Udfyld spørgeskemaet.
2. Skriv navn/ID på båndet og de to testark og aflever det hele samlet i kuverten.

Læsetest

Denne test indeholder følgende:

- Tre tekster med arabiske og latinske pseudo-ord. Skal læses højt.
- Ni pseudo-ord med arabiske bogstaver. Skal læses højt.
- To små historier på arabisk med spørgsmål på dansk om tekstens indhold. Du bestemmer selv, om du vil læse højt.

Fremgangsmåde:

Start:
1. Tænd for båndoptageren (*i Odense: tryk "attendance"*).
2. Isæt båndet – side A.
3. Påbegynd optagelse (*i Odense: tryk "drill"*).

Test A:
1. Vend det første testark og læs med det samme.
2. Læs teksten højt.
3. Gentag med de næste to testark – hold evt. en lille pause mellem hvert ark.

Test C:
1. Vend testarket.
2. Læs de ni ord højt i den rækkefølge, tallene angiver.
3. Sluk båndet og vend det til side B.
4. Spol båndet tilbage.

Test B:
1. Påbegynd optagelse (*i Odense: tryk "drill"*).
2. Vend testarket.
3. Nulstil timeren på pulten foran dig.
4. Læs teksten så hurtigt som muligt (men langsomt nok til at du forstår den).
5. Når du er færdig, så sluk båndet og skriv din tid fra timeren på **svararket**.
6. Aflever historien til Gunna eller læg den uden for rækkevidde.
7. Udfyld svararket. Har du ingen anelse om det rigtige svar, så lad være med at gætte – bare spring det over.
8. Gentag med den næste historie (husk at nulstille timeren)

Afslutning:
1. Udfyld spørgeskemaet.
2. Skriv navn/ID på båndet og de to testark og aflever det hele samlet i kuverten.

اختبار في القراءة

هذا الاختبار يتألف من التالي:

3 نصوص بكلمات بدون معنى مكتوبة بالعربية وبالأوروبية. اقرأ بصوت عالٍ.

9 كلمات بدون معنى مكتوبة بالعربية. اقرأ بصوت عالٍ.

4 قصص قصيرة بالعربية وأسئلة حول المعنى لكل منها. اقرأ بدون صوت أو بصوت عالٍ كما تريد.

يجب تسجيل وقت قراءتك لكل قصة.

تعليمات الإجراءات:

لتكون جاهزا: أ. افتح المسجل

ب. أعد الشريط للتسجيل

ج. ابدأ التسجيل

اختبار A: أ. اِقلب ورقة الاختبار الأول واِبدأ القراءة في الحال. اِقرأ النص بصوت عالٍ.

ب. ردد العملية للاختبارات التالية. لك الحق في الاستراحة بعد كل ورقة.

اختبار C: أ. اقلب ورقة الاختبار.

ب. اقرأ الكلمات في الأرقام المتتابعة.

ج. توقف عن التسجيل و اقلب الشريط إلى الجانب B. أرجع الشريط إلى البداية من هذا الجانب

اختبار B: أ. اِبدأ التسجيل و اقلب ورقة الاختبار.

ب. اضبط موقّت المسجل على صفر واِقرأ النص بسرعة كما تستطيع

ج. عندما تنتهي: توقف عن التسجيل و اكتب وقت قراءتك على ورقة الإجابة.

د. ضع القصة بعيدا عن النظر و لا تنظر إليها مرة ثانية حينما تجيب عن الأسئلة.

ه. عندما تنتهي من إجابة الأسئلة عن القصة الأولى، ردد العملية مع القصص الثلاث الأخرى.

لا تنس وضع الموقّت على الصفر قبل قراءة كل قصة.

الختام: المرجو ملء استمارة الأسئلة.

كذلك المرجو كتابة الاسم (أو أي اسم أخر للتعريف) على شريط الكاسيت و أوراق الاختبار الأربعة ثم وضعها في الظرف.

وَلاَلَ لُبْمْ نَشِيهُمْ بَلَثَ تَالُوتَ مَنِكًا سَالُوا أَكى غَهُ الْنلْكُ وَمَبْنُ أَهَس بِالْمُمْكِ وَوَمْ كُولَ

سَفَةً الْمَاشَ غِمْلَ اسْبَفَاهُ مَلْيُنْمْ وَبَادَهُ بَسْغَةً فُسِي الْدلْمِ وَالْغِسْمِ يُوْنِي لُلْبُهُ يَشَأَ وَاسِكٌ دَلِيمٌ

نَلُبوا إِسْتَنْبَهَ يُتَينُ مَدِكًا مِسْنُمْ يَتُوسُ شَالُوتٍ بِالسمْلِكَ اِمْتِلاَسٍ.

وَمَاقَ لَظُمْ نَذْيُهُمْ وَلَثْ قَامُوتَ جَلِذًا قَانُوا أَزى ذُهُ الشْبطُ كَجَيْنَا أَحَق ِبالْمُظْكِ وَطَمْ يُؤْقَ

ذَفَةً الصاضَ ظْيفَ ِاسْبَزَاحُ مَطْيُبمْ وَصَاقُهُ جَهْقَةً ظُسِي الْرلْمِ وَالْنثْكِ يُوْخي طُعْبَهُ يَزَاً وَاعِقٌ

جَضِيمٌ نَزَبُوا ِإسْتَقْذَخَ يُحَيجُ مَذِحًا مِقُنْمْ يَقُوشُ ضَاعُوتٍ ِبالضمْلِكَ ِاعْتِجَاضٍ.

Koruse katum nodikom salas fidura halfian kafo linna paso kirapso luhenso lotukko meljoski kafam kinsa ostafut hekaso lurki kentanaro mivutam lisumera kestaltaf lossa femosi dodakelti hamsi korfatu horan kasium rakifon karasu futiskana tekainim kafirak fidkom katono lakofit haldomsik hastikaf.

الْجَرِيدَة

قَرَأَ مُحْمُودٌ فِي الصفْحَةِ الأُولَى بِالْجَرِيدَةِ الْعُنْوَانَ التَالِي: "رِسَالَةٌ هَامةٌ مِنْ الْمَلِكِ السعُودِي إِلَى الْمَلِكِ الأُرْدُنِي عَنْ مُشْكِلَةِ فِلسْطِينَ".

وَفِي صَفْحَةِ أَخْبَارِ التجَارَةِ وَالْمَالِ قَرَأَ: "دَخَلَتْ الْمَرْأَةُ فِي الْخَلِيجِ الْعَرَبِي مَيْدَانَ الْعَمَلِ فِي الْبُنُوكِ وإِن الْمُدِيرَاتِ وَالْمُوَظفَاتِ فِي بَعْضِ الْبُنُوكِ الآنَ جَمِيعًا مِنَ النسَاءِ".

وَقَرَأَ مُحْمُودٌ فِي صَفْحَةِ الرِيَاضَةِ: "تَجْرِي الْيَوْمَ مُبَارَاةُ كُرَةِ الْقَدَمِ بَيْنَ فَرِيقِ جَامِعَةِ الْمَلِكِ عَبْدِ الْعَزِيزِ وَفَرِيقِ جَامِعَةِ الرِيَاضِ فِي الساعَةِ الثالِثَةِ وَالنصْفِ مَسَاءً عَلَى مَلْعَبِ الْجَامِعَةِ فِي جُدةَ ".

وَقَرَأَ فِي صَفْحَةِ الإِعْلاَنَاتِ: " الْمُوَاطِنُ الْبَاكِسْتَانِي رَشِيد أَحْمَد ضَاعَ جَوَازَ سَفَرِهِ رَقْمْ ٤٥٦٤٦. نَرْجُو الاِتصَالَ بِالسفَارَةِ الْبَاكِسْتَانِيةِ رَقْمْ الْهَاتِفِ ٢٤٣٥٦٧٥٤٩".

وَأَخِيرًا قَرَأَ قِصةَ الْيَوْمِ: دَخَلَ شُرَطِي مَكْتَبَ الضابِطِ ، وَهُوَ يَصِيحُ: سَرَقَ اللصُوصُ سَيَارَةَ الشرْطَةِ !!! فَسَأَلَ الضابِطُ: وَمَاذَا فَعَلْتَ أَنتَ؟ وَجَابَ الشرَطِي: الحَمْدُ للهِ كَتَبْتُ رَقَمَهَا !

الجريدة

قرأ محمود في الصفحة الأولى بالجريدة العنوان التالي: "رسالة هامة من الملك السعودي إلى الملك الأردني عن مشكلة فلسطين".

وفي صفحة أخبار التجارة والمال قرأ: "دخلت المرأة في الخليج العربي ميدان العمل في البنوك وإن المديرات والموظفات في بعض البنوك الآن جميعا من النساء".

وقرأ محمود في صفحة الرياضة: "تجري اليوم مباراة كرة القدم بين فريق جامعة الملك عبد العزيز وفريق جامعة الرياض في الساعة الثالثة والنصف مساء على ملعب الجامعة في جدة ".

وقرأ في صفحة الإعلانات: " المواطن الباكستاني رشيد أحمد ضاع جواز سفره رقم ٤٥٦٤٦. نرجو الاتصال بالسفارة الباكستانية رقم الهاتف ٢٤٣٥٦٧٥٤٩".

وأخيرا قرأ قصة اليوم: دخل شرطي مكتب الضابط ، وهو يصيح: سرق اللصوص سيارة الشرطة !!! فسأل الضابط: وماذا فعلت أنت؟ وجاب الشرطي: الحمد لله كتبت رقمها !

Navn: _____ Tid for læsning: _____

Spørgsmål:

Hvad læste Mahmud om på forsiden?

Et vigtigt officielt brev om Palæstina-problemet ☐

Et arabisk topmøde om Palæstina-problemet ☐

Den saudiske konge på officielt besøg i Jordan ☐

Hvad læste han på siden om handel og økonomi?

Om en strejke blandt de bankansatte i Den Arabiske Golf ☐

Om arbejdsløse kvinder i Den Arabiske Golf ☐

Om kvinder på arbejdsmarkedet i Den Arabiske Golf ☐

Hvilke sportsnyheder læste han om?

En international fodboldlandskamp ☐

En fodboldkamp mellem to saudiske universiteter ☐

En fodboldkamp mellem to lokale klubhold fra Riyadh ☐

Hvad læste han ellers om?

At den pakistanske ambassade søgte en medarbejder ☐

At den pakistanske ambassadør havde mistet sin hund ☐

At en pakistansk statsborger havde tabt sit pas ☐

Hvad læste han til sidst?

En lille humoristisk historie ☐

Dagens horoskop ☐

Vejrudsigten for den kommende uge ☐

الاسم _____ وقت للقراءة : _____ TEST BII · NIVEAU 1 · TEKST A

الأسئلة:

عن ماذا قرأ محمود في الصفحة الأولى ؟

- ☐ رسالة رسمية عن مشكلة فلسطين
- ☐ قمة عربية عن مشكلة فلسطين
- ☐ الملك السعودي في زيارة رسمية للأردن

ماذا قرأ في صفحة أخبار التجارة والمال ؟

- ☐ عن إضراب موظفي البنوك في الخليج العربي
- ☐ عن النساء بدون عمل في الخليج العرب
- ☐ عن النساء في ميدان العمل في الخليج العربي

عن أي أخبار رياضية قرأ محمود ؟

- ☐ مباراة دولية لكرة القدم
- ☐ مباراة كرة القدم بين جامعتين سعوديتين
- ☐ مباراة كرة القدم بين فريقين محليين من الرياض

ما هي الأشياء الأخرى التي قرأ عنها محمود ؟

- ☐ إن السفارة الباكستانية تطلب موظف
- ☐ إن السفير الباكستاني ضاع منه كلبه
- ☐ إن مواطنا باكستانيا ضاع جوازا سفره

ماذا قرأ أخيرا ؟

- ☐ قصة ممتعة قصيرة
- ☐ أبراج الحظ اليوم
- ☐ الحالة الجوية للأسبوع القادم

فَاطِمَة

فَاطِمَة اِمْرَأَةٌ مِصْرِيَةٌ تَسْكُنُ فِي الْقَاهِرَةِ مَعَ أُسْرَتِهَا. هِيَ مُوَظفَةٌ فِي مَكْتَبِ الْبَرِيدِ الْمَرْكَزِي فِي الْقَاهِرَةِ. زَوْجُهَا اِسْمُهُ أَحْمَد وَهُوَ مُهَنْدِسٌ فِي شَرِكَةٍ دَوْلِيةٍ كَبِيرَةٍ. لِفَاطِمَة وأَحْمَد أَرْبَعَةُ أَوْلاَدٍ: وَلَدَان وَبِنْتَان.

تَصْحُو فَاطِمَة كُل يَوْمٍ قَبْلَ طُلُوعِ الشمْسِ وَتُصَلي وَبَعْدَ ذَلِكَ تُعِد الْفُطُورَ لِزَوْجِهَا وَأَوْلاَدِهَا. فِي الساعَةِ السابِعَةِ يَخْرُجُ أَحْمَد وَفَاطِمَة إِلَى عَمَلِهِمَا، وَيَذْهَبُ الأَوْلاَدُ إِلَى مَدَارِسِهِمْ. بَعْدَ الظهْرِ تَرْجِعُ الأُسْرَةُ إِلَى الْبَيتِ وَتَتَنَاوَلُ الْغَدَاءَ ، ثُم تَسْتَرِيحُ قَبْلَ الرُجُوعِ إِلَى الْعَمَلِ وَالدرَاسَةِ.

تَذْهَبُ فَاطِمَة كُل يَوْمٍ تَقْرِيبًا إِلَى السوقِ. تَشْتَرِي هُنَاكَ اللَحْمَ وَالْخُضَرَ وَالْخُبْزَ وَالْحَلِيبَ. وَتَذْهَبُ مَرةً وَاحِدَةً فِي الأُسْبُوعِ إِلَى الْبقالِ لِتَشْتَرِي الْجُبْنَ وَالزَيْتُونَ وَالْبَيْضَ والْمَكَرُونَةَ والأُرْزَ.

تَرْجِعُ فَاطِمَة إِلَى الْبَيتِ فِي الْمَسَاءِ وَتُعِد الْعَشَاءَ. يُذَاكِرُ الأَوْلاَدُ دُرُوسَهُمْ وَيَقْرَأُ أَحْمَد الْجَرِيدَةَ. بَعْدَ الأَكْلِ تَجْلِسُ الأُسْرَةُ مَعًا أَمَامَ التِلِيفِزْيُون بَعْضَ الْوَقْتِ.

فاطمة

فاطمة امرأة مصرية تسكن في القاهرة مع أسرتها. هي موظفة في مكتب البريد المركزي قي القاهرة. زوجها اسمه أحمد وهو مهندس في شركة دولية كبيرة. لفاطمة وأحمد أربعة أولاد: ولدان وبنتان.

تصحو فاطمة كل يوم قبل طلوع الشمس وتصلي وبعد ذلك تعد الفطور لزوجها وأولادها. في الساعة السابعة يخرج أحمد وفاطمة إلى عملهما، ويذهب الأولاد إلى مدارسهم. بعد الظهر ترجع الأسرة إلى البيت وتتناول الغداء ، ثم تستريح قبل الرجوع إلى العمل والدراسة.

تذهب فاطمة كل يوم تقريبا إلى السوق. تشتري هناك اللحم والخضر والخبز والحليب. وتذهب مرة واحدة في الأسبوع إلى البقال لتشتري الجبن والزيتون والبيض والمكرونة والأرز.

ترجع فاطمة إلى البيت في المساء وتعد العشاء. يذاكر الأولاد دروسهم ويقرأ أحمد الجريدة. بعد الأكل تجلس الأسرة معا أمام التليفزيون بعض الوقت.

Navn: _____ Tid for læsning: _____

Spørgsmål:

Hvad laver Fatima?

Hun arbejder på Cairos hovedpostkontor ☐

Hun er hjemmegående husmor ☐

Hun er sekretær i et internationalt firma ☐

Hvor mange børn har hun?

To sønner og en datter ☐

To sønner og to døtre ☐

To døtre og en søn ☐

Hvad laver familien i frokostpausen?

De spiser på arbejdspladsen/i skolen ☐

De tilbereder frokosten sammen i familiens køkken ☐

De samles hjemme, spiser og hviler lidt ☐

Hvordan arrangerer Fatima sine indkøb?

Hun køber ind i et supermarked ☐

Hun går på markedet dagligt og til købmanden en gang om ugen ☐

Hun går til købmanden dagligt og på markedet en gang om ugen ☐

Hvad laver familien om aftenen?

Børnene laver lektier, far læser avis og mor laver mad ☐

Børnene spiller computerspil mens forældrene ser TV ☐

Børnene hygger sig med deres venner mens forældrene ser TV ☐

الاسم _____ وقت للقراءة : _____ TEST BII · NIVEAU 1 · TEKST B

الأسئلة:

ماذا تعمل فاطمة ؟

- ☐ هي تعمل في مكتب البريد المركزي في القاهرة
- ☐ هي تجلس في البيت
- ☐ هي سكرتيرة في شركة دولية

كم أولاد عندها ؟

- ☐ عندها ولدان وبنت
- ☐ عندها ولدان وبنتان
- ☐ عندها بنت

ماذا تفعل الأسرة خلال استراحة الغذاء ؟

- ☐ يأكلون في العمل/المدرسة
- ☐ يطبخون الغداء جميعا في مطبخهم
- ☐ يجتمعون في البيت، يأكلون ويستريحون قليلا

كيف تشتري فاطمة الأشياء المنزلية ؟

- ☐ تتسوق في السوبرماركت
- ☐ تذهب يوميا إلى السوق و إلى البقال مرة في الأسبوع
- ☐ تذهب إلى البقال يوميا و إلى السوق مرة في الأسبوع

ماذا تفعل الأسرة في الليل ؟

- ☐ الأولاد يفعلون واجبهم، الأب يقرأ الجريدة والأم تطبخ
- ☐ الأولاد يلعبون بالعاب الكمبيوتر والوالدان يشاهدان التلفزيون
- ☐ الأولاد يستمتعون بالجلسة مع الأصدقاء والوالدان يشاهدان التلفزيون

جُحَا أكَلَ السَمَكَ

عَادَ جُحَا مِنْ عَمَلِهِ، وَمَر بِالسوق، فَرَأَى بَائِعَ السمَكِ يَعْرِضُ سَمَكًا كَبِيرًا لِلْبَيْعِ. فَرَغِبَ جُحَا فِي أَنْ يَأْكُلَ سَمَكًا، فَأَخْرَجَ مَا فِي كِيسِهِ مِنْ نُقُودٍ وَاشْتَرَى سَمَكاً.

حَمَلَ جُحَا السمَكَ إِلَى بَيْتِهِ وَقَالَ لِزَوْجَتِهِ: لَقَدْ اشْتَرَيتُ سَمَكًا لَذِيذاً. فَفَرِحَتْ زَوْجَتُهُ كَثِيرًا لِأَنهَا تُحِب السمَكَ كَذَلِكَ، وَعِنْدَمَا جَلَسَتْ فِي الْمَطْبَخِ لِتَنظفَهُ قَالَتْ فِي نَفْسِهَا: إِن هَذَا السمَكَ لَذِيذٌ جِدًا - أَنا أُرِيدُ أَنْ أَكُلَهُ وَحْدِي... فَقَالَتْ لِجُحَا: لَقَدْ تَعِبْتَ كَثِيرًا فِي الْعَمَلِ الْيوم يَا زَوْجِي الْعَزِيزَ... لِمَ لاَ تَسْتَرِحُ قَلِيلاً حَتى يَكُونَ الْطَعَامَ جَاهِزًا ؟ قَالَ جُحَا: لاَ بَأْسَ، سَأَنامُ قَلِيلاً.

فَنَامَ جُحَا. وَجَلَسَتْ زَوْجَتُهُ تَأْكُلُ الْسَمَكَ فِي سُرُورٍ حَتى شَبِعَتْ وَلَمْ يَبْقَ سِوَى الْقَلِيلَ مِنَ الْسَمَكِ. ثُم أَخَذَتْ قَلِيلاً مِنَ الْزَيتِ فِي الْطَبَقِ وَلَطخَتْ بِهِ لِحْيَةَ وفَمَ وَيَدَيَ جُحَا، وَتَرَكَتْ بَعْضَ قِطَعِ الْخُبْزِ وَمَا تَبقى مِنَ الْسَمَكِ عَلَى الْطَاوِلَةِ وَانْتَظَرَتْ. وَعِنْدَمَا اسْتَيْقَظَ جُحَا مِنَ الْنَوْمِ قَالَتْ زَوْجَتُهُ: أَيْنَ الْطَعَامُ ؟ هَلْ أَكَلْتَهُ وَحْدَكَ ؟ فَأَجَابَ جُحَا: مَا أَكَلْتُ أَبدًا. فَقَالَتْ زَوْجَتُهُ: أَتَنكِرُ أَنكَ أَكَلْتَ الْسَمَكَ، وَيَدَيْكَ وَلِحْيَتكَ مُلَطخَةً بِالأَكْلِ ؟ فَظَن جُحَا أَنهُ أَكَلَ وَنَسِيَ فَقَالَ: وَهَلْ نِمْتُ دُونَ أَنْ أَغْسِلَ يَدَي ؟ وَاللهِ، مَا رَأَيتُ سَمَكً أَلَذ مِنْ هَذَا الْسَمَكِ !

ثُم غَسَلَ يَدَيْهِ وَعَادَ إِلَى الْنَوْمِ.

جحا أكل السمك

عاد جحا من عمله، ومر بالسوق، فرأى بائع السمك يعرض سمكا كبيرا للبيع. فرغب جحا في أن يأكل سمكا، فأخرج ما في كيسه من نقود واشترى سمكا.

حمل جحا السمك إلى بيته وقال لزوجته: لقد اشتريت سمكا لذيذا. ففرحت زوجته كثيرا لأنها تحب السمك كذلك، وعندما جلست في المطبخ لتنظفه قالت في نفسها: إن هذا السمك لذيذ جدا - أنا أريد أن أكله وحدي... فقالت لجحا: لقد تعبت كثيرا في العمل اليوم يا زوجي العزيز... لم لا تسترح قليلا حتى يكون الطعام جاهزا ؟ قال جحا: لا بأس، سأنام قليلا.

فنام جحا. وجلست زوجته تأكل السمك في سرور حتى شبعت ولم يبق سوى القليل من السمك. ثم أخذت قليلا من الزيت في الطبق ولطخت به لحية وفم ويدي جحا، وتركت بعض قطع الخبز وما تبقى من السمك على الطاولة وانتظرت. وعندما استيقظ جحا من النوم قالت زوجته: أين الطعام ؟ هل أكلته وحدك ؟ فأجاب جحا: ما أكلت أبدا. فقالت زوجته: أتنكر أنك أكلت السمك، ويديك ولحيتك ملطخة بالأكل ؟ فظن جحا أنه أكل ونسى فقال: وهل نمت دون أن أغسل يدي ؟ والله، ما رأيت سمك ألذ من هذا السمك !

ثم غسل يديه وعاد إلى النوم.

Navn: _____ **Tid for læsning: _____**

Spørgsmål:

Hvem købte Joha en fisk af?

En fisker, som han tilfældigt mødte på vej fra arbejde ☐

En fiskehandler på markedspladsen ☐

En nabo, der var fritidsfisker ☐

Hvad tænkte Johas kone mens hun gjorde fisken klar?

At det er hårdt arbejde at rense fisk ☐

At en enkelt fisk ikke var nok til et måltid til dem begge ☐

At fisken var så lækker, at hun havde lyst til at spise den for sig selv ☐

Hvad gjorde Joha imens?

Han tog på konens opfordring en lur efter en hård dag på arbejdet ☐

Han faldt i søvn fordi madlavningen trak ud ☐

Han lagde sig foran fjernsynet og faldt i søvn ☐

Hvad gjorde konen inden Joha vågnede?

Hun dækkede bord og vaskede op ☐

Hun smurte madrester på hans hænder og ansigt ☐

Hun gik i byen og købte ekstra fisk ☐

Hvad gjorde Joha efter at han vågnede?

Han satte sig til bords sammen med konen for at spise ☐

Han blev sur på konen, fordi hun havde spist det hele selv ☐

Han troede, at han allerede havde spist og lagde sig til at sove igen ☐

الاسم _____ **وقت للقراءة :** _____

الأسئلة:

من اشترى جحا السمك؟

- ☐ من بائع السمك في طريقه إلى العمل
- ☐ من بائع السمك في السوق
- ☐ من جاره الذي ِاصطاد السمك في وقت الفراغ

ماذا فكرت زوجة جحا عندما كانت تنظف السمك ؟

- ☐ إن تنظيف السمك عمل شاق
- ☐ إن سمكة واحدة لا تكفي لكلاهما
- ☐ إن السمكة لذيذة جدا وهي أرادت أكلها وحدها

ماذا فعل جحا في **نفس الوقت ؟**

- ☐ ذهب ليرتاح من العمل بطلب من زوجته
- ☐ نام لأن طهي السمك يأخذ وقتا طويلا
- ☐ جلس أمام التلفزيون ونام

ماذا فعلت زوجة جحا قبل أن يستيقظ؟

- ☐ وضعت الأكل على الطاولة وغسلت الأواني
- ☐ وضعت بقايا الأكل على يدي جحا و وجهه
- ☐ ذهبت إلى المدينة لتشتري سمكة أخرى

ماذا فعل جحا بعد أن استيقظ من النوم ؟

- ☐ جلس على الطاولة يأكل مع زوجته
- ☐ غضب كثيرا لأن زوجته أكلت كل السمكة
- ☐ اعتقد أنه أكل السمكة ونام بعد ذلك

جُحَا حَزِينٌ عَلى حِمَارِهِ

كَانَتْ زَوْجَةُ جُحَا مَعْرُوفَةً بِأَنَّهَا امْرَأَةٌ صَعْبَةٌ. وَفِي يَوْمٍ عَلِمَ النَاسُ أَن زَوْجَةَ جُحَا مَاتَتْ فَقَالُوا: لَقَدْ تَخَلصَ جُحَا مِنْ زَوْجَتِهِ الْمِسْكِينَة - لَقَدْ قَتَلَهَا جُحَا ، وَلاَبُد أَنْ يُعَاقَبَ عَلَى ذَلِكَ. قَالَ جُحَا لَهُم أَن مَوتَ زَوْجَتِهِ كَانَ قَضَاءً وقَدَرًا وأَنهُ بَرِئٌ.

عِنْدَمَا تَأَكدَ النَاسُ مِنْ بَرَاءَةِ جُحَا اجْتَمَعُوا وَذَهَبُوا إِلَى بَيْتِهِ وَقَدمُوا أَعْذَارَهُمْ. فَوَجَدُوهُ حَزِينًا جِدا وَتَعَجبُوا مِنْ ذَلِكَ. فَقَالَ جُحَا لَهُمْ: عَلَى الرغْمِ مِمَا عَانَيتُ مِنْ هَذِهِ الْمَرْأَةِ ومِنْ لِسَانِهَا فَإِني حَزِنْتُ عَلَيْهَا. فَالْحَيَاةُ مَعَ امْرَأَةٍ صَعْبَةٍ خَيْرٌ مِنَ الْوَحْدَةِ. قَالَ أَحَدُ الأَصْدِقَاءِ: يَا رَجُلُ، لاَ تَحْزَنْ وَلاَ تُفَكرْ فِي الْوَحْدَة، فَإِن النسَاءَ كَثِيرَاتٌ وَيَكْفِي أَنْ تُشِيرَ بِإِصْبَعِكَ إِلَى إِحْدَى بَنَاتِنَا أَو أَخَوَاتِنَا وَنَحْنُ نُقَدمُهَا إِلَيكَ وَسَنُعْطِيكَ إِيَاهَا لِلزوَاجِ. قَالَ جُحَا: سَأُفَكرُ فِي الأَمْرِ.

وَمَرتْ أَيامٌ فَمَاتَ حِمَارُ جُحَا. فَحَزِنَ جُحَا عَلَى حِمَارِهِ حُزْنًا شَدِيدًا حَتى إِنهُ مَرِضَ مِنْ شِدةِ حُزنِهِ عَلَيهِ، فَذَهَبَ النَاسُ لِزِيَارَتِهِ. وَقَالُوا فِي تَعَجبٍ: مَا لَكَ يَا جُحَا ؟ فَأَنتَ لَمْ تَحْزَنْ عَلَى زَوْجَتِكَ كَمَا حَزِنْتَ عَلَى حِمَارِكَ !

قَالَ جُحَا: عِنْدَمَا مَاتَتْ زَوْجَتِي قَدمَ لِي كُل مِنْكُمْ زَوْجَةً جَدِيدَةً. وَلَكِن عِنْدَمَا مَاتَ حِمَارِي لَمْ أَجِدْ وَاحِدًا مِنْكُمْ يَقُولُ إِن عِنْدِي لَكَ حِمَارٌ جَدِيدٌ - فَأَي أَصْدِقَاءٍ أَنْتُمْ ؟

جحا حزين على حماره

كانت زوجة جحا معروفة بأنها امرأة صعبة. وفي يوم علم الناس أن زوجة جحا ماتت فقالوا: لقد تخلص جحا من زوجته المسكينة - لقد قتلها جحا ، ولابد أن يعاقب على ذلك. قال جحا لهم أن موت زوجته كان قضاء وقدرا وأنه برئ.

عندما تأكد الناس من براءة جحا اجتمعوا وذهبوا إلى بيته وقدموا أعذارهم. فوجدوه حزينا جدا وتعجبوا من ذلك. فقال جحا لهم: على الرغم مما عانيت من هذه المرأة ومن لسانها فإني حزنت عليها. فالحياة مع امرأة صعبة خير من الوحدة. قال أحد الأصدقاء: يا رجل، لا تحزن ولا تفكر في الوحدة، فإن النساء كثيرات ويكفي أن تشير بإصبعك إلى إحدى بناتنا أو أخواتنا ونحن نقدمها إليك وسنعطيك إياها للزواج. قال جحا: سأفكر في الأمر.

ومرت أيام فمات حمار جحا. فحزن جحا على حماره حزنا شديدا حتى إنه مرض من شدة حزنه عليه، فذهب الناس لزيارته. وقالوا في تعجب: ما لك يا جحا ؟ فأنت لم تحزن على زوجتك كما حزنت على حمارك !

قال جحا: عندما ماتت زوجتي قدم لي كل منكم زوجة جديدة. ولكن عندما مات حماري لم أجد واحدا منكم يقول إن عندي لك حمار جديد - فأي أصدقاء أنتم ؟

Navn: _____ **Tid for læsning: _____**

Spørgsmål:

Hvad sagde folk, da Johas kone døde?

At det var trist, men hun var jo så gammel og syg ☐

At det var en skam, for hun var nu så smuk ☐

At det nok var Joha, der havde slået hende ihjel ☐

Hvordan tog Joha konens død?

Han var ganske godt tilfreds, for hun havde været en strid kone ☐

Han var trist, for ensomheden var værre end konen havde været ☐

Han var trist, for hun havde været den sødeste kone, man kan få ☐

Hvad gjorde Johas venner efter konens død?

De tilbød ham selskab for at han ikke skulle føle sig alene ☐

De tilbød ham at lade deres døtre og søstre lave mad til ham ☐

De tilbød ham ægteskab med deres døtre og søstre ☐

Hvordan reagerede Joha da hans æsel døde?

Han blev endnu mere trist end da konen døde ☐

Han var så trist over konens død, at han ikke tog sig særligt af det ☐

Han blev trist, men kom hurtigt over det ☐

Hvad kritiserede Joha sine venner for til sidst?

At de ikke besøgte ham ofte nok ☐

At de ikke tilbød ham et nyt æsel ☐

At de ikke tilbød ham hjælp til at finde en ny kone ☐

الاسم _____ **وقت للقراءة :** _____

الأسئلة:

ماذا قال الناس عندما ماتت زوجة جحا ؟

- ☐ للأسف ولكن كانت عجوز ومريضة
- ☐ للأسف لقد كانت جميلة جدا
- ☐ من المحتمل أن جحا كان السبب في موتها

كيف كان رد فعل جحا لموت زوجته ؟

- ☐ كان مسرورا فقد كانت امرأة صعبة
- ☐ كان حزينا لأن الوحدة أصعب من العيش مع الزوجة
- ☐ كان حزينا فلقد كانت أفضل الزوجات

ماذا فعل أصدقاء جحا بعد موت زوجته ؟

- ☐ عرضوا عليه صداقته حتى لا يحس بالوحدة
- ☐ عرضوا عليه بناتهم و أخواتهم لطهي الأكل له
- ☐ عرضوا عليه الزواج من بناتهم و أخواتهم

كيف كان رد فعل جحا عندما مات حماره ؟

- ☐ حزن اكثر مما حزن على زوجته
- ☐ كان حزينا على زوجته حتى أنه لم يبال بموت الحمار
- ☐ كان حزينا ولكن شُفِي بسرعة

على ماذا عاتب جحا أصدقاءه ؟

- ☐ لأنهم لم يزوروه كثيرا
- ☐ لأنهم لم يعرضوا عليه حمارا جديدا
- ☐ لأنهم لم يساعدوه في إيجاد زوجة جديدة

TEST C

1	انبفص	4	يتشارل	7	استفداك
2	استهفل	5	انكشال	8	مكادغة
3	مصبوم	6	طاهث	9	امتفال

Spørgeskema

Dette spørgeskema anvendes i forbindelse med læsetests i arabisk, som er en del af et forskningsprojekt ved Center for Mellemøststudier, Syddansk Universitet i Odense. Projektet har til formål at afdække særlige vanskeligheder ved at lære at læse på arabisk som fremmedsprog.

Det er frivilligt at medvirke i forskningsprojektet. Hvis du ikke ønsker at medvirke, skal du kun skrive dit navn under skemaets pkt. 1 og krydse "nej" under pkt. 5 om samtykke. Selvom du ikke vil medvirke kan du deltage i testforløbet, som under alle omstændigheder trener din læsefærdighed i arabisk og dermed vil være at betragte som en del af dit undervisningsforløb.

Alle oplysninger behandles naturligvis fortroligt og resultater vil udelukkende blive fremstillet i anonymiseret form.

1. Personlige oplysninger

Navn: _____ (evt. anden identifikation)

Fødselsår: _____

Køn: ☐ Mand ☐ Kvinde

2. Dit kendskab til arabisk

Hvornår påbegyndte du dette arabiskstudium? _____

Kunne du noget arabisk før du påbegyndte dette studium? ☐ ja ☐ nej

Hvis ja: Før jeg begyndte studiet…

skrev jeg arabisk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt
læste jeg arabisk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt
talte jeg standardarabisk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt
talte jeg dialekt	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt
forstod jeg standardarabisk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt
forstod jeg dialekt	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt

Hvis du talte/forstod dialekt, skriv venligst hvilken: _____

Har du studeret arabisk andre steder sideløbende med studiet her (fx kurser i udlandet)? ☐ ja ☐ nej

Hvis ja, uddyb nedenfor: Sted, kursusform, niveau og periode/varighed:

_____

_____

_____

3. Dit kendskab til andre sprog

Hvor godt behersker du andre sprog ?

Engelsk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	☐ intet
Tysk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	☐ intet
Fransk	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	☐ intet
_____	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	
_____	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	
_____	☐ meget godt	☐ godt	☐ lidt	☐ meget lidt	

4. Indtryk og erfaringer vedrørende læsning på arabisk

Sæt kryds ved det svar, der passer bedst til dine indtryk og erfaringer med at læse på arabisk:

	Meget enig	Enig	Mindre enig	Slet ikke enig
Det er svært at læse på arabisk	☐	☐	☐	☐
Jeg føler at jeg læser meget langsomt på arabisk	☐	☐	☐	☐
De tekster vi læser i undervisningen passer til mit niveau	☐	☐	☐	☐
De tekster vi læser i undervisningen er oftest for svære	☐	☐	☐	☐

	Meget enig	Enig	Mindre enig	Slet ikke enig
Læseretningen kan være forvirrende	☐	☐	☐	☐
De fremmede sætningskonstruktioner gør læsning svært	☐	☐	☐	☐
Jeg er stadig usikker på nogle af bogstaverne	☐	☐	☐	☐
Det er sværere at læse, når teksten ikke er vokaliseret	☐	☐	☐	☐
Jeg tror at mit største problem er mit ordforråd	☐	☐	☐	☐
Jeg staver mig gennem ordene	☐	☐	☐	☐
Læsning på arabisk er så svært at det kan være demotiverende	☐	☐	☐	☐
Læsning på arabisk er en motiverende udfordring	☐	☐	☐	☐
Jeg bruger ordbogen meget, når jeg læser	☐	☐	☐	☐
Jeg gætter mig til en del, når jeg læser noget jeg ikke forstår	☐	☐	☐	☐

Jeg synes at det sværeste ved at læse på arabisk er:

_____

_____

_____

5. Samtykke

Jeg giver hermed mit samtykke til at disse oplysninger og mit testresultat må indgå som empirisk materiale i Gunna Funder Hansens ph.d.–projekt om det arabiske skriftsystems indflydelse på læsefærdighed på arabisk hos danske arabiskstuderende.

☐ ja ☐ nej Underskrift: _____

Det udfyldte spørgeskema bedes afleveret sammen med lydbåndet og testbesvarelse. Check venligst at både bånd, besvarelser og spørgeskema er påført navn eller anden identifikation.

Tusind tak for din medvirken !

استمارة الأسئلة

إن جدول الأسئلة هذا سيستعمل في إطار مشروع اختبار في القراءة باللغة العربية و هو جزء من مشروع بحث بجامعة جنوب الدانمارك. الغرض من المشروع هو الإحاطة ببعض الصعوبات في القراءة باللغة العربية كلغة أجنبية. كل المعلومات ستعالج بشكل سري ولن يفصح عن أي اسم في النتائج.

1. المعلومات الشخصية.

الاسم (أو أي اسم أخر للتعريف): _____

سنة الميلاد: _____ الجنس: ذكر☐ أنثى ☐

2. الدراسة.

أي دراسة تتابع حاليا ؟ _____

في أي سنة بدأت هذه الدراسة ؟ _____

هل تابعت أي دراسة عليا أخرى من قبل ؟ نعم ☐ لا ☐

هل سكنت ببلد أجنبي من قبل ؟ نعم ☐ لا ☐

إذا كان الجواب نعم، المرجو إعطاء تفاصيل أكثر بذكر المكان، الزمان – الفترة – و إذا ما كانت هناك إقامة من أجل الدراسة:

_____

_____

_____

_____

3. معرفتك للغات الأجنبية.

إلى أي حد تتقن اللغات الأجنبية ؟

الفرنسية:	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐
الإنجليزية:	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐
الإسبانية:	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐
_____ :	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐
_____ :	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐
_____ :	جيد جدا ☐	جيد متوسط ☐	ضعيف ☐	لا شئ ☐

4. التعبير و الخبرة فيما يخص القراءة باللغات الأوربية.

أعتقد بأن الصعوبات في القراءة باللغات الأوربية هي:

_____

_____

_____

5. الموافقة.

أنني أوافق على استعمال هذه المعلومات و نتائج الاختبار فقط في إطار ما هو علمي للسيدة Gunna Funder Hansen من أجل بحثها في مشروع الدكتوراه حول تأثير نظم الكتابة العربية على قدرة القراءة بالعربية لدي الطلبة الدنمركيين بقسم اللغة العربية.

التوقيع نعم ☐ لا ☐

المرجو وضع جدول الأسئلة المملوء, شريط الكاسيت وأجوبة الاختبارات – مع كتابة الاسم أو أي اسم تعريفي أخر - داخل الظرف.

- ولكم جزيل الشكر على مساعدتكم -

APPENDIX 2

ANALYSEARK

FEJLREGISTRERING

testperson nr. _____

AI	UDELADT	ÆNDRET	TILFØJET
BOGSTAVFEJL			
ikke relateret fejl			
grafisk/fonetisk fejl			
repetition			
ombytning			
lang vokal			
hamza			
DIAKRITISK FEJL			
kort vokal (- endelse)			
enkel endelse			
tanwiin-endelse			
shadda			

AII	UDELADT	ÆNDRET	TILFØJET
BOGSTAVFEJL			
ikke relateret fejl			
grafisk/fonetisk fejl			
repetition			
ombytning			
lang vokal			
hamza			
DIAKRITISK FEJL			
kort vokal (- endelse)			
enkel endelse			
tanwiin-endelse			
shadda			

AIII	UDELADT	ÆNDRET	TILFØJET
vokalfejl			
konsonantfejl			
ombytning			
repetition			

Genkender: bestemt artikel _____ sol/måne-bogstaver _____ endelsen وا _____
Grafisk/fonetiske fejl:

TESTPERSON NR: _____ NIVEAU: _____ ANALYSEDATO: _____

TEST AI:
antal bogstavfejl: ____ antal diakritiske fejl: _____ antal realfejl: _____
tid: _____

TEST AII:
antal bogstavfejl: ____ antal diakritiske fejl: _____ antal realfejl: _____
tid: _____

TEST AIII:
antal vokalfejl: ____ antal konsonantfejl: _____ antal ombytninger _____ antal fejl i alt: _____
tid: _____

TEST BI:
antal korrekte besvarelser: _____ tid: _____
højtlæsning: ☐ slet ikke ☐ noget ☐ alt Tekst: ☐ A ☐ B

TEST BII:
antal korrekte besvarelser: _____ tid: _____
højtlæsning: ☐ slet ikke ☐ noget ☐ alt Tekst: ☐ A ☐ B

TEST C:
fejl i ord nr.: ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9
Antal fejl: _____

APPENDIX 3

STATISTISK MATERIALE

Appendix 3.1. DATASÆT: Side 1 af 4

ID	Hold	Niveau	Fødselsår	Køn	AI tid	AI realfejl	AI bogstavfejl	AI diakritfejl	AII tid	AII realfejl	AII bogstavfejl	AII diakritfejl	AIII tid	AIII fejl
1	AU 2000	1	1977	k	.	.	.	.	.	.	.	.	.	.
2	AU 2000	1	1946	k	.	.	.	.	.	.	.	.	.	.
3	AU 2000	1	1941	m	214	18	27	6	229	26	38	8	42	0
4	AU 2000	1	1977	k	.	.	.	.	.	.	.	.	.	.
5	AU 2000	1	1974	k	.	.	.	.	.	.	.	.	.	.
6	AU 2000	1	1968	k	226	11	19	24	228	34	33	16	40	0
7	AU 2000	1	1941	k	256	17	22	10	321	21	26	8	29	0
8	NEG 2001	1	1978	k	106	21	22	29	115	28	43	34	27	1
9	NEG 2001	1	1969	k	164	12	17	27	168	34	38	26	46	1
10	NEG 2001	1	1979	m	110	8	12	20	110	20	28	19	35	2
11	NEG 2001	1	1970	m	140	6	11	15	150	16	19	7	38	0
12	NEG 2001	1	1977	k	279	11	13	9	150	26	37	18	38	2
13	NEG 2001	1	1974	m	175	10	22	16	144	36	47	23	37	1
14	NEG 2001	1	1980	k	261	7	18	15	194	35	41	10	73	6
15	NEG 2001	1	1980	m	293	12	19	4	299	26	30	5	33	2
16	KBH 2000	2	1974	m	137	2	3	12	126	7	10	6	49	2
17	KBH 2000	2	1977	k	96	3	5	3	98	6	12	3	41	1
18	KBH 2000	2	1974	k	.	.	.	.	.	.	.	.	.	.
19	KBH 2000	2	1977	k	.	.	.	.	.	.	.	.	.	.
20	KBH 2000	2	1976	m	83	2	2	3	106	2	4	8	33	0
21	KBH 2000	2	1973	k	108	4	13	2	102	13	25	5	55	2
22	KBH 2000	2	1979	m	67	2	8	5	80	4	8	3	36	0
23	NEG 2001	1	1979	k	151	8	9	4	142	25	31	21	35	1
24	NEG 2001	1	1981	m	245	19	28	30	229	24	38	17	33	4
25	NEG 1999	2	1975	m	142	4	6	17	150	14	20	24	51	4
26	KBH 2001	1	1976	m	.	.	.	.	.	.	.	.	.	.
27	KBH 2001	1	1971	k	.	.	.	.	.	.	.	.	.	.
28	KBH 2001	1	1977	k	132	12	28	13	149	14	34	10	40	0
29	KBH 2001	1	1980	m	168	2	14	7	178	7	17	12	55	3
30	KBH 2001	1	1975	k	127	2	8	4	145	12	18	6	48	0
31	KBH 2001	1	1976	k	.	.	.	.	.	.	.	.	.	.
32	KBH 2001	1	1979	k	109	0	6	10	128	16	27	5	36	1

Appendix 3.1. DATASÆT: Side 2 af 4

ID	BI tid	BI forst	BII tid	BII forst	C korrekte	sp_1	sp_2	sp_3	sp_4	sp_5	sp_6	sp_7	sp_8	sp_9	sp_10	sp_11	sp_12	sp_13	sp_14
1	720	3	757	3	.	2	2	3	2	3	2	2	2	3	2	3	2	3	2
2	1400	2	900	2	.	1	1	3	1	2	2	1	1	1	1	1	3	4	2
3	1068	1	628	1	1	1	2	2	2	4	4	4	4	3	3	3	2	1	1
4	912	0	850	1	.	1	1	3	2	4	.	2	.	2	.	.	.	3	1
5	703	1	529	0	2	2	2	2	3	4	2	2	3	1	3	4	1	3	2
6	602	3	526	1	5	1	2	1	2	2	3	2	4	3	2	1	4	4	3
7	1252	1	845	2	5	1	1	3	2	4	3	4	1	1	1	3	1	3	2
8	467	1	322	2	1	2	1	3	4	4	4	3	2	3	1	3	2	3	1
9	392	4	495	3	1	2	2	2	3	4	2	3	2	2	3	2	2	2	2
10	421	2	524	2	3	2	2	1	4	3	3	3	2	3	2	3	1	3	2
11	443	2	552	2	4	1	1	2	3	3	2	3	3	1	2	2	2	1	3
12	1066	3	638	4	4	2	2	3	3	4	2	3	2	1	3	2	3	1	2
13	1080	2	732	2	2	2	1	1	4	4	1	4	2	1	3	3	2	1	3
14	903	0	528	4	1	2	1	2	3	4	3	4	1	1	2	3	2	2	2
15	1512	0	1082	0	2	2	3	2	3	4	3	4	4	2	4	4	2	2	3
16	643	4	555	5	5	2	1	3	3	4	3	2	1	3	1	3	1	4	2
17	358	5	449	5	5	2	3	2	4	4	4	3	1	3	.	4	2	4	1
18	844	0	633	4	.	2	1	2	3	4	2	2	1	1	3	4	2	3	2
19	536	3	420	4	.	2	3	2	4	3	2	1	3	1	1	3	2	2	3
20	438	5	358	5	9	3	3	1	4	4	3	3	4	2	3	4	1	3	2
21	1100	4	1050	3	3	2	2	2	3	4	3	3	2	2	2	3	3	3	3
22	300	5	234	5	8	2	2	2	3	3	2	2	2	3	2	3	3	2	3
23	485	4	391	3	3	3	2	1	4	3	3	3	1	3	2	3	2	3	3
24	752	2	564	2	2	3	2	3	2	3	3	3	2	2	2	3	2	4	2
25	705	5	472	5	4	2	2	2	3	4	4	4	4	1	3	4	3	4	1
26	701	2	1028	4	3	1	1	2	3	3	2	3	2	1	1	3	2	3	2
27	482	4	1025	2	.	1	2	1	3	4	2	4	1	2	1	4	1	2	4
28	1028	4	806	3	2	2	3	1	3	4	3	4	2	2	1	4	1	3	2
29	614	2	602	3	3	2	2	2	3	4	3	2	1	1	1	2	2	2	3
30	937	1	442	5	4	3	2	1	3	4	2	3	4	1	3	4	2	1	2
31	.	2	.	4	.	2	1	2	4	4	3	3	3	1	2	4	2	3	3
32	1342	1	1105	5	3	2	2	1	3	4	2	3	1	2	3	4	1	2	2

Appendix 3.1. DATASÆT: Side 3 af 4

ID	Hold	Niveau	Fødselsår	Køn	AI tid	AI realfejl	AI bogstavfejl	AI diakritfejl	AII tid	AII realfejl	AII bogstavfejl	AII diakritfejl	AIII tid	AIII fejl
33	KBH 2001	1	1975	k	149	15	30	14	168	13	28	9	86	3
34	KBH 2001	1	1973	m	117	1	6	1	130	5	6	7	44	0
35	KBH 2001	1	1979	k	243	28	48	24	254	21	35	19	61	3
36	NEG 1994	3	1974	m	110	6	9	21	110	6	7	25	67	2
37	NEG 1994	3	1974	k	109	3	4	13	102	4	4	9	43	2
38	KBH 1988	3	1966	m	96	0	1	0	94	5	9	0	50	1
39	AKH 2003	4	1986	k	36	1	3	7	39	0	0	6	37	1
40	AKH 2001	4	1983	k	35	0	0	2	49	1	2	2	61	0
41	AKH 2003	4	.	k	38	0	3	15	46	7	8	11	52	18
42	AKH 2003	4	.	k	49	3	4	9	50	6	7	8	57	8
43	AKH 2003	4	1976	m	46	2	2	5	48	3	3	10	53	11
44	AKH 2003	4	1983	k	37	0	0	1	42	6	6	1	40	2
45	AKH 2003	4	1986	k	31	2	2	3	34	8	10	8	39	3
46	AKH 2002	4	1984	k	.	.	.	.	.	.	.	.	.	.
47	AKH 2003	4	1984	m	.	.	.	.	.	.	.	.	.	.
48	AKH 2003	4	1984	k	.	.	.	.	.	.	.	.	.	.
49	AKH 2003	4	1985	m	.	.	.	.	.	.	.	.	.	.
50	AKH 2003	4	1984	m	.	.	.	.	.	.	.	.	.	.
51	AKH 2002	4	1985	m	.	.	.	.	.	.	.	.	.	.
52	AKH 2003	4	1984	k	52	2	2	4	46	7	8	2	56	0
53	AKH 2003	4	1983	k	48	0	0	4	49	9	9	5	67	1
54	AKH 2003	4	1984	m	55	3	8	10	57	5	6	2	74	10
55	AKH 2003	4	1985	k	.	.	.	.	.	.	.	.	.	.
56	AKH 2003	4	1982	k	.	.	.	.	.	.	.	.	.	.
57	AKH 2003	4	1985	k	32	0	0	4	41	6	6	2	42	4
58	AKH 2000	4	1983	m	32	1	1	6	36	7	7	11	43	3
59	AKH 2003	4	1985	k	41	1	1	0	44	2	3	2	54	4
60	AKH 2003	4	1987	m	52	1	1	10	56	6	4	14	48	3
61	AKH 2003	4	.	k	51	1	1	2	59	3	4	0	51	2
62	AKH 2003	4	1984	k	.	.	.	.	.	.	.	.	.	.
63	AKH 2003	4	1984	m	.	.	.	.	.	.	.	.	.	.

Appendix 3.1. DATASÆT: Side 4 af 4

ID	BI tid	BI forst	BII tid	BII forst	C korrekte	sp_1	sp_2	sp_3	sp_4	sp_5	sp_6	sp_7	sp_8	sp_9	sp_10	sp_11	sp_12	sp_13	sp_14
33	1028	2	701	4	1	2	1	.	2	4	3	3	2	1	2	3	2	1	2
34	614	3	508	2	2	1	1	1	4	.	.	.	.	.	.	.	.	.	.
35	1348	4	616	0	3	1	1	3	2	2	1	2	1	2	2	2	2	2	2
36	599	4	532	1	9	2	1	3	2	4	2	2	3	2	2	2	3	2	1
37	323	5	392	5	5	3	3	2	3	4	3	4	3	1	3	3	2	2	2
38	368	5	330	5	6	1	2	3	2	4	4	3	3	3	2	1	3	3	1
39	160	5	170	5	7														
40	170	5	160	5	7														
41	260	5	270	4	7														
42	120	4	120	5	9														
43	210	4	170	5	8														
44	130	5	140	5	9														
45	120	5	120	5	8														
46	140	5	130	4	9														
47	180	5	170	5	7														
48	140	5	160	5	8														
49	140	5	150	5	9														
50	100	5	140	4	7														
51	170	4	160	4	.														
52	.	5	.	5	7														
53	.	5	.	5	8														
54	.	5	.	4	8														
55	.	4	.	5	.														
56	.	4	.	5	.														
57	.	5	.	2	7														
58	.	5	.	5	8														
59	.	5	.	3	9														
60	.	5	.	5	7														
61	.	4	.	5	8														
62	.	4	.	4	.														
63	.	5	.	5	7														

```
                                                  ___  ____  ____  ____  ____tm
                                                 /    /    /    /    /    /
                                                /   /   /   /   /   /   /
                                               Statistics/Data Analysis

                                                User: Gunna Funder Hansen
                                  Project: Danske lørneres læsning på arabisk{space -13}

      log:  W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Stata datalog
  log type:  smcl
 opened on:   4 Jun 2004, 13:49:14

. do "C:\DOCUME~1\GUNNAF~1\LOKALE~1\Temp\STD00000000.tmp"

. *               ****************************************
. *                                ANALYSE
. *               ****************************************
.
.
.
.
.
.
. ****************************************
. * RESPONDENTERNE
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. * stud= ID nr.
. * hold= universitetsby - immatrikuleret år
. * niveau
. * fodselsaar
. * kon= m eller k
.
. tabulate  niveau, missing
```

Niveau	Freq.	Percent	Cum.
1	27	42.86	42.86
2	8	12.70	55.56
3	3	4.76	60.32
4	25	39.68	100.00
Total	63	100.00	

```
.
. * aldersdistribution:
.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)
(3 missing values generated)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

. tabulate aldersgrp niveau, mi
```

	Niveau				
aldersgrp	1	2	3	4	Total
fodt for/i 1970	6	0	1	0	7
1971-1975	6	4	2	0	12
1976-1980	14	4	0	1	19
1981-1985	1	0	0	18	19
1986-1990	0	0	0	3	3
.	0	0	0	3	3
Total	27	8	3	25	63


```
. tabstat fodsel, s(n mean sd min p50 max) by(niveau) missing
```

Summary for variables: fodselsaar
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
1	27	1972.296	11.2483	1941	1976	1981
2	8	1975.625	1.995531	1973	1975.5	1979
3	3	1971.333	4.618802	1966	1974	1974
4	22	1983.909	2.113654	1976	1984	1987
Total	60	1976.95	9.399648	1941	1978.5	1987

```
.
. * konsdistribution
.
. tab niveau kon, missing
```

	Kon		
Niveau	k	m	Total
1	18	9	27
2	4	4	8
3	1	2	3
4	16	9	25
Total	39	24	63

```
.
.
.
. ****************************************
. * KODNING AF NIVEAUER
. ****************************************
.
. gen niveau_grp = niveau

. replace niveau grp = 2 if niveau_grp == 3
(3 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. gen niveau_sproggrp = niveau

. replace niveau sproggrp = 1 if niveau == 2
(8 real changes made)

. replace niveau_sproggrp = 1 if niveau == 3
(3 real changes made)
```

```
. label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"

. label values niveau_sproggrp niveau_sproggrpl

.
.
. ****************************************
. * KONTROLANALYSER
. ****************************************
.
.
. * TEST A:
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. gen niveau_grp = niveau

. replace niveau_grp = 2 if niveau_grp == 3
(3 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. gen niveau_sproggrp = niveau

. replace niveau_sproggrp = 1 if niveau == 2
(8 real changes made)

. replace niveau sproggrp = 1 if niveau == 3
(3 real changes made)

. label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"

. label values niveau_sproggrp niveau_sproggrpl

.
. save help, replace
(note: file help.dta not found)
file help.dta saved

.
. rename ai_tid A_tid1

. rename aii_tid A_tid2

. rename aiii_tid A_tid3

.
. keep hold stud niveau* A_tid*

.
. reshape long A tid, i(stud) j(task)
(note: j = 1 2 3)
```

Data	wide	->	long
Number of obs.	**63**	->	**189**
Number of variables	**8**	->	**7**
j variable (3 values)		->	**task**
xij variables:			
A_tid1 A_tid2 A_tid3		->	**A_tid**


```
.
. label define taskl 1 "AI" 2 "A2" 3 "A3"

. label values task taskl

. sort stud task

.
. * tempfile atid
. save atid, replace
file atid.dta saved

.
. use "help", clear

.
. rename ai_realfejl A1_err1

. rename ai_bogstavfejl A1_err2

. rename ai_fejl_diakrit A1_err3

. rename aii_realfejl A2_err1

. rename aii_bogstavfejl A2_err2

. rename aii_diakritfejl A2_err3

. rename aiii_fejl A3_err1

.
. keep stud hold niveau* A1_err* A2_err* A3_err*

.
. reshape long A1_err A2_err A3_err, i(stud) j(type)
(note: j = 1 2 3)
(note: A3_err2 not found)
(note: A3_err3 not found)
```

Data	wide	->	long
Number of obs.	**63**	->	**189**
Number of variables	**12**	->	**9**
j variable (3 values)		->	**type**
xij variables:			
A1_err1 A1_err2 A1_err3		->	**A1_err**
A2_err1 A2_err2 A2_err3		->	**A2_err**
A3_err1 A3_err2 A3_err3		->	**A3_err**

```
. label define errl 1 "realfejl" 2 "bogstavfejl" 3 "diakritfejl"

. label values type errl

. rename A1_err A_err1
```

```
. rename A2_err A_err2

. rename A3_err A_err3

.
. reshape long A_err, i(stud type) j(task)
(note: j = 1 2 3)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                      189   ->     567
Number of variables                   9   ->       8
j variable (3 values)                     ->   task
xij variables:
                   A_err1 A_err2 A_err3   ->   A_err
-----------------------------------------------------------------------------

. label define taskl 1 "Uden fremmede fonemer" 2 "Med fremmede fonemer" 3 "Latinsk alphabe

. label values task taskl

. sort stud task

.
. save help, replace
file help.dta saved

.
. merge stud task using "atid"
(label taskl already defined)
(label niveau grpl already defined)
(label niveau_sproggrpl already defined)

.
. drop _merge

.
. sort task niveau niveau_grp niveau_sproggrp stud

.
. gen task_grp = task

. replace task_grp = 1 if task == 2
(189 real changes made)

. label define task_grpl 1 "Arabisk skrift" 2 "Latinsk skrift"

. label values task_grp task_grpl

.
. * Korrelationskoefficient - hastighed mod alle fejl:
. correlate A_err A_tid if task_grp==1 & niveau_sproggrp==1
(obs=174)

             |   A_err    A_tid
-------------+------------------
       A_err |  1.0000
       A_tid |  0.3667   1.0000
```



```
.
. * Korrelationskoefficient - hastighed mod bogstavfejl:
. correlate A_err A_tid if  niveau_sproggrp==1 & task_grp==1 & type == 2
(obs=58)

             |   A_err    A_tid
-------------+------------------
       A err |  1.0000
       A_tid |  0.4809   1.0000


.
. * Korrelationskoefficient - hastighed mod diakritiske fejl:
. correlate A_err A_tid if  niveau_sproggrp==1 & task_grp==1 & type == 3
(obs=58)

             |   A_err    A_tid
-------------+------------------
       A_err |  1.0000
       A_tid |  0.1440   1.0000


.
.
. * TEST B:
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. gen niveau_sproggrp = niveau

. replace niveau sproggrp = 1 if niveau == 2
(8 real changes made)

. replace niveau_sproggrp = 1 if niveau == 3
(3 real changes made)

. label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"

. label values niveau_sproggrp niveau_sproggrpl

.
. rename bi_tid B_tid1

. rename bii_tid B_tid2

. rename bi_forstaaelse B_forst1

. rename bii_forstaaelse B_forst2

. keep stud niveau_sproggrp B_tid* B_forst*

.
. reshape long B_tid B_forst, i(stud) j(task)
(note: j = 1 2)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                       63   ->     126
Number of variables                   6   ->       5
j variable (2 values)                     ->   task
xij variables:
                          B_tid1 B_tid2   ->   B_tid
                       B_forst1 B_forst2   ->   B_forst
-----------------------------------------------------------------------------
```

```
.
. label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"

. label values task taskl

.
. * Korrelationskoefficient - hastighed mod korrekte svar:
. correlate B_tid B_forst if niveau_sproggrp == 1
(obs=74)
```

	B_tid	B_forst
B tid	1.0000	
B_forst	-0.4214	1.0000

```
.
.
. * KONTROL AF KØN OG ALDER GENNEMFØRT UNDER HVER DELTEST
.
.
.
. ****************************************
. * TEST A - TID
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. gen niveau_grp = niveau

. replace niveau grp = 2 if niveau_grp == 3
(3 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. * AI:
. tabstat ai_tid, s(n mean sd min p50 max) by(niveau) missing

Summary for variables: ai_tid
     by categories of: niveau (Niveau)
```

niveau	N	mean	sd	min	p50	max
1	20	183.25	62.61316	106	166	293
2	6	105.5	29.6968	67	102	142
3	3	105	7.81025	96	109	110
4	15	42.33333	8.440266	31	41	55
Total	44	119.2727	76.61786	31	109	293

```
. * AII:
```



```
. tabstat aii_tid, s(n mean sd min p50 max) by(niveau) missing

Summary for variables: aii_tid
     by categories of: niveau (Niveau)
```

niveau	N	mean	sd	min	p50	max
1	20	181.55	59.87661	110	159	321
2	6	110.3333	24.41038	80	104	150
3	3	102	8	94	102	110
4	15	46.4	7.385314	34	46	59
Total	44	120.3409	73.22153	34	110	321

```
. * AIII:
. tabstat aiii_tid, s(n mean sd min p50 max) by(niveau) missing

Summary for variables: aiii_tid
     by categories of: niveau (Niveau)
```

niveau	N	mean	sd	min	p50	max
1	20	43.8	14.82033	27	39	86
2	6	44.16667	8.818541	33	45	55
3	3	53.33333	12.34234	43	50	67
4	15	51.6	10.5884	37	52	74
Total	44	47.15909	12.87316	27	43.5	86

```
.
. label var ai_tid "AI"

. label var aii_tid "AII"

. label var aiii_tid "AIII"

.
. label define niveaul 1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4"

. label values niveau niveaul

.
. rename ai_tid A_tid1

. rename aii_tid A_tid2

. rename aiii_tid A_tid3

.
. keep hold niveau stud fodselsaar kon niveau_grp A_tid*

.
. reshape long A tid, i(stud) j(task)
(note: j = 1 2 3)
```

Data	wide	->	long
Number of obs.	63	->	189
Number of variables	9	->	8
j variable (3 values)		->	task
xij variables:			
	A_tid1 A_tid2 A_tid3	->	A_tid

```
.
. label define taskl 1 "AI" 2 "AII" 3 "AIII"

. label values task taskl

. sort stud task

. save help, replace
file help.dta saved

.
.
. * ANALYSE PÅ TVÆRS AF NIVEAU:
.
. * AI:
. xi: regress A_tid i.niveau_grp if task == 1, robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    41) =   76.10
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6830
                                                       Root MSE      =  44.177

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Iniveau_g~2 |  -77.91667    16.1211    -4.83   0.000    -110.4739   -45.35945
_Iniveau_g~4 |  -140.9167   14.30392    -9.85   0.000     -169.804   -112.0293
       _cons |     183.25   14.13666    12.96   0.000     154.7004    211.7996
------------------------------------------------------------------------------

. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    41) =   76.10
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |         63   8.049937     7.83   0.000     46.74282    79.25718
------------------------------------------------------------------------------

.
. * AII:
. xi: regress A_tid i.niveau_grp if task == 2, robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    41) =   85.75
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6871
                                                       Root MSE      =  41.943
```



```
------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Iniveau_g~2 |  -73.99444   15.02556    -4.92   0.000    -104.3392   -43.64971
_Iniveau_g~4 |    -135.15   13.65285    -9.90   0.000    -162.7225   -107.5775
       _cons |     181.55   13.51881    13.43   0.000     154.2482    208.8518
------------------------------------------------------------------------------

. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    41) =   85.75
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  61.15556   6.830181     8.95   0.000     47.36173    74.94939
------------------------------------------------------------------------------

.
. * AIII:
. xi: regress A_tid i.niveau_grp if task == 3, robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    41) =    1.68
                                                       Prob > F      =  0.1994
                                                       R-squared     =  0.0732
                                                       Root MSE      =  12.692

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Iniveau_g~2 |   3.422222   4.755119     0.72   0.476    -6.180936    13.02538
_Iniveau_g~4 |        7.8   4.322364     1.80   0.078    -.9291908    16.52919
       _cons |       43.8   3.346101    13.09   0.000     37.04241    50.55759
------------------------------------------------------------------------------

. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    41) =    1.68
            Prob > F =    0.1994

.
```

```
. * AIII - alle med dansk som modersmaal i samme niveaugruppe:
. gen niveau_sproggrp = niveau

. replace niveau sproggrp = 1 if niveau_grp == 2
(33 real changes made)

. label define niveau_sproggrpl 1 "modersmål dansk" 4 "modersmål arabisk"

. label values niveau_sproggrp niveau_sproggrpl

.
. xi: regress A tid i.niveau sproggrp if task == 3, robust
i.niveau_spro~p   _Iniveau_sp_1-4     (naturally coded; _Iniveau_sp_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  1,    42) =    3.32
                                                       Prob > F      =  0.0755
                                                       R-squared     =  0.0630
                                                       Root MSE      =  12.609

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Iniveau_s~4 |   6.737931   3.697357     1.82   0.076    -.7236377     14.1995
       _cons |   44.86207   2.522352    17.79   0.000     39.77176    49.95238
------------------------------------------------------------------------------

. test _Iniveau_sp_4

 ( 1)  _Iniveau_sp_4 = 0

       F(  1,    42) =    3.32
            Prob > F =    0.0755

.
.
. * ANALYSE PÅ TVÆRS AF TESTTYPE:
.
. * Niveau 1:
. xi: regress A tid i.task if niveau grp == 1, cluster(stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  2,    19) =   52.42
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6357
Number of clusters (stud) = 20                         Root MSE      =  50.745

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Itask_2 |       -1.7   8.854393    -0.19   0.850    -20.23246    16.83246
     Itask 3 |   -139.45   14.45148    -9.65   0.000    -169.6973   -109.2027
       _cons |    183.25   14.24424    12.86   0.000     153.4365    213.0635
------------------------------------------------------------------------------
```



```
. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,    19) =   52.42
            Prob > F =    0.0000

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |     137.75   14.04124     9.81   0.000     108.3614    167.1386
------------------------------------------------------------------------------

.
. * Niveau 2/3:
. xi: regress A_tid i.task if niveau_grp == 2, cluster(stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      27
                                                       F(  2,     8) =   41.58
                                                       Prob > F      =  0.0001
                                                       R-squared     =  0.7092
Number of clusters (stud) = 9                          Root MSE      =  18.973

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Itask_2 |   2.222222   3.743856     0.59   0.569    -6.411126    10.85557
    _Itask_3 |  -58.11111   6.991691    -8.31   0.000    -74.23398   -41.98824
       _cons |   105.3333   8.257712    12.76   0.000     86.29101    124.3757
------------------------------------------------------------------------------

. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,     8) =   41.58
            Prob > F =    0.0001

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   60.33333   6.671353     9.04   0.000     44.94917     75.7175
------------------------------------------------------------------------------
```

```
.
. * Niveau 4:
. xi: regress A tid i.task if niveau grp == 4, cluster(stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      45
                                                       F(  2,    14) =   11.65
                                                       Prob > F      =  0.0011
                                                       R-squared     =  0.1627
Number of clusters (stud) = 15                         Root MSE      =   8.905

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Itask_2 |   4.066667   1.196417     3.40   0.004     1.500608    6.632726
     Itask 3 |   9.266667   2.136421     4.34   0.001       4.6845    13.84883
       _cons |   42.33333   2.230551    18.98   0.000     37.54928    47.11739
------------------------------------------------------------------------------

. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,    14) =   11.65
            Prob > F =    0.0011

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |       -5.2   2.086171    -2.49   0.026    -9.674391   -.7256085
------------------------------------------------------------------------------

.
.
. * ANALYSE PÅ TVÆRS AF SKRIFTSYSTEM - AI OG AII SAMMEN VS. AIII:
.
. gen task_grp = task

. replace task grp = 1 if task == 2
(63 real changes made)

. label define task_grpl 1 "Arabisk skrift" 3 "Latinsk skrift"

. label values task_grp task_grpl

.
. * Niveau 1 - arabisk vs. latinsk skrift
. xi: regress A_tid i.task_grp if niveau_grp == 1, cluster(stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  1,    19) =  106.58
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6357
Number of clusters (stud) = 20                         Root MSE      =  50.311
```



```
------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Itask_grp_3 |     -138.6    13.4253   -10.32   0.000    -166.6995   -110.5005
       _cons |      182.4   13.10013    13.92   0.000     154.9811    209.8189
------------------------------------------------------------------------------

. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,    19) =  106.58
            Prob > F =    0.0000

.
. * Niveau 2/3 - arabisk vs. latinsk skrift
. xi: regress A_tid i.task_grp if niveau_grp == 2, cluster(stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      27
                                                       F(  1,     8) =   84.59
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7084
Number of clusters (stud) = 9                          Root MSE      =  18.614

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Itask_grp_3 |  -59.22222   6.439222    -9.20   0.000    -74.07109   -44.37335
       _cons |   106.4444   7.267193    14.65   0.000     89.68627    123.2026
------------------------------------------------------------------------------

. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,     8) =   84.59
            Prob > F =    0.0000

.
. * Niveau 4 - arabisk vs. latinsk skrift
. xi: regress A_tid i.task_grp if niveau_grp == 4, cluster(stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      45
                                                       F(  1,    14) =   13.06
                                                       Prob > F      =  0.0028
                                                       R-squared     =  0.1315
Number of clusters (stud) = 15                         Root MSE      =  8.9632

------------------------------------------------------------------------------
             |               Robust
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_Itask_grp_3 |   7.233333   2.001247     3.61   0.003     2.941085    11.52558
       _cons |   44.36667   1.985113    22.35   0.000     40.10902    48.62431
------------------------------------------------------------------------------
```

```
. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,    14) =   13.06
            Prob > F =    0.0028

.
.
. * ALDERS OG KØNS INDFLYDELSE:
.
. drop if niveau == 4
(75 observations deleted)

. drop if task==3
(38 observations deleted)

.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

.
. xi: regress A tid i.aldersgrp if niveau==1
i.aldersgrp       _Ialdersgrp_1970-1985(naturally coded; _Ialdersgrp_1970 omitted)

      Source |       SS       df       MS              Number of obs =      40
-------------+------------------------------           F(  3,    36) =    2.58
       Model |  25224.375     3    8408.125           Prob > F      =  0.0688
    Residual | 117411.225    36  3261.42292           R-squared     =  0.1768
-------------+------------------------------           Adj R-squared =  0.1082
       Total |  142635.6    39  3657.32308           Root MSE      =  57.109

------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 Ialder~1975 |     -65.225   27.08912    -2.41   0.021    -120.1643   -10.28573
 Ialder~1980 |      -31.05   22.11817    -1.40   0.169    -75.90773    13.80773
 Ialder~1985 |        27.4   44.23634     0.62   0.540    -62.31545    117.1155
       _cons |       209.6   18.05941    11.61   0.000     172.9738    246.2262
------------------------------------------------------------------------------

. test _Ialdersgrp_1975 _Ialdersgrp_1980 _Ialdersgrp_1985

 ( 1)  _Ialdersgrp_1975 = 0
 ( 2)  _Ialdersgrp_1980 = 0
 ( 3)  _Ialdersgrp_1985 = 0

       F(  3,    36) =    2.58
            Prob > F =    0.0688

.
. xi: regress A tid i.kon
i.kon             _Ikon_1-2           (_Ikon_1 for kon==k omitted)

      Source |       SS       df       MS              Number of obs =      58
-------------+------------------------------           F(  1,    56) =    0.82
       Model |  3196.84729     1  3196.84729           Prob > F      =  0.3696
    Residual |  218857.429    56  3908.16837           R-squared     =  0.0144
-------------+------------------------------           Adj R-squared = -0.0032
       Total |  222054.276    57  3895.68905           Root MSE      =  62.515
```



```
------------------------------------------------------------------------------
       A_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     _Ikon_2 |  -14.85714    16.4271    -0.90   0.370    -47.76458     18.0503
       _cons |        166   11.41369    14.54   0.000     143.1356    188.8644
------------------------------------------------------------------------------

.
.
.
. ****************************************
. * TEST A - FEJL
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. label define niveaul 1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4"

. label values niveau niveaul

.
. rename ai_realfejl A1_err1

. rename ai_bogstavfejl A1_err2

. rename ai_fejl_diakrit A1_err3

. rename aii_realfejl A2_err1

. rename aii_bogstavfejl A2_err2

. rename aii_diakritfejl A2_err3

. rename aiii_fejl A3_err1

.
. keep stud kon fodselsaar niveau A1_err* A2_err* A3_err*

.
. reshape long A1_err A2_err A3_err, i(stud) j(type)
(note: j = 1 2 3)
(note: A3_err2 not found)
(note: A3_err3 not found)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                       63   ->     189
Number of variables                  11   ->       8
j variable (3 values)                     ->   type
xij variables:
                 A1_err1 A1_err2 A1_err3   ->   A1_err
                 A2_err1 A2_err2 A2_err3   ->   A2_err
                 A3_err1 A3_err2 A3_err3   ->   A3_err
-----------------------------------------------------------------------------
```

```
. label define errl 1 "realfejl" 2 "bogstavfejl" 3 "diakritfejl"

. label values type errl

. rename A1_err A_err1

. rename A2_err A_err2

. rename A3_err A_err3

.
. reshape long A err, i(stud type) j(task)
(note: j = 1 2 3)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                      189   ->     567
Number of variables                   8   ->       7
j variable (3 values)                     ->   task
xij variables:
                   A_err1 A_err2 A_err3   ->   A_err
-----------------------------------------------------------------------------

. label define taskl 1 "Uden fremmede fonemer" 2 "Med fremmede fonemer" 3 "Latinsk alphabe

. label values task taskl

. sort stud task

. save help, replace
file help.dta saved

.
. * AI - uden fremmede fonemer:
. bysort type: tabstat A_err if task == 1, s(n mean sd min p50 max) by(niveau) missing
```

-> type = realfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	11	7.254763	0	11	28
Niveau 2	6	2.833333	.9831921	2	2.5	4
Niveau 3	3	3	3	0	3	6
Niveau 4	15	1.133333	1.060099	0	1	3
Total	44	5.977273	6.787432	0	3	28

-> type = bogstavfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	18.95	10.13838	6	18.5	48
Niveau 2	6	6.166667	3.970726	2	5.5	13
Niveau 3	3	4.666667	4.041452	1	4	9
Niveau 4	15	1.866667	2.099887	0	1	8
Total	44	10.40909	10.66033	0	7	48

-> type = diakritfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	14.1	8.978923	1	13.5	30
Niveau 2	6	7	6.09918	2	4	17
Niveau 3	3	11.33333	10.59874	0	13	21
Niveau 4	15	5.466667	4.085981	0	4	15
Total	44	10	8.183932	0	8	30

```
.
. * AII - med fremmede fonemer:
. bysort type: tabstat A_err if task == 2, s(n mean sd min p50 max) by(niveau) missing
```

-> type = realfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	21.95	9.139129	5	22.5	36
Niveau 2	6	7.666667	4.844241	2	6.5	14
Niveau 3	3	5	1	4	5	6
Niveau 4	15	5.066667	2.65832	0	6	9
Total	44	13.09091	10.46993	0	7.5	36

-> type = bogstavfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	30.7	10.01105	6	32	47
Niveau 2	6	13.16667	7.859177	4	11	25
Niveau 3	3	6.666667	2.516611	4	7	9
Niveau 4	15	5.533333	2.79966	0	6	10
Total	44	18.09091	13.99139	0	11	47

-> type = diakritfejl

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	14	7.980239	5	11	34
Niveau 2	6	8.166667	7.985403	3	5.5	24
Niveau 3	3	11.33333	12.66228	0	9	25
Niveau 4	15	5.6	4.452928	0	5	14
Total	44	10.15909	8.014357	0	8	34

```
.
. * AII - latinsk skrift:
. tabstat A_err if task == 3 & type == 1, s(n mean sd min p50 max) by(niveau) missing
```

Summary for variables: A_err
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
Niveau 1	20	1.5	1.638356	0	1	6
Niveau 2	6	1.5	1.516575	0	1.5	4
Niveau 3	3	1.666667	.5773503	1	2	2
Niveau 4	15	4.666667	5.009515	0	3	18
Total	44	2.590909	3.452792	0	2	18

```
.
.
. * ANALYSE PÅ TVÆRS AF NIVEAU
.
. * Niveau 2 og 3 sammen:
.
. gen niveau_grp = niveau

. replace niveau grp = 2 if niveau_grp == 3
(27 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. * AI - bogstavfejl
. xi: regress A_err i.niveau_gr if task == 1 & type == 2, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =   28.67
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5640
Number of clusters (stud) = 44                         Root MSE      =  7.2089
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-13.28333	2.603173	-5.10	0.000	-18.53313	-8.033535
_Iniveau_g~4	-17.08333	2.35246	-7.26	0.000	-21.82752	-12.33915
_cons	18.95	2.289022	8.28	0.000	14.33375	23.56625

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =   28.67
            Prob > F =    0.0000
```



```
. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	3.8	1.353267	2.81	0.007	1.070876	6.529124

```
.
. * AII - bogstavfejl
. xi: regress A_err i.niveau_gr if task == 2 & type == 2, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =   57.06
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7125
Number of clusters (stud) = 44                         Root MSE      =  7.6826
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-19.7	3.238508	-6.08	0.000	-26.23107	-13.16893
_Iniveau_g~4	-25.16667	2.37323	-10.60	0.000	-29.95274	-20.38059
_cons	30.7	2.260272	13.58	0.000	26.14173	35.25827

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =   57.06
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	5.466667	2.429504	2.25	0.030	.567104	10.36623

```
.
. * AI - diakritfejl
. xi: regress A_err i.niveau_gr if task == 1 & type == 3, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =    7.23
                                                       Prob > F      =  0.0020
                                                       R-squared     =  0.2313
Number of clusters (stud) = 44                         Root MSE      =  7.3481
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-5.655556	3.169892	-1.78	0.081	-12.04825	.7371401
_Iniveau_g~4	-8.633333	2.285724	-3.78	0.000	-13.24293	-4.023732
_cons	14.1	2.027241	6.96	0.000	10.01168	18.18832

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =    7.23
            Prob > F =    0.0020

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	2.977778	2.655811	1.12	0.268	-2.378175	8.333731

```
.
. * AII - diakritfejl
. xi: regress A_err i.niveau_gr if task == 2 & type == 3, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =    7.79
                                                       Prob > F      =  0.0013
                                                       R-squared     =  0.2226
Number of clusters (stud) = 44                         Root MSE      =  7.2367
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-4.777778	3.461908	-1.38	0.175	-11.75938	2.203825
_Iniveau_g~4	-8.4	2.137848	-3.93	0.000	-12.71138	-4.088618
_cons	14	1.801761	7.77	0.000	10.3664	17.6336

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =    7.79
            Prob > F =    0.0013

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	3.622222	3.172148	1.14	0.260	-2.775023	10.01947


```
.
. * AI - realfejl
. xi: regress A_err i.niveau_gr if task == 1 & type == 1, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =   20.59
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4757
Number of clusters (stud) = 44                         Root MSE      =  5.0331
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-8.111111	1.728057	-4.69	0.000	-11.59607	-4.626151
_Iniveau_g~4	-9.866667	1.660713	-5.94	0.000	-13.21581	-6.517519
_cons	11	1.637964	6.72	0.000	7.69673	14.30327

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =   20.59
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	1.755556	.6150592	2.85	0.007	.5151705	2.995941

```
.
. * AII - realfejl
. xi: regress A_err i.niveau_gr if task == 2 & type == 1, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =   30.13
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6140
Number of clusters (stud) = 44                         Root MSE      =  6.6616
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-15.17222	2.455026	-6.18	0.000	-20.12325	-10.22119
_Iniveau_g~4	-16.88333	2.174752	-7.76	0.000	-21.26914	-12.49753
_cons	21.95	2.063412	10.64	0.000	17.78873	26.11127

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =   30.13
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	1.711111	1.497116	1.14	0.259	-1.308112	4.730334

```
.
. * AIII - fejl
. xi: regress A_err i.niveau_gr if task == 3 & type == 1, cluster (stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  2,    43) =    2.82
                                                       Prob > F      =  0.0704
                                                       R-squared     =  0.1913
Number of clusters (stud) = 44                         Root MSE      =  3.1798
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	.0555556	.546591	0.10	0.920	-1.04675	1.157861
_Iniveau_g~4	3.166667	1.346316	2.35	0.023	.4515627	5.881771
_cons	1.5	.3699044	4.06	0.000	.7540167	2.245983

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    43) =    2.82
            Prob > F =    0.0704

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-3.111111	1.355606	-2.29	0.027	-5.844952	-.3772703


```
.
.
. * Niveau 2 separat - Niveau 3 udeladt:
. use "help", clear

. drop if niveau==3
(27 observations deleted)

.
. * AI - bogstavfejl
. xi: regress A_err i.niveau if task == 1 & type == 2, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =   28.40
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5591
Number of clusters (stud) = 41                         Root MSE      =  7.4224
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-12.78333	2.76234	-4.63	0.000	-18.36623	-7.200437
_Iniveau_4	-17.08333	2.358783	-7.24	0.000	-21.85061	-12.31606
_cons	18.95	2.295174	8.26	0.000	14.31128	23.58872

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =   28.40
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	4.3	1.63056	2.64	0.012	1.004515	7.595485

```
.
. * AII - bogstavfejl
. xi: regress A_err i.niveau if task == 2 & type == 2, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =   57.37
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7091
Number of clusters (stud) = 41                         Root MSE      =  7.8183
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-17.53333	3.793718	-4.62	0.000	-25.20072	-9.865943
_Iniveau_4	-25.16667	2.379608	-10.58	0.000	-29.97603	-20.3573
_cons	30.7	2.266347	13.55	0.000	26.11954	35.28046

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =   57.37
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	7.633333	3.127647	2.44	0.019	1.312123	13.95454

```
.
. * AI - diakritfejl
. xi: regress A_err i.niveau if task == 1 & type == 3, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =    7.10
                                                       Prob > F      =  0.0023
                                                       R-squared     =  0.2635
Number of clusters (stud) = 41                         Root MSE      =  7.1663
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-7.1	3.115508	-2.28	0.028	-13.39668	-.8033234
_Iniveau_4	-8.633333	2.291867	-3.77	0.001	-13.26537	-4.001297
_cons	14.1	2.03269	6.94	0.000	9.99178	18.20822

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =    7.10
            Prob > F =    0.0023

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	1.533333	2.587545	0.59	0.557	-3.69629	6.762956


```
.
. * AII - diakritfejl
. xi: regress A_err i.niveau if task == 2 & type == 3, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =    7.68
                                                       Prob > F      =  0.0015
                                                       R-squared     =  0.2587
Number of clusters (stud) = 41                         Root MSE      =  6.8948
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-5.833333	3.580432	-1.63	0.111	-13.06966	1.40299
_Iniveau_4	-8.4	2.143594	-3.92	0.000	-12.73237	-4.067634
_cons	14	1.806603	7.75	0.000	10.34872	17.65128

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =    7.68
            Prob > F =    0.0015

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	2.566667	3.299524	0.78	0.441	-4.10192	9.235253

```
.
. * AI - realfejl
. xi: regress A_err i.niveau if task == 1 & type == 1, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =   22.25
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4724
Number of clusters (stud) = 41                         Root MSE      =  5.1824
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-8.166667	1.685891	-4.84	0.000	-11.57398	-4.759355
_Iniveau_4	-9.866667	1.665177	-5.93	0.000	-13.23212	-6.501218
_cons	11	1.642367	6.70	0.000	7.680653	14.31935

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =   22.25
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	1.7	.4693667	3.62	0.001	.7513746	2.648625

```
.
. * AII - realfejl
. xi: regress A_err i.niveau if task == 2 & type == 1, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =   30.08
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5994
Number of clusters (stud) = 41                         Root MSE      =  6.8886
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-14.28333	2.792339	-5.12	0.000	-19.92686	-8.639805
_Iniveau_4	-16.88333	2.180598	-7.74	0.000	-21.29049	-12.47618
_cons	21.95	2.068958	10.61	0.000	17.76848	26.13152

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =   30.08
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	2.6	1.997745	1.30	0.201	-1.437594	6.637594


```
.
. * AIII - fejl
. xi: regress A_err i.niveau if task == 3 & type == 1, cluster (stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      41
                                                       F(  2,    40) =    2.81
                                                       Prob > F      =  0.0720
                                                       R-squared     =  0.1873
Number of clusters (stud) = 41                         Root MSE      =  3.3001
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-5.89e-16	.6944275	-0.00	1.000	-1.40349	1.40349
_Iniveau_4	3.166667	1.349934	2.35	0.024	.4383481	5.894985
_cons	1.5	.3708986	4.04	0.000	.7503859	2.249614

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    40) =    2.81
            Prob > F =    0.0720

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-3.166667	1.424577	-2.22	0.032	-6.045844	-.2874889

```
.
. * test AIII: alle med dansk som modersmaal i samme niveaugruppe:
. use "help", clear

. gen niveau_sproggrp = niveau

. replace niveau_sproggrp = 1 if niveau == 2
(72 real changes made)

. replace niveau_sproggrp = 1 if niveau == 3
(27 real changes made)

. label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"

. label values niveau_sproggrp niveau_sproggrpl

.
. xi: regress A_err i.niveau_sproggrp if task == 3 & type == 1, cluster (stud) robust
i.niveau_spro~p   _Iniveau_sp_1-4     (naturally coded; _Iniveau_sp_1 omitted)

Regression with robust standard errors                 Number of obs =      44
                                                       F(  1,    43) =    5.78
                                                       Prob > F      =  0.0205
                                                       R-squared     =  0.1913
Number of clusters (stud) = 44                         Root MSE      =  3.1418
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_s~4	**3.149425**	**1.309434**	**2.41**	**0.021**	**.5087007**	**5.79015**
_cons	**1.517241**	**.2806754**	**5.41**	**0.000**	**.9512055**	**2.083277**

```
. test _Iniveau_sp_4

 ( 1)  _Iniveau_sp_4 = 0

       F(  1,    43) =    5.78
            Prob > F =    0.0205

.
.
. * ANALYSE PÅ TVÆRS AF TESTTYPE
.
. use "help", clear

. gen niveau_grp = niveau

. replace niveau_grp = 2 if niveau_grp == 3
(27 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. * Niveau 1 - bogstavfejl
. xi: regress A_err i.task if niveau_grp == 1 & type == 2, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      40
                                                       F(  1,    19) =   25.98
                                                       Prob > F      =  0.0001
                                                       R-squared     =  0.2636
Number of clusters (stud) = 20                         Root MSE      =  10.075
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Itask 2	**11.75**	**2.305333**	**5.10**	**0.000**	**6.924883**	**16.57512**
Itask 3	**(dropped)**					
_cons	**18.95**	**2.296647**	**8.25**	**0.000**	**14.14306**	**23.75694**

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    19) =   25.98
            Prob > F =    0.0001
```



```
.
. * Niveau 2/3 - bogstavfejl
. xi: regress A_err i.task if niveau_grp == 2 & type == 2, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      18
                                                       F(  1,     8) =    7.59
                                                       Prob > F      =  0.0249
                                                       R-squared     =  0.1969
Number of clusters (stud) = 9                          Root MSE      =  5.7118
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	**5.333333**	**1.936043**	**2.75**	**0.025**	**.8688093**	**9.797857**
_Itask_3	**(dropped)**					
_cons	**5.666667**	**1.30836**	**4.33**	**0.003**	**2.649583**	**8.68375**

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,     8) =    7.59
            Prob > F =    0.0249

.
. * Niveau 4 - bogstavfejl
. xi: regress A_err i.task if niveau_grp == 4 & type == 2, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      30
                                                       F(  1,    14) =   16.90
                                                       Prob > F      =  0.0011
                                                       R-squared     =  0.3703
Number of clusters (stud) = 15                         Root MSE      =  2.4746
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	**3.666667**	**.8920155**	**4.11**	**0.001**	**1.753484**	**5.57985**
_Itask_3	**(dropped)**					
_cons	**1.866667**	**.5517854**	**3.38**	**0.004**	**.6832047**	**3.050129**

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    14) =   16.90
            Prob > F =    0.0011

.
. * Niveau 1 - diakritfejl
```

```
. xi: regress A_err i.task if niveau_grp == 1 & type == 3, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      40
                                                       F(  1,    19) =    0.00
                                                       Prob > F      =  0.9502
                                                       R-squared     =  0.0000
Number of clusters (stud) = 20                         Root MSE      =  8.4943
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	-.1	1.581275	-0.06	0.950	-3.409646	3.209646
_Itask_3	(dropped)					
_cons	14.1	2.033994	6.93	0.000	9.842801	18.3572

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    19) =    0.00
            Prob > F =    0.9502

.
. * Niveau 2/3 - diakritfejl
. xi: regress A_err i.task if niveau_grp == 2 & type == 3, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      18
                                                       F(  1,     8) =    0.27
                                                       Prob > F      =  0.6148
                                                       R-squared     =  0.0025
Number of clusters (stud) = 9                          Root MSE      =  8.3208
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Itask_2	.7777778	1.485592	0.52	0.615	-2.648003	4.203559
Itask_3	(dropped)					
_cons	8.444444	2.571846	3.28	0.011	2.513758	14.37513

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,     8) =    0.27
            Prob > F =    0.6148

.
. * Niveau 4 - diakritfejl
. xi: regress A_err i.task if niveau_grp == 4 & type == 3, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      30
                                                       F(  1,    14) =    0.02
                                                       Prob > F      =  0.8923
                                                       R-squared     =  0.0003
Number of clusters (stud) = 15                         Root MSE      =  4.2734
```


A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	.1333333	.9673467	0.14	0.892	-1.941419	2.208086
_Itask_3	(dropped)					
_cons	5.466667	1.07367	5.09	0.000	3.163874	7.769459

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    14) =    0.02
            Prob > F =    0.8923

.
. * Niveau 1 - realfejl: arabisk og latinsk
. xi: regress A_err i.task if niveau_grp == 1 & type == 1, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  2,    19) =   53.03
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6136
Number of clusters (stud) = 20                         Root MSE      =  6.8029
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	10.95	2.111614	5.19	0.000	6.530342	15.36966
_Itask_3	-9.5	1.633007	-5.82	0.000	-12.91792	-6.082076
_cons	11	1.650429	6.66	0.000	7.545613	14.45439

```
. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,    19) =   53.03
            Prob > F =    0.0000

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	20.45	2.042101	10.01	0.000	16.17583	24.72417

```
.
. * Niveau 2/3 - realfejl: arabisk og latinsk
```

```
. xi: regress A_err i.task if niveau_grp == 2 & type == 1, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      27
                                                       F(  2,     8) =   16.10
                                                       Prob > F      =  0.0016
                                                       R-squared     =  0.4400
Number of clusters (stud) = 9                          Root MSE      =   2.651
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	3.888889	1.27999	3.04	0.016	.9372266	6.840551
_Itask_3	-1.333333	.5204165	-2.56	0.034	-2.533416	-.1332507
_cons	2.888889	.5868501	4.92	0.001	1.53561	4.242168

```
. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,     8) =   16.10
            Prob > F =    0.0016

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	5.222222	1.10775	4.71	0.002	2.667747	7.776697

```
.
. * Niveau 4 - realfejl: arabisk og latinsk
. xi: regress A_err i.task if niveau_grp == 4 & type == 1, cluster (stud) robust
i.task            _Itask_1-3          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      45
                                                       F(  2,    14) =   15.96
                                                       Prob > F      =  0.0002
                                                       R-squared     =  0.2318
Number of clusters (stud) = 15                         Root MSE      =   3.331
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	3.933333	.7505906	5.24	0.000	2.323477	5.54319
_Itask_3	3.533333	1.294164	2.73	0.016	.7576279	6.309039
_cons	1.133333	.2801576	4.05	0.001	.532455	1.734212


```
. test _Itask_2 _Itask_3

 ( 1)  _Itask_2 = 0
 ( 2)  _Itask_3 = 0

       F(  2,    14) =   15.96
            Prob > F =    0.0002

. lincom _Itask_2 - _Itask_3

 ( 1)  _Itask_2 - _Itask_3 = 0
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	.4	1.415432	0.28	0.782	-2.635799	3.435799

```
.
.
. * ANALYSE PÅ TVÆRS AF SKRIFTSYSTEMER - REALFEJL AI OG AII sAMMEN VS. AIII
.
. gen task_grp = task

. replace task_grp = 1 if task == 2
(189 real changes made)

. label define task_grpl 1 "Arabisk skrift" 3 "Latinsk skrift"

. label values task_grp task_grpl

.
. * Niveau 1 - arabisk vs. latinsk skrift
. xi: regress A_err i.task_grp if niveau_grp == 1 & type == 1, cluster (stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  1,    19) =   99.05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4380
Number of clusters (stud) = 20                         Root MSE      =  8.1336
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_grp_3	-14.975	1.504659	-9.95	0.000	-18.12429	-11.82571
_cons	16.475	1.538525	10.71	0.000	13.25483	19.69517

```
. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,    19) =   99.05
            Prob > F =    0.0000
```

```
.
. * Niveau 2/3 - arabisk vs. latinsk skrift
. xi: regress A_err i.task_grp if niveau_grp == 2 & type == 1, cluster (stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      27
                                                       F(  1,     8) =   32.98
                                                       Prob > F      =  0.0004
                                                       R-squared     =  0.2140
Number of clusters (stud) = 9                          Root MSE      =  3.0772
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_grp_3	-3.277778	.5707911	-5.74	0.000	-4.594024	-1.961531
_cons	4.833333	.8582929	5.63	0.000	2.854106	6.81256

```
. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,     8) =   32.98
            Prob > F =    0.0004

.
. * Niveau 4 - arabisk vs. latinsk skrift
. xi: regress A_err i.task_grp if niveau_grp == 4 & type == 1, cluster (stud) robust
i.task_grp        _Itask_grp_1-3      (naturally coded; _Itask_grp_1 omitted)

Regression with robust standard errors                 Number of obs =      45
                                                       F(  1,    14) =    1.48
                                                       Prob > F      =  0.2439
                                                       R-squared     =  0.0405
Number of clusters (stud) = 15                         Root MSE      =  3.6791
```

A_err	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Itask grp 3	1.566667	1.287949	1.22	0.244	-1.195708	4.329041
_cons	3.1	.3765555	8.23	0.000	2.292369	3.907631

```
. test _Itask_grp_3

 ( 1)  _Itask_grp_3 = 0

       F(  1,    14) =    1.48
            Prob > F =    0.2439

.
.
.
. * ALDERS OG KØNS INDFLYDELSE:
```



```
.
. drop if niveau == 4
(225 observations deleted)

. drop if task==3
(114 observations deleted)

.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

.
. xi: regress A_err i.aldersgrp if type == 1 & niveau==1
i.aldersgrp       _Ialdersgrp_1970-1985(naturally coded; _Ialdersgrp_1970 omitted)
```

Source	SS	df	MS
Model	320.925	3	106.975
Residual	3465.05	36	96.2513889
Total	3785.975	39	97.0762821

```
Number of obs =      40
F(  3,    36) =    1.11
Prob > F      =  0.3572
R-squared     =  0.0848
Adj R-squared =  0.0085
Root MSE      =  9.8108
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Ialder~1975	-7.75	4.653661	-1.67	0.105	-17.18806	1.688062
_Ialder~1980	-3.15	3.799698	-0.83	0.413	-10.85615	4.556146
_Ialder~1985	2	7.599397	0.26	0.794	-13.41229	17.41229
_cons	19.5	3.102441	6.29	0.000	13.20796	25.79204

```
. test _Ialdersgrp_1975 _Ialdersgrp_1980 _Ialdersgrp_1985

 ( 1)  _Ialdersgrp_1975 = 0
 ( 2)  _Ialdersgrp_1980 = 0
 ( 3)  _Ialdersgrp_1985 = 0

       F(  3,    36) =    1.11
            Prob > F =    0.3572

.
. xi: regress A_err i.kon
i.kon             _Ikon_1-2           (_Ikon_1 for kon==k omitted)
```

Source	SS	df	MS
Model	920.634921	1	920.634921
Residual	20159.3651	172	117.205611
Total	21080	173	121.849711

```
Number of obs =     174
F(  1,   172) =    7.85
Prob > F      =  0.0057
R-squared     =  0.0437
Adj R-squared =  0.0381
Root MSE      =  10.826
```

A_err	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Ikon_2	-4.603175	1.642434	-2.80	0.006	-7.845096	-1.361253
_cons	17.22222	1.141177	15.09	0.000	14.96971	19.47474

```
.
.
.
. ****************************************
. *TEST B - TID
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. rename bi_tid B_tid1

. rename bii_tid B_tid2

. rename bi_forstaaelse B_forst1

. rename bii_forstaaelse B_forst2

. keep stud fodselsaar kon niveau B_tid* B_forst*

. reshape long B_tid B_forst, i(stud) j(task)
(note: j = 1 2)
```

Data	wide	->	long
Number of obs.	63	->	126
Number of variables	8	->	7
j variable (2 values)		->	task
xij variables:			
	B_tid1 B_tid2	->	B_tid
	B_forst1 B_forst2	->	B_forst

```
.
. label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"

. label values task taskl

.
. for num 1/2: tabstat B_tid if task == X, s(n mean sd min p50 max) by(niveau) missing
```

-> tabstat B_tid if task == 1, s(n mean sd min p50 max) by(niveau) missing

Summary for variables: B_tid
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
1	26	856.6154	336.6868	392	827.5	1512
2	8	615.5	267.0987	300	589.5	1100
3	3	430	148.0777	323	368	599
4	13	156.9231	42.6975	100	140	260
Total	50	610.52	397.6057	100	567.5	1512

-> tabstat B_tid if task == 2, s(n mean sd min p50 max) by(niveau) missing

Summary for variables: B_tid
by categories of: niveau (Niveau)

niveau	N	mean	sd	min	p50	max
1	26	680.6154	217.477	322	622	1105
2	8	521.375	245.0259	234	460.5	1050
3	3	418	103.4795	330	392	532
4	13	158.4615	38.04518	120	160	270
Total	50	503.62	286.7562	120	516	1105

```
.
.
. * ANALYSE PÅ TVÆRS AF NIVEAU:
.
. * Niveau 2/3 samlet:
. gen niveau_grp = niveau

. replace niveau grp = 2 if niveau_grp == 3
(6 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

. save help, replace
file help.dta saved

.
. * Vokaliseret
. xi: regress B_tid i.niveau_gr if task == 1, robust cluster(stud)
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      50
                                                       F(  2,    49) =   66.69
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5515
Number of clusters (stud) = 50                         Root MSE      =  271.88
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-291.7063	99.47918	-2.93	0.005	-491.6172	-91.7954
_Iniveau_g~4	-699.6923	67.80513	-10.32	0.000	-835.9518	-563.4328
_cons	856.6154	66.78191	12.83	0.000	722.4121	990.8187

```
. quietly test _Iniveau_gr_2

. test _Iniveau_gr_4, accumulate

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    49) =   66.69
            Prob > F =    0.0000
```

```
. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

B_tid	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	407.986	74.6592	5.46	0.000	257.9527	558.0193

```
.
. * Uvokaliseret:
. xi: regress B_tid i.niveau_gr if task == 2, robust cluster(stud)
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      50
                                                       F(  2,    49) =   80.35
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5868
Number of clusters (stud) = 50                         Root MSE      =   188.2
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Iniveau_g~2	-187.4336	77.12977	-2.43	0.019	-342.4316	-32.4355
Iniveau_g~4	-522.1538	44.38586	-11.76	0.000	-611.3506	-432.9571
_cons	680.6154	43.13662	15.78	0.000	593.9291	767.3017

```
. quietly test _Iniveau_gr_2

. test _Iniveau_gr_4, accumulate

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    49) =   80.35
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

B_tid	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	334.7203	64.78866	5.17	0.000	204.5226	464.918

```
.
.
. * BI og BII sammen:
.
.
. * Niveau 2/3 samlet:
```



```
.
. xi: regress B_tid i.niveau_gr, robust cluster(stud)
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =     100
                                                       F(  2,    49) =   89.18
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5382
Number of clusters (stud) = 50                         Root MSE      =  239.63
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	-239.5699	82.38687	-2.91	0.005	-405.1325	-74.00732
_Iniveau_g~4	-610.9231	49.45238	-12.35	0.000	-710.3014	-511.5448
_cons	768.6154	48.28399	15.92	0.000	671.5851	865.6457

```
. quietly test _Iniveau_gr_2

. test _Iniveau_gr_4, accumulate

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    49) =   89.18
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

B_tid	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	371.3531	67.60508	5.49	0.000	235.4957	507.2106

```
.
.
. * Niveau 2 separat - Niveau 3 udeladt:
. drop if niveau == 3
(6 observations deleted)

.
. xi: regress B_tid i.niveau, robust cluster(stud)
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      94
                                                       F(  2,    46) =   86.08
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5444
Number of clusters (stud) = 47                         Root MSE      =   243.9
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	-200.1779	97.73218	-2.05	0.046	-396.9026	-3.45321
_Iniveau_4	-610.9231	49.51757	-12.34	0.000	-710.5968	-511.2494
_cons	768.6154	48.34764	15.90	0.000	671.2966	865.9341

```
. quietly test _Iniveau_2

. test _Iniveau_4, accumulate

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    46) =   86.08
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

B_tid	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	410.7452	85.60713	4.80	0.000	238.427	583.0634

```
.
.
.
. * ANALYSE PÅ TVÆRS AF TESTTYPE:
.
. use "help", clear

.
. * Niveau 1:
. xi: regress B_tid i.task if niveau_gr == 1, robust cluster(stud)
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      52
                                                       F(  1,    25) =    9.97
                                                       Prob > F      =  0.0041
                                                       R-squared     =  0.0911
Number of clusters (stud) = 26                         Root MSE      =  283.42
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	-176	55.73615	-3.16	0.004	-290.7907	-61.20925
_cons	856.6154	66.68674	12.85	0.000	719.2715	993.9593

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    25) =    9.97
            Prob > F =    0.0041

.
. * Niveau 2/3 sammen:
. xi: regress B_tid i.task if niveau_gr == 2, robust cluster(stud)
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      22
                                                       F(  1,    10) =    5.64
                                                       Prob > F      =  0.0390
                                                       R-squared     =  0.0255
Number of clusters (stud) = 11                         Root MSE      =  232.74
```


B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	-71.72727	30.21598	-2.37	0.039	-139.0527	-4.401877
_cons	564.9091	76.8256	7.35	0.000	393.731	736.0872

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    10) =    5.64
            Prob > F =    0.0390

.
. * Niveau 2 alene:
. drop if niveau == 3
(6 observations deleted)

. xi: regress B_tid i.task if niveau == 2, robust cluster(stud)
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      16
                                                       F(  1,     7) =    6.55
                                                       Prob > F      =  0.0376
                                                       R-squared     =  0.0371
Number of clusters (stud) = 8                          Root MSE      =   256.3
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	-94.125	36.7712	-2.56	0.038	-181.0751	-7.174923
_cons	615.5	97.74811	6.30	0.000	384.3624	846.6376

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,     7) =    6.55
            Prob > F =    0.0376

.
. * Niveau 4:
. xi: regress B_tid i.task if niveau_gr == 4, robust cluster(stud)
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      26
                                                       F(  1,    12) =    0.08
                                                       Prob > F      =  0.7806
                                                       R-squared     =  0.0004
Number of clusters (stud) = 13                         Root MSE      =  40.438
```

B_tid	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	1.538462	5.401378	0.28	0.781	-10.23013	13.30705
_cons	156.9231	12.08635	12.98	0.000	130.5892	183.257

```
. test _Itask_2

 ( 1)  _Itask_2 = 0

       F(  1,    12) =    0.08
            Prob > F =    0.7806

.
.
. * ALDERS OG KØNS INDFLYDELSE:
.
. drop if niveau == 4
(50 observations deleted)

.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

.
. xi: regress B_tid i.aldersgrp if niveau==1
i.aldersgrp       _Ialdersgrp_1970-1985(naturally coded; _Ialdersgrp_1970 omitted)

      Source |       SS       df       MS              Number of obs =      52
-------------+------------------------------           F(  3,    48) =    0.24
       Model |  65628.5256     3  21876.1752           Prob > F      =  0.8672
    Residual |  4353415.78    48  90696.1621           R-squared     =  0.0149
-------------+------------------------------           Adj R-squared = -0.0467
       Total |  4419044.31    51  86647.9276           Root MSE      =  301.16

------------------------------------------------------------------------------
       B_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 Ialder~1975 |  -26.83333   122.9473    -0.22   0.828    -274.0353    220.3687
_Ialder~1980 |    40.1859   105.1015     0.38   0.704    -171.1348    251.5066
_Ialder~1985 |  -100.5833   230.0132    -0.44   0.664     -563.056    361.8893
       _cons |   758.5833   86.93684     8.73   0.000     583.7851    933.3816
------------------------------------------------------------------------------

. test _Ialdersgrp_1975 _Ialdersgrp_1980 _Ialdersgrp_1985

 ( 1)  _Ialdersgrp_1975 = 0
 ( 2)  _Ialdersgrp_1980 = 0
 ( 3)  _Ialdersgrp_1985 = 0

       F(  3,    48) =    0.24
            Prob > F =    0.8672

.
. xi: regress B_tid i.kon
i.kon             _Ikon_1-2           (_Ikon_1 for kon==k omitted)

      Source |       SS       df       MS              Number of obs =      68
-------------+------------------------------           F(  1,    66) =    1.91
       Model |  165322.624     1  165322.624           Prob > F      =  0.1711
    Residual |  5699098.36    66  86349.9752           R-squared     =  0.0282
-------------+------------------------------           Adj R-squared =  0.0135
       Total |  5864420.99    67  87528.6714           Root MSE      =  293.85
```



```
------------------------------------------------------------------------------
       B_tid |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     _Ikon_2 |  -101.4634   73.32874    -1.38   0.171    -247.8689    44.94219
       _cons |   760.3095    45.3426    16.77   0.000     669.7801     850.839
------------------------------------------------------------------------------

.
.
.
. ****************************************
. *TEST B - FORSTAAELSE
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. rename bi_tid B_tid1

. rename bii_tid B_tid2

. rename bi_forstaaelse B_forst1

. rename bii_forstaaelse B_forst2

. keep stud fodselsaar kon niveau B_tid* B_forst*

. reshape long B_tid B_forst, i(stud) j(task)
(note: j = 1 2)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                       63   ->     126
Number of variables                   8   ->       7
j variable (2 values)                     ->   task
xij variables:
                          B_tid1 B_tid2   ->   B_tid
                       B_forst1 B_forst2   ->   B_forst
-----------------------------------------------------------------------------

.
. label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"

. label values task taskl

.
. bysort task: tabstat B_forst, s(n mean sd min p50 max) by(niveau) missing

-------------------------------------------------------------------------------
-> task = Vokaliseret

Summary for variables: B_forst
     by categories of: niveau (Niveau)

  niveau |         N      mean        sd       min       p50       max
---------+----------------------------------------------------------------
       1 |        27  2.074074  1.268734         0         2         4
       2 |         8     3.875  1.726888         0       4.5         5
       3 |         3  4.666667  .5773503         4         5         5
       4 |        25      4.72  .4582576         4         5         5
---------+----------------------------------------------------------------
   Total |        63   3.47619  1.634873         0         4         5
-------------------------------------------------------------------------------
-> task = Uvokaliseret
```

```
Summary for variables: B_forst
     by categories of: niveau (Niveau)
```

niveau	N	mean	sd	min	p50	max
1	27	2.444444	1.42325	0	2	5
2	8	4.5	.7559289	3	5	5
3	3	3.666667	2.309401	1	5	5
4	25	4.56	.7681146	2	5	5
Total	63	3.603175	1.54021	0	4	5

```
.
.
. * ANALYSE PÅ TVÆRS AF NIVEAU:
.
. * Niveau 2 og Niveau 3 sammen:
. use "help", clear

.
. * Vokaliseret:
. xi: regress B_forst i.niveau_grp if task == 1, cluster(stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      63
                                                       F(  2,    62) =   51.14
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5788
Number of clusters (stud) = 63                         Root MSE      =  1.0786
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Iniveau g~2	2.016835	.5089974	3.96	0.000	.9993637	3.034306
Iniveau g~4	2.645926	.2621977	10.09	0.000	2.1218	3.170052
_cons	2.074074	.2455208	8.45	0.000	1.583285	2.564863

```
. quietly test _Iniveau_gr_2

. test _Iniveau_gr_4, accumulate

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    62) =   51.14
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

B_forst	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-.6290909	.4552638	-1.38	0.172	-1.53915	.2809684


```
.
. * Uvokaliseret:
. xi: regress B_forst i.niveau_grp if task == 2, cluster(stud) robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      63
                                                       F(  2,    62) =   22.61
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4356
Number of clusters (stud) = 63                         Root MSE      =  1.1762
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_g~2	1.828283	.4650571	3.93	0.000	.898647	2.757919
_Iniveau_g~4	2.115556	.3156678	6.70	0.000	1.484545	2.746566
_cons	2.444444	.2754221	8.88	0.000	1.893883	2.995005

```
. quietly test _Iniveau_gr_2

. test _Iniveau_gr_4, accumulate

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    62) =   22.61
            Prob > F =    0.0000

. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

B_forst	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-.2872727	.4052278	-0.71	0.481	-1.097311	.522766

```
.
.
. * Niveau 2 separat og Niveau 3 udeladt:
. drop if niveau == 3
(6 observations deleted)

.
. * Vokaliseret:
. xi: regress B_forst i.niveau if task == 1, cluster(stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  2,    59) =   51.17
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5780
Number of clusters (stud) = 60                         Root MSE      =  1.0904
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	1.800926	.6354293	2.83	0.006	.5294349	3.072417
_Iniveau_4	2.645926	.2625261	10.08	0.000	2.120612	3.171239
_cons	2.074074	.2458282	8.44	0.000	1.582173	2.565975

```
. quietly test _Iniveau_2

. test _Iniveau_4, accumulate

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    59) =   51.17
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

B_forst	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-.845	.5931503	-1.42	0.160	-2.031891	.3418909

```
.
. * Uvokaliseret:
. xi: regress B_forst i.niveau if task == 2, cluster(stud) robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      60
                                                       F(  2,    59) =   23.58
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.4807
Number of clusters (stud) = 60                         Root MSE      =  1.1147
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Iniveau 2	2.055556	.3766124	5.46	0.000	1.301956	2.809155
Iniveau 4	2.115556	.3160631	6.69	0.000	1.483115	2.747996
_cons	2.444444	.275767	8.86	0.000	1.892636	2.996253

```
. quietly test _Iniveau_2

. test _Iniveau_4, accumulate

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    59) =   23.58
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

B_forst	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-.06	.2993959	-0.20	0.842	-.6590898	.5390898


```
.
.
. * ANALYSE PÅ TVÆRS AF TESTTYPE: UVOKLAISERET VS. VOKALISERET SKRIFT:
.
. * Niveau 2 og niveau 3 samlet:
. use "help", clear

.
. * Niveau 1:
. xi: regress B_forst i.task if niveau_grp == 1, cluster(stud) robust
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      54
                                                       F(  1,    26) =    1.05
                                                       Prob > F      =  0.3158
                                                       R-squared     =  0.0192
Number of clusters (stud) = 27                         Root MSE      =  1.3482
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	.3703704	.3620784	1.02	0.316	-.3738925	1.114633
_cons	2.074074	.2465046	8.41	0.000	1.567377	2.580772

```
.
. * Niveau 2/3 sammen:
. xi: regress B_forst i.task if niveau_grp == 2, cluster(stud) robust
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      22
                                                       F(  1,    10) =    0.13
                                                       Prob > F      =  0.7307
                                                       R-squared     =  0.0046
Number of clusters (stud) = 11                         Root MSE      =  1.3981
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	.1818182	.5136163	0.35	0.731	-.9625901	1.326226
_cons	4.090909	.4676299	8.75	0.000	3.048965	5.132853

```
.
. * Niveau 2 alene:
. xi: regress B_forst i.task if niveau == 2, cluster(stud) robust
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      16
                                                       F(  1,     7) =    1.29
                                                       Prob > F      =  0.2941
                                                       R-squared     =  0.0591
Number of clusters (stud) = 8                          Root MSE      =   1.333
```

B_forst	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itask_2	.625	.5511178	1.13	0.294	-.6781866	1.928187
_cons	3.875	.6319764	6.13	0.000	2.380613	5.369387

```
.
. * Niveau 4:
. xi: regress B_forst i.task if niveau grp == 4, cluster(stud) robust
i.task            _Itask_1-2          (naturally coded; _Itask_1 omitted)

Regression with robust standard errors                 Number of obs =      50
                                                       F(  1,    24) =    0.70
                                                       Prob > F      =  0.4096
                                                       R-squared     =  0.0164
Number of clusters (stud) = 25                         Root MSE      =  .63246

------------------------------------------------------------------------------
             |               Robust
     B_forst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Itask_2 |       -.16   .1906349    -0.84   0.410    -.5534511    .2334511
       _cons |       4.72   .0926013    50.97   0.000      4.52888     4.91112
------------------------------------------------------------------------------

.
.
.
. * ALDERS OG KØNS INDFLYDELSE:
.
. drop if niveau == 4
(50 observations deleted)

.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

.
. xi: regress B_forst i.aldersgrp if niveau==1
i.aldersgrp       _Ialdersgrp_1970-1985(naturally coded; _Ialdersgrp_1970 omitted)

      Source |       SS       df       MS              Number of obs =      54
-------------+------------------------------           F(  3,    50) =    0.22
       Model |  1.27513228     3  .425044092           Prob > F      =  0.8796
    Residual |  95.0952381    50  1.90190476           R-squared     =  0.0132
-------------+------------------------------           Adj R-squared = -0.0460
       Total |  96.3703704    53  1.81830887           Root MSE      =  1.3791

------------------------------------------------------------------------------
     B_forst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 Ialder~1975 |   .3333333   .5630134     0.59   0.556    -.7975124    1.464179
 Ialder~1980 |   .3571429   .4758332     0.75   0.456    -.5985962    1.312882
_Ialder~1985 |  -2.25e-16   1.053302    -0.00   1.000    -2.115619    2.115619
       _cons |          2   .3981106     5.02   0.000     1.200371    2.799629
------------------------------------------------------------------------------

. test _Ialdersgrp_1975 _Ialdersgrp_1980 _Ialdersgrp_1985

 ( 1)  _Ialdersgrp_1975 = 0
 ( 2)  _Ialdersgrp_1980 = 0
 ( 3)  _Ialdersgrp_1985 = 0

       F(  3,    50) =    0.22
            Prob > F =    0.8796
```



```
.
. xi: regress B_forst i.kon
i.kon             _Ikon_1-2           (_Ikon_1 for kon==k omitted)

      Source |       SS       df       MS              Number of obs =      76
-------------+------------------------------           F(  1,    74) =    0.26
       Model |  .684820748     1  .684820748           Prob > F      =  0.6096
    Residual |  192.736232    74  2.60454367           R-squared     =  0.0035
-------------+------------------------------           Adj R-squared = -0.0099
       Total |  193.421053    75  2.57894737           Root MSE      =  1.6139

------------------------------------------------------------------------------
     B_forst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     _Ikon_2 |   .1942029   .3787329     0.51   0.610    -.5604389    .9488447
       _cons |    2.73913   .2379507    11.51   0.000     2.265003    3.213257
------------------------------------------------------------------------------

.
.
.
. ****************************************
. * TEST C
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. tab c_korrekte niveau, missing nol

           |                 Niveau
    C_fejl |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
         1 |         5          0          0          0 |         5
         2 |         6          0          0          0 |         6
         3 |         6          1          0          0 |         7
         4 |         3          1          0          0 |         4
         5 |         2          2          1          0 |         5
         6 |         0          0          1          0 |         1
         7 |         0          0          0          9 |         9
         8 |         0          1          0          7 |         8
         9 |         0          1          1          5 |         7
         . |         5          2          0          4 |        11
-----------+--------------------------------------------+----------
     Total |        27          8          3         25 |        63

. sort niveau c_korrekte

. by niveau: tabstat c_korrekte, s(n mean sd min p50 max) missing

------------------------------------------------------------------------------
-> niveau = 1

    variable |         N      mean        sd       min       p50       max
-------------+----------------------------------------------------------------
  c_korrekte |        22  2.590909  1.259595         1       2.5         5
------------------------------------------------------------------------------

------------------------------------------------------------------------------
-> niveau = 2

    variable |         N      mean        sd       min       p50       max
-------------+----------------------------------------------------------------
  c_korrekte |         6  5.666667   2.33809         3         5         9
------------------------------------------------------------------------------
```

-> niveau = 3

variable	N	mean	sd	min	p50	max
c_korrekte	**3**	**6.666667**	**2.081666**	**5**	**6**	**9**

-> niveau = 4

variable	N	mean	sd	min	p50	max
c_korrekte	**21**	**7.809524**	**.8135753**	**7**	**8**	**9**

```
.
. * KORREKTE SVAR PÅ TVÆRS AF NIVEAUER:
.
. * Niveau 2 og 3 sammen:
.
. gen niveau_grp = niveau

. replace niveau_grp = 2 if niveau_grp == 3
(3 real changes made)

. label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"

. label values niveau_grp niveau_grpl

.
. xi: regress c korrekte i.niveau grp, robust
i.niveau_grp      _Iniveau_gr_1-4     (naturally coded; _Iniveau_gr_1 omitted)

Regression with robust standard errors                 Number of obs =      52
                                                       F(  2,    49) =  129.85
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7790
                                                       Root MSE      =  1.3136
```

c_korrekte	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Iniveau g~2	**3.409091**	**.7555874**	**4.51**	**0.000**	**1.890681**	**4.927501**
Iniveau g~4	**5.218615**	**.3238983**	**16.11**	**0.000**	**4.567717**	**5.869513**
_cons	**2.590909**	**.2702848**	**9.59**	**0.000**	**2.047751**	**3.134067**

```
. test _Iniveau_gr_2 _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 = 0
 ( 2)  _Iniveau_gr_4 = 0

       F(  2,    49) =  129.85
            Prob > F =    0.0000
```



```
. lincom _Iniveau_gr_2 - _Iniveau_gr_4

 ( 1)  _Iniveau_gr_2 - _Iniveau_gr_4 = 0
```

c_korrekte	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	**-1.809524**	**.727815**	**-2.49**	**0.016**	**-3.272123**	**-.3469249**

```
.
.
. * Niveau 2 alene:
.
. drop if niveau == 3
(3 observations deleted)

.
. xi: regress c_korrekte i.niveau, robust
i.niveau          _Iniveau_1-4        (naturally coded; _Iniveau_1 omitted)

Regression with robust standard errors                 Number of obs =      49
                                                       F(  2,    46) =  129.50
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.7992
                                                       Root MSE      =  1.2674
```

c_korrekte	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iniveau_2	**3.075758**	**.9392044**	**3.27**	**0.002**	**1.185237**	**4.966278**
_Iniveau_4	**5.218615**	**.324507**	**16.08**	**0.000**	**4.565416**	**5.871813**
_cons	**2.590909**	**.2707928**	**9.57**	**0.000**	**2.045831**	**3.135987**

```
. test _Iniveau_2 _Iniveau_4

 ( 1)  _Iniveau_2 = 0
 ( 2)  _Iniveau_4 = 0

       F(  2,    46) =  129.50
            Prob > F =    0.0000

. lincom _Iniveau_2 - _Iniveau_4

 ( 1)  _Iniveau_2 - _Iniveau_4 = 0
```

c_korrekte	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	**-2.142857**	**.9169254**	**-2.34**	**0.024**	**-3.988532**	**-.297182**

```
.
.
```

```
.
. * ALDERS OG KØNS INDFLYDELSE:
.
. drop if niveau == 4
(25 observations deleted)

.
. gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)

. label define alderl 1970 "fodt for/i 1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-
> 990"

. label values aldersgrp alderl

.
. xi: regress c_korrekte i.aldersgrp if niveau==1
i.aldersgrp       _Ialdersgrp_1970-1985(naturally coded; _Ialdersgrp_1970 omitted)
```

Source	SS	df	MS
Model	2.99090909	3	.996969697
Residual	30.3272727	18	1.68484848
Total	33.3181818	21	1.58658009

Number of obs =	22
F(3, 18) =	0.59
Prob > F =	0.6284
R-squared =	0.0898
Adj R-squared =	-0.0619
Root MSE =	1.298

c_korrekte	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Ialder~1975	-1	.8209381	-1.22	0.239	-2.724727	.724727
_Ialder~1980	-.6545455	.7000984	-0.93	0.362	-2.125398	.8163067
_Ialder~1985	-1.2	1.421907	-0.84	0.410	-4.187315	1.787315
_cons	3.2	.5804909	5.51	0.000	1.980434	4.419566

```
. test _Ialdersgrp_1975 _Ialdersgrp_1980 _Ialdersgrp_1985

 ( 1)  _Ialdersgrp_1975 = 0
 ( 2)  _Ialdersgrp_1980 = 0
 ( 3)  _Ialdersgrp_1985 = 0

       F( 3,    18) =    0.59
            Prob > F =    0.6284

.
. xi: regress c korrekte i.kon
i.kon             _Ikon_1-2           (_Ikon_1 for kon==k omitted)
```

Source	SS	df	MS
Model	4.7474359	1	4.7474359
Residual	100.502564	26	3.86548323
Total	105.25	27	3.89814815

Number of obs =	28
F(1, 26) =	1.23
Prob > F =	0.2779
R-squared =	0.0451
Adj R-squared =	0.0084
Root MSE =	1.9661

c_korrekte	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Ikon_2	.825641	.7450126	1.11	0.278	-.7057543	2.357036
_cons	2.866667	.5076405	5.65	0.000	1.823197	3.910137


```
.
.
.
. ****************************************
. *SPØRGESKEMA-UNDERSØGELSE
. ****************************************
.
. use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.

.
. foreach i of numlist 1/14 {
  2. tabulate sp `i', missing
  3. tab sp_`i' niveau, missing
  4. }
```

Sp_1	Freq.	Percent	Cum.
1	11	17.46	17.46
2	22	34.92	52.38
3	5	7.94	60.32
.	25	39.68	100.00
Total	63	100.00	

Sp_1	Niveau 1	2	3	4	Total
1	10	0	1	0	11
2	14	7	1	0	22
3	3	1	1	0	5
.	0	0	0	25	25
Total	27	8	3	25	63

Sp_2	Freq.	Percent	Cum.
1	15	23.81	23.81
2	17	26.98	50.79
3	6	9.52	60.32
.	25	39.68	100.00
Total	63	100.00	

Sp_2	Niveau 1	2	3	4	Total
1	12	2	1	0	15
2	13	3	1	0	17
3	2	3	1	0	6
.	0	0	0	25	25
Total	27	8	3	25	63

Sp_3	Freq.	Percent	Cum.
1	10	15.87	15.87
2	16	25.40	41.27
3	11	17.46	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_3	Niveau 1	2	3	4	Total
1	9	1	0	0	10
2	9	6	1	0	16
3	8	1	2	0	11
.	1	0	0	25	26
Total	27	8	3	25	63

Sp_4	Freq.	Percent	Cum.
1	1	1.59	1.59
2	10	15.87	17.46
3	18	28.57	46.03
4	9	14.29	60.32
.	25	39.68	100.00
Total	63	100.00	

Sp_4	Niveau 1	2	3	4	Total
1	1	0	0	0	1
2	8	0	2	0	10
3	12	5	1	0	18
4	6	3	0	0	9
.	0	0	0	25	25
Total	27	8	3	25	63

Sp_5	Freq.	Percent	Cum.
2	3	4.76	4.76
3	8	12.70	17.46
4	26	41.27	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_5	Niveau 1	2	3	4	Total
2	3	0	0	0	3
3	6	2	0	0	8
4	17	6	3	0	26
.	1	0	0	25	26
Total	27	8	3	25	63

Sp_6	Freq.	Percent	Cum.
1	2	3.17	3.17
2	14	22.22	25.40
3	15	23.81	49.21
4	5	7.94	57.14
.	27	42.86	100.00
Total	63	100.00	

Sp_6	Niveau 1	2	3	4	Total
1	2	0	0	0	2
2	10	3	1	0	14
3	11	3	1	0	15
4	2	2	1	0	5
.	2	0	0	25	27
Total	27	8	3	25	63

Sp_7	Freq.	Percent	Cum.
1	2	3.17	3.17
2	10	15.87	19.05
3	16	25.40	44.44
4	9	14.29	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_7	Niveau 1	2	3	4	Total
1	1	1	0	0	2
2	6	3	1	0	10
3	12	3	1	0	16
4	7	1	1	0	9
.	1	0	0	25	26
Total	27	8	3	25	63

Sp_8	Freq.	Percent	Cum.
1	11	17.46	17.46
2	12	19.05	36.51
3	7	11.11	47.62
4	6	9.52	57.14
.	27	42.86	100.00
Total	63	100.00	

Sp_8	Niveau 1	2	3	4	Total
1	8	3	0	0	11
2	10	2	0	0	12
3	3	1	3	0	7
4	4	2	0	0	6
.	2	0	0	25	27
Total	27	8	3	25	63

Sp_9	Freq.	Percent	Cum.
1	16	25.40	25.40
2	11	17.46	42.86
3	10	15.87	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_9	Niveau 1	2	3	4	Total
1	12	3	1	0	16
2	8	2	1	0	11
3	6	3	1	0	10
.	1	0	0	25	26
Total	27	8	3	25	63

Sp_10	Freq.	Percent	Cum.
1	9	14.29	14.29
2	14	22.22	36.51
3	11	17.46	53.97
4	1	1.59	55.56
.	28	44.44	100.00
Total	63	100.00	

Sp_10	Niveau 1	2	3	4	Total
1	7	2	0	0	9
2	10	2	2	0	14
3	7	3	1	0	11
4	1	0	0	0	1
.	2	1	0	25	28
Total	27	8	3	25	63

Sp_11	Freq.	Percent	Cum.
1	3	4.76	4.76
2	6	9.52	14.29
3	16	25.40	39.68
4	11	17.46	57.14
.	27	42.86	100.00
Total	63	100.00	

Sp_11	Niveau 1	2	3	4	Total
1	2	0	1	0	3
2	5	0	1	0	6
3	11	4	1	0	16
4	7	4	0	0	11
.	2	0	0	25	27
Total	27	8	3	25	63

Sp_12	Freq.	Percent	Cum.
1	8	12.70	12.70
2	20	31.75	44.44
3	7	11.11	55.56
4	1	1.59	57.14
.	27	42.86	100.00
Total	63	100.00	

Sp_12	Niveau 1	2	3	4	Total
1	6	2	0	0	8
2	16	3	1	0	20
3	2	3	2	0	7
4	1	0	0	0	1
.	2	0	0	25	27
Total	27	8	3	25	63

Sp_13	Freq.	Percent	Cum.
1	6	9.52	9.52
2	11	17.46	26.98
3	14	22.22	49.21
4	6	9.52	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_13	Niveau 1	2	3	4	Total
1	6	0	0	0	6
2	7	2	2	0	11
3	10	3	1	0	14
4	3	3	0	0	6
.	1	0	0	25	26
Total	27	8	3	25	63

Sp_14	Freq.	Percent	Cum.
1	7	11.11	11.11
2	19	30.16	41.27
3	10	15.87	57.14
4	1	1.59	58.73
.	26	41.27	100.00
Total	63	100.00	

Sp_14	Niveau 1	2	3	4	Total
1	3	2	2	0	7
2	15	3	1	0	19
3	7	3	0	0	10
4	1	0	0	0	1
.	1	0	0	25	26
Total	27	8	3	25	63

```
.
.
.
. ***
.
.
. erase help.dta

. clear

.
end of do-file

.
.
.
.
.
.
. log close
      log:  W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Stata datalog
 log type:  smcl
closed on:   4 Jun 2004, 13:50:14
```

```
                                                  ___  ____  ____  ____  ____tm
                                                 /    /    /    /    /    /
                                                /   /   /   /   /   /   /
                                               Statistics/Data Analysis

                                                 User: Gunna Funder Hansen
                         Project: Danske lørneres læsning på arabisk{space -13}
tabstat fodsel, s(n mean sd min p25 p50 p75 max) by(niveau) missing


*               ****************************************
*                              GRAFER
*               ****************************************




****************************************
* RESPONDENTERNE
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt

* stud= ID nr.
* hold= universitetsby - immatrikuleret år
* niveau
* fodselsaar
* kon= m eller k

* aldersdistribution:

gen aldersgrp = recode(fodsel,1970,1975,1980,1985,1990)
label define alderl 1970 "-1970" 1975 "1971-1975" 1980 "1976-1980" 1985 "1981-1985" 1990 "
label values aldersgrp alderl
tabulate aldersgrp, mi
tabstat fodsel, s(n mean sd min p25 p50 p75 max) by(niveau) missing

graph hbar (count) stud, over(aldersgrp) over(niveau, relabel(1 "Niveau 1" 2 "Niveau 2" 3
>  "Niveau 4")) bar(1, bfcolor(black)) bar(2, bfcolor(dkgreen)) bar(3, bfcolor(green)) bar
> dgreen)) bar(5, bfcolor(lime)) nofill asyvars percentages stack ytitle(Procent af respon
> ert niveau) legend(off)


* konsdistribution

graph hbar (count) stud, over(kon) over(niveau, relabel(1 "Niveau 1" 2 "Niveau 2" 3 "Nivea
> )/*
> */ bar(1, bfcolor(lime)) bar(2, bfcolor(green)) bar(3, bfcolor(midgreen)) bar(4, bfcolor
> */ bar(5, bfcolor(lime)) nofill asyvars percentages stack ytitle(Procent af respondenter
> )/*
> */ legend(off)


* Sammen:

graph combine "C:\Documents and Settings\Gunna Funder Hansen\Skrivebord\kon.gph" "C:\Docum
> tings\Gunna Funder Hansen\Skrivebord\alder.gph", cols(1) imargin(medlarge)



****************************************
* KONTROLANALYSER
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt

* HOLD-TID-FEJL:
```



```
* Niveau 1 - tid:
twoway (scatter ai_tid bi_tid if niveau == 1 & hold == "AU 2000", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter ai_tid bi_tid if niveau == 1 & hold == "KBH 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs5)) /*
> */ (scatter ai_tid bi_tid if niveau == 1 & hold == "NEG 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter aii_tid bii_tid if niveau == 1 & hold == "AU 2000", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter aii_tid bii_tid if niveau == 1 & hold == "KBH 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs5)) /*
> */ (scatter aii_tid bii_tid if niveau == 1 & hold == "NEG 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)), /*
> */ xlabel(250(250)1500, grid) ylabel(100(50)300) ytitle("Omkodningshastighed") xtitle("L
> */ legend(off)

* Niveau 1 - fejl:
twoway (scatter ai_bogstavfejl ai_fejl_diakritfejl if niveau == 1 & hold == "AU 2000", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter ai_bogstavfejl ai_fejl_diakritfejl if niveau == 1 & hold == "KBH 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs5)) /*
> */ (scatter ai_bogstavfejl ai_fejl_diakritfejl if niveau == 1 & hold == "NEG 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter aii_bogstavfejl aii_diakritfejl if niveau == 1 & hold == "AU 2000", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)) /*
> */ (scatter aii_bogstavfejl aii_diakritfejl if niveau == 1 & hold == "KBH 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs5)) /*
> */ (scatter aii_bogstavfejl aii_diakritfejl if niveau == 1 & hold == "NEG 2001", /*
> */ sort msymbol(smcircle) msize(vlarge) mlcolor(gs5) mfcolor(gs14)), /*
> */ xlabel(, grid) ytitle("Bogstavfejl") xtitle("Diakritfejl")/*
> */ legend(off)

graph combine "C:\Documents and Settings\Gunna Funder Hansen\Skrivebord\bogstavfejl-diakri
> ocuments and Settings\Gunna Funder Hansen\Skrivebord\omkodningshastighed-læsehastighed",
> ze(20)


* SAMMENLIGNING TID OG KORREKTHED - STATA 8:


* Test A:

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt

gen niveau_grp = niveau
replace niveau_grp = 2 if niveau_grp == 3
label define niveau_grpl 1 "niveau 1" 2 "niveau 2/3" 4 "niveau 4"
label values niveau_grp niveau_grpl

gen niveau sproggrp = niveau
replace niveau sproggrp = 1 if niveau == 2
replace niveau_sproggrp = 1 if niveau == 3
label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"
label values niveau_sproggrp niveau_sproggrpl

save help, replace

rename ai_tid A_tid1
rename aii tid A tid2
rename aiii_tid A_tid3

keep hold stud niveau* A_tid*

reshape long A_tid, i(stud) j(task)
```

```
label define taskl 1 "AI" 2 "A2" 3 "A3"
label values task taskl
sort stud task

* tempfile atid
save atid, replace

use "help", clear

rename ai_realfejl A1_err1
rename ai_bogstavfejl A1_err2
rename ai_fejl_diakrit A1_err3
rename aii_realfejl A2_err1
rename aii_bogstavfejl A2_err2
rename aii_diakritfejl A2_err3
rename aiii_fejl A3_err1

keep stud hold niveau* A1_err* A2_err* A3_err*

reshape long A1_err A2_err A3_err, i(stud) j(type)
label define errl 1 "realfejl" 2 "bogstavfejl" 3 "diakritfejl"
label values type errl
rename A1_err A_err1
rename A2_err A_err2
rename A3_err A_err3

reshape long A_err, i(stud type) j(task)
label define taskl 1 "Uden fremmede fonemer" 2 "Med fremmede fonemer" 3 "Latinsk alphabet"
label values task taskl
sort stud task

save help, replace

merge stud task using "atid"

drop _merge

sort task niveau niveau_grp niveau_sproggrp stud

gen task_grp = task
replace task_grp = 1 if task == 2
label define task_grpl 1 "Arabisk skrift" 2 "Latinsk skrift"
label values task_grp task_grpl

* Alle danskere bogstavfejl vs. tid både med og uden fremmede fonemer:
* twoway scatter A_err A_tid if niveau_sproggrp==1 & type == 2 & task_grp == 1, /*
> */ || lfit A_err A_tid if niveau_sproggrp==1 & type == 2 & task_grp == 1||,/*
> */ xlabel(, grid) xtitle(omkodningshastighed) ytitle(bogstavfejl) legend (off) title("Al
> AI + AII")


* Test B:

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt

gen niveau_sproggrp = niveau
replace niveau_sproggrp = 1 if niveau == 2
replace niveau_sproggrp = 1 if niveau == 3
label define niveau_sproggrpl 1 "modersmaal dansk" 4 "modersmaal arabisk"
label values niveau_sproggrp niveau_sproggrpl

rename bi_tid B_tid1
rename bii_tid B_tid2
rename bi_forstaaelse B_forst1
rename bii_forstaaelse B_forst2
keep stud niveau B_tid* B_forst*
```



```
reshape long B_tid B_forst, i(stud) j(task)

label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"
label values task taskl

* Alle danskere forståelse vs. tid både med og uden vokaler:
twoway scatter B_forst B_tid if niveau_sproggrp==1,  /*
> */ || lfit B_forst B_tid if niveau_sproggrp==1||, /*
> */ xlabel(250(250)1500, grid) xtitle(læsehastighed) ytitle("læseforståelse") legend (off
> ke lørnere BI + BII")



****************************************
* TEST A - TID
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


* PÅ TVÆRS AF NIVEAU - STATA 7:

label var ai_tid "AI"
label var aii_tid "AII"
label var aiii_tid "AIII"

label define niveaul 1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4"
label values niveau niveaul

dotplot  ai_tid, by(niveau) center vert mean bar ylabel(0 100 to 300)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\ai_ti

dotplot aii_tid, by(niveau) center vert mean bar ylabel(0 100 to 300)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aii_t

dotplot aiii_tid, by(niveau) center vert mean bar ylabel(0 20 to 100)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aiii_


* PÅ TVÆRS AF TESTTYPE - STATA 7:

rename ai_tid A_tid1
rename aii_tid A_tid2
rename aiii_tid A_tid3

keep hold niveau stud niveau_grp A_tid*
reshape long A_tid, i(stud) j(task)

label define taskl 1 "AI" 2 "AII" 3 "AIII"
label values task taskl
sort stud task
save help, replace

dotplot A_tid if niveau==1, by(task) center vert mean bar ylabel(0 100 to 300)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\a_tid
> lace)

dotplot A_tid if niveau==2, by(task) center vert mean bar ylabel(0 50 to 150)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\a_tid
> lace)

dotplot A_tid if niveau==3, by(task) center vert mean bar ylabel(0 40 to 120)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\a_tid
> lace)
```

```
dotplot A_tid if niveau==4, by(task) center vert mean bar ylabel(0 20 to 80)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\a_tid
> lace)


* TVÆRS AF TESTTYPE OG NIVEAU - STATA 8:

graph bar (mean) ai tid aii tid aiii tid, over(niveau)/*
> */ legend(label (1 "AI") label (2 "AII") label (3 "AIII" ) rows (1)) bar(1, bfcolor(blue
> red)) bar(3, bfcolor(yellow))/*
> */ title("Omkodningshastighed i gennemsnit" "for alle testtyper og niveauer")/*
> */ subtitle(.) ytitle (Omkodningshastighed i sekunder) ylabel(, angle(0))/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\a_tid
> ce)


* BOX PLOTS STATA 8 - NOTE #2
graph box A tid if task==3,/*
> */ medtype(cline) medline( lcolor(white) lpattern(solid) lwidth(medium) )/*
> */ over(niveau_sproggrp) title("Omkodningshastighed af latinske pseudoord") /*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\"aiii
> h", replace)



****************************************
* TEST A - FEJL
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


* PÅ TVÆRS AF NIVEAU - STATA 7:

label define niveaul 1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4"
label values niveau niveaul

dotplot ai_bogstavfejl, by(niveau) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\ai_bo
> place)

dotplot aii bogstavfejl, by(niveau) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aii_b
> eplace)

dotplot ai fejl diakritfejl, by(niveau) center vert mean bar ylabel(0 10 to 40)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\ai_di
> place)

dotplot aii diakritfejl, by(niveau) center vert mean bar ylabel(0 10 to 40)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aii_d
> eplace)

dotplot ai_realfejl, by(niveau) center vert mean bar ylabel(0 10 to 40)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\ai_re
> ce)

dotplot aii_realfejl, by(niveau) center vert mean bar ylabel(0 10 to 40)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aii_r
> ace)

dotplot aiii_fejl, by(niveau) center vert mean bar ylabel(0 4 to 20)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\aii_r
> ace)
```



```
* PÅ TVÆRS AF TASK - KUN REALFEJL NØDVENDIG - STATA 7:

rename ai_realfejl realfejl_A1
rename ai bogstavfejl bogstavfejl A1
rename ai_fejl_diakrit diakritfejl_A1
rename aii_realfejl realfejl_A2
rename aii bogstavfejl bogstavfejl A2
rename aii_diakritfejl diakritfejl_A2
rename aiii_fejl realfejl_A3

keep stud niveau real* bogstav* diakrit*

reshape long realfejl A bogstavfejl A diakritfejl_A, i(stud) j(task)
label define Al 1 "AI" 2 "AII" 3 "AIII"
label values task Al

rename bogstavfejl A bogstavfejl
rename diakritfejl A diakritfejl
rename realfejl_A realfejl

dotplot bogstavfejl if niveau==1, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\bogst
> ", replace)

dotplot bogstavfejl if niveau==2, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\bogst
> ", replace)

dotplot bogstavfejl if niveau==3, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\bogst
> ", replace)

dotplot bogstavfejl if niveau==4, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\bogst
> ", replace)


dotplot diakritfejl if niveau==1, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\diakr
> ", replace)

dotplot diakritfejl if niveau==2, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\diakr
> ", replace)

dotplot diakritfejl if niveau==3, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\diakr
> ", replace)

dotplot diakritfejl if niveau==4, by(task) center vert mean bar ylabel(0 10 to 50)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\diakr
> ", replace)


dotplot realfejl if niveau==1, by(task) center vert mean bar ylabel(0 10 to 40)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\realf
> replace)

dotplot realfejl if niveau==2, by(task) center vert mean bar ylabel(0 5 to 20)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\realf
> replace)

dotplot realfejl if niveau==3, by(task) center vert mean bar ylabel(0 5 to 20)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\realf
> replace)
```

```
dotplot realfejl if niveau==4, by(task) center vert mean bar ylabel(0 5 to 20)/*
> */ saving ("W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Test A\realf
> replace)


****************************************
*TEST B - TID
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


* PÅ TVÆRS AF NIVEAU:

rename bi_tid B_tid1
rename bii_tid B_tid2
rename bi forstaaelse B forst1
rename bii forstaaelse B forst2
keep stud niveau B_tid* B_forst*

dotplot B tid1, by(niveau) center vert mean bar ylabel(0 500 to 1500)
dotplot B_tid2, by(niveau) center vert mean bar ylabel(0 500 to 1500)

reshape long B_tid B_forst, i(stud) j(task)

label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"
label values task taskl
save help, replace

dotplot B_tid, by(niveau) center vert mean bar ylabel(0 500 to 1500)


* PÅ TVÆRS AF TESTTYPE

dotplot B_tid if niveau==1, by(task) center vert mean bar ylabel(0 500 to 1500)
dotplot B tid if niveau==2, by(task) center vert mean bar ylabel(0 400 to 1200)
dotplot B tid if niveau==3, by(task) center vert mean bar ylabel(0 200 to 800)
dotplot B_tid if niveau==4, by(task) center vert mean bar ylabel(0 100 to 300)



****************************************
*TEST B - FORSTAAELSE
****************************************

use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


* PÅ TVÆRS AF NIVEAU: BUBBLE PLOT - STATA 8

rename bi tid B tid1
rename bii_tid B_tid2
rename bi_forstaaelse B_forst1
rename bii_forstaaelse B_forst2
keep stud niveau B_tid* B_forst*

reshape long B_tid B_forst, i(stud) j(task)

label define taskl 1 "Vokaliseret" 2 "Uvokaliseret"
label values task taskl

bysort niveau task: egen mean = mean(B_forst)

save help, replace
```



```
program drop _all
program define bubble
args var1 var2 cond var1tit var2tit tit
use "help", clear
keep if task == `cond'
gen sortindex = .
bysort `var1' `var2': replace sortindex = _N
twoway (scatter `var2' `var1'[w=sortindex], /*
>         */msymbol(Oh) msize(small) title("`tit'") /*
>         */note("Cirklernes størrelse afspejler antallet af observationer i den relevante
>         */ ylab(0(1)5.1) xlab(0(1)4.1) ytitle("`var2tit'") xtitle("`var1tit'")/*
>         */ scheme(s1mono))
drop sortindex
end

bubble niveau B_forst 1 "Niveau" "B: Forståelse (vokaliseret)" "BI"
bubble niveau B_forst 2 "Niveau" "B: Forståelse (uvokaliseret)" "BII"


* PÅ TVÆRS AF TESTTYPE - STATA 7:

* BUBBLEPLOTS - STATA 8:

gen sortindex = .
bysort niveau task B_forst: replace sortindex = _N

twoway (scatter B_forst task[w=sortindex] if niveau ==1, /*
*/ msymbol(Oh) msize(small)), xscale(range(0 3)) /*
*/ xlabel(0.5 " " 1 "Vokaliseret" 1.5 " " 2 "Uvokaliseret" 2.5 " ", noticks valuelabel)/*
*/ note("Cirklernes størrelse afspejler antallet af observationer i den relevante kategori
*/ title("Niveau 1") ylab(0(1)5.2) ytitle(" ") xtitle(" ")/*
*/ scheme(s1mono)

twoway (scatter B_forst task[w=sortindex] if niveau ==2, /*
*/ msymbol(Oh) msize(small)), xscale(range(0 3)) /*
*/ xlabel(0.5 " " 1 "Vokaliseret" 1.5 " " 2 "Uvokaliseret" 2.5 " ", noticks valuelabel)/*
*/ note("Cirklernes størrelse afspejler antallet af observationer i den relevante kategori
*/ title("Niveau 2") ylab(0(1)5.2) ytitle(" ") xtitle(" ")/*
*/ scheme(s1mono)

twoway (scatter B_forst task[w=sortindex] if niveau ==3, /*
*/ msymbol(Oh) msize(small)), xscale(range(0 3)) /*
*/ xlabel(0.5 " " 1 "Vokaliseret" 1.5 " " 2 "Uvokaliseret" 2.5 " ", noticks valuelabel)/*
*/ note("Cirklernes størrelse afspejler antallet af observationer i den relevante kategori
*/ title("Niveau 3") ylab(0(1)5.2) ytitle(" ") xtitle(" ")/*
*/ scheme(s1mono)

twoway (scatter B_forst task[w=sortindex] if niveau ==4, /*
*/ msymbol(Oh) msize(small)), xscale(range(0 3)) /*
*/ xlabel(0.5 " " 1 "Vokaliseret" 1.5 " " 2 "Uvokaliseret" 2.5 " ", noticks valuelabel)/*
*/ note("Cirklernes størrelse afspejler antallet af observationer i den relevante kategori
*/ title("Niveau 4") ylab(0(1)5.2) ytitle(" ") xtitle(" ")/*
*/ scheme(s1mono)


* DOTPLOTS - STATA 7:
dotplot B_forst if niveau == 1, by(task) center vert mean bar ylabel(0 1 to 5)
dotplot B_forst if niveau == 2, by(task) center vert mean bar ylabel(0 1 to 5)
dotplot B_forst if niveau == 3, by(task) center vert mean bar ylabel(0 1 to 5)
dotplot B_forst if niveau == 4, by(task) center vert mean bar ylabel(0 1 to 5)



****************************************
* TEST C
****************************************
```

```
use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


* BUBBLE PLOT - STATA 8:

tab c_korrekte niveau, missing nol
sort niveau c korrekte
by niveau: tabstat c_korrekte, s(n mean sd min p25 p50 p75 max) missing

program drop _all
program define bubble
args var1 var2 var1tit var2tit tit
gen sortindex = .
bysort `var1' `var2': replace sortindex = _N
twoway (scatter `var2' `var1' [w=sortindex], /*
>         */msymbol(Oh) msize(small) title("`tit'") /*
>         */note("Cirklernes størrelse afspejler antallet af observationer i den relevante
>         */ ylab(0(1)8.1) xlab(0(1)4.1) ytitle("`var2tit'") xtitle("`var1tit'")/*
>         */ scheme(s1mono))
drop sortindex
end

bubble niveau c_korrekte "Niveau" "C: antal korrekte mønstre" "Title"



***********************************************
*SPØRGESKEMA-UNDERSØGELSE
***********************************************

* Sp. 1: "Det er svært at læse på arabisk"
* Sp. 2: "Jeg føler, at jeg læser meget langsomt på arabisk"
* Sp. 3: "De tekster vi læser i undervisningen, passer til mit niveau"
* Sp. 4: "De tekster vi læser i undervisningen, er oftest for svære"
* Sp. 5: "Læseretningen kan være forvirrende"
* Sp. 6: "De fremmede sætningskonstruktioner gør læsning svært"
* Sp. 7: "Jeg er usikker på nogle af bogstaverne"
* Sp. 8: "Det er sværere at læse, når teksten ikke er vokaliseret"
* Sp. 9: "Jeg tror, at mit største problem er mit ordforråd"
* Sp. 10: "Jeg staver mig igennem ordene"
* Sp. 11: "Læsning på arabisk er så svært, at det kan være demotiverende"
* Sp. 12: "Læsning på arabisk er en motiverende udfordring"
* Sp. 13: "Jeg bruger ordbogen meget, når jeg læser"
* Sp. 14: "Jeg gætter mig til en del, når der er noget, jeg ikke forstår"


foreach i of numlist 1/14 {
tabulate sp_`i', missing
tab sp_`i' niveau, missing
}


use "W:\Mellemøst-Disk II\Gunna Funder Hansen\Ph.D\Empiri\Resultater\Phd_Testresultater.dt


label define niveaul 1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4"
label values niveau niveaul

keep stud niveau sp*

reshape long sp_, i(stud) j(task)
```



```
label define taskl 1 "Det er svært at læse på arabisk" /*
> */ 2 "Jeg føler, at jeg læser meget langsomt på arabisk" /*
> */ 3 "De tekster vi læser i undervisningen, passer til mit niveau" /*
> */ 4 "De tekster vi læser i undervisningen, er oftest for svære" /*
> */ 5 "Læseretningen kan være forvirrende" /*
> */ 6 "De fremmede sætningskonstruktioner gør læsning svært" /*
> */ 7 "Jeg er usikker på nogle af bogstaverne" /*
> */ 8 "Det er sværere at læse, når teksten ikke er vokaliseret" /*
> */ 9 "Jeg tror, at mit største problem er mit ordforråd" /*
> */ 10 "Jeg staver mig igennem ordene" /*
> */ 11 "Læsning på arabisk er så svært, at det kan være demotiverende" /*
> */ 12 "Læsning på arabisk er en motiverende udfordring" /*
> */ 13 "Jeg bruger ordbogen meget, når jeg læser"/*
> */ 14 "Jeg gætter mig til en del, når der er noget, jeg ikke forstår"

label values task taskl
sort stud task

graph hbar (count) stud, over(sp_, relabel(1 "Meget enig" 2 "Enig" 3 "Mindre enig" 4 "Slet
> ) over(niveau, relabel(1 "Niveau 1" 2 "Niveau 2" 3 "Niveau 3" 4 "Niveau 4")) by(task, co
> ff)) bar(1, bfcolor(black)) bar(2, bfcolor(green)) bar(3, bfcolor(midgreen)) bar(4, bfco
> bel(name, color(white) position(inside) orientation(horizontal) justification(left) alig
>  nofill asyvars percentages stack ytitle(Procent af respondenterne på hvert niveau) lege



erase help.dta

clear
```

APPENDIX 4

ØVRIGE DATA

	Niveau 1																				Niveau 2					
Respondent	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25
Udelukkende grafisk lighed																										
ن ← ب	3	1					2	1			1	1														
ب ← ن												1								1						1
ن ← ت							1											1								
ت ← ن																										
ن ← ي	2																									
ي ← ن																										
ب ← ت																										
ت ← ب																								1		
ب ← ي								1	2		1															
ي ← ب									1		1		1					1								
ن ← ث								1																		
ث ← ن																										
ت ← ي							1																1			
ي ← ت										1																
ج ← خ				1		2										1	4	1	1				2	1		1
خ ← ج									1			1												1		
ج ← ح					1																					1
ح ← ج																										
ص ← ض																										
ض ← ص									2		2															
ظ ← ص																										
ص ← ظ												1														
ع ← غ													1													
غ ← ع											1															
ف ← ق																										
ق ← ف																	1	1								
ه ← ة														3						3						
ة ← ه	2	1																								
ى ← ي				1																2						
ي ← ى		1																								
ذ ← ز										1		1														
ز ← ذ																										
ز ← ن																				1						
ن ← ز																										
ز ← ض																										
ض ← ز											1															
غ ← ف	2																									
ف ← غ																										
ذ ← ف								1																		
ف ← ذ																										
ذ ← غ																										
غ ← ذ																										
ذ ← ن			1																							
ن ← ذ																										
ذ ← خ																										
خ ← ذ					1																					
ق ← ت					1																					
ت ← ق																										
ة ← ق												1				1										
ق ← ة																										
ث ← ش		1	1	2			1			1							1							1		
ش ← ث	1	2										1	1													
ا ← ل	2		2	2	2	1			4				1					1		3		1				
ل ← ا								1														1				
ل ← ك				1		1																				
ك ← ل																										
د ← ل					1																					1
ل ← د																					1					
م ← و														1												

	Niveau 3			Niveau 4																	Antal resp. m. denne fejltype			
	36	37	38	39	40	41	42	43	44	45	52	53	54	57	58	59	60	61	63	Niveau:	1	2	3	4
Udelukkende grafisk lighed																								
ن ← ب	1																				6		1	
ب ← ن										1											2	1		1
ن ← ت																					2			
ت ← ن																								
ن ← ي																					1			
ي ← ن																								
ب ← ت																	1							1
ت ← ب																						1		
ب ← ي																					3			
ي ← ب																			1		4			1
ن ← ث																					1			
ث ← ن																								
ت ← ي																					1	1		
ي ← ت																					1			
ج ← خ			2																		6	3	1	
خ ← ج																					2	1		
ج ← ح																					1	1		
ح ← ج																								
ص ← ض																								
ض ← ص																					2			
ظ ← ص																								
ص ← ظ																					1			
ع ← غ							1									2					1			2
غ ← ع																					1			
ف ← ق																								
ق ← ف												1									2			1
ه ← ة																					2			
ة ← ه																					2			
ى ← ي		1																			2		1	
ي ← ى																					1			
ذ ← ز																					2			
ز ← ذ									2	1	3	2						1						5
ز ← ن																					1			
ن ← ز																								
ز ← ض																								
ض ← ز																					1			
غ ← ف																					1			
ف ← غ																								
ذ ← ف																					1			
ف ← ذ																								
ذ ← غ																								
غ ← ذ																								
ذ ← ن																					1			
ن ← ذ																								
ذ ← خ																								
خ ← ذ																					1			
ق ← ت																					1			
ت ← ق																								
ة ← ق		1																			2		1	
ق ← ة																								
ث ← ش																					6	1		
ش ← ث																					4			
ا ← ل																					9	1		
ل ← ا																					1	1		
ل ← ك																					2			
ك ← ل																								
د ← ل																					1	1		
ل ← د																						1		
م ← و																					1			

	Niveau 1																				Niveau 2					
	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25
Udelukkende fonetisk lighed																										
ق ← خ				1	1						1															
ق ← ك	7	2			5			5					1			2				1						
ق ← g			4																							
ق ← خ																										
خ ← ق						1																				
خ ← ه											1	1	2	2		2	1	1								
خ ← ق																										
ح ← ه		2		1			2	1	1											2				3		
ح ← ق																										
ك←خ					1											1										
ك←ق													2				1				2					
ه←ح						1												1								
ه←خ			1																							
ج ← س																										
ج ← ش		2																								
ج ← ز		1																								
ز ← س							1	1	2							1		1								
ز ← ش		1																								
ز ← ج			1																							
ز ← ص								1																		
ز ← ث			1					1																		
ش←ج					1	1						1		1												
ش←ز																										
س ← ز											1	2					1									
س ← ج																										
س ← ز																										
ص ← ز																		1								
ض←د		1					1				1		2	1		1		1				1		1		1
ض←ذ													1													
ض←ت						1																				
ض←ث								1	1																	
ط←ت		1			1		1	1	2																	
ط←د												1			1											
ط←ذ																										
ظ←ذ									1			1					1									
ظ←د						1											1									
ظ←ت					2		1		1			1														
ظ ← ث																1								2		
ر←غ					1																					
غ←ر								1			1			3				3								
ذ←ض						1																				
ذ ← ث									2																	
ا←ع			1					1					2													1
ع←ا	1	1							1																	
ع←ء																										
ة←ء	1	1		1								1														
ة←-		1	3			4																				
ة ← ا									2	4			2							1						

	Niveau 3			Niveau 4																			
	36	37	38	39	40	41	42	43	44	45	52	53	54	57	58	59	60	61	63	Niveau:	1	2	3	4
Udelukkende fonetisk lighed																								
ق ← خ																					3			
ق ← ك		1																			7		1	
ق ← g																					1			
ق ← خ																								
خ ← ق																					1			
خ ← ه																					7			
خ ← ق																								
ح ← ه																					6	1		
ح ← ق																								
ك←خ																					2			
ك←ق												1									2	1		1
ه←ح																					2			
ه←خ																					1			
ج ← س																								
ج ← ش																					1			
ج ← ز																					1			
ز ← س	1																				5		1	
ز ← ش																					1			
ز ← ج																					1			
ز ← ص																					1			
ز ← ث																					2			
ش←ج																					4			
ش←ز																								
س ← ز																					3			
س ← ج																								
س ← ز																								
ص ← ز																					1			
ض←د																					7	3		
ض←ذ																					1			
ض←ت																					1			
ض←ث																					2			
ط←ت																					5			
ط←د	1																				2		1	
ط←ذ																								
ظ←ذ																					3			
ظ←د																					2			
ظ←ت																					4			
ظ ← ث																					1	1		
ر ← غ											1										1			1
غ←ر																					4			
ذ←ض																					1			
ذ ← ث																					1			
ا←ع																					3	1		
ع←ا																					3			
ع←ء																								
ة←ء																					3			
ة←-																					3			
ة ← ا								1			1										4			2

	Niveau 1																				Niveau 2					
	3	6	7	8	9	10	11	12	13	14	15	23	24	28	29	30	32	33	34	35	16	17	20	21	22	25
Både grafisk og fonetisk lighed																										
ت ← ث																		1							1	
ث ← ت				1									1				1									
خ ← ح											1						1									
ح ← خ						1	1						1		2			1								
د←ذ													1							1						
ذ←د		2					3				4			1				1							1	
س ← ش	2	3	2		4					1						1		1		7					1	
ش ←س																										
ص←س													2			1		1								
س←ص											1															
ض←ظ												1		1			1		1	1						
ظ←ض																										
ض←ط								1		1										1						
ط←ض																										
ط ←ظ												1								1						
ظ←ط		4		1	1	1	1	4	1	1	4		1	3												1
غ ← خ													1			1										
خ ← غ								1																		
Fejl i latinske bogstaver																										
f → v																										
v → f																										
f → d										1																
r → f																										
t → f																										
t → k																										
t → b																										
l → t																										
k → h																										
h → k						1						1														
m → n																										
vokal					1				1	1			1		1		1			2		1		1		1

	Niveau 3			Niveau 4																			
	36	37	38	39	40	41	42	43	44	45	52	53	54	57	58	59	60	61	63	Niveau:	1	2	3	4
Både grafisk og fonetisk																								
ت ← ث							1	3											2		1	1		3
ث ← ت																					3			
خ ← ح									1			1	1	1	1		1		1		2			7
ح ← خ																					5			
د←ذ		4						1								1					2		1	2
ذ←د			1		1	2	2	1	2	1					3				4		5	1	1	8
س ← ش	1	1																			8	1	2	
ش ←س																								
ص←س																					3			
س←ص																					1			
ض←ظ	2		1									2	2	4			2	1			5		3	5
ظ←ض		1	1			3	1		1		1		1	1	1		1	1	2				2	10
ض←ط																					3			
ط←ض																								
ط ←ظ																					2			
ظ←ط												1			1				1		11	1		3
غ ← خ																					2			
خ ← غ						1															1			1
Fejl i latinske bogstaver																								
f → v									1															1
v → f						1																		1
f → d																								
r → f													1											1
t → f						1																		1
t → k													1											1
t → b																		1						1
l → t			1																					
k → h						1				2														2
h → k																								
m → n							2																	1
vokal						8	2	4	1			1	3					1						

Besvarelser i spørgeskemaets åbne svarkategori

(Tal refererer til respondent-ID)

"Jeg synes at det sværeste ved at læse på arabisk er:"

1: *"Simpelthen bare, at det er så fremmed fra alt hvad man kender i lyd og billede"*
2: *"At det er så svært at huske ordene, så der er meget få "holdepunkter" i form af gloser"*
3: *"At have tid til det"*
4: *"At genkende ord, dvs. ofte udtaler jeg det så forkert, at jeg ikke genkender det"*
5: *"At mit ordforråd er for ringe"*
6: *"Man er helt fortabt, hvis man ikke kan få bogstaverne til at give mening, dvs. man SKAL være sikker på alle bogstaver for at kunne forstå et ords betydning. Selvom jeg synes jeg er godt med i arabisk, laver mange afleveringsopgaver osv., så er det stadig svært for mig at læse – især er det svært at læse ord og tekster, man ikke kender i forvejen. Nu synes jeg faktisk at vokaliserede tekster er sværere (de virker mere forvirrende) end uvokaliserede tekster"*
7: *"At få hastigheden op. At få lært ordene, for så går det bedre!"*
8: *"Reklametekst eller fx avistekst, hvor skriften kan være lidt "udtværet". Det er meget lettere at læse fx i lærebogen"*
9: *"At læse højt"*
10: *"Når der optræder nye ord... Det hjælper meget at man har set ordet en enkelt gang før"*
11: *"At man i stedet for at opfatte ord som helheder staver sig igennem fra venstre og går i stå pga. konsonanternes ophobning og fordi vokal- og sukuun [ikke-vokal] -forholdet er uklart. Hvis man kan Læse et ord er det ofte fordi man genkender det og derfor kender dets lyd og vokalisering"*
12: *"Mangel på eget ordforråd og endnu ikke stort kendskab til grammatikken – og sætningsstrukturen"*
13: *"At "gætte" korte vokaler. Det er nødvendigt for mig at stave mig igennem ordene"*
14: *"At forstå sætningernes opbygning"*
15: *"At 'gætte' de korte vokaler ved ukendte ord"*
16: *"Jeg synes generelt at højtlæsning ved første gennemlæsning er meget svært. De anderledes sætningskonstruktioner"*
17: *"At det er et helt andet alfabet end man er vant til...* :-)
18: *"Ordforråd"*
19: *"At huske nye ord og vokaler"*
20: *"Ordforråd"*
21: *"At sproget er så fremmed fra de europæiske sprog som jeg kender i forvejen, det er ikke muligt at gætte et ord ud fra noget som jeg kender i forvejen"*
22: *"Manglende ordforråd"*
23: *"Manglende vokalisering midt i ord"*
25: *"Øvelse"*
26: *"Jeg har et lille ordforråd"*
28: "*At huske alle de grammatiske regler mht. kasus, defektive, assimilerende osv. verber"*
29: *"Vokalisering, alle ord er næsten 100% fremmede, på europæiske sprog kan man gætte sig til meget"*
30: *"Mit ordforråd + fremmede sætningskonstruktioner, der kan være svære at gennemskue"*
31: *"At jeg endnu ikke læser ordbilleder – så jeg må stave. At det ikke er umiddelbart synligt, hvor trykket er i ordet. At mit ordforråd er meget begrænset"*
32: *"Sætningskonstruktion og at gennemskue nye ord"*
33: *"At ens ordforråd er begrænset"*

36: *"Mit manglende ordforråd"*
37: *"Det meget varierede ordforråd, samt manglende rutine (mister tålmodigheden med mig selv)*
28: *"Fortrolighed med at læse – dvs. læsehastighed"*

أعتقد بأن الصعوبات في القراءة باللغات الأوربية هي:
(***"Jeg synes at det svære ved at læse på europæiske sprog er:***")

40: "النطق ببعض الحروف أجده صعبا و مختلف عن لغتي الام"
("Udtalen af nogle af bogstaverne er vanskelig og anderledes end på modersmålet")
41: "تتجلى في كونا ليست لغة الام لهذا نعاني من بعض الصعوبات التي تزول طالما هناك ارادة لتعليمهٍ"
("Det er jo ikke modersmålet. Derfor har vi nogle vanskeligheder, der forsvinder med tiden, så længe der er vilje til at lære et europæisk sprog")
42: "متعلقة بالنطق"
("Det afhænger af udtalen")
43: "أسهل بكثير من القراءة باللغة العربية"
("Det er meget lettere at læse end arabisk")
44: "النطق مختلف من لغة الى اخرى"
("Udtalen er forskellig fra det ene sprog til det andet")
45: "تكمن الصعوبات في القراءة باللغات الاوروبية في انفراد كل لغة بنطق خاص با"
("Vanskelighederne ved at læse på de europæiske sprog ligger især i den særlige udtale på hvert sprog")
47: "التركيب والصرف والتحويل"
("Sætningskonstruktioner og bøjninger")
50: "صعوبة التعبير و التطبيق سيما في انحصار حدود بعض اللغات. على سبيل المثال لا الحصر، الذين يدرسون الالمانية ليس لهم امكانية التحدث با سوى في نطاق دروسهم مع صعوبة التداول با خارجا"
("Vanskeligheder med udtryk og deres brug, specielt i inden for nogle sprog. For eksempel har man, hvis man studerer tysk, kun mulighed for at tale sproget i undervisningen, mens det er vanskeligt at få brugt det i andre sammenhænge")
51: "في اللغة الفرنسية الصعوبات تكمن في كثرة قواعدها المختلفة جدا. أما اللغة الانجليزية فالمشكل الوحيد عندي يكمن في نطق الكلمات"
("På fransk ligger vanskelighederne i de mange forskellige grammatiske regler, men på engelsk er mit eneste problem udtalen af ordene")
52: "كيفية نطق الكلملت جيدا و ان يتفهم الناس ماذا تقول بسرعة"
("At udtale ordene rigtig og forstå folk, der taler hurtigt")
53: "في مشكل معنى الكلمات *Vocabulaire*"
("Problemer med at forstå ordene - ordforråd")
54: "اختلاف الحروف و النطق"
("De anderledes bogstaver og lyde")
55: "كيفية نطق الكلمات"
("Hvordan man udtaler ordene")

56: "الكتابة و القراءة"

("Skrivning og læsning")

57: "النطق. فالحروف يختلف نطقها من بلد لاخر"

("Udtalen. For udtalen af bogstaverne er forskellig fra et land til et andet.")

59: "الكلمات التى يجب فهمها"

("At forstå ordene")

60: "النطق و ال ال المعرفي"

("Udtale og generel forståelse")

61: "الحروف تختلف و نطق الكلمات صعب"

("Bogstaverne er anderledes og udtalen af ordene er svær")

62: "الشرح اللغوي"

("Sproglige forklaringer")

63: "النطق غير متطابق مع الكتابة كما ان هناك الكثير من الحالات الشادة عن القواعد"

("Udtalen er ikke i overensstemmelse med skriften, ligesom der er mange grammatiske uregelmæssigheder ")