

Påvisning af ordblindhed ud fra eye tracking-optagelser på dansk

MARINA BJÖRNSDÓTTIR, NORA HOLLENSTEIN & ETHAN WEED

ABSTRACT

Ordblindhed er en indlæringsforstyrrelse, som påvirker omkring 10-20 % af verdens befolkning. I Danmark benyttes en officiel landsdækkende test til at diagnosticere ordblindhed, Ordblindetesten, som er lavet til at afdække fonologiske afkodningsvanskeligheder. Samtidig udvikles flere teknologiske diagnosticeringsmetoder på internationalt plan, og blandt disse er *eye tracking* specielt interessant, da denne metode giver adgang til mange informationer om kognitiv processering. Formålet med denne undersøgelse er at vurdere, i hvilket omfang maskinlæringsklassificering af eye tracking-data kan bruges til at diagnosticere ordblindhed i Danmark. Dataindsamlingen består af eye tracking-optagelser af ordblinde og ikke-ordblinde læsere under naturlig læsning af danske tekster. Her benytter vi en dybdegående psykolingvistisk analyse, hvor læsemønstrene fra de to grupper af deltagere sammenlignes. Denne viser en signifikant forskel i udvalgte målbare eye tracking-komponenter. På baggrund af denne forskel i datasættet mellem de to grupper af læsere udvalgte vi maskinlæringsmetoder til at klassificere læsemønstrene. Dette resulterede i en nøjagtighed på 85 %, hvilket indikerer, at selvom der stadigvæk er en vis usikkerhed, er vores metoder i stand til at opdage målbare forskelle i øjenbevægelserne hos ordblinde og ikke-ordblinde læsere, og de kan dermed på sigt tænkes ind i udredningsprocessen.

EMNEORD: eye tracking; maskinlæring; ordblindhed; læsevanskeligheder; psykolingvistik

1 INTRODUKTION

Ordblindhed (eller dysleksi) er en specifik indlæringsforstyrrelse, som angiveligt påvirker omkring 10-20 % af verdens befolkning (Rello & Ballesteros 2015, Kaiser 2020). Kendetegn for ordblindhed inkluderer besvær ved læsning, skrivning samt ordafkodning og er ikke relateret til intelligens (Perera et al. 2018, Rauschenberger et al. 2017). Det er vigtigt at opdage ordblindhed så tidligt som muligt, da vanskelighe-

der med diagnosen har en række negative konsekvenser, herunder lavt selvværd, angst og depression (Perera et al. 2018, Schulte-Körne 2010). Manglen på effektive hjælpeværktøjer, læringsfaciliteter og assistance med de fonologiske udfordringer, der karakteriserer ordblindhed, kan i høj grad udfordre en ordblind persons velfærd i uddannelses- samt erhvervssammenhænge. Diagnosticering af ordblindhed kan være en svær proces, da indikationer på ordblindhed samt medfølgende vanskeligheder forekommer i forskellige grader (Eckert 2004). Derfor bliver ordblindhed ofte betragtet som en usynlig diagnose (Rello & Ballesteros 2015). Ifølge Dysleksiforeningen i Danmark lider op til 7-8 % af landets befolkning af ordblindhed i et eller andet omfang. I de seneste år har bekymringer om den danske nationale ordblindetest, Ordblindetesten, og dens effektivitet fået opmærksomhed i medierne og i forskningssammenhænge, fordi testen i mange tilfælde ikke resulterer i påvist ordblindhed eller læsevanskeligheder hos adskillige testpersoner (Stanek 2021, Ottosen et al. 2022). Derudover får mange personer med ordblindhed først en officiel diagnose som teenagere eller i voksenalderen. Børne- og Undervisningsministeriet rapporterer, at screening med Ordblindetesten foretages på uddannelsesinstitutioner, og målgruppen er således personer, der er i gang med en uddannelse (Børne- og Undervisningsministeriet 2021). Dette kan være problematisk, da mange individer over 30 år typisk ikke er under uddannelse, og det kan derfor være vanskeligt for dem at blive testet for ordblindhed, trods behovet for dette.

Metoder, som anvendes til at påvise ordblindhed, består typisk af standardiserede vurderinger såsom læse- og stavetests. Singleton et al. (2009) forklarer, at computeriserede screeningmetoder for ordblindhed er veletablerede for børn i Storbritannien, men at udviklingen af ordblindetests for voksne kan være mere udfordrende end for børn, da mange voksne med ordblindhed ikke viser tydelige tegn på nedsatte læsefærdigheder. Dette skyldes, at mange voksne med ordblindhed har udviklet strategier, som hjælper dem med at skjule disse vanskeligheder, og af denne årsag bliver ordblindheden ofte ikke opdaget (Singleton et al. 2009). Dette tyder på, at der er et behov for at udvikle eller kombinere teknikker og metoder, der kan identificere karakteristiske træk ved ordblindhed hos voksne.

Mindre anvendte metoder for påvisning af ordblindhed, som endnu kun er i eksperimentelle faser, er eye tracking (måling af øjenbevægelser) under læsning og skrivning, som beskrives nærmere i afsnit 2, samt hjerneskaning, herunder (f)MRI, EEG, PET og MEG (Kaisar 2020, Perera et al. 2018). Her har eye tracking-metoden den fordel at være prismæssigt mere overkommelig og hurtigere til at indsamle data i forhold til hjerneskaning. Eye tracking er en sensorteknologi, der ofte anvendes i kognitive studier, og den kan bruges til at identificere nøglekomponenter i kognitive processer, såsom beslutningstagning, opmærksomhed og sprogforståelse. Da eye tracking giver essentiel indsigt i menneskelig adfærd, er metoden ikke kun begrænset til medicinske og kognitive undersøgelser, men har vist sig også at være yderst anvendelig på tværfaglige områder, såsom sprogteknologi og psykolingvistik (Płużyczka 2018). Endvidere er en stor fordel ved eye tracking-metoden, at den bygger på læserens naturlige adfærd og ikke nødvendigvis kræver, at læseren aktivt skal løse en opgave eller en problemstilling (Nilsson Benfatto et al. 2016). Optagelser af øjenbevægelsesmønstre fra eye tracking-metoden indeholder informationer om læserens blik, mens de læser, hvilket kan give meget indsigtsfulde informationer om underliggende kognitive processer (Rayner 1998). Disse kan imidlertid være komplekse at analysere. Ved hjælp af maskinlæringsteknikker er det muligt at identificere mønstre og tendenser i eye tracking-data, da sådanne teknikker er avancerede statistiske metoder. En begrænset mængde af forskning har undersøgt, hvorvidt maskinlæringsklassifikation af læsemønstre målt med eye tracking-teknologier kan bruges til at screene for ordblindhed (Rello & Ballesteros 2015, Nilsson Benfatto et al. 2016, Prabha & Bhargavi 2020, Asvestopoulou et al. 2019, Raatikainen et al. 2021). Disse undersøgelser har alle vist, at der er signifikante forskelle i øjenmønstre hos ordblinde og ikke-ordblinde læsere. Så vidt vi ved, er denne metode endnu ikke blevet anvendt til at påvise ordblindhed med dansk som målsprog.

Formålet med denne artikel er at bygge en maskinlæringsmodel, der kan klassificere eye tracking-læsemønstre som enten tilhørende en ordblind eller en ikke-ordblind læser, samt vurdere, hvorvidt denne metode kan bruges til at diagnosticere ordblindhed i Danmark. Dette vil vi opnå ved at bruge en binær klassifikation af eye tracking-data fra både

ordblinde samt ikke-ordblinde deltagere, mens de læste danske tekster i et så naturligt setting som muligt.

Artiklen er en del af et større projekt, nemlig det voksende korpus ”The Copenhagen Corpus of Eye Tracking Recordings from Natural Reading of Danish Texts” (CopCo) (Hollenstein et al. 2022). Eye tracking-optagelser fra ikke-ordblinde deltagere stammer fra CopCo-projektet, og denne artikel benytter den samme forsøgsopsætning og de samme processer og protokoller for ordblinde deltagere som CopCo-projektet under dataindsamlingen.

1.1 Udredning af ordblindhed i Danmark

I 2015 lancerede Børne- og Undervisningsministeriet en national, elektronisk ordblindetest, kaldt Ordblindetesten. Testen består af tre delprøver, som omfatter følgende: a) afkodning af nonsensord, hvor deltageren lytter til en lydoptagelse af et nonsensord og derefter skal vælge den korrekte stavemåde; b) afkodning af pseudo-homofoner baseret på identificering af et nonsensord, der lyder som et rigtigt dansk ord, blandt fire nonsensord, der ikke lyder som et rigtigt ord (f.eks. nonsensordet ”åst”, der lyder som ”ost” ved udtalelse); c) en ordforrådsopgave, hvor deltageren ud fra en lydoptagelse af et ord skal vælge et billede – med det formål at undersøge, om deltageren har det nødvendige ordforråd til at kunne gennemføre afkodningen af pseudo-homofoner i den forrige opgave. Alle disse delprøver gennemføres elektronisk og består af multiple choice-svarmuligheder. Testen fokuserer således på fonologiske afkodningsevner, og ud fra deltagerens præstation i delprøverne vil resultatet fremkomme som enten grønt, gult eller rødt. Et grønt resultat betyder ’ikke-ordblind’, og et gult resultat indikerer ’usikker fonologisk afkodning’, hvilket vil sige, at deltageren har fonologiske vanskeligheder, men ikke i samme omfang som ordblinde på samme uddannelsesniveau. Et rødt resultat betyder, at deltageren er ordblind. Personer, der kategoriseres i den gule kategori i Ordblindetesten, kræver yderligere vurdering og evaluering på basis af fx tekstforståelse, læsehastighed og stave- og ordforrådstest, der er tilpasset deltagerens alder og uddannelsesniveau (Center for Læseforskning et al. 2020). Selvom en grundig evaluering af deltagere, der kategoriseres i den gule kategori, er påkrævet med henblik på at vurdere deres behov

for redskaber og anden støtte, afviser den officielle vejledning for Ordblindetesten, at disse individer er ordblinde (Center for Læseforskning et al. 2020).

I et interview fra 2021 (Stanek 2021) fortæller en dansk læsevejleder og lærer, at mange unge elever, der klassificeres i den gule kategori, efterfølgende bliver diagnosticeret med ordblindhed, nogle gange flere år senere. Det betyder, at de elever, der ikke bliver diagnosticeret som ordblinde, men som har fonologiske vanskeligheder, ikke får samme adgang til hjælpeværktøjer. Hun tilføjer, at testen ifølge standarderne kun skal tages én gang, men at hun ser en tydelig tendens til, at mange børn først bliver diagnosticeret som ordblinde, efter de har videreudviklet deres fonologiske færdigheder, samt har modtaget mere uddannelse. Det sker, selvom testen ved en tidligere lejlighed ikke har kategoriseret dem som ordblinde. Dette indikerer, at der er et stort behov for en opfølgning på testene og eventuelt udvikling af en mere raffineret test med færre falsk-negative resultater. Ikke længe efter interviewet vurderede Ottosen et al. (2022), at der mangler klinisk evaluering af den danske nationale ordblindetest, og forklarer at testen ideelt bør indeholde opgaver med større variation, da den udelukkende er baseret på en standardscore fra fonologisk afkodning. Som følge heraf afskæres nogle universitetsstuderende med ordblindhed eller andre uidentificerede læsevanskeligheder fra at modtage læsestøtte, da de nødvendige hjælperedskaber er begrænset til studerende med en dokumenteret ordblindediagnose. Her anfører Ottosen et al. (2022), at Ordblindetesten er for begrænset, idet den definerer ordblindhed på baggrund af en kognitiv vurdering af evnen til at tilknytte grafemer til hørte fonemer, hvilket vil sige, at tilgangen udelukkende fokuserer på fonologiske vanskeligheder. I modsætning til dette defineres ordblindhed på internationalt plan som vanskeligheder ved læsning af ord trods sprogkendskab som svarer til ikke-ordblinde læsere. Derved defineres ordblindhed på et mere adfærdsmæssigt niveau sammenlignet med Ordblindetesten (Ottosen et al. 2022). I forbindelse med den pågældende undersøgelse fortalte enkelte deltagere i eksperimentet, at de har taget mere end én ordblindescreeningstest, men at de har fået forskellige resultater. Dette inkluderer tilfælde, hvor deltagerne fik en score, der tydede på, at de havde fonologiske vanskeligheder uden at

være ordblinde, men efter at have gentaget testen i de efterfølgende år er de blevet diagnosticeret som ordblinde. Overordnet set belyser dette afsnit problemstillinger, der stærkt tyder på, at der er behov for at anvende et mere inkluderende og præcist diagnostisk værktøj som kan bruges til at udrede ordblindhed i Danmark.

2 FORHOLDET MELLEM ØJENBEVÆGELSER OG LÆSE-VANSKELIGHEDER

Eye tracking under naturlig læsning kan give oplysninger om øjenbevægelser, som typisk inkluderer fikseringer (fast blik på et enkelt sted), sakkader (hurtige bevægelser mellem fikseringer) og regressioner (blik på et sted, der allerede er blevet fikseret på). En fiksering varer i gennemsnit omkring 200-250 ms, og en sakkade er i gennemsnit omkring 7-9 bogstavmellemrum (Rayner, 1998). Undersøgelser viser, at omkring 80 % af alle sanseindtryk til hjernen leveres via øjnene (Płużyczka 2018). Tidligere forskning (Rayner 1998, Henderson 2013) viser, at information om øjenbevægelser under læsning indeholder karakteriseringer af igangværende visuelle og kognitive processer; disse processer har nemlig direkte indflydelse på øjenbevægelserne. Disse er også stærkt relateret til identificering af information om fx opmærksomhed under læsning, da denne ofte kan analyseres ved måling af sakkaders hyppighed og varighed (Rayner 1998). Forskning har gentagne gange vist, at de største forskelle i øjenbevægelser mellem ordblinde læsere og typiske læsere konsekvent findes i sakkader og fikseringer; ordblinde læsere opviser et højere antal fikseringer og regressioner, længere fikseringsvarighed samt korte og hyppige sakkader (Pirozzolo & Rayner 1979, Rayner 1986, Biscaldi et al. 1998). Dette afspejler de vanskeligheder, som ordblinde læsere typisk oplever, da mange tekster kan være tunge og tidskrævende for dem at læse. Tidlige teorier om forholdet mellem øjenbevægelser og ordblindhed blev udviklet af Rubino & Minden (1973), som konstaterede, at øjenbevægelser er relaterede til læsefærdigheder eller mangel på disse, selvom de ikke afgør, om dette er årsagen eller virkningen. Denne konklusion blev draget ud fra deres resultater, som viste, at der var signifikante forskelle blandt deres udvalgte gruppe af læsere – børn med læsehandicap og børn uden læsehandicap – på forskellige målbare øjenbevægelseskomponenter. Disse komponenter var bl.a. fikseringer,

regressioner og læste ord pr. minut. Et par år senere hævdede Pirozzolo & Rayner (1979), at de vanskeligheder, der almindeligvis observeres ved ordblindhed – herunder øget fikseringsvarighed og antal fikseringer – skyldes ordblindhed. Denne hypotese udelukker således, at adfærd afspejlet i øjenbevægelser og læseopgaver skulle være årsagen til ordblindhed, men snarere er effekten af det. I sin undersøgelse, der rapporterer resultaterne af et eye tracking-forsøg med begyndende læsere og dygtige læsere, giver Rayner (1986) forklaringerne på, at øjenbevægelser ikke er årsagen til langsom læsning, men snarere afspejler langsommere kognitive processer. I denne undersøgelse udviser de uerfarne læsere øjenbevægelsesegenskaber svarende til ordblindes: lange fikseringsvarigheder, mange fikseringer og regressioner og korte sakkader i forhold til dygtige læsere. Idéen om, at øjenbevægelser hos ordblinde personer afspejler langsommere kognitive processer er herefter blevet populær og er almindeligt anerkendt i lignende undersøgelser (Fischer & Weber 1990, Hyönä & Olson 1995, Rayner 1998, Henderson 2013, Rello & Ballesteros 2015, Nilsson Benfatto et al. 2016, Raatikainen et al. 2021), typisk med hensyn til de læsevanskeligheder og kognitive vanskeligheder, der afspejles under naturlig læsning. Hyönä & Olson (1995) forklarer endvidere, at der er rigeligt belæg for at øjenfikseringsmønstre fra naturlig læsning afspejler sprogprocesser.

2.1 Eye tracking som metode til at opspore ordblindhed

Som vist i det foregående afsnit er der masser af evidens for, at øjenbevægelser afspejler kognitive processer og derfor fungerer som en indikator for underliggende tilstande, herunder ordblindhed. Bellocchi et al. (2013) beskriver eye tracking under naturlig læsning som en nyttig metode til at identificere og karakterisere ordblinde læsers afkodningsstrategier og læsemønstre – og kan således give det datagrundlag, der kan bruges til at sammenligne ordblinde læsere med ikke-ordblinde læsere. Smyrnakis et al. (2017) beskriver eye tracking som et værktøj, der giver visuelle opmærksomhedsdata, som mange af de eksisterende ordblindescreeninger mangler. På grundlag af eye tracking af ordblinde og ikke-ordblinde læsere rapporterer De Luca et al. (1999), at fikseringerne hos ordblinde deltagere var omkring 20 % længere end de registrerede fikseringer i kontrolgruppen.

2.2 Maskinlæringsklassifikation af eye tracking-data

Tidligere undersøgelser, som har benyttet maskinlæringsklassificering af eye tracking-data til at påvise ordblindhed, har primært benyttet binær klassifikation af øjenmønstre, dvs. at det identificeres, hvorvidt det er mest sandsynligt, at et øjenbevægelsesmønster er fra en ordblind eller en ikke-ordblind læser (Rello & Ballesteros 2015, Nilsson Benfatto et al. 2016, Prabha & Bhargavi 2020, Asvestopoulou et al. 2019, Raatikainen et al. 2021). I en sådan undersøgelse udgivet af Rello & Ballesteros (2015) havde metoden en præcision på 80,18 %. Nilsson Benfatto et al. (2016) benyttede samme maskinlæringsmodel og opnåede en imponerende nøjagtighed på 95,6 % ved hjælp af 48 øjenbevægelseskomponenter. Eksempler på sådanne komponenter er blandt andet gennemsnitlig antal samt varighed af fiksationer og sakkader. Selvom denne undersøgelse bekræftede, at maskinlæringsklassifikation af eye tracking-mønstre er en effektiv metode til at påvise ordblindhed, foreslog forfatterne yderligere tiltag i forbindelse med ordblindhedsdiagnosticering, som maskinlæringsmodellen ikke omfatter. De foreslår her en diagnostisk opfølgning efter et positivt screeningsresultat. Formålet med denne opfølgning er at få en mere grundig forståelse af det enkelte individ og vedkommendes kognitive profil. Dette tyder på, at klassificering baseret på øjenbevægelser kan være utilstrækkeligt som enkeltstående screeningsværktøj, selvom det giver nyttig indsigt og nøjagtig påvisning af læsevanskeligheder. Den pågældende undersøgelse anvender en eksperimentel metode til kvantitativ dataindsamling, hvilket er yderst brugbart i studier af menneskelig adfærd. Til dataanalysen anvendes statistisk analyse af de indsamlede data. Disse tilgange blev tilpasset til at identificere karakteristiske træk ved de mest signifikante forskelle mellem ordblinde læsere og ikke-ordblinde læsere. De mest signifikante træk indebærer antal, varighed og placering af fiksationer, sakkader og regressioner under læsning. Dette bekræftede, at maskinlæringsklassifikation af disse forskelle i deltagernes læsemønstre er en effektiv metode at identificere ordblinde læsere fra ikke ordblinde læsere.

3 DATAINDSAMLING

De data, der blev anvendt til dette forsøg, består af eye tracking-optagelser fra CopCo-datasættet (Hollenstein et al. 2022), som består

af eye tracking-optagelser fra 18 ikke-ordblindede voksne læsere med dansk som modersmål. Optagelserne blev indsamlet under naturlig læsning af danske tekster i form af transskriberede taler. Forsøget i denne artikel bygger videre på CopCo-projektets forsøg og bidrager til datasættet med data fra 18 ordblindede voksne læsere med dansk som modersmål.

3.1 Deltagere

Vi brugte data fra 18 ikke-ordblindede deltagere i eye tracking-eksperimentet fra CopCo-datasættet. Ydermere deltog 19 ordblindede læsere i eksperimentet i forbindelse med den pågældende undersøgelse, og deraf er data fra 18 deltagere blevet brugt til analyse. Alle deltagere har dansk som modersmål og er mellem 20 og 64 år. Minimumsalderen for at deltage i forsøget var 18 år, da forsøget var begrænset til voksne læsere. Alle deltagere har normalt syn eller syn, der er korrigeret til normalt ved hjælp af briller eller kontaktlinser. Derudover gav alle deltagere samtykke til deltagelse og til brug af deres data til forskning. Der blev registreret oplysninger om deltagernes alder, køn, antal læste taler under forsøget, gennemsnitlig læsetid pr. skærm og nøjagtighed af korrekt besvarede forståelsesspørgsmål. Dette er vist i tabel 1. Ikke-ordblindede læsere med rapporterede opmærksomhedsforstyrrelser blev ikke brugt til videre analyse, da opmærksomhedsforstyrrelser har høj komorbiditet med ordblindhed.

TABEL 1. OVERSIGT OVER DELTAGERE MED ORDBLINDHED

Deltager	Nøjagtighed	Læste tekster	Ord pr. minut	Alder	Køn	Diag. alder
Læsere med ordblindhed						
P23	1.00	2	200,0	33	K	16
P24	0.80	2	203,7	64	K	9
P25	0.82	4	142,0	20	K	16
P26	0.57	2	86,7	32	M	12
P27	0.71	4	137,4	53	M	48
P28	0.93	4	173,3	25	K	15
P29	0.73	3	143,3	25	K	21
P30	0.93	4	179,0	61	M	50
P31	0.75	2	61,9	20	M	15
P33	0.86	2	59,3	30	K	8
P34	0.62	2	107,4	56	K	9
P35	0.71	4	285,1	24	K	19
P36	0.40	2	58,5	23	K	11
P37	0.58	4	270,7	25	K	23
P38	0.75	2	115,5	30	M	29
P39	1.00	1	160,2	32	K	17
P40	0.92	4	173,3	29	M	7
P41	0.88	4	154,9	51	K	50
Gns.	0,78	2,9	150,7	35,1	67,7 %K	20,8
Læsere uden ordblindhed						
Gns.	0,81	4,4	276,8	30,7	78 %K	-

Gennemsnitlige data for læsere uden ordblindhed angives nederst til sammenligning. *Nøjagtighed* indikerer andelen af korrekte svar på forståelsesspørgsmål til de læste taler. *Diag. alder* refererer til deltagerens alder ved diagnosticering af ordblindhed.

3.1.1 Ikke-ordblinde læsere

Blandt de 18 ikke-ordblinde deltagere var fire deltagere mænd (22 %), og 14 læsere var kvinder (78 %). Deres højest fuldførte uddannelsesniveau ligger mellem gymnasiet og ph.d., og de har ingen kendte læsevaneskeligheder. Antallet af taler læst af deltagerne varierer fra 1 til 7 taler, men de fleste af deltagerne læste fire taler eller mere.

3.1.2 Ordblinde læsere

Blandt de 18 ordblinde læsere var seks deltagere mænd (33,3 %), og 12 deltagere var kvinder (67,7 %). Det højest fuldførte uddannelsesniveau blandt de ordblinde læsere spænder fra folkeskole til kandidatuddannelse. De ordblinde deltagere blev bedt om at udføre en leksikalisk vurderingstest, bestående af en læseopgave og en nonsensord-afkodningsopgave. De ordblinde deltagere læste i gennemsnit færre taler under testen sammenlignet med de ikke-ordblinde læsere, med et interval på 1 til 4 læste taler.

3.2 Læsemateriale

Det læsemateriale, der blev brugt til forsøget, bestod af transskriberede danske taler tilgået fra Danske Talers arkiv (<https://dansketalers.dk>). Alle taler blev korrekturlæst, inden de blev tilføjet til forsøget. De udvalgte taler tilhørte alle en af følgende kategorier: politiske taler, sankt-hans-båltaler og dimissionstaler. Alle talerne blev holdt i perioden 2010-2019 og blev udvalgt på baggrund af, at de er engagerende, ikke blot for publikum, men også for læserne i forsøget. Taler, der blev vurderet til at være mindre interessante for læserne, blev sorteret fra. Endvidere var kønsfordelingen af de personer, der holdt talerne, afbalanceret, og både aldersgrupperne og de steder, hvor talerne blev holdt, var også afbalanceret. Der blev udvalgt i alt 46 taler af forskellig længde (fra 37 sætninger og op til 134 sætninger). Oplysninger om antal sætninger, ord, samt talernes læsbarhed, kan ses i tabel 2.

TABEL 2. STATISTIK OVER LÆSEMATERIALE

	Min.	Maks.	Gns.	I alt
Sætninger pr. dokument	37	134	92,4	1849
Ord pr. dokument	391	1056	603,6	7361
LIX pr. dokument	25,4	50,1	37,2	-

Dokument refererer til dokumenter bestående af de anvendte taler for læsematerialet. *LIX* er læsbarhedsscoren.

Læsbarheden (LIX) af de enkelte taler er beregnet ud fra deres lixtal, som er baseret på længden af ord og sætninger i en tekst (Björnsson 1968). For typiske læsere anses en tekst med et lixtal på mellem 25 og

34 for let, hvorimod en tekst, der scorer mere end 55, anses for svær og svarer til en akademisk tekst. Talerne i forsøget havde et lixtal på mellem 26,43 og op til 50,14. For at forhindre, at deltagere læste automatisk og uden at forholde sig til tekstens indhold, blev omkring 20 % af afsnittene efterfulgt af forståelsesspørgsmål, der indeholdt mere end 100 tegn. Disse bestod af enten ja- eller nej-spørgsmål og var strengt relateret til indholdet af det foregående afsnit, der var blevet læst. Spørgsmålene kunne ikke springes over. Der var 68 spørgsmål i alt, hvoraf 30 havde det rigtige svar "nej", og 38 havde det rigtige svar "ja". Afsnittene med tilhørende spørgsmål blev valgt tilfældigt.

3.3 Eksperimentets opsætning

Der findes flere metoder til eye tracking-eksperimenter. Det pågældende forsøg fulgte forsøgsopsætningen for CopCo-projektet, som bruger en EyeLink 1000 Plus infrarød, videobaseret eye tracker. Denne teknologi er udviklet af SR-Research og er knyttet til softwaren SR Experiment Builder, som blev brugt til at sætte forsøgsprotokollet op elektronisk. Den anvendte skærm var 27 tommer og har skærmmålene 590 x 335 mm og en opløsning på 1920 x 1080 pixels. Den estimerede afstand mellem monitoren og deltagerne under optagelserne var 85 cm. Forud for forsøget blev en kalibrering af øjnene udført og valideret, dvs. at deltagerens pupiller bliver korrekt identificeret af eye trackeren. Efter hver blok, som bestod af to taler, blev kalibreringen valideret igen. Kun deltagerens højre øje blev optaget, men i enkelte tilfælde blev venstre øje på grund af kalibreringsbesvær optaget i stedet. I de tilfælde, hvor kvaliteten af valideringen var mindre end 'god', dvs. enten 'rimelig' eller 'dårlig', blev der foretaget omkalibrering.

Deltagerne blev instrueret i at sidde foran monitoren og hvile deres hage på en justerbar hagestøtte, der skulle begrænse deltagerens hoved- og kropsbevægelser. En plade med tre taster blev brugt: for at gå videre til næste skærbillede skulle deltagerne klikke på 'E' (for 'Enter'), og for at besvare forståelsesspørgsmålene skulle de klikke enten 'ja' eller 'nej'. Hvert skærbillede blev efterfulgt af en rød fikseringsprik i midten af skærmen. For at fortsætte skulle deltagerne holde deres blik på den røde prik og samtidig trykke på 'E'-knappen. Forsøget bestod af blokke, hvor en blok omfatter to taler. Ikke-ordblinde del-

tagere i eksperimentet læste typisk 2-3 blokke (4-6 taler), hvorimod de ordblinde deltagere typisk læste 1-2 blokke (2-4 taler), afhængigt af den enkelte deltagers læsehastighed. Der blev vurderet, at forsøget ville vare ca. 1-1,5 time.

Der var ingen tidsbegrænsning for læsedelen af forsøget, hverken for læsning af talerne eller for forståelsesspørgsmålene. Dette sikrede, at deltagerne kunne læse i deres eget tempo. Efter hver blok fik læseren en pause og blev spurgt, om de ville fortsætte forsøget. Typisk varede forsøget i 1-1,5 time. Forsøgets instruktioner bestod af sort skærm med hvid tekst, og disse dukkede op i starten af forsøget såvel som efter en blok, hvor deltagerne blev instrueret om, at de nu måtte holde en pause, før kalibreringen skulle revalideres. Talerne var opdelt i paragraffer, som så vidt muligt fulgte den oprindelige opdeling af talen, og disse fremstod på en lysegrå baggrund med sort tekst. Tekstens skrifttype var *Consolas*, og skriftstørrelsen var på 16. Et eksempel på den præsenterede læsestimulus kan ses i figur 1. De samme betingelser gjaldt for de skærbilleder, der indeholdt forståelsesspørgsmålene, som havde overskriften "Spørgsmål". Hvert afsnit bestod af maksimum 10 linjer og blev præsenteret med tredobbelt linjeafstand, så læserne kunne læse sætningerne tydeligt. De øverste og nederste tekstmargener var 140 pixels, og venstre og højre tekstmargener var 200 pixels.

FIGUR 1. EKSEMPEL PÅ EN TEKSTPARAGRAF (SKÆRM) OG TILHØRENDE SPØRGSMÅL

Fælles for os er, at vi er en regnbueminoritet i forhold til flertallets skab - men det betyder ikke at majoriteten har ret. Vi vil udefra informere om, at der er andre måder at se tingene på, tilbyde et alternativt syn på livet og verden, som kan hjælpe nogle - isæt børn og unge - ud af et skab, de mistrives i.

Spørgsmål:

Mener taleren at minoriteter bør tie med deres verdenssyn?

3.4 Leksikalske vurderingstests

Alle ordblinde deltagere i forsøget udførte to leksikalske vurderingstest, der bruges som kontroltest. Formålet med testene var at have et komparativt benchmark for en leksikalsk vurdering, der ikke er relateret til øjenbevægelserne hos deltagerne med ordblindhed. Nergård-Nilssen og Eklund (2018) vurderede i deres psykometriske evaluering, at pseudohomofontests (dvs. afkodning af nonsensord, der lyder som rigtige ord) har en høj pålidelighed, og at sådanne tests indeholder evalueringer, der giver en præcis identifikation af læsere med ordblindhed. Et eksempel på en pseudohomofon er *åst* – dvs. et ord, der ikke eksisterer på dansk, men lyder som det eksisterende danske ord *øst*. På grund af denne evaluering, samt at en pseudohomofonopgave anvendes i den danske ordblindetest, blev en pseudohomofontest udvalgt som en af de leksikalske vurderingstest i denne undersøgelse. Da læseforståelsestest er meget pålidelige og kan give stor indsigt i et individs læsefærdigheder, blev en læseforståelsestest benyttet som supplement. Denne fungerede som en supplerende leksikalsk vurderingstest (se 3.4.2).

3.4.1 Pseudohomofontest

Den udvalgte test, der blev benyttet som pseudohomofonkontroltest, bestod af testmateriale udviklet af Center for Læseforskning ved Københavns Universitet, og den var oprindeligt tiltænkt som en del af en diagnostisk læsetest for voksne. Standarderne for testene er angivet for forskellige grupper af læsere i den officielle dokumentation for testen (Center for Læseforskning et al. 2020). Testen omfatter 38 delopgaver, og hver opgave består af fire nonsensord, hvoraf et af ordene overholder dansk fonotaks, når det udtales. Der gives fem minutter til, at deltageren kan udføre så mange opgaver som muligt. I denne del skal deltageren udvælge, hvilket af de fire angivne ord vedkommende synes lyder som et rigtigt ord. Sværhedsgraden af de 38 opgaver stiger gradvist for hver opgave. For at kunne forbinde pseudohomofonerne med det korrekte kendte danske ord er det vigtigt for deltageren at kende ordet. Testen indebærer ikke en ordforrådstest for at sikre dette, men der anvendes hyppige, almindelige dagligdagsord på dansk i testen. Hermed antages det, at voksne læsere med dansk som modersmål kender ordene, og manglende ordforråd bør derfor ikke være en faktor, når testen ud-

føres af denne målgruppe. Eksempler på de anvendte ord er: *åst* for *ost*, *sbise* for *spise*, *fabræg* for *fabrik* og *jælpe* for *hjælpe*.

3.4.2 Læseforståelsestest

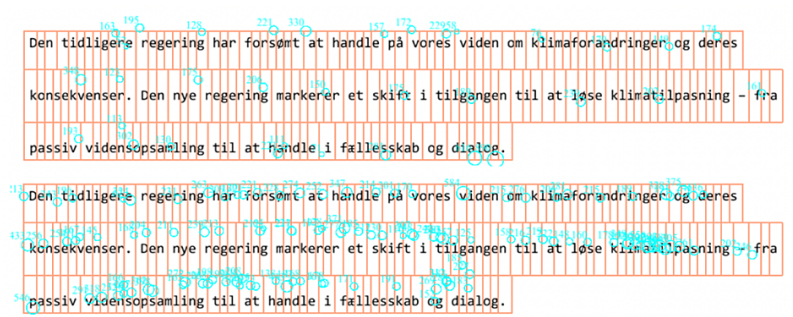
Den udvalgte test til læseforståelse er udviklet af Center for Læseforskning ved Københavns Universitet og kaldes ‘Selvtest af læsning’. Der er tale om en online dansk læsetest for voksne, som kan tilgås på <https://selvtest.nu/>. Da læseopgaven er en online selvevalueringstest, der ikke kræver log-in, støtte fra en forsøgsleder eller opfyldelse af særlige adgangskrav, blev deltagerne uden ordblindhed kontaktet efter deres deltagelse i eye tracking-forsøget og bedt om at tage testen. Målet for dette var at danne en kontrolgruppe. Ti deltagere uden ordblindhed indsendte deres score som et bidrag til dette forsøg. Det oprindelige formål med læsetesten er at give voksne let adgang til en uformel evaluering af deres læsefærdigheder (Jensen et al. 2014). Testen tager ti minutter og er således relativt kort, men informativ. Opgaverne i testen består af tre varianter af *cloze tests*, som er tests, hvor deltagerne skal vælge et manglende ord i en sætning, fx ‘Det havde regnet i nogle [*dage, øjeblikke, lande*]’.

4 METODE TIL KLASSIFICERING

4.1 Dataforbehandling

Forbehandlingen af de indsamlede data fulgte den samme metode som ved forbehandling af data i CopCo-datasættet, hvor relevante øjenbevægelseskompontener blev udtrukket fra de indsamlede eye tracking-data inden for hvert interesseområde. Interesseområder er definerede områder på et givent skærbillede, markeret som rektangler. Dette kan ses i figur 2. I dette tilfælde er de valgte områder hvert enkelt anslag i afsnittet. Markeringen af disse områder udføres i DataViewer-softwaren fra SR Research.

FIGUR 2. FIKSERINGER OPTAGET FRA HHV. IKKE-ORDBLIND LÆSER (ØVERST) OG ORDBLIND LÆSER (NEDERST) FRA DEN SAMME SÆTNING



De røde rektangler definerer interesseområder, dvs. hvert anslag (uden mellemrum) i teksten. De blå cirkler er fikseringer på de angivne områder med målte millisekunder for hver fiksering angivet over den pågældende cirkel.

4.2 Komponenter

Øjenbevægelseskomponenterne fra eye tracking-optagelserne er oprindeligt data for hvert anslag i teksten, men er blevet analyseret samlet for hvert ord i teksten. Otte komponenter blev udvalgt til dataanalysen og -klassificeringen, da disse har vist sig i høj grad at afspejle kognitive processer relateret til ordblindhed.

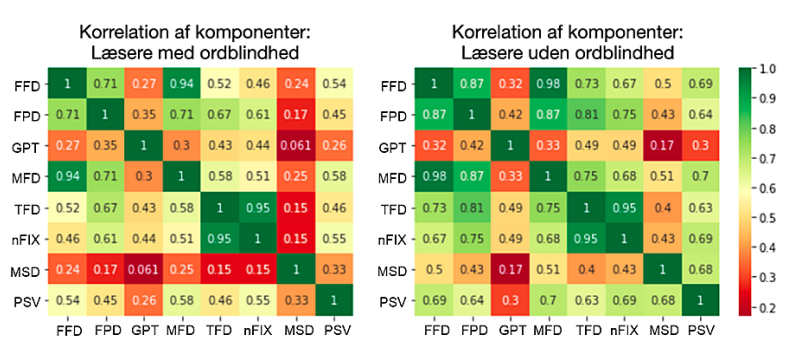
Disse komponenter er følgende:

- **FFD (*first fixation duration*)**: Varigheden af den første fiksering af det pågældende ord, målt i millisekunder.
- **FPD (*first pass duration*)**: Summen af varigheden af fikseringer på det angivne ord ved første observation af ordet, indtil der fikseres på et andet ord, målt i millisekunder. Dette vil sige, fra når læseren først fikserer på ordet, til læseren fikserer på et andet ord - dette er uafhængigt af hvilken retning læserens efterfølgende fiksering er i. Måler man FPD af ordet "stol" i sætningen "den røde stol er høj", måles der fra den første gang, læseren fikserer på et område i ordet "stol", før læseren hopper til et andet ord i sætningen.
- **GPT (*go-past time*)**: Summen af varigheden af fikseringer samt regressioner (dvs. genfikseringer) fra første fiksering på et givent ord, før der fikseres på et ord til højre for ordet, målt i

millisekunder. Dette vil sige, at fikseringer på ord til venstre for ordet også er inkluderet i varigheden, da genlæsning tælles med. I sætningen ”den røde stol er høj”, vil GPT for ”stol” inkludere regressioner på f.eks. ”den” og ”røde”, så længe at man ikke hopper videre til ”er” eller et andet ord til højre for ”stol”.

- **TFD (*total fixation duration*)**: Summen af varigheden af alle fikseringer på det pågældende ord, i millisekunder.
- **MFD (*mean fixation duration*)**: Den gennemsnitlige varighed af alle fikseringer på det pågældende ord, målt i millisekunder.
- **nFIX (*number of fixations*)**: Antallet af fikseringer på det pågældende ord.
- **MSD (*mean saccade duration*)**: Den gennemsnitlige varighed af alle sakkader, der stammer fra det pågældende ord, målt i millisekunder.
- **PSV (*peak saccade velocity*)**: Tophastighed af blik, målt i visuel grad pr. sekund, af alle sakkader, der udgår fra det pågældende ord.

FIGUR 3. KORRELATIONSMATRICER AF DE ANVENDTE ØJENBEVÆGELSESKOMPONENTER



Til venstre vises korrelationsmatrix over komponenterne fra ordblinde læsere, mens matricen til højre illustrerer komponenterne fra de ikke-ordblinde læsere. Jo tættere korrelationsscoren er på 1,0 (grøn farve), desto højere er korrelationen mellem de angivne komponenter. En lav korrelationsscore, dvs. tæt på 0 (rød farve), indikerer en lav korrelation. Komponenternes forkortelser er defineret i afsnit 4.2.

Korrelationerne af de anvendte komponenter for læsere med og uden ordblindhed er vist i figur 3. Disse illustrerer, at komponenternes korrelation generelt er højere for læsere uden ordblindhed sammenlignet med læsere

med ordblindhed. Dette indikerer, at data varierer mere blandt ordblinde læsere, hvilket tyder på, at de ordblinde læses læsemønstre omfatter en større variation. De højest korrelerede komponenter er dem, der er relateret til fikseringer, hvor de højest korrelerede er fikseringsvarighed og gennemsnitlig fikseringsvarighed (FFD og MFD på hhv. 0.94 og 0.98) samt samlet fikseringsvarighed og antal fikseringer (TFD og nFIX på hhv. 0.95 for begge). En t-test-analyse blev udført for at sammenligne komponenterne fra læsere med og uden ordblindhed, hvilket viser, at alle otte træk viser en signifikant forskel mellem grupperne ($p < 0,0001$).

4.3 Maskinlæringsmetoder

For at klassificere eye tracking-komponenterne brugte vi maskinlæringsmetoderne Support Vector Machine (SVM) og Random Forest (RF). Begge disse metoder sammenligner data, bestående af øjenbevægelseskomponenterne, for at forudsige, hvilken 'klasse' øjenbevægelsesmønstret tilhører, dvs. enten en 'ordblind' eller 'ikke-ordblind'. Dette sker ved at træne maskinlæringsmodellen med 80 % af de eksisterende data, så modellen kan analysere tendenser fra begge grupper af læsere. Herefter bliver øjenbevægelseskomponenterne for de resterende 20 % af dataene testet på modellen, hvilket foregår ved, at modellen forsøger at definere hvert tilfælde som enten tilhørende en ordblind eller ikke-ordblind læser. Dette foregår ved en sandsynlighedsberegning, som modellerne er bygget til at foretage.

5 RESULTAT

Da der er væsentligt flere data fra de ikke-ordblinde læsere end fra de ordblinde læsere, grundet et væsentligt højere antal læste tekster, blev klassifikationerne som vist i tabel 3 udført med balancerede data, dvs. samme datamængde fra begge grupper af læsere. Vi trænede to typer af maskinlærings teknikker (SVM og RF) med øjenbevægelseskomponenter bestående af både komponenter sammenlagt til sætningsniveau og komponenter sammenlagt til trial-niveau. En trial svarer til tekst præsenteret på skærmen og svarer nogenlunde til et afsnit i den originale tekst. Derudover trænede vi tre forskellige datakombinationer på hhv. sætnings- og trialniveau: a) sammenlagte gennemsnitsværdier alene, b) sammenlagte gennemsnits- og standardafvigelsesværdier og

c) sammenlagte gennemsnits-, standardafvigelses- og maksimumsværdier.

Resultaterne fra klassifikationen af disse kan ses i tabel 3. Her kan det ses, at Random Forest-modellen opnår den højeste nøjagtighed, trænet på gennemsnit, standardafvigelser og maksimumsværdier på trial-niveau-komponenter. Denne gav en nøjagtighed på 0,85. Generelt præsterede modellerne bedre med trial-niveau-data end med sætningsniveau-data, hvilket tyder på, at klassifikationen præsterer bedre med længere tekster end korte tekster.

TABEL 3. RESULTATER FRA MASKINLÆRINGSKLASSIFIKATIONEN

Maskinlæringsmodel	Trial-niveau		Sætningsniveau	
	Præcision	F1-score	Præcision	F1-score
SVM	0,80	0,80	0,72	0,71
SVM + SD	0,81	0,81	0,71	0,71
SVM + SD + Max	0,81	0,81	0,72	0,72
RF	0,83	0,83	0,72	0,72
RF + SD	0,85	0,85	0,73	0,72
RF + SD + Max	0,85	0,85	0,73	0,73

Maskinlæringsmodellerne SVM og RF alene svarer til de otte udvalgte øjenbevægelseskomponenter kun sammenlagt til gennemsnitsværdier. Her står SD for standardafvigelser, og Max står for maksimumsværdier. Dvs. kombinationer med '+ SD' for hhv. SVM og RF består både af komponenternes gennemsnitsværdier samt standardafvigelser. Ved tilføjelse af den tredje kombination for hhv. SVM og RF består kombinationen af komponenternes gennemsnitsværdier, standardafvigelser og maksimumsværdier.

5.1 Misklassifikation

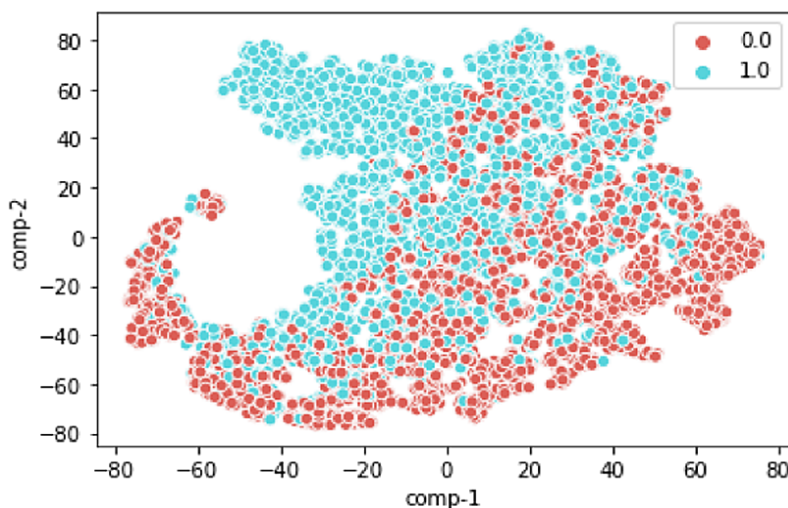
Ved en analyse af maskinlæringsmodellernes klassifikationer var det muligt at observere andelen af misklassifikationer. Misklassifikationerne er de tilfælde, hvor modellerne klassificerer øjenbevægelsesmønstrene forkert, dvs. ved at forudsige, at data fra en ordblind læser tilhører data fra en ikke-ordblind læser og omvendt. Med andre ord giver misklassifikationerne indsigt i modellernes falsk-positive og falsk-negative forudsigelser, samt om hvilke læsers læsemønstre er vanskelige at klassificere korrekt. De læsere, der havde de højeste misklassifikationsrater, var konsistent de samme læsere fra begge grupper, uanset om det gjaldt sætningsniveau eller trial-niveau. Ud fra dette analyserede vi potentielle

korrelationer mellem disse misklassifikationsrater og andre faktorer og fandt frem til, at der er en stærk korrelation mellem misklassifikationerne og læste ord pr. minut ($\rho = 0.79$, $p < 0.0001$). Derudover fandt vi en korrelation mellem misklassifikationsrater og læseforståelsesresultater ($\rho = 0.71$, $p = 0.0008$). Der var dog ikke en korrelation mellem raterne og resultaterne fra pseudohomofontesten. Dette betyder, at data fra ordblinde deltagere med hurtigere læsehastighed og højere læseforståelsesscore, sammenlignet med gennemsnittet fra de ordblinde læsere, er mere tilbøjelige til at blive kategoriseret som 'ikke-ordblind' af maskinlæringsmodellerne, da deres øjenbevægelseskomponenter minder mere om data fra ikke-ordblinde læsere.

5.2 Clusteranalyse

For at sammenligne de to grupper af læsere har vi endvidere benyttet en cluster-teknik, som svarer til gruppering af data på baggrund af flerdimensionelle, målbare ligheder og forskelle mellem datapunkter. Dette vil sige, at en model laver en gruppering af øjenbevægelseskomponenterne og deler disse op i to clusters. Dette illustreres i figur 4, hvor datapunkterne for begge læsegrupper er spredt. De blå punkter svarer til data fra ordblinde læsere, og de røde svarer til data fra ikke-ordblinde læsere.

FIGUR 4. IDENTIFICEREDE CLUSTERS FOR DE TO LÆSEGRUPPER



Figuren illustrerer to identificerede clusters. De røde punkter svarer til data fra ikke-ordblinde læsere (0.0), mens de blå svarer til data fra ordblinde læsere (1.0).

Her ses det, at der er betydeligt overlap mellem de to grupper af læsere, og der er især mange punkter tilhørende ordblinde læsere blandt punkter, hvor ikke-ordblinde læsere er i overtal. Dog er der også et bredt område med data, som udelukkende dækker ordblinde læsere. Dette viser, at selvom der er store forskelle i læsemønstrene mellem de to grupper af læsere, er der også mange data fra ordblinde læsere, som er nærmest identiske med data fra ikke-ordblinde læsere.

5.3 Resultat fra de leksikale vurderingstests

For at skelne de forskelle mellem de to grupper af læsere, der rækker ud over eye tracking-komponenter, blev der udført en analyse af resultaterne fra de to leksikalske vurderingstests. Pseudohomofontesten blev udført af alle ordblinde deltagere i forsøget og sammenlignet med standarder fra det officielle testmateriale angående grænseværdier. Denne sammenligning er illustreret i tabel 4. Tabellen viser, at scorerne fra læsere med ordblindhed i den pågældende undersøgelse svarer til de scorer, der definerer gruppen ”voksne på læsekursus”. Selvom scorerne fra den pågældende undersøgelse er højere end de standarder, der er registreret for personer, der oplever leksikale eller fonologiske vanskeligheder, svarer scorerne for alle percentiler til omkring halvdelen af standarderne fra voksne uden læsevanskeligheder.

TABEL 4. STANDARDER FOR PSEUDOHOMOFONTESTEN

Læsegruppe	Antal	Nøjagtighed
Voksne uden læsevanskeligheder	72	66 %
Voksne i ordblindeundervisning	46	23 %
Voksne læsekursister	167	31 %
Læsere fra denne undersøgelse	18	33 %

Antal svarer til antal læsere i den pågældende læsegruppe. *Nøjagtigheden* er andelen af antal korrekte svar ud af 38 delopgaver. De tre øverste rækker er angivet i den officielle testdokumentation fra Center for Læseforskning under Københavns Universitet (https://laes.hum.ku.dk/test/find_der_der_lyder_som_et_ord/standarder/).

Læseforståelsestesten blev udført af alle deltagerne med ordblindhed og ti af deltagerne uden ordblindhed. For at analysere disse resultater blev der udført en dobbeltsidet variansanalyse (t-test), som er en statistisk

vurdering af, hvorvidt der er signifikant forskel mellem standarderne fra de ordblind og de ikke-ordblind læsere. Dette resulterede i en signifikant forskel, $p < 0,0001$, mellem de scorer, der blev opnået af de to grupper af læsere. En score på 0-3,4 indikerer, at læseren kan opleve, at mange tekster er svære og tidskrævende at læse, og en score på 3,5-3,9 indikerer, at læseren kan finde nogle tekster svære og/eller tidskrævende at læse. Scorer på 4 og derover indikerer gode læsefærdigheder. Den gennemsnitlige score for gruppen af læsere med ordblindhed er 3,5, mens den gennemsnitlige score for læsere uden ordblindhed er 5,7. Disse forskelle er også synlige i minimums- og maksimumsværdierne, hvor minimumsværdien for læserne uden ordblindhed er 4,4, og dermed ligger den inden for kriterierne for ”gode læsefærdigheder”. Scorerne for læsere med ordblindhed spænder fra 0,7 til 5,2, hvilket indikerer stor variation i dataene.

6 DISKUSSION OG KONKLUSION

Det specifikke formål med denne undersøgelse var at bygge maskinlæringsmodeller til klassificering af eye tracking-data optaget fra ordblind og ikke-ordblind læsere. Dette er, så vidt vi ved, det første forsøg på at påvise ordblindhed ud fra eye tracking-komponenter med dansk som målsprog. Resultaterne indikerer, at ordblind læsere med højere læsehastighed samt højere læseforståelsesresultater har højere sandsynlighed for at blive misklassificeret, da deres øjenbevægelsesmønstre under læsning minder mere om øjenbevægelsesmønstrene for læsere uden ordblindhed. Manglende korrelation mellem misklassifikationer og resultater fra pseudohomofontesten tyder på, at den sidstnævnte muligvis kan være mere effektiv i forhold til at identificere de tendenser, der er karakteristiske for ordblind læsere. Ud fra dette lader de anvendte deltests i Ordblindetesten, som fokuserer på fonologiske afkodningsopgaver, til at være meget effektive. Dog er det vigtigt at gøre opmærksom på, at de ordblind deltagere i denne undersøgelse er diagnosticeret som ordblind på grundlag af Ordblindetesten, hvilket betyder, at standarderne og definitionerne af ordblindhed i denne undersøgelse er definerede af standarderne fra Ordblindetesten. Dette fører videre til det grundlæggende spørgsmål: Hvordan sættes standarderne for definitionen af ordblindhed?

En pålidelig vurderingsmetode spiller en vigtig rolle for personer med ordblindhed – ikke blot for screening for ordblindhed, men også for disse individers generelle velbefindende. En af grundene til dette er, at en officiel diagnose kan være adgangsgivende til de rigtige værktøjer og assistance til uddannelsesmæssige, professionelle og dagligdags formål. Denne artikel har redegjort for nødvendigheden af videreudvikling af de eksisterende metoder og understreget behovet for yderligere forskning og undersøgelser af ordblindhed. Ud fra de anvendte vurderingstests i denne undersøgelse viser det sig, at pseudohomofontests er et værktøj med høj pålidelighed samt en god evne til at skelne ordblinde læsere fra ikke-ordblinde læsere. Dog er der blevet argumenteret for, at testen med fordel kan suppleres med yderligere opgaver eller protokoller, og dette gælder især, fordi de vanskeligheder, som opleves af ordblinde personer, kan variere meget og opleves forskelligt.

I denne undersøgelse opnåede metoden med maskinlæringsklassifikation en nøjagtighed på 85 %, hvilket betyder, at metoden ikke alene kan benyttes som et screeningsværktøj, men at den opdager signifikante og målbare forskelle hos ordblinde og ikke-ordblinde læsere, som ultimativt kan tænkes ind i udredningsprocessen for ordblindhed. Eye tracking-metoden giver dermed ikke blot indsigt i kognitive processer såsom læsevanskeligheder, men åbner også mulighed for dynamiske og målbare processer. Dette indebærer mulighed for at teste læsning under forhold, der svarer til dem, der naturligt forekommer i det daglige liv. Dette kan hjælpe med til at give en mere nuanceret forståelse af læseprocesser hos personer med ordblindhed sammenlignet med, hvad vi kan opnå ved opgaver, der består af afkodning af nonsensord og pseudohomofoner eller ordlister. Læsning er en kompliceret, dynamisk proces, og herunder udgør fonem-grafem-forbindelser blot en del af de underliggende processer.

Den tekniske tilgang i denne undersøgelse ved brug af maskinlæringsklassifikation for påvisning af ordblindhed ud fra øjenbevægelsesmønstre kan bidrage til at udvide vores forståelse af læsemønstre hos ordblinde læsere, såvel som til at fremme teknologier, der kan være nyttige til at opdage ordblindhed. Præcise kriterier for ordblindhed er stadig vanskelige at standardisere grundet de forskellige grader af diagnosens symptomer og indikatorer, hvorfor ordblindhed fortje-

ner mere forskningsmæssig opmærksomhed. Eye tracking kan være et springbræt til at opnå mere pålidelige screeningsmetoder for ordblindhed.

Marina Björnsdóttir
Center for Sprogteknologi (CST)
Københavns Universitet
marina.bjorns@gmail.com

Nora Hollenstein
Center for Sprogteknologi (CST)
Københavns Universitet
nora.hollenstein@hum.ku.dk

Ethan Weed
Afdeling for Lingvistik, Kognitionsvidenskab og Semiotik
Aarhus Universitet
ethan@au.dk

LITTERATUR

- Asvestopoulou, T. et al. 2019. DysLexML: Screening Tool for Dyslexia Using Machine Learning. DOI: 10.48550/arXiv.1903.06274.
- Bellocchi, S. et al. 2013. I can read it in your eyes: What eye movements tell us about visuo-attentional processes in developmental dyslexia. *Research in Developmental Disabilities* 34(1). 452–460. DOI: 10.1016/j.ridd.2012.09.002.
- Biscaldi, M., Gezeck, S., & Stuhr, V. 1998. Poor saccadic control correlates with dyslexia. *Neuropsychologia* 36(11). 1189–1202. DOI: 10.1016/S0028-3932(97)00170-X.
- Björnsson, C. 1968. Läsbarhet. Stockholm, Sverige: Liber.
- Børne- og Undervisningsministeriet. (2021). *Systematisk test af voksne ordblinde halter efter børn og unge*. Børne- og Undervisningsministeriet. 5. oktober 2021. <https://www.uvm.dk/aktuelt/nyheder/uvm/2021/okt/211005-systematisk-test-af-voksne-ordblinde-halter-efter-boern-og-unge> (tilgået 20. december 2022).

- Center for Læseforskning, Københavns Universitet, Institut for Uddannelse og Pædagogik, Socialstyrelsen, & Undervisningsministeriet. 2020. Vejledning til Ordblindetesten (version 8). Børne- og Undervisningsministeriet.
- De Luca, M. et al. 1999. Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia* 37(12). 1407–1420. DOI: 10.1016/S0028-3932(99)00038-X.
- Eckert, M. 2004. Neuroanatomical markers for dyslexia: a review of dyslexia structural imaging studies. *The Neuroscientist* 10(4). 362–371. DOI: 10.1177/1073858404263596.
- Fischer, B., & Weber, H. 1990. Saccadic reaction times of dyslexic and age-matched normal subjects. *Perception* 19(6). 805–818. DOI: 10.1068/p190805.
- Henderson, J. M. 2013. Eye movements. D. Reisberg (red.), *The Oxford Handbook of Cognitive Psychology*, 69–82. Oxford University Press. DOI: 10.1093/oxford-hb/9780195376746.013.0005.
- Hollenstein, N., Barrett, M., & Björnsdóttir, M. 2022. The Copenhagen Corpus of Eye Tracking Recordings from Natural Reading of Danish Texts. DOI: 10.48550/arXiv.2204.13311.
- Hyönä, J., & Olson, R. K. 1995. Eye fixation patterns among dyslexic and normal readers: effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(6). 1430–1440. DOI: 10.1037/0278-7393.21.6.1430.
- Jensen, K. L., Gellert, A. S., & Elbro, C. 2014. Rapport om udvikling og afprøvning af selvtest af læsning – en selvtest af voksnes læsefærdigheder på nettet. Center for Læseforskning, Københavns Universitet.
- Kaisar, S. 2020. Developmental dyslexia detection using machine learning techniques: A survey. *ICT Express* 6(3). 181–184. DOI: 10.1016/j.icte.2020.05.006.
- Nergård-Nilssen, T., & Eklund, K. 2018. Evaluation of the psychometric properties of “the Norwegian screening test for dyslexia”. *Dyslexia* 24(3). 250–262.
- Nilsson Benfatto, M. et al. 2016. Screening for dyslexia using eye tracking during reading. *PLOS ONE* 11(12). DOI: 10.1371/journal.pone.0165508.
- Ottosen, H. F. et al. 2022. Identifying dyslexia at the university: assessing phonological coding is not enough. *Annals of Dyslexia* 72(1). 147–170.
- Perera, H., Shiratuddin, M.F., & Wong, K.W. 2018. Review of EEG-based pattern classification frameworks for dyslexia. *Brain Informatics* 5(2). 1–14. DOI: 10.1186/s40708-018-0079-9.

- Pirozzolo, F.J., & Rayner, K. 1979. The neural control of eye movements in acquired and developmental reading disorders. *Studies in Neurolinguistics*. 97–123. DOI: 10.1016/B978-0-12-746304-9.50009-4.
- Plużyczka, M. 2018. The first hundred years: A history of eye tracking as a research method. *Applied Linguistics Papers* 25/4. 101–116. DOI: 10.32612/uw.25449354.2018.4.pp.101-116.
- Prabha, A. J., & Bhargavi, R. 2020. Predictive model for dyslexia from fixations and saccadic eye movement events. *Computer Methods and Programs in Biomedicine* 195. DOI: 10.1016/j.cmpb.2020.105538
- Raatikainen, P. et al. 2021. Detection of developmental dyslexia with machine learning using eye movement data. *Array* 12. DOI: 10.1016/j.array.2021.100087.
- Rauschenberger, M. et al. 2017. Towards the prediction of dyslexia by a web-based game with musical elements. *Proceedings of the 14th International Web for All Conference*. 1–4. DOI: 10.1145/3058555.3058565.
- Rayner, K. 1986. Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology* 41(2). 211–236. DOI: 10.1016/0022-0965(86)90037-8.
- Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3). 372–422. DOI: 10.1037/0033-2909.124.3.372.
- Rello, L., & Ballesteros, M. 2015. Detecting readers with dyslexia using machine learning with eye tracking measures. *Proceedings of the 12th International Web for All Conference*. 1–8. DOI: 10.1145/2745555.2746644.
- Rubino, C., & Minden, H. 1973. An analysis of eye-movements in children with a reading disability. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*. DOI: 10.1016/S0010-9452(73)80030-9.
- Schulte-Körne, G. 2010. The prevention, diagnosis, and treatment of dyslexia. *Deutsches Ärzteblatt International* 107(41)- 718–726. DOI: 10.3238/arztebl.2010.0718.
- Singleton, C., Horne, J., & Simmons, F. 2009. Computerised screening for dyslexia in adults. *Journal of Research in Reading* 32(1). 137–152. DOI: 10.1111/j.1467-9817.2008.01386.x.

- Smyrnakis, I. 2017. Radar: A novel fast-screening method for reading difficulties with special focus on dyslexia. *PLOS ONE* 12(8). DOI: 10.1371/journal.pone.0182597.
- Stanek, H. 2021. Kritiske læsevejledere: National test fanger ikke alle ordblinde blandt de yngste elever. *Folkeskolen*. 10. december 2021. <https://www.folkeskolen.dk/danskundervisning-ordblindlaesesvag-skolepolitik/kritiske-laesevejledere-national-test-fanger-ikke-alle-ordblinde-blandt-de-yngste-elever/2071880> (tilgået 20. december 2023).